Bayesian estimation of self-similarity exponent

DISSERTATION
ZUR ERLANGUNG DES AKADEMISCHEN GRADES
DOKTOR DER NATURWISSENSCHAFTEN (DR. RER. NAT.)
IN DER WISSENSCHAFTSDISZIPLIN NICHTLINEARE DYNAMIK

NATALLIA MAKARAVA

INSTITUT FÜR PHYSIK
FAKULTÄT MATHEMATIK UND NATURWISSENSCHAFTEN
UNIVERSITÄT POTSDAM

2012

Published online at the
Institutional Repository of the University of Potsdam:
URL http://opus.kobv.de/ubp/volltexte/2013/6409/
URN urn:nbn:de:kobv:517-opus-64099
http://nbn-resolving.de/urn:nbn:de:kobv:517-opus-64099

Contents

Kurzfassung

Die Abschätzung des Selbstähnlichkeitsexponenten hat in den letzten Jahrzehnten an Aufmerksamkeit gewonnen und ist in vielen wissenschaftlichen Gebieten und Disziplinen zu einem intensiven Forschungsthema geworden. Reelle Daten, die selbsähnliches Verhalten zeigen und/oder durch den Selbstähnlichkeitsexponenten (insbesondere durch den Hurst-Exponenten) parametrisiert werden, wurden in verschiedenen Gebieten gesammelt, die von Finanzwissenschaften über Humanwissenschaften bis zu Netzwerken in der Hydrologie und dem Verkehr reichen. Diese reiche Anzahl an möglichen Anwendungen verlangt von Forschern, neue Methoden zu entwickeln, um den Selbstähnlichkeitsexponenten abzuschätzen, sowie großskalige Abhängigkeiten zu erkennen.

In dieser Arbeit stelle ich die Bayessche Schätzung des Hurst-Exponenten vor. Im Unterschied zu früheren Methoden, erlaubt die Bayessche Herangehensweise die Berechnung von Punktschätzungen zusammen mit Konfidenzintervallen, was von bedeutendem Vorteil in der Datenanalyse ist, wie in der Arbeit diskutiert wird. Zudem ist diese Methode anwendbar auf kurze und unregelmäßig verteilte Datensätze, wodurch die Auswahl der möglichen Anwendung, wo der Hurst-Exponent geschätzt werden soll, stark erweitert wird. Unter Berücksichtigung der Tatsache, dass der Gauß'sche selbstähnliche Prozess von bedeutender Interesse in der Modellierung ist, werden in dieser Arbeit Realisierungen der Prozesse der fraktionalen Brown'schen Bewegung und des fraktionalen Gauß'schen Rauschens untersucht. Zusätzlich werden Anwendungen auf reelle Daten, wie Wasserstände des Nil und fixierte Augenbewegungen, diskutiert.

Abstract

Estimation of the self-similarity exponent has attracted growing interest in recent decades and became a research subject in various fields and disciplines. Real-world data exhibiting self-similar behavior and/or parametrized by self-similarity exponent (in particular Hurst exponent) have been collected in different fields ranging from finance and human sciencies to hydrologic and traffic networks. Such rich classes of possible applications obligates researchers to investigate qualitatively new methods for estimation of the self-similarity exponent as well as identification of long-range dependencies (or long memory).

In this thesis I present the Bayesian estimation of the Hurst exponent. In contrast to previous methods, the Bayesian approach allows the possibility to calculate the point estimator and confidence intervals at the same time, bringing significant advantages in data-analysis as discussed in this thesis. Moreover, it is also applicable to short data and unevenly sampled data, thus broadening the range of systems where the estimation of the Hurst exponent is possible. Taking into account that one of the substantial classes of great interest in modeling is the class of Gaussian self-similar processes, this thesis considers the realizations of the processes of fractional Brownian motion and fractional Gaussian noise. Additionally, applications to real-world data, such as the data of water level of the Nile River and fixational eye movements are also discussed.

Chapter 1

Introduction

In 1951, in the work *"Long-term storage capacity of reservoirs"* [62], Hurst reported for the first time his studies of the hydrologic data. There, he observed the property of the data to be dependent on the scalings in time and space. To estimate the parameter which measures the intensity of this type of scaling, he proposed the method called rescaled range statistical analysis (R/S analysis). Nowadays, this parameter is known as the Hurst exponent (H).

In the last two decades, a significant amount of techniques and methods have been proposed in the literature for robust and accurate estimation of the Hurst exponent. Depending on the domain they work in, these methods can be divided into:

– *time domain:* the R/S analysis [88], the modified R/S analysis [79], the detrended fluctuation analysis [105, 106] and its variants [13, 25, 68, 69, 129], the detrended moving average analysis [9, 57], the aggregated variance method [117], *etc.*;

– *frequency (spectral) domain:* the periodogram method [50], the Whittle method [128], *etc.*;

– *wavelet domain (time-scale methods):* the wavelet-based estimator algorithm [12, 48, 116], the Abry-Veitch Daubechies wavelet-based estimator [4, 6], the wavelet maximum likelihood method [93], *etc*.

Some of the available methods also allow the estimation of the confidence intervals for the Hurst exponent [23, 24, 33, 70], whereas others propose a new and radically different approach to the estimation problem (*e.g.* [72]). In order to identify the proper estimator for particular data sets, large number of comparison studies can be found in the literature: *i.e.* Bardet *et al.* [14] presented the results of the comparison between the wavelet and the Whittle estimators, and between the log-periodogram and quadratic variations estimators. The comparison between eight different techniques for the evaluation of the Hurst exponent based on the fractional Brownian motion, Gaussian white noise and fractional autoregressive integrated moving-average model can be found in work of Esposti *et al.* [45].

This significantly high research interest in the estimation of the Hurst exponent can be explained by the fact that many data sets gathered from real-world models have self-similar properties. Examples include geologic [55], hydrologic [121] and finance [15, 17, 92, 107] data; data from human sciences [53, 60, 64], traffic networks [32, 75], turbulence [51, 63], DNA sequences [10] and other data types.

However, most of the existing methods elaborate either a point estimator of the Hurst exponent, or a confidence interval indicating the Hurst estimates. In order to overcome this problem, I propose in this thesis to approach the estimation task from another perspective: I use Bayesian inference, whose flexibility and generality in dealing with very complex problems is well established. The direct quantification of uncertainty, the possibility for coping with multiple parameters and the natural way to provide a degree of belief about the assumed model, while it yields to associated confidence intervals, makes the Bayesian approach effective and practical, and allows to develop an estimation method that provides as outcome a robust and accurate estimator, characterized by an increased efficiency in comparison to other modern methods. The idea was initially considered in the work of Conti *et al.* [34] for the estimation of the Hurst exponent in telecom networks. However, in contrast to this, I propose a more general model for the estimation of the Hurst exponent in terms of a linear mixed model, where the random part is represented by a fractional Brownian motion [82].

The choice of the model type used in this study lies in the possibility to model real-world data using a fractional Brownian motion (fBm). These properties can be asserted for many processes in biology [28, 30], biomechanics [31], eye-movement data [42], aggregated traffic [101] and number of other applications. Besides this, taking into account that "the existence of trends in time series generated by physical and biological systems is so common that it is almost unavoidable" [61], I consider a linear trend in model. Even though the current analysis is limited to a linear trend, it is possible to extend the introduced technique to more general trends, such as the seasonal variations or higher order polynomials. Therefore, I propose to keep the global trend in a joint inversion for all parameters in the model instead of removing it on the first step. Although I am mainly interested in the estimation of the Hurst exponent, the technique presented in this thesis is suitable for estimation of the remaining parameters of the model as well.

The theoretical basis of self-similar processes and Bayesian theory essential to the current study I overview in *Chapter 2*. A mathematical representation of the model is given in *Chapter 3*. Here, I synthesize data sets of fractional Brownian motion that can be sampled evenly as well as unevenly, present numerical results and a comparison of the Bayesian method with the detrended fluctuation analysis. The applicability of the estimation method for non-Gaussian data from Rosenblatt processes and α-stable Lévy motion is also shown [84].

The real-world data is in most of the cases stochastic: they contain almost always a noise component which is inherent for the measurement process. The influence of this measurement noise on the estimation process of the Hurst exponent is the central topic of *Chapter 4*. I evaluate the robustness of the Hurst exponent estimator based on the Bayesian approach, considering the noise contained in the data. I provide a possibility to determine the amount of the noise according to the amplitude of the data which allows correct estimation of the Hurst exponent. I show that even in the case of an explicit consideration of noise in the model which allows for a better estimation of the internal parameters, the percentage ratio between noise and amplitude should not be more than 5%.

From another viewpoint, fractional Gaussian noise is often found to explain quite well the heterogeneity and the multiscale properties of geophysical time series, and in particular those from hydrology. Thus instead of using the integrated time series and then applying an estimator for the Hurst exponent, I propose to use the noise signal directly [18]. I extend the Bayesian approach in *Chapter 5* to the stationary process of fractional Gaussian noise. This kind of model allows to discriminate between the short and long term dependencies. This difference has observable consequences since, for instance, a long memory process explain the patterns observed in sedimentary deposits in river runoff areas (see for e.g. [97]). For this reason it is important to devise the proposed technique to estimate the underlying self-similarity exponent and to provide a method to quantify the uncertainty of the measurements. A direct application of this method is done to the data of the annual minimum water level of the Nile River in *Section 5.4*.

Engbert and Kliegl [42] showed recently that realizations of a fractional Brownian motion have properties similar to those observed in fixational eye movements. On the other hand, the Hurst exponent which parameterizes the fractional Brownian motion, indicates the correlation behavior and allows the detection of the long-range dependencies inside fixational eye movements. This type of data has been investigated so far using the approaches such as detrended fluctuation and standard deviation analysis [31] in the work of Mergenthaler and Engbert [95], where a change of the scaling behavior from a persistent behavior on short time scales to an anti-persistent behavior on long time scales was observed. I investigate the estimation problem of the Hurst exponent for fixational eye movements data based on a sequence of increments of a discrete process over some scale in *Chapter 6*. This allows to examine the value of the Hurst exponent on different levels according to the scale factor. With this approach, the Hurst exponent remains constant for a sampled sequence of a fractional Brownian motion at all scale levels. In order to test this property, I examine both experimental and simulated data of fixational eye movements [83]. In the first case, I show the existence of the scale level starting from which the data acquires an anti-persistent behavior and the signal turns into a short-memory process. The invariance of the Hurst exponent for different trials for

the same participant will be shown. I also validate the proposed method on simulated time series, mimicking the fixational eye movements based on a time-delayed model proposed in [95] and an integrated model introduced in [44]. I speculate that the integrated model describes data more closely, since the obtained results show qualitatively the same behavior.

Chapter 2

Theoretical background

> It is the theory that decides what can be observed.
>
> Albert Einstein

The estimation of the self-similarity exponent from observations which is the subject of the study in the next chapters is based on the Bayesian theory. In this chapter I discuss the basic characteristics of self-similar processes and their relation to long-range dependency. I further focus on the fractional Brownian motion and its basic properties, since it is one of the most important self-similar process. The fundamental concepts of the Bayesian theory are overviewed as well.

2.1 Self-similar processes

I consider the probability space $(\Omega, \mathcal{F}, \mathbb{P})$, where Ω is the sample space of all possible outcomes, $\mathcal{F}$ is a σ-algebra of the measurable subsets of Ω, whose elements are the events for which it is possible to obtain information, and a probability measure $\mathbb{P} : \mathcal{F} \to [0, 1]$, where $0 \leq \mathbb{P}(A) \leq 1$ is the probability that the event $A \in \mathcal{F}$ occurs. In this definition, $\mathcal{F}$ is a σ-algebra on Ω if it is a collection of subsets of Ω such that $\emptyset$ and Ω belong to $\mathcal{F}$, the complement of a set in $\mathcal{F}$ belongs to $\mathcal{F}$, and a countable union intersection of sets in $\mathcal{F}$ belongs to $\mathcal{F}$. For the probability measure $\mathbb{P}$ on $\mathcal{F}$ the following statements hold true: $\mathbb{P}(\emptyset) = 0, \mathbb{P}(\Omega) = 1$, and for any sequence A_n of the pairwise disjoint sets (meaning that $A_i \cap A_j = \emptyset$ for $i \neq j$)

$$\mathbb{P}\left(\cup_{n=1}^{\infty} A_n\right) = \lim \sum_{n=1}^{\infty} \mathbb{P}(A_n). \tag{2.1}$$

A function $X : \Omega \to \mathbb{R}$ with the property that $\{\omega \in \Omega : X(\omega) \leq x\} \in \mathcal{F}$ for each $x \in \mathbb{R}$ is said to be $\mathcal{F}$-measurable and known as a random variable. Its distribution function $F : \mathbb{R} \to [0, 1]$ is given by $F(x) = \mathbb{P}(X \leq x)$.

A stochastic process is a family of random variables $\{X(t,\omega), t \in T, \omega \in \Omega\}$ defined on a common probability space $(\Omega, \mathcal{F}, \mathbb{P})$, indexed by the time variable t, where t varies over an index set T [74]. Since I consider the stochastic process as a collection of random numbers, meaning that $t \in [0,\infty)$ is fixed, in the following I simplify this notation to $\{X(t), t \in T\}$. The index t represents the time, whereas we refer to $X(t)$ as the "state" of the process at time t. Therefore, the random variables $X(t)$ take values in a set which is called the state space. In this thesis I use the so-called real-valued processes with the state space $\mathbb{R}$.

There are two different ways in which stochastic process $\{X(t), t \in T\}$ can be formulated:

• *Analytical definition:* Let $X(t) = \phi(t, A_1, A_2, \ldots, A_k)$, where ϕ is a function of t and a set of random variables $A_1, A_2, \ldots, A_k$ $(k \geq 1)$ defined on a probability space independent on t, and ϕ is a random variable on this space.

• *Definition in terms of finite-dimensional distributions:* Define $\{X(t), t \in T\}$ by specifying the joint distribution functions:

$$F(x_1, x_2, \ldots, x_n; t_1, t_2, \ldots, t_n) = \mathbb{P}\{X(t_1) \leq x_1, X(t_2) \leq x_2, \ldots X(t_n) \leq x_n\} \tag{2.2}$$

for $t_1, t_2, \ldots, t_n \in T$, and $n \geq 1$. These joint distribution functions satisfy the following consistency conditions:
a) symmetry: for every permutation $(j_1, j_2, \ldots, j_n)$ of $(1, 2, \ldots, n)$

$$F(x_{j_1}, x_{j_2}, \ldots, x_{j_n}; t_{j_1}, t_{j_2}, \ldots, t_{j_n}) = F(x_1, x_2, \ldots, x_n; t_1, t_2, \ldots, t_n). \tag{2.3}$$

b) compatibility: for $m < n$

$$\begin{aligned} F(x_1, x_2, \ldots, x_m, \infty, \infty, \ldots, \infty; t_1, t_2, \ldots, t_n) \\ = F(x_1, x_2, \ldots, x_m; t_1, t_2, \ldots, t_m). \end{aligned} \tag{2.4}$$

A real-valued stochastic process $\{X(t), t \in T\}$ is self-similar with parameter $H \in]0,1[$ if, for any $a > 0$, the finite-dimensional distributions of process $\{X(at), t \in T\}$ are identical to the finite-dimensional distributions of $\{a^H X(t), t \in T\}$, *i.e.* for any $a > 0$:

$$\{X(at), t \in T\} =_d \{a^H X(t), t \in T\}, \tag{2.5}$$

where the sign $=_d$ denotes the equivalence of the finite-dimensional distributions [115]:

$$\mathbb{P}(X(at_0) \leq x_0, \ldots, X(at_n) \leq x_n) = \mathbb{P}(a^H X(t_0) \leq x_0, \ldots, a^H X(t_n) \leq x_n) \tag{2.6}$$

for every $t_0, \ldots, t_n$ and every $x_0, \ldots, x_n$ in T. Here, T could be discrete or a continuum like $T = \mathbb{R}, [0,\infty)$.

Thus, from the mathematical aspect, a more correct term for these processes would be "statistical self-similar" processes. However, the term "self-similar" became well-spread and I restrict in the thesis to this terminology.

Next, I list some basic properties of self-similar processes [58]:

• The value H is called Hurst exponent.

• Lamperti [73] has shown that the Hurst exponent is unique:

Theorem 2.1.1. *Let $\{X(t), t \in [0, \infty)\}$ be a self-similar non trivial process (i.e. not almost surely constant), whose paths are stochastic continuous at $t = 0$ (i.e. $\mathbb{P}(|X_h - X_0| \geq \epsilon) \to 0$ for all $\epsilon > 0$ and $h \to 0$), then there is exactly one $H \geq 0$.*

• Equality in distribution expresses the invariance of scales and not any change of the corresponding paths.

• Self-similarity means that zooming in the dimension with a factor a implies an enlarging of the space vector by a factor a^H.

• Choosing $t_0 = 0$ and $H > 0$ the relation $X(a0) = a^H X(0)$ reveals that $X(0) = 0$ almost surely. Under the assumption of the Theorem 2.1.1 the stronger property holds:

$$H = 0 \Leftrightarrow X(t) = X(0) \text{ almost surely.} \tag{2.7}$$

Therefore, the self-similar processes under consideration are the processes with $H > 0$, whose paths are stochastic continuous at 0.

• The process $\{X(t), t \in T\}$ has stationary increments if $\{X(t), t \in T\}$ and $\{X(t + h) - X(h), t \in T\}$ have the same finite dimensional distributions for every $h > 0$. Self-similar process with stationary increments and existing first moment satisfies for all $t \geq 0$

$$2^H \mathbb{E}(X(t)) = \mathbb{E}(X(2t)) = \mathbb{E}(X(2t) - X(t)) + \mathbb{E}(X(t)) = 2\mathbb{E}(X(t)). \tag{2.8}$$

Thus $\mathbb{E}(X(t)) = 0$ for all $t \in T$.

• If a self-similar process with stationary increments has s finite second moment, then

$$\mathbb{E}(X^2(t)) = t^{2H}\sigma^2. \tag{2.9}$$

When $\sigma^2 = 1$, the process $\{X(t), t \in T\}$ is called a standard process.

2.1.1 Fractional Brownian motion

Gaussian processes are one of the richest and best understood classes of processes in probability theory. A process is Gaussian if all of its finite-dimensional distributions are multivariate Gaussian distributions. A fractional Brownian motion (fBm) is the only one self-similar process with Hurst exponent $H \in]0, 1[$ in the class of Gaussian processes. It was initially implicitly considered by Kolmogorov in his study of the "spirals of Wiener" [71], whereas its definition and properties were later reported by Mandelbrot and Van Ness in [87].

A fractional Brownian motion $\{B_t^H, t \in \mathbb{R}\}$ is a continuous and centered Gaussian process for which the covariance at time points t and u is given by the explicit expression:

$$\mathbb{E}(B_t^H B_u^H) = \frac{\sigma^2}{2}(|t|^{2H} + |u|^{2H} - |t-u|^{2H}), \tag{2.10}$$

where $\sigma^2 = VarB_1^H$. It is called a standard fBm if $\sigma^2 = 1$. Moreover, since the covariance function in **Eq.** 2.10 is homogeneous of order $2H$, fBm $\{B_t^H, t \in \mathbb{R}\}$ is a self-similar process with the Hurst exponent H. All self-similar Gaussian processes parameterized by the Hurst exponent have the covariance function **Eq.** 2.10 and differ only by a multiplicative constant.

Thus fractional Brownian motion $\{B_t^H, t \in \mathbb{R}\}$ is a zero mean Gaussian process that starts at zero:

$$\mathbb{E}(B_t^H) = 0, \quad B_0^H = 0, \tag{2.11}$$

and has homogeneous increments, *i.e.* its increments $B_{t+s}^H - B_t^H$ constitute a stationary process with $s \neq 0$. That means that its joint probability distribution does not change when shifted in time or space:

$$B_{t+s}^H - B_s^H =_d B_t^H - B_0^H. \tag{2.12}$$

The paths of the fBm have the following properties [119]:

i) The fBm $\{B_t^H, t \in \mathbb{R}\}$ admits a version whose sample paths are almost surely Hölder continuous of order strictly less than H, *i.e.* there exist constants $\alpha \geq 1$, $\beta \geq 0$ and $k \geq 0$ such that

$$\mathbb{E}[|B_t^H - B_s^H|^\alpha] \leq k|t-s|^{1+\beta} \tag{2.13}$$

for all $s, t \in \mathbb{R}$.

ii) The fBm sample path is not differentiable. In fact, for every $t_0 \in [0, \infty)$

$$\limsup_{t \to t_0} |\frac{B_t^H - B_{t_0}^H}{t - t_0}| \to \infty \tag{2.14}$$

with probability one.

The Hurst exponent determines the properties of the observed process: higher values of H indicate a smoother behavior, less volatility, and less roughness so the paths of the fBm get less zigzagged as the Hurst exponent goes from 0 to 1. It is commonly agreed to distinguish H-parameterized processes into three classes:

• for $H \in]0, \frac{1}{2}[$ the behavior of the process can be described as an anti-persistent behavior;

• for $H = \frac{1}{2}$ the process turns into usual Brownian motion since **Eq.** 2.10 reduces to:

$$\mathbb{E}(B_t^{\frac{1}{2}} B_u^{\frac{1}{2}}) = \begin{cases} min(t,u) & \text{if } t \text{ and } u \text{ have the same signs,} \\ 0 & \text{if } t \text{ and } u \text{ have opposite signs;} \end{cases} \tag{2.15}$$

• for $H \in]\frac{1}{2}, 1[$ the behavior of the process can be described as a persistent behavior.

The anti-persistence in time exists, in the sense that the increments of the process are negatively correlated with correlations quickly decaying to zero, and the process becomes a so-called short-memory process:

$$\sum_{n=1}^{\infty} \mathbb{E}[|(B_1^H - B_0^H)(B_{n+1}^H - B_n^H)|] < \infty. \tag{2.16}$$

In contrast to this, the persistent behavior is characterized by positive covariance of two consecutive increments and occurrence of long-range dependencies [114, 115]. The increment sequence $B_{n+1}^H - B_n^H$ exhibits a long-range dependence if its covariance function satisfies

$$\lim_{n\to\infty} \frac{\mathbb{E}[(B_1^H - B_0^H)(B_{n+1}^H - B_n^H)]}{cn^{-\gamma}} = 1 \tag{2.17}$$

for some constant c and $\gamma \in (0,1)$. In this case, the dependence between $B_1^H - B_0^H$ and $B_{n+1}^H - B_n^H$ decays slowly as n tends to infinity and

$$\sum_{n=1}^{\infty} \mathbb{E}[(B_1^H - B_0^H)(B_{n+1}^H - B_n^H)] = \infty. \tag{2.18}$$

The extreme case of $H = 1$ is usually eliminated due to the fact that for $H = 1$, fBm corresponds merely to a line ($B_t^1 = tB_1^1$) with random slope B_1^1.

The difference in the correlation structure is visible in **Fig.** 2.1, where the different realizations of fBm of length $N = 500$ for $H = 0.05$, $H = 0.5$ and $H = 0.95$ are shown.

In general, the Hurst exponent is directly related to the Hausdorff dimension. Next, I discuss and characterize this relationship in more details. The traditional definition of the Hausdorff dimension D by Falconer [46] is: for any given set $F \subset \mathbb{R}^n$ and non-empty subset $U \subset \mathbb{R}^n$

$$D = dimF = \inf\{s \geq 0, \mathcal{H}^s(F) = 0\} = \sup\{s : \mathcal{H}^s(F) = \infty\}, \tag{2.19}$$

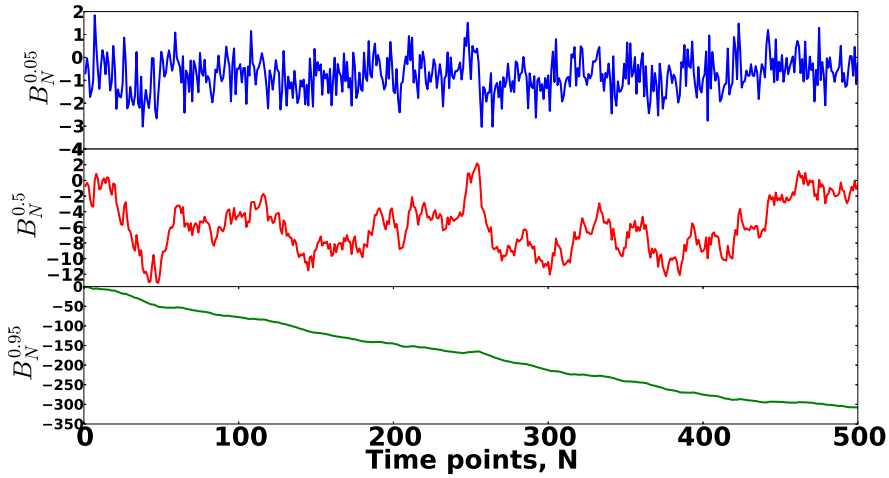

Figure 2.1: Simulation of fractional Brownian motion of length $N = 500$ with Hurst exponent $H = 0.05$ (blue line), $H = 0.5$ (red line) and $H = 0.95$ (green line).

where

$$\mathcal{H}^s(F) = \lim_{\delta \to 0} \inf \left\{ \sum_{i=1}^{\infty} |U_i|^s : F \subset \cup_{i=1}^{\infty} U_i \quad \text{with} \quad 0 \leq |U_i| \leq \delta \quad \text{for each} \quad i \right\}.$$

The Hausdorff dimension for the graph (t, B_t^H) is equal to $2 - H$ [86]. That means if the Hurst exponent is close to 0, the paths of fBm zigzags in order to cover the unit square. Thus for the case of a fractional Brownian motion and a fractional Gaussian noise model, the fractal dimension D and the Hurst exponent H linearly depend on each other. Gneiting and Schlather [52] have presented a Cauchy class of stationary Gaussian processes $\{Z(t), t \in \mathbb{R}\}$ with the correlation function

$$\mathbb{E}[(Z(t)Z(t+\tau)] = (1 + |\tau|^{\alpha})^{-\beta/\alpha}, \quad \tau \in \mathbb{R}, \tag{2.20}$$

where any combination of the parameters $\alpha \in (0, 2]$ and $\beta > 0$ is permissible. For such stochastic models, the fractal dimension D and the Hurst exponent H may vary independently of each other.

2.2 Bayesian approach

In 1763, in his work *"An Essay towards solving a Problem in the Doctrine of Chance"* [16], Bayes gave the solution to the inverse probability problem: what is the probability of success given the results of independent trials. He proposed to answer this question in the following manner:

i) specify the prior distribution of the parameter $\mathbb{P}(\theta)$ which represents the state of knowledge (or ignorance) about θ before having analyzed the given data.

ii) calculate the posterior density (commonly known as the Bayes' theorem):

$$\mathbb{P}(\theta|x) = \frac{\mathbb{P}(x|\theta)\mathbb{P}(\theta)}{\int \mathbb{P}(x|\theta)\mathbb{P}(\theta)d\theta}; \tag{2.21}$$

iii) use the posterior distribution to quantify the information about the parameter of interest (θ).

The function $\mathbb{P}(x|\theta)$ (when regarded as a function of θ for fixed x) is called a likelihood function $\mathcal{L}(\theta|x)$. This means that a θ for which $\mathbb{P}(x|\theta)$ is large is more "likely" to be the true θ than a θ for which $\mathbb{P}(x|\theta)$ is small, in that x would be a more plausible occurrence if $\mathbb{P}(x|\theta)$ were large. Further in the thesis I will use the notation $\mathcal{L}(\theta)$ to indicate the likelihood function for parameter θ.

In this way the Bayesian inference obeys what is sometimes called the likelihood principle: in making inferences or decisions about θ after x is observed, all relevant experimental information is contained in the likelihood function for the observed x. Furthermore, two likelihood functions contain the same information about θ if they are proportional to each other (as functions of θ).

In practice, the parameter θ is unknown. In order to infer θ from the data, two main methods are commonly used: frequentist (classical) approach and the Bayes' approach. Next, I provide a discussion about both approaches in order to underline the strength and advantages of the Bayesian approach used further in this thesis.

In the frequentist approach the probability is represented by a long-run frequency of a repeatable event in which the uncertainty is due to the randomness. Here, there are no probability statements about the parameters, since they are considered as unknown constants. All uncertainty is measured only through the significance levels (the p-value approach [1] advocated by Fisher [47] and the α-level approach [2] advocated by Newman and Pearson [99]) and confidence intervals. For instance, if the frequentist approach gives a 95% confidence interval, this means only that in a repeated sampling, 95% of realized intervals cover the true parameter θ. This however, does not mean that there is 95% probability that θ will be in this interval. In contrast to this, in the Bayesian

[1] The traditional approach is to use the ideas of Fisher, Neyman and Pearson, wherein one states a primary, or *null*, hypothesis H_0, and an *alternative* hypothesis H_1. After determining an appropriate test statistic $T(x)$, one then computes the observed significance, or p-value, of the test as

$$p - \text{value} = \mathbb{P}\{T(x) \text{ more "extreme" then } T(x_{obs})|\theta, H_0\},$$

where "extremeness" is in direction of the alternative hypothesis. If the p-value is less than some prespecified Type I error rate, H_0 is rejected; otherwise, not [27].

[2] A test $T(x)$ of $H_0 : \theta \in \Theta_0$ versus $H_1 : \theta \in \Theta_1$ is said to have size α if

$$\sup_{\theta \in \Theta_0} \mathbb{E}_\theta[T(x)] = \alpha - \text{level}.$$

interpretation, a 95% confidence interval means that the probability to find parameter θ in this interval is 0.95.

Now, based on the work of Wagenmakers [125] I articulate several general problems associated with the frequentist approach:

• *Frequentist inference does not prescribe which estimator is the best*

Point estimation involves calculation of a single "best guess" of some quantity of interest from data and we assume there is some "true" value θ for this quantity, which is fixed but unknown. There is a large amount of possible estimators, thus the question arises how to determine which estimator to use? The common procedure here is to use an unbiased estimator.

An estimator $\varepsilon(\cdot)$ based on data x is unbiased when for all θ

$$\int_{\Omega} \varepsilon(x)\mathbb{P}(x|\theta)dx = \theta, \tag{2.22}$$

where Ω indicates as before the sample space.

However, this choice of the estimator is based on all possible data sets that could be observed, thus making the procedure biased on the previous knowledge of the researcher. Additionally, estimator which is unbiased for some parameter may be biased for some nonlinear transformation of it. The examples when unbiased estimators perform uniformly worse than biased one can be found in [102].

• *Frequentist inference generally does not condition on the observed data*

Berger and Wolpert stated in [19] that "a particular procedure is decided upon for use, and the accuracy of the evidence from an experiment is identified with the long run behavior of the procedure, were the experiment repeatedly performed."

• *Frequentist inference does not apply to non-nested models*[3]

Frequentist inference is not well suitable for the comparison of for example two non-nested models due to the fact that the choice of the null hypothesis in such case is unclear. This can lead to rejecting or accepting both models.

• *Frequentist inference does not quantify statistical evidence*

This point involves the tradition that the p-values reflect the strength of evidence against the null hypothesis. However, there are exist some studies that the p-postulate is false due to the following fact: given the same p-values for different experiments, the studies with small sample size provide more evidence against the null hypothesis than studies with large sample size (see [3, 11, 113]).

• *Frequentist inference depends on data that were never observed and the intention*

[3]Two models are non-nested, either partially or strictly, if one model cannot be reduced to the other model by imposing a set of linear restrictions on the parameter vector; otherwise they called nested [29].

with which the data were collected

Wagenmakers [124] asserts that "p-value is a tail-area integral, and this integral is effectively over data that are not observed but only hypothesized. The probability of these hypothesized data depends crucially on the possibly unknown subjective intentions of the researcher who carried out the experiment. These hypothetical data are data expected under null hypothesis, without which it is impossible to construct the sampling distribution of the test statistic."

In contrast to frequentist inference, in Bayes' approach, the probability of inferring parameter is concerned with the uncertainty of the knowledge, rather than with the variability of the outcome. Thus all the parameters assume to be random variables, bringing to the Bayesian approach many advantages over the frequentist counterpart. I discuss next the main characteristics of the Bayesian approach [125]:

• *Bayesian inference is coherent*

In contrast to the frequentist inference, Bayesian theory is based on limited amount of axioms about rational decision making. Therefore, the number of possible ways to obtain an appropriate response is limited and there is no need for any ad-hoc solutions to remedy procedures that give internally inconsistent results.

• *Bayesian inference allows marginalization*

Bayesian inference allows to deal only with the parameters of interest by integrating out the other parameters, according to the law of total probability. The resulting marginal posterior distributions may be identical to the one from the frequentist counterparts, but this only holds true in a few special cases.

• *Bayesian inference is extended to non-nested models*

Due to the possibility to marginalize the parameters of interest, Bayesian inference does not make a fundamental distinction between nested and non-nested models. This allows to expand the amount of possible applications.

• *Bayesian inference allows to incorporate prior knowledge*

Bayesian inference is open for any prior knowledge about parameters in the model. By incorporating prior information into the optimization procedure, a smooth transition between fixed and fitted parameters can be achieved.

• *Bayesian inference allows to obtain evidence in favor of the null hypothesis*

Quantifying the evidence by comparing the marginal probability of the data given one hypothesis, to the marginal probability of the data given another hypothesis, Bayesian inference is free from rejection the null hypothesis as it has no special status here.

• *Bayesian subjectivity is open to inspection*

A main criticism addressed to Bayesian inference is that it is subjective. However, the frequentist inference is also not as objective as one may think: both

of them have subjective elements. However, Bayesian subjectivity is open to inspection, whereas frequentist is not.

• *Bayesian inference is being simplified by computerizing statistical thought*

Until recent times application of the Bayesian inference was limited by technical requirements necessary to solve the complex tasks which Bayesian statistics deal with. Nowadays a large amount of Bayesian software programs have been developed to operate with these problems. To incorporate multiple data and efficiently implement Bayesian inference such programs as BUGS, BOA, SAS, S-Plus and others could be used.

Moreover, the Bayesian inference allows to receive results that are closely connected to the ones the researchers are interested in. Testing both the Bayesian and the frequentist inference, Berger [20] concluded that "based on the philosophical foundations of the approaches, Bayesian models are best suited to addressing hypotheses, conjectures, or public-policy goals, while the frequentist approach is best suited to those epidemiological studies which can be considered "experiments", i.e. testing constructed sets of data". However, the Bayesian inference is more flexible for considering hierarchical nonlinear models, where one experiment can be affected by several parameters which, for individual people are assumed to be drawn from group-level distribution. For non-Bayesian approaches such multi-level structures can not be considered in most cases.

Based on the given discussion I assumed the Bayesian approach to be more suitable for the estimation of the Hurst exponent. Thus, in the following I discuss the Bayes' theorem **Eq**. 2.21 in detail.

2.2.1 Prior distributions

From Bayesian perspective, all the parameters in the model are random variables, and have both prior and posterior distributions. Therefore, the Bayesian method requires prior probabilities to be given as explicit values. The choice of the prior distribution is usually based on the information about the unknown parameter. However, in most of the cases there is no prior information given. Thus, it is necessary to consider general theories in order to construct prior distributions for the given task in hand. Next, I describe the three possibilities for the choice of the prior [27, 49, 102].

2.2.1.1 Elicited (informative) priors

The case of informative priors allows definition of priors when we can use prior information from previous studies, our personal opinions about the investigating system or models themself suggest theoretical parameters values. Here, we have a specific, definite numerical information about a variable that is crucial to estimation of the model. If parameter θ is deterministic, we limit consideration to manageable collection of θ_i values deemed "possible", and

subsequently to assign probability masses to these values in such a way that their sum equals to one. If θ is continuous instead consider the single points, rather assign the masses to intervals on real line, resulting in a histogram prior for θ (even in a case of continuously, the computation of the posterior distribution required to be done numerically). Alternatively, we might simply assume that the prior belongs to a distribution family.

2.2.1.2 Conjugate priors

In general, the posterior density $\mathbb{P}(\theta|x)$ is not easy to calculate. In order to reduce the computational costs, very often the choice of the prior is restricted by some familiar distributional family. Using this line of thinking, Raiffa and Schlaifer [108] developed the so-called conjugate priors:

Let $\mathcal{F}$ denote the class of density functions $\mathbb{P}(x|\theta)$. A class $\mathcal{P}$ of prior distributions is said to be a conjugate family for $\mathcal{F}$ if $\mathbb{P}(\theta|x)$ is an element of the class $\mathcal{P}$ for all $\mathbb{P}(x|\theta) \in \mathcal{F}$ and $\mathbb{P}(\theta) \in \mathcal{P}$. This leads to a posterior distribution belonging to the same distribution family as the prior.

A conjugate family can frequently be determined by examining the likelihood function $\mathbb{P}(x|\theta)$ and choosing, as a conjugate family, the class of distributions with the same functional form as the likelihood function. This is a convenient way to obtain a posterior distribution a known parametric form. However, it is still necessary to check whether the prior information is appropriately represented by the distribution family.

2.2.1.3 Noninformative priors

The necessity to use noninformative priors arises when the prior distributions have no population basis and we want to have the prior distributions that will play a minimal role in the posterior distribution. In that way we make the analysis as object as possible and "let data speak for itself".

Thus, the inference has to be driven only by the likelihood function. Therefore, the best candidate for the prior will be an uniform prior. Uniform distribution is used to represent a variable that is known to lie in an interval and equally likely to be found anywhere in the interval. However, since it is not invariant under reparametrization, it is necessary to ask on which scale should the uniform distribution be specified? The general solution to this problem is to use Jeffreys' noninformative priors which are based on considering one-to-one transformations of the parameter [65].

In particular, the Jeffreys' rule offers that the prior distribution for parameter θ is approximately noninformative, if it is taken proportional to the square root of the information-matrix determinant:

$$\mathbb{P}(\theta) = \sqrt{detI(\theta)}, \tag{2.23}$$

where $I(\theta)$ is the Fisher information matrix, having ij-th element:

$$I_{i,j}(\theta) = -E\left[\frac{\partial^2}{\partial\theta_i\partial\theta_j}\log\mathbb{P}(x|\theta)\right]. \tag{2.24}$$

It is motivated by considering one-to-one transformations of the parameter: $\gamma = \kappa(\theta)$. By transformation of variables, the prior density $\mathbb{P}(\theta)$ is equivalent, in terms of expressing the same beliefs, to the prior density on γ:

$$\mathbb{P}(\gamma) = \mathbb{P}(\theta)\left|\frac{d\theta}{d\gamma}\right| = \mathbb{P}(\theta)|\kappa'(\theta)|^{-1}. \tag{2.25}$$

Determining the prior density $\mathbb{P}(\theta)$ should yield an equivalent result if applied to the transformed parameter. In general, the flat prior is defined as the Jeffreys' prior for location parameters where as the inverse prior denotes the Jeffreys' prior for scale parameters.

It will frequently happen that the natural noninformative prior is an improper prior, meaning that the prior can not be normalized (does not integrate to 1). In some examples (see [49]) the improper priors can be used in the Bayesian analysis and the posterior density in fact will be proper. However, the posterior distributions obtained from improper prior distributions must be interpreted with great care and one must always check if the posterior distribution has a finite integral.

2.2.2 Bayes' theorem

Having specified the prior distribution by one of the proposed methods (*Section 2.2.1*), now we can calculate the posterior distribution of the parameter. We start the analysis without having any knowledge neither about the parameter θ nor the data x. Our uncertainty is thus determined via $\mathbb{P}(\theta, x)$ [36]. Assuming that x is the observed quantity, we can find conditional probability using:

$$\mathbb{P}(\theta|x) = \frac{\mathbb{P}(\theta, x)}{\mathbb{P}(x)}. \tag{2.26}$$

Next, using the marginalization integral $\mathbb{P}(x) = \int \mathbb{P}(\theta, x)d\theta$ we can rewrite **Eq.** 2.26 in the following form:

$$\mathbb{P}(\theta|x) = \frac{\mathbb{P}(\theta, x)}{\int \mathbb{P}(\theta, x)d\theta}. \tag{2.27}$$

Here, value $\mathbb{P}(\theta, x)$ can be calculated using $\mathbb{P}(\theta, x) = \mathbb{P}(x|\theta)\mathbb{P}(\theta)$. Summarizing all together, one obtains the Bayes' theorem:

$$\mathbb{P}(\theta|x) = \frac{\mathbb{P}(x|\theta)\mathbb{P}(\theta)}{\int \mathbb{P}(x|\theta)\mathbb{P}(\theta)d\theta}. \tag{2.28}$$

Thus, the complete information about the parameter θ given the data is obtained. Considering the fact that the denominator in **Eq.** 2.28 is strictly posi-

tive, normalization constant which does not depend on θ, the equation may be rewritten as:

$$\begin{aligned} Posterior &\propto Likelihood \times Prior \\ \mathbb{P}(\theta|x) &\propto \mathbb{P}(x|\theta) \times \mathbb{P}(\theta), \end{aligned} \tag{2.29}$$

where $\propto$ denotes that two quantities are proportional to each other. Hence, the posterior distribution can be found by combining the likelihood and the prior distribution of the parameter.

I would like to stress the fact that the Bayesian statistics provides intuitive and significant inferences: taking into account all available information (due to the inclusion of the prior information), Bayesian methods can answer complex questions.

Chapter 3

Estimation of the self-similarity exponent

> The only relevant thing is uncertainty - the extent of our knowledge and ignorance. The actual fact of whether or not the events considered are in some sense determined, or known by other people, and so on, is of no consequence.
>
> Bruno de Finetti

As shown in *Chapter 2*, the Bayesian method could be used as a powerful tool to obtain all information about parameters in the model. I developed a novel method for estimation of the Hurst exponent based on Bayesian theory [82]. In this chapter I discuss in details the Bayesian method for the estimation of the Hurst exponent in terms of a linear mixed model [38]. I start with considering the following model:

$$X_t = \lambda B_t^H + a\,t + b. \tag{3.1}$$

Here, I use the assumption that the underlying process B_t^H is a fractional Brownian motion parameterized by the Hurst exponent H, λ is the amplitude ($\lambda > 0,\ \lambda \in \mathbb{R}$), $a \in \mathbb{R}$ is the slope and $b \in \mathbb{R}$ is the offset. The model contains a linear trend in order to describe real data more accurately. As shown further in this chapter, the estimation method based on Bayesian analysis allows to estimate the Hurst exponent besides the remaining parameters in the model. I also consider the data with gaps and show that random gaps influence the estimation of the parameters only weakly. The method is compared by means of Monte-Carlo simulations to the most popular technique, the detrended fluctuation analysis. The performance of the proposed method for non-Gaussian data is further tested on Rosenblatt process. Moreover, the limitation of the

method to Lévy flights with stable distribution are also discussed. In addition, I apply the proposed estimation method to the Dow-Jones Industrial Average closing values for the years 1900-1993 and produce a time-dependent posterior distribution of the Hurst exponent.

3.1 Bayesian inversion

Let D be the set of N observations,

$$D = \{(t_k, \mathcal{Y}_k = X_{t_k})\}, \quad k = 1, \ldots, N,$$

where X_{t_k} denotes the data value at the time point t_k. Since very often real time series are not evenly sampled, here it is not assumed that the time points are on a regular grid. Thus, one of the advantage of the method I propose here is the robustness with respect to the data sampling.

Next, let F denote the $N \times 2$ system matrix with entries

$$F_{k,1} = 1, \quad F_{k,2} = t_k. \tag{3.2}$$

Defining β as a 2×1 vector of the slope and the offset, the model (**Eq.** 3) can be rewritten in a more compact form as the following linear mixed effects model (see [38]):

$$\mathcal{Y} = \lambda u + F\beta, \tag{3.3}$$

where $u \sim \mathcal{N}(0, \mathbf{\Sigma}_H)$ represents the random part, $F\beta$ represents the fixed effects and the Hurst exponent indicates a hyperparameter. Here u is a Gaussian random variable with zero mean and covariance matrix $\mathbf{\Sigma} = \mathbf{\Sigma}(H) = \mathbf{\Sigma}_H$ with entries

$$\mathbf{\Sigma}_{i,j} = \mathbb{E}(B^H_{t_i} B^H_{t_j}) = \frac{1}{2}(|t_i|^{2H} + |t_j|^{2H} - |t_i - t_j|^{2H}). \tag{3.4}$$

For a fixed parameter β and observation points $\{t_k\}$, the vector of outcomes $\mathcal{Y}_k$ is a N-dimensional Gaussian random vector with the mean value $F\beta$ and the variance $\lambda^2\mathbf{\Sigma}_H$: $\mathcal{N}(F\beta, \lambda^2\mathbf{\Sigma}_H)$. Hence, the likelihood of observing $D = \{(t_k, \mathcal{Y}_k = X_{t_k})\}$ with $\mathcal{Y} = \{\mathcal{Y}_k\}$ for a fixed set of parameters and observation points is

$$\mathcal{L}(H, \lambda, \beta, \{t_k\}) = \frac{1}{(2\pi)^{N/2}\lambda^N |\mathbf{\Sigma}_H|^{1/2}} e^{-\frac{(\mathcal{Y}-F\beta)^T \mathbf{\Sigma}_H^{-1} (\mathcal{Y}-F\beta)}{2\lambda^2}}. \tag{3.5}$$

Here, $|\mathbf{\Sigma}_H|$ denotes the determinant of the matrix $\mathbf{\Sigma}_H$. Since in this case no knowledge about the parameters exists, I estimate λ using Jeffreys-like priors (see *Section 2.2.1*): $\mathbb{P}(\lambda) \sim \lambda^{-1}$, and use the flat prior to model β and H, *i.e.* $\mathbb{P}(\beta) \sim 1$, $\mathbb{P}(H) = \chi_{[0,1]}$ [122]. Now, applying the Bayes' theorem **Eq.** 2.28, [16],

the posterior distribution for the unknown parameters can be expressed as:

$$\mathbb{P}(H,\lambda,\beta|D) = C\frac{1}{\lambda^{N+1}|\mathbf{\Sigma}_H|^{1/2}}e^{\frac{-(\mathcal{Y}-F\beta)^T\mathbf{\Sigma}_H^{-1}(\mathcal{Y}-F\beta)}{2\lambda^2}}. \tag{3.6}$$

Here, C is a normalization constant which allows to obtain a probability density of the form

$$\int_0^1 dH \int_0^\infty d\lambda \int_{-\infty}^{+\infty} d\beta\, \mathbb{P}(H,\lambda,\beta|D) = 1. \tag{3.7}$$

Hence, **Eq.** 3.6 can be rewritten in the following form:

$$\mathbb{P}(H,\lambda,\beta|D) = C\,\lambda^{-N-1}|\mathbf{\Sigma}_H|^{-1/2}e^{-\frac{R^2(H)}{2\lambda^2}}e^{-\frac{(\beta-\beta^*(H))^T F^T\mathbf{\Sigma}_H^{-1}F(\beta-\beta^*(H))}{2\lambda^2}} \tag{3.8}$$

with

$$\begin{aligned} \beta^*(H) &= (F^T\mathbf{\Sigma}_H^{-1}F)^{-1}F^T\mathbf{\Sigma}_H^{-1}\mathcal{Y}, & (3.9)\\ R^2(H) &= (\mathcal{Y}-F\beta^*(H))^T\,\mathbf{\Sigma}_H^{-1}\,(\mathcal{Y}-F\beta^*(H))\,, & (3.10) \end{aligned}$$

where $\beta^*(H)$ and $R^2(H)$ are parameters of the model that most likely occur and $R(H)$ is the residuum around the linear trend. Here, $R(H)$ is measured in terms of the Mahalanobis distance, induced by the covariance matrix $\mathbf{\Sigma}_H$. This is a statistical distance which differs from Euclidean distance in that it takes into account the correlations of the data set and it is scale-invariant [81]. In this form, the term $\left(e^{-\frac{(\mathcal{Y}-F\beta)^T\mathbf{\Sigma}_H^{-1}(\mathcal{Y}-F\beta)}{2\lambda^2}}\right)$, the Gaussian part of the posterior distribution, is clearly visible. Therefore the marginal distribution of (H,λ) and H, obtained by the integration over β and λ can be computed using Gaussian integrals with

$$\mathbb{P}(H,\lambda|D) = C|\mathbf{\Sigma}_H|^{-1/2}|F^T\mathbf{\Sigma}_H^{-1}F|^{-1/2}\lambda^{1-N}e^{-\frac{R^2(H)}{2\lambda^2}}. \tag{3.11}$$

Thus, integration over λ results in

$$\mathbb{P}(H|D) = C|\mathbf{\Sigma}_H|^{-1/2}|F^T\mathbf{\Sigma}_H^{-1}F|^{-1/2}R^{2-N}(H), \tag{3.12}$$

or, using the decimal logarithmic scale,

$$\begin{aligned} -2\log\mathbb{P}(H,\lambda|D) &= \log|\mathbf{\Sigma}_H| + \log|F^T\mathbf{\Sigma}_H^{-1}F| + \\ &+ 2(N-1)\log\lambda + \frac{R^2(H)}{\lambda^2}, & (3.13)\\ -2\log\mathbb{P}(H|D) &= \log|\mathbf{\Sigma}_H| + \log|F^T\mathbf{\Sigma}_H^{-1}F| + \\ &+ 2(N-2)\log R(H). & (3.14) \end{aligned}$$

The observations enter the equation only through the residuum $R(H)$ and the time points of the measurements through F and $\mathbf{\Sigma}_H$.

Next, the point-estimator for H is taken at the location of the maximal value of the posterior density:

$$\hat{H} = \text{argmax}\, \mathbb{P}(H|D). \tag{3.15}$$

The marginalized posterior density (at $\hat{H}$) of the amplitude λ is obtained by integration over β in its explicit expression and reads

$$\mathbb{P}(\lambda|\hat{H}, D) = C\lambda^{1-N} e^{-\frac{R^2(\hat{H})}{2\lambda^2}}. \tag{3.16}$$

This equation allows to derive the following marginal maximal posterior estimator for the amplitude λ

$$\hat{\lambda}^2 = \frac{R^2(H)}{N-1}. \tag{3.17}$$

In the same way, the profiled posterior distribution of the offset and the slope parameters $\beta = [b, a]^T$ can be estimated as a Gaussian distribution with mode at $\beta^*(H)$ and a covariance matrix

$$\Gamma = F^T \mathbf{\Sigma}_{\hat{H}}^{-1} F. \tag{3.18}$$

However, the full marginal posterior distributions of β and λ involve a numerical integration over H.

An estimation of the amplitude and the Hurst exponent for the test model is shown in **Fig.** 3.1, whereas **Fig.** 3.2 depicts the estimations of the slope and the offset for the same model. Here, I generated data sets of length $N = 100$ at every time point t_k for $k = 1, \ldots, 100$.

In general, several methods exist to generate numerically fBm data as required in the system I study here [39, 67]. I note that here, as well as in the following chapters of this thesis, the synthetic data I use for model testing is based on the Cholesky decomposition of $\mathbf{\Sigma}_H = L^t L$ [56]. I furthermore apply $L^t e$, with $e = [e_1, \ldots, e_N]^t$ being a random vector of N independent standard Gaussian random variables.

Next, in order to identify the robustness of the proposed method with respect to unevenly sampled data set, I use the same model and draw randomly $N = 36$ time points from $t_k, k = 1, \ldots, 100$ as depicted in **Fig.** 3.5. I note here that the data is not interpolated (**Fig.** 3.5). The estimations of the interior parameters using the Bayesian method (**Eq.** 3.13) for unevenly sampled data are shown in **Fig.** 3.3, 3.4.

In both cases, the intersection of the white dotted lines show the exact true value of the parameters in the figures: $H = 0.1,\ \lambda = 0.8,\ \beta = [10, 0.2]^T$. The one dimensional projections of the posterior densities $p(\alpha|data)$, $p(H|data)$ and $p(\beta|data)$ show the confidence bounds for the estimates. Thus, this allows for estimation of the Hurst exponent with confidence intervals. More-

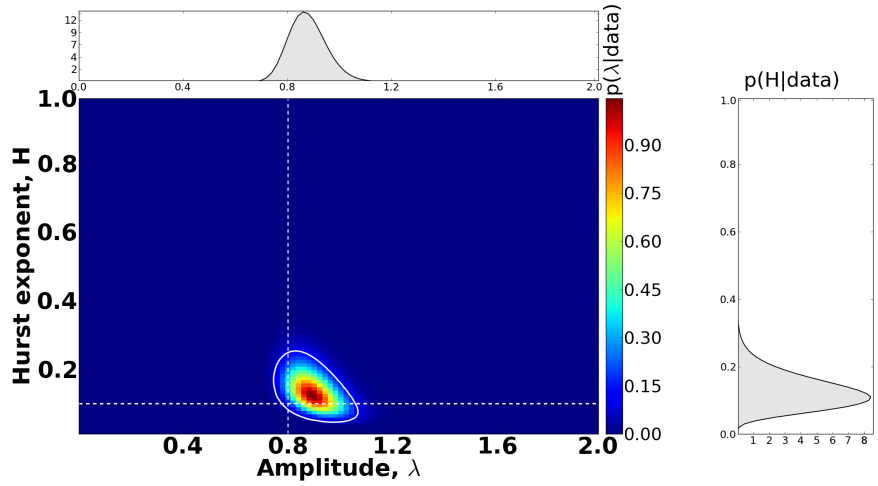

Figure 3.1: The Bayesian estimation of amplitude and Hurst exponent, for the model $X_{t_k} = 0.8B^{0.1}_{t_k} + 0.2\,t_k + 10$ with $N = 100$ at time points $t_k, k = 1, \ldots, 100$. The contour plots present the one dimensional projections of the posterior densities. The intersection of white lines denotes the true values of the parameters.

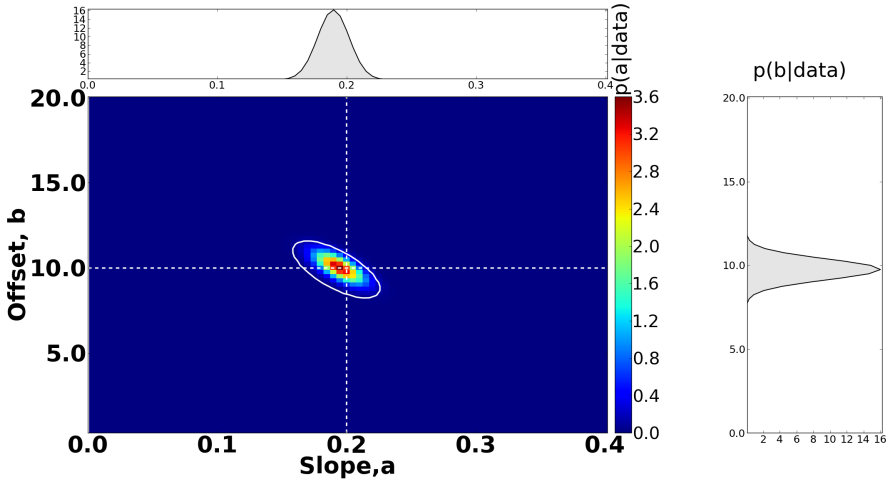

Figure 3.2: The Bayesian estimation of slope and offset for the model given in **Fig.** 3.1.

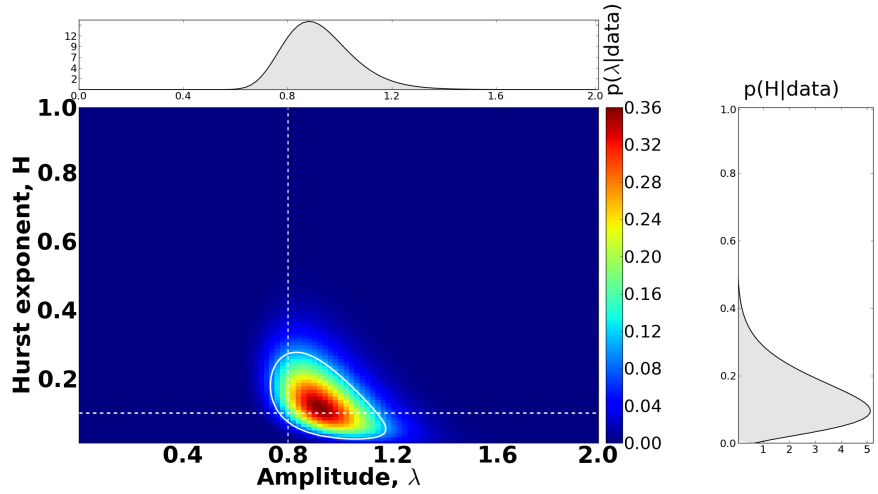

Figure 3.3: The Bayesian estimation of amplitude and Hurst exponent for unevenly sampled data with the model $X_{t_k} = 0.8B^{0.1}_{t_k} + 0.2\,t_k + 10$ of total length $N = 36$ at randomly chosen time points $t_k, k = 1, \ldots, 100$.

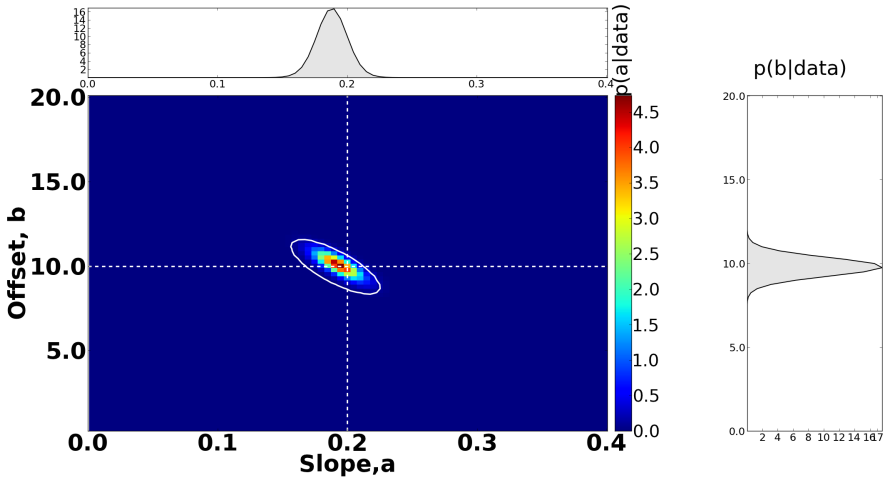

Figure 3.4: The Bayesian estimation of slope and offset for the model given in **Fig.** 3.3.

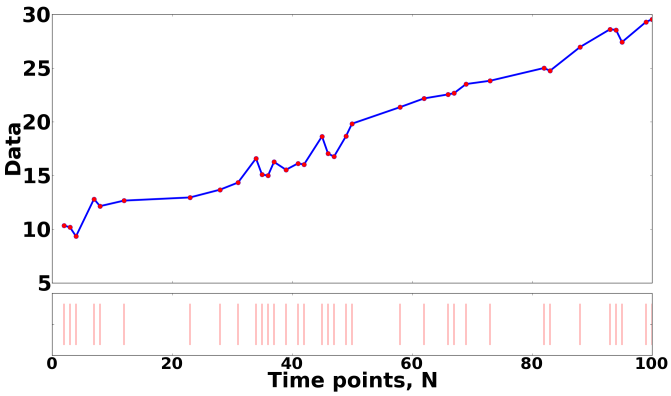

Figure 3.5: Simulation of fractional Brownian motion for the model $X_{t_k} = 0.8B^{0.1}_{t_k} + 0.2\,t_k + 10$ of total length $N = 36$ at randomly chosen time points $t_k, k = 1, \ldots, 100$ (red dots).

over, the method I propose estimates other parameters of the model. The true and estimated values for both evenly and unevenly sampled data are given in **Table** 3.1. As shown, the gaps do not significantly change the quality of the estimation. Clearly, this can only hold within some limit. This means that for irregular, but evenly distributed data points, the estimation will essentially be of the same quality as for regular sampled data. However, if lacunary distributions of data are considered, an influence of the data distribution to the quality of the Bayesian inversion is to be expected.

In order to check the sensitivity of the estimates to data loss, I consider the same model $X_{t_k} = 0.8B^{0.1}_{t_k} + 0.2\,t_k + 10$ in time points $t_k, k = 1, \ldots, 100$, randomly producing gaps with irregular sampling steps. Thus, I loose from 0% up to 100% information about the parameters in the model. Next I obtained 100 data realizations in the described way and calculated for each of them the posterior density $\mathbb{P}(H|D)$. The obtained averaged posterior densities are depicted in **Fig**. 3.6. As shown, starting from 27% of the original data set, the maximum of the averaged posterior densities falls in the interval, $[0.0294, 0.11215]$, con-

Table 3.1: The estimation of Hurst exponent, amplitude, slope and offset for the model $X_{t_k} = 0.8B^{0.1}_{t_k} + 0.2\,t_k + 10$.

	Hurst exponent $\hat{H}$	Amplitude $\hat{\lambda}$	Fixed effects $\hat{\beta} = [b, a]^T$
Evenly sampled data	0.115	0.89	$[9.75, 0.1975]^T$
Unevenly sampled data	0.135	0.87	$[10, 0.1925]^T$
True value	0.1	0.8	$[10, 0.2]^T$

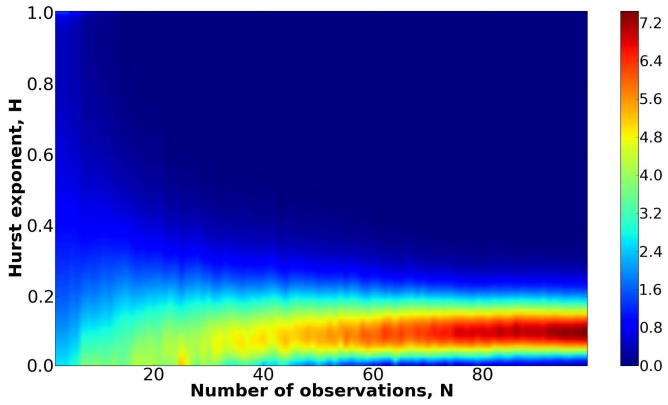

Figure 3.6: The averaged posterior densities for the model $X_{t_k} = 0.8B^{0.1}_{t_k} + 0.2\,t_k + 10$ with 100 realizations.

taining more than 75% of the distribution. Thus it can be concluded that the proposed method is an efficient tool for the Hurst exponent estimation while remaining relatively robust to data loss.

3.2 Comparison of the Bayesian estimation and the detrended fluctuation analysis

In this section, I compare the performance of the proposed method with the well-known detrended fluctuation analysis (DFA) by Peng *et. al.* [105]. The authors in [105] investigated the local scaling properties in DNA nucleotides which contain coding and non-coding regions in order to show that the patchiness (mixed coding and non-coding regions) of the DNA by itself cannot account for long-range power law correlations found in DNA dominated by non-coding regions. Later, they applied the DFA method to heartbeat data as well [106], where they aimed at a detailed description of the DFA algorithm. I underline here the fact that the DFA method (as well as its modificated versions [13, 25, 68, 69, 129]) is a widely used method for the Hurst exponent estimation. The DFA has been also used for the estimation of the Hurst exponent from fixational eye movements data [77, 95], a research question which will be discussed in details in *Chapter 6*.

The DFA algorithm can be summarized as follows:

1. For a given time series $\{X(t), t = 1, \ldots, N\}$ we divide $\{1, \ldots, N\}$ into $\left[\frac{N}{n}\right]$ boxes of equal size n.

2. In each box the least-square regression $X_n(t)$ for $t = 1, \ldots, N$, which represents the linear trend of the process in the box, is subtracted from the time series:
$$Y_n(t) = X(t) - X_n(t). \tag{3.19}$$

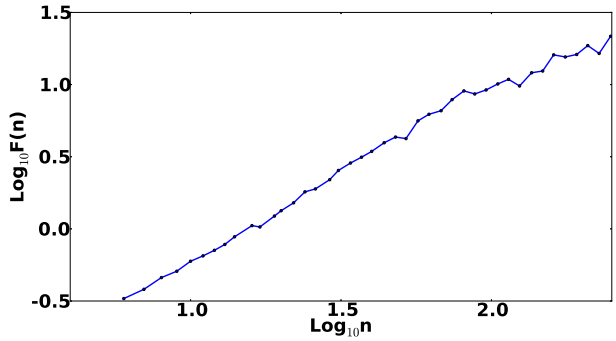

Figure 3.7: The scaling relation between the size of fluctuations $F(n)$ and the number of points n in artificial fractional Brownian motion of length $N = 1,000$ with $H = 0.3$.

3. For a given box size n, the detrended fluctuation function is the standard deviation of $Y_n(t)$ calculated by

$$F(n) = \sqrt{\frac{1}{n\left[\frac{N}{n}\right]} \sum_{t=1}^{n\left[\frac{N}{n}\right]} (Y_n(t))^2}. \tag{3.20}$$

In order to establish the relationship between $F(n)$ and the box size n, the algorithm is repeated for several box sizes which represent all time scales of the given system. Typically, $F(n)$ scales with the box size as $F(n) \sim n^{-H}$. A linear relationship in $log_{10}F(n) - log_{10}(n)$ indicates the presence of a power law fractal scaling. In **Fig.** 3.7, the existence of a linear relationship in artificial fractional Brownian motion of length $N = 1,000$ with $H = 0.3$ is shown. Thus next, I implemented the DFA method as proposed in [54].

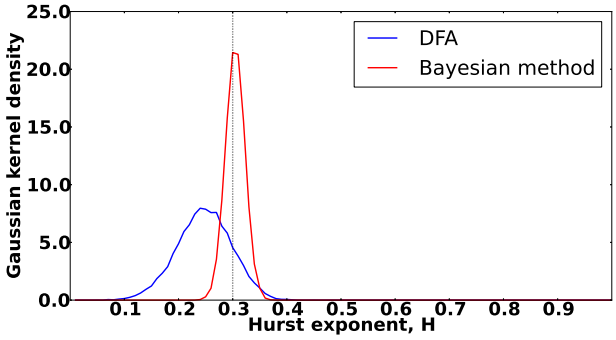

Figure 3.8: Gaussian kernel density of the estimator of the Hurst exponent for $H = 0.3$, $N = 1,000$ with 25,000 realizations by Bayesian method (red line) and DFA (blue line).

Table 3.2: The interval for $\hat{H}$ with $p \geq 95\%$ for artificial fBm with H=0.3.

Method	Interval for $\hat{H}$ with $p \geq 95\%$
DFA	$[0.159, 0.3247]$
Bayesian approach	$[0.2748, 0.3342]$
True value	0.3

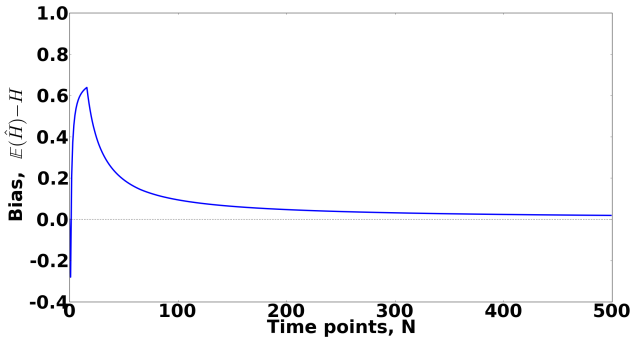

Figure 3.9: The validation test for Bayesian estimation with 500 realizations for the Hurst exponent $H = 0.3$.

Next, I use the DFA method as commonly used point estimator to estimate the Hurst exponent and compare the obtained results with the Bayesian estimation method I propose in this thesis (**Eq.** 3.15). In order to access the bias and the variance of this estimator, I performed Monte-Carlo simulations for the fixed true Hurst exponent ($H = 0.3$). In particular, I generated $25,000$ realizations of fBm and apply for each of them the DFA algorithm to estimate H. Then I repeat the analysis with the maximal posterior estimator $\hat{H}$ of the Bayesian method. The **Fig.** 3.8 shows the comparison between the methods. It clearly indicates that the Bayesian estimation method proposed here performs significantly better than the DFA algorithm. Moreover, not only its expectation value is closer to the true value, but also its variance is considerably smaller. The intervals containing more than 95% of the distribution for both methods are given in **Table** 3.2.

Furthermore, in **Fig.** 3.9 I quantify the bias of the maximal posterior estimator $\hat{H}$, $\mathbb{E}(\hat{H}) - H$, as a function of the number of data points in the case of evenly sampled data. Here, the expectation value is estimated using 500 random realizations at each data point for a Hurst exponent $H = 0.3$. Also the bias quickly decays for large N.

3.2.1 Bayesian comparison

As I have already shown, the Bayesian approach allows to obtain a full distribution of the posterior probabilities for the estimated value of the Hurst

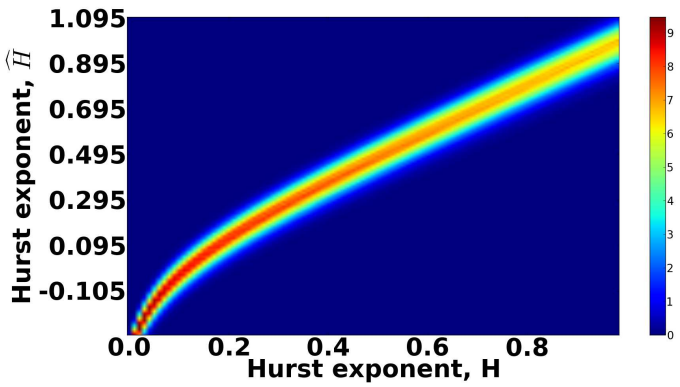

Figure 3.10: Projection of the space of the likelihood functions.

exponent H that parameterizes the data. Thus, using the Bayesian approach I analyze next the posterior information about the Hurst exponent that is contained in the $\mathcal{DFA}$ values. To that matter, I use the DFA approach to transform the data into a new quantity, the $\mathcal{DFA}$ value, and show how this may imply a loss of information on H. In order to perform this task, it is necessary to calculate the conditional probability $\mathbb{P}(H|\mathcal{DFA})$ which, using the Bayes' theorem, can be written as

$$\mathbb{P}(H|\mathcal{DFA}) = C\,\mathbb{P}(\mathcal{DFA}|H)\mathbb{P}(H). \tag{3.21}$$

Again, I assume a flat prior for H and use Monte-Carlo simulations to produce the distribution of $\mathcal{DFA}$ that occurs by random fluctuations for each given value of H. **Fig.** 3.10 shows the likelihood function obtained from $25,000$ realizations for separate H values varying in $[0,01, 0.99]$ with a step size of 0.01. Here, the number of data points was fixed to $1,000$. Now, having the information about the true value of H, I can process of the posterior density of

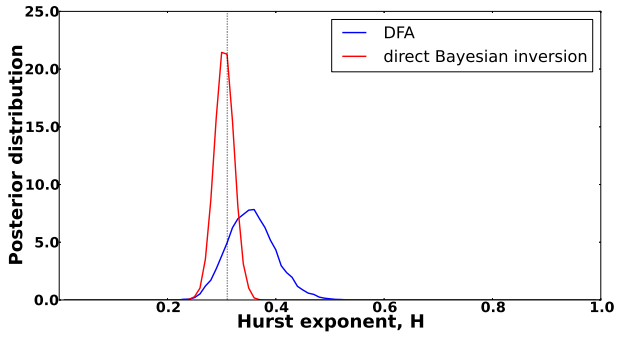

Figure 3.11: The posterior distribution for the Hurst exponent given the DFA value $H = 0.310496$ performed by the DFA-estimator (blue line) and direct Bayesian inversion (red line).

the Hurst exponent given by the $\mathcal{DFA}$ value. In **Fig.** 3.11, the posterior distribution for the Hurst exponent $H = 0.310496$ obtained via the detrended fluctuation algorithm is plotted. The method based on Bayesian theory outperforms the detrended fluctuation estimation in the calculation of the posterior information about H: the posterior information contained in the detrended fluctuation analysis is less sharp than the posterior information contained in the full Bayesian inversion.

3.3 Application to Dow-Jones Industrial Average closing values

Next, I apply the proposed Bayesian method for the estimation of the Hurst exponent for real-world data, in particular to analyze of the Dow-Jones Industrial Average closing values (DJIA), as obtained in the period from 1900 to 1993. Data for this purpose is depicted in **Fig.** 3.12. (The time series is available at http://lib.stat.cmu.edu/datasets/djdc0093.) I choose the Dow-Jones values as a good example of a time series with fractal structure and because the particular data set has been analyzed with various methods and techniques [107]. This allows me to compare the Bayesian method for the estimation of the Hurst exponent with previous results.

As necessary in many cases when dealing with real-world data, certain data preprocessing steps need to be haven. In particular, it is necessary to remove first the existent spikes from the data set. The time series consists of 25753 data points. The Bayesian approach estimation was applied to the given data set by years, with shift equal to one data point, and each year consisting of approximately 274 data points.

The same data set in the period from 1900 to 1951 was studied in the work of Pilgram and Kaplan in [107]. To estimate the Hurst exponent, they applied five different methods: fast Fourier transform based regression, band-integrated regression, R/S analysis as proposed by Hurst, detrended fluctuation analysis and maximum likelihood estimation. The results which they

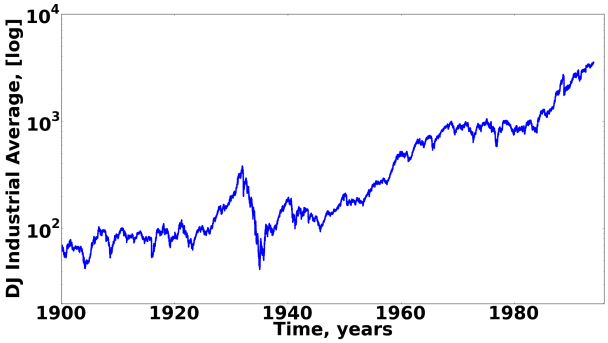

Figure 3.12: Dow-Jones Industrial Average closing values, the years 1900-1993.

Table 3.3: The estimation of the Hurst exponent for the Dow-Jones Industrial Average daily closing prices data from 1900 to 1951 years from the work of Pilgram and Kaplan [107].

Method	Hurst exponent, $\hat{H}$
Fast Fourier transform based regression	0.47375
Band-integrated regression	0.49625
Hurst R/S analysis	0.49375
Detrended fluctuation analysis	0.49
Maximum likelihood estimation	0.48625
This study	0.51

obtained for H are shown in **Table** 5.1, where the estimation computed by proposed Bayesian approach is also presented. Moreover, the distribution of the Hurst exponent obtained by the Bayesian approach for this time period is depicted in **Fig.** 3.13 with blue line.

Next, applying the same approach to the whole data set I received an point estimation of the Hurst exponent for the years between 1900-1993 to be equal to $H = 0.515$. The full distribution is shown in **Fig.** 3.13 with red line. Moreover, to see how the Hurst exponent is changing with time, I show the projection of its Bayesian estimation in **Fig.** 3.14. Comparing the results obtained for 1900-1951 and 1900-1993 years, we can see the shift of the distribution of the Hurst exponent to persistent area, *i.e.* the data possesses the property to have a long-memory for the last period. Thus, as shown here, the Bayesian method can also be applied to real data and estimate a value in the same range as previous estimates.

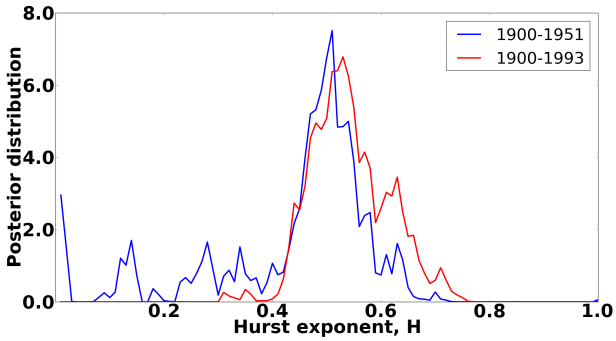

Figure 3.13: The sum of local posterior densities computed for sub time series of 274 data points with change of 1 data point of DJIA closing values for 1900-1951 (blue line) and 1900-1993 (red line) years.

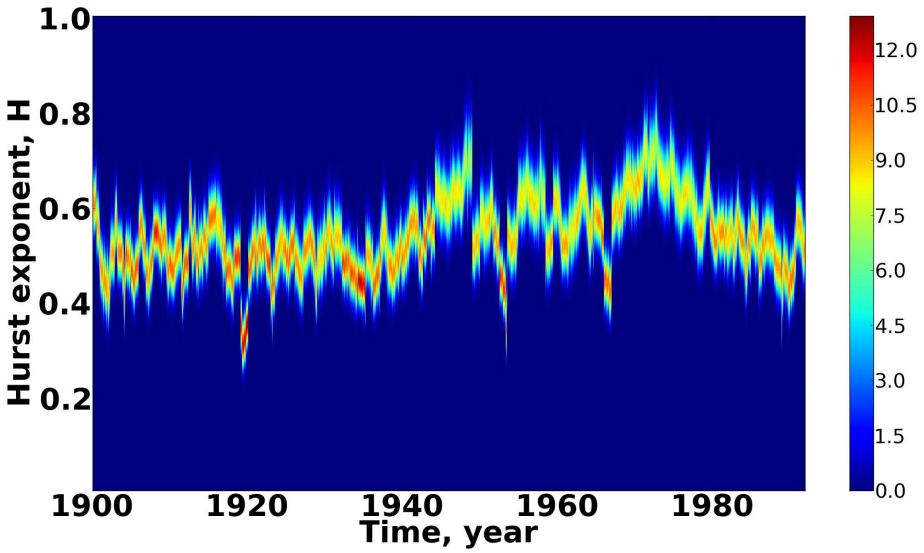

Figure 3.14: Projection of Bayesian estimation of the Hurst exponent of DJIA closing values from 1900 to 1993 years.

3.4 Estimation of the Hurst exponent for non-Gaussian data of Rosenblatt and α-stable Lévy processes

In order to investigate the efficiency of the Bayesian method, I consider in this section the Rosenblatt process as an example of a H-self-similar process, whose finite-dimensional distributions are non-Gaussian. This kind of process was initially introduced by Rosenblatt [112], where the current name is given by Taqqu in 1975 [118]. By definition (see [112, 118]), the Rosenblatt process is H-self-similar process with Hurst exponent $H \in (1/2, 1)$ and stationary increments. It can be written in explicit form as a stochastic integral

$$Z(t) = a(H) \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \left(\int_{0}^{t} (s - y_1)_{+}^{-\frac{2-H}{2}} (s - y_2)_{+}^{-\frac{2-H}{2}} ds \right) dB_{y_1} dB_{y_2}, \qquad (3.22)$$

where $\{B_y, y \in \mathbb{R}\}$ is a usual Brownian motion, $s \in [0, \infty)$, $a(H)$ is a positive normalization constant chosen as $\mathbb{E}(Z(1)^2) = 1$ and $x_+ = max\{x, 0\}$. The Rosenblatt process is also known as Hermite process of the second order (note that the fractional Brownian motion is of the first order) [80, 120].

Thus, in order to produce synthetic data for estimation purpose, I use here the wavelet-based synthesis of the Rosenblatt process as proposed and kindly provided by Abry and Pipiras [5]. An example of realization of the Rosenblatt process with Hurst exponent $H = 0.8$ and length $N = 513$ is shown in **Fig.** 5.4. As before, I generated 25,000 realizations of the Rosenblatt process of length $N = 1,025$ (due to the generation procedure) and applied the DFA and the

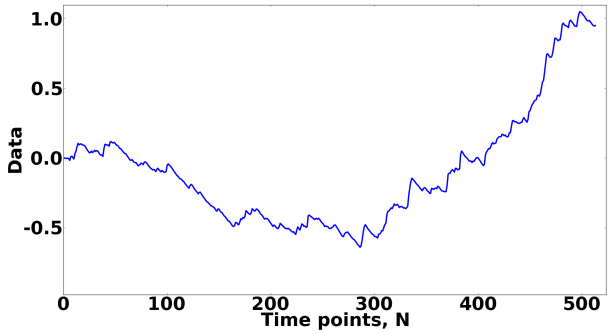

Figure 3.15: Simulation of the Rosenblatt process for $N = 513$ and $H = 0.8$.

Table 3.4: The interval for $\hat{H}$ with $p \geq 95\%$ for artificial Rosenblatt process with H=0.8.

Method	Interval for $\hat{H}$ with $p \geq 95\%$
DFA	$[0.5047, 0.745]$
Bayesian approach	$[0.7629, 0.8919]$
True value	0.8

Bayesian approach as point estimators in order to estimate the Hurst exponent from non-Gaussian data and test the efficiency of both methods. In **Fig.** 3.16 I present the resulting Gaussian kernel density for both methods, whereas the numerical results are reported in **Table** 3.4. Also this analysis shows that the variance of the Bayesian point estimator is clearly smaller than the one examined by the DFA. Moreover, the detrended fluctuation analysis strongly underestimates the Hurst exponent whereas the Bayesian method performs this task more reliably.

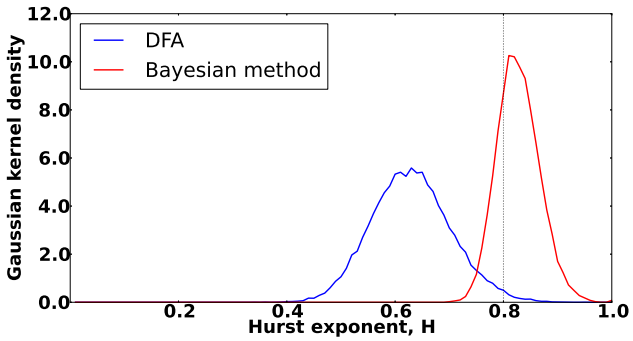

Figure 3.16: Gaussian kernel density for a point-estimator of the Hurst exponent $H = 0.8$ of length $N = 1,025$ with 25,000 realizations by DFA (blue line) and direct Bayesian method (red line).

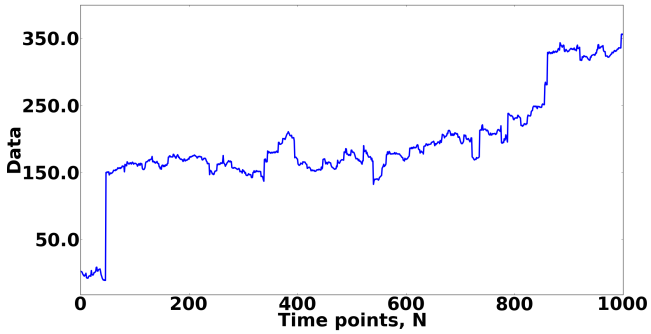

Figure 3.17: Simulation of a α-stable Lévy motion for $N = 1,000$ and $\alpha = 1.2$.

Now, I discuss the limitation of the proposed Bayesian method to a process which is self-similar, but has no properties resembling a Gaussian process, such the α-stable Lévy motion. I therefore perform the analysis with Lévy process following a stable distribution as discussed next [26, 96, 115, 127].

A random variable X is stable, or has a stable distribution, if and only if $X = aZ + b$ where $a > 0, b \in \mathbb{R}$ and Z is a random variable with characteristic function [100]:

$$\mathbb{E}\left(e^{iuZ}\right) = \begin{cases} e^{(-|u|^{\alpha}[1-i\beta\tan\frac{\pi\alpha}{2}(sign(u))])} & \alpha \neq 1 \\ e^{(-|u|[1+i\beta\frac{2}{\pi}(sign(u))\log|u|])} & \alpha = 1 \end{cases}$$

with α being the index of stability, $0 < \alpha \leq 2$, and β - skewness parameter, $-1 \leq \beta \leq 1$.

These distributions are symmetric around zero whenever $\beta = 0$ and $b = 0$. Symmetric α-stable motion is called α-stable Lévy motion. Moreover, the dependency between the stability index α of α-stable Lévy motion and the Hurst exponent H of fBm has been identified as $H = \frac{1}{\alpha}$ [126]. I show the example of simulated α-stable Lévy motion for the stability index $\alpha = 1.2$ and sample length $N = 1,000$ in **Fig.** 3.17.

In order to see how both methods (Bayesian method and DFA) estimate the self-similarity parameter in this case, I simulated 25,000 realizations of α-stable Lévy motion with length $N = 1,000$ with a stability index $\alpha = 1.2$ or self-similarity exponent $H = \frac{5}{6}$. The results are presented in **Fig.** 3.18, I observed that both, the Bayesian method and the DFA estimate the value for the self-similarity exponent H close to 0.5 ($\alpha \approx 2$).

Performing the same estimation approach for different values of α, I obtained the same results for both methods (Hurst exponent H close to 0.5). Thus, it can be stated that both methods are not suitable for the analysis of the self-similarity exponent from α-stable Lévy motion.

Bayesian inversion is carried out based on the assumption that the underlying process is fractional Brownian motion. Therefore, it is expected that the

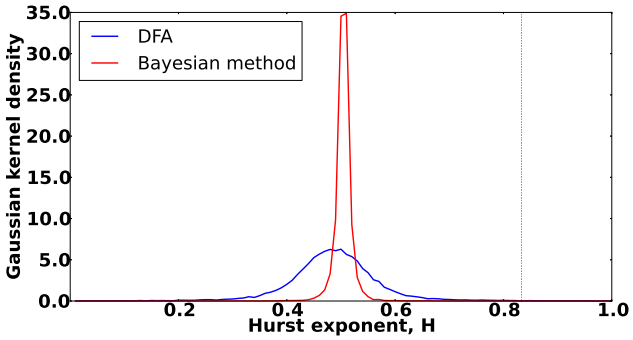

Figure 3.18: Gaussian kernel density of the Hurst exponent for $\alpha = 1.2$ ($H = \frac{5}{6}$), $N = 1,000$ with 25,000 realizations by Bayesian method (red line) and DFA (blue line).

proposed method for the estimation of the Hurst exponent fails to α-stable Lévy motion due to their different nature. Moreover, the Bayesian method has been developed for analysis of time series which do not have large jumps as those in Lévy process. The only case where the proposed estimator can infer the value of the Hurst exponent reliably, even in case of a Lévy process is for value $\alpha \to 2$ ($\alpha = 2$ is a special case of the Gaussian distribution). Therefore, in order to check whether the estimation can be applied to different data sets with unknown nature, I propose the following methods:

a) Taking into account that Bayesian estimator shows the same result (Hurst exponent close to $H = 0.5$) for whole set of $\alpha \in (0, 2)$, one should check the view of the original data set before applying the method. Due to the fact that techniques for detection stable data usually have very little theory be-

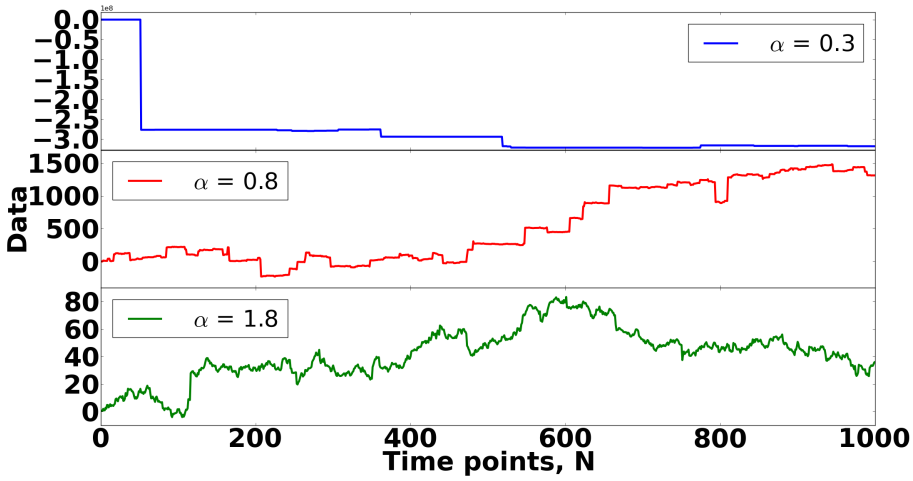

Figure 3.19: α-stable Lévy motion for $\alpha = 0.3, 0.8$ and 1.8.

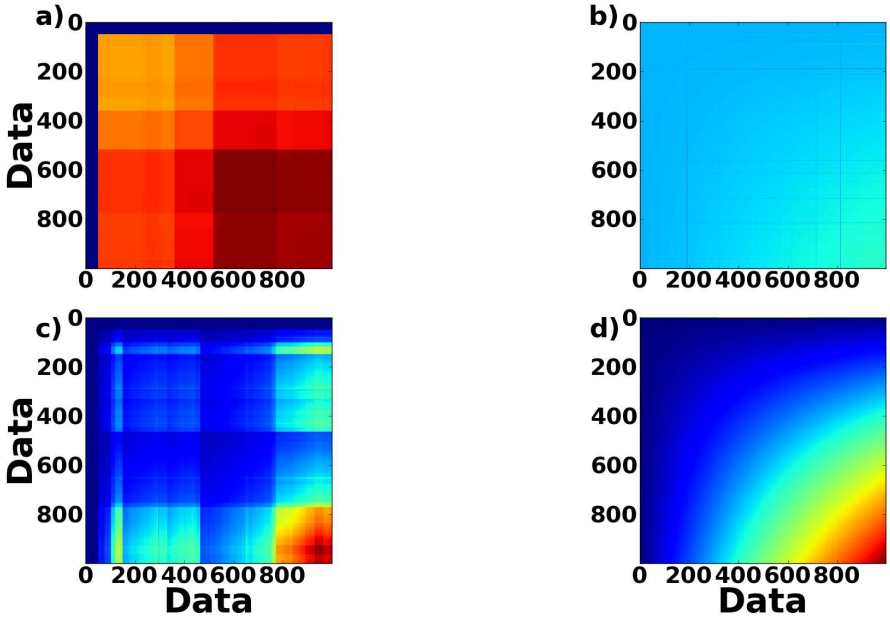

Figure 3.20: Covariance function for α-stable Lévy motion with a) $\alpha = 0.3$, b) $\alpha = 0.8$, c) $\alpha = 1.2$ and d) $\alpha = 1.8$.

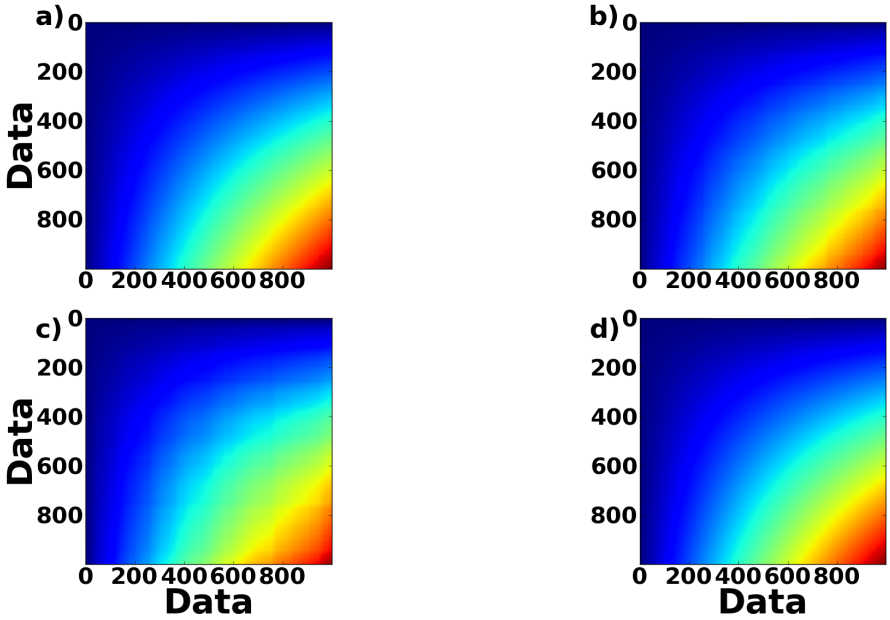

Figure 3.21: Covariance function for fBm with a) $H = 0.2$, b) $H = 0.5$, c) $H = 0.9$ and d) Rosenblatt process with $H = 0.8$.

hind them, the most simplest and most obvious way is, therefore, the visual inspection of the data [7]. The jump Lévy process shows significant different behavior for different α values, as shown in **Fig.** 3.19. Here, the several observations dominate the rest. The only case where the time series of jump Lévy process are similar to fBm is for $\alpha \to 2$, for which the Hurst exponent $H \sim 0.5$. Therefore, one possible method how to detect time series for which the estimator can be applied is to check visually their similarity to time series obtained from fBm.

b) However, based only on the graphical inspection it would be difficult to distinguish between the pure stable case and the case of Gaussian data with the occasional outlier. Therefore, the other qualitative measure is to calculate the covariance function. Such, if the data comes from the population with finite variance, the resulting variance should converge to a finite value. In other case, the variance diverges and its graph has large jumps [7]. **Fig.** 3.20 and 3.21 show the covariance functions for α-stable Lévy motion for $\alpha = 0.3, 0.8, 1.2$ and 1.8 as well as for fractional Brownian motion for $H = 0.2, 0.5, 0.9$ and Rosenblatt process for $H = 0.8$ consequently. It can be seen that the covariance functions of Lévy process is similar to ones of fBm and Rosenblatt processes only in a case $\alpha = 1.8$ or $H \approx 0.55$ (only in case the H-value close to 0.5).

3.5 Discussion

I proposed here a Bayesian approach for the estimation of the Hurst exponent for self-similar processes. Moreover, a formulation in terms of linear mixed models was used to incorporate a linear trend instead of removing it in a first step. This is a necessary step since the real-world data most often contains it [61]. Using synthetic data sets, I performed a comparison study with the detrended fluctuation analysis, method which is most-commonly used for the estimation of the Hurst exponent. I showed that the proposed Bayesian method applied as a point estimator outperforms the DFA. Moreover, the Bayesian method is not dependent on the data sampling, which is especially valuable when working with real-world data.

The Bayesian approach I proposed here is capable of estimating the Hurst exponent from the non-Gaussian data. That was tested on the Rosenblatt process. The results show that the Bayesian method outperforms the standard DFA analysis in this case as well. Additionally, I also discussed possible limitations of proposed method on the α-stable Lévy motion. In the case of Lévy process, the data cannot be well-approximated by a true fBm, and my method, in its present formulation, fails. Therefore, it does not recover the true self-similarity exponent. This limitation has been encountered for the DFA method as well. The case of Bayesian inference for infinite variance stable distributions can be found in work of Ravishanker and Qiou [109].

In addition, I illustrated the performance of my method using the stock market data of the Dow-Jones Industrial Average closing values. The obtained

results indicate that the proposed Bayesian approach outperforms all standard techniques used so far for this data set. Also, the shift of the Hurst exponent to persistent area (thus indicates long-range dependency) is visible. Although it was assumed previously that DJIA values can be well-represented by a Gaussian process, recent studies have shown that such kind of price movements has power low tails. However, it has been also shown that the tails of the power low observed for financial data are more narrow that the ones of Lévy process [15], thus allows me to classify the received estimation values for DJIA as a reliable one.

Chapter 4

Estimation of the Hurst exponent from noisy data

White Gaussian noise remains the same dull hum.

Benoit Mandelbrot

When dealing with real-world data, it is of significant importance to consider the influence of noise. To that extent, I generalize the proposed Bayesian method (*Section 3.1*) in order to evaluate its robustness with respect to the noise intensity contained in the data. Previously, Lennratz *et al.* [76] studied the Gaussian distributed records where the data consists of long-term correlated component characterized by a correlation exponent and a white noise component. However, they considered the data in the absence of short-range correlation and trends. In contrast to this, in [61] the scaling behavior of noisy signals was quantified under effect of different kinds of trends. In this chapter I consider artificial data of fractional Brownian motion under the influence of noise. I evaluate the robustness of the Bayesian estimation with respect to the noise-amplitude ratio. Here, a definite numerical information regarding a model variable - crucial to the estimation task is obtained. Moreover, I identify the corresponding noise-amplitude level that allows to receive the correct estimation of the Hurst exponents in various cases.

4.1 Definition of model

I consider the following model

$$\mathcal{Y} = \lambda u + F\beta + \xi, \tag{4.1}$$

where as before $\lambda > 0$ is the amplitude, $\lambda \in \mathbb{R}$, β represents the fixed effects with system matrix F. The random part of the model $u = B_t^H, \quad u \sim N(0, \mathbf{\Sigma}_H)$

is characterized by the covariance matrix:

$$\mathbf{\Sigma}_H = \mathbf{\Sigma}_{i,j} = \mathbb{E}(B_{t_i}^H B_{t_j}^H) = \frac{1}{2}(|t_i|^{2H} + |t_j|^{2H} - |t_j - t_i|^{2H}).$$

and ξ represents a white Gaussian noise: the random term is assumed to be Gaussian distributed with zero mean $\mathbb{E}(\xi) = 0$ and correlation $\mathbb{E}(\sigma\sigma\prime) = \sigma^2$, ($\sigma$ denotes the noise intensity). In the current case, I consider the following noise distribution: $\xi \sim N(0, \lambda^2\sigma^2 I)$.

The probability density for the observation takes the following form:

$$\mathcal{Y} \sim \mathcal{N}(F\beta, \lambda^2(\mathbf{\Sigma}_H + \sigma^2 I)) = \mathcal{N}(F\beta, \lambda^2(\mathbf{\Omega}_{H,\sigma})), \tag{4.2}$$

with $\mathbf{\Sigma}_H + \sigma^2 I = \mathbf{\Omega}_{H,\sigma}$. Thus, the likelihood function is given via:

$$\mathcal{L}(H, \beta, \lambda, \sigma) = \frac{1}{(2\pi)^{N/2}\lambda^N |\mathbf{\Omega}_{H,\sigma}|^{1/2}} e^{-\frac{(\mathcal{Y}-F\beta)^T \mathbf{\Omega}_{H,\sigma}^{-1}(\mathcal{Y}-F\beta)}{2\lambda^2}}, \tag{4.3}$$

where the $|\mathbf{\Omega}_{H,\sigma}|$ denotes the determinant of the matrix $\mathbf{\Omega}_{H,\sigma}$. Since no knowledge about the parameters of the model exists, I specify the prior information in the following way:

– I set flat priors for the parameters H and β:

$$\mathbb{P}(H) = \chi_{[0,1]}, \;\; \mathbb{P}(\beta) \sim 1 \tag{4.4}$$

– and Jeffreys-like priors [65] for amplitude λ and noise intensity σ:

$$\mathbb{P}(\lambda) \sim \frac{1}{\lambda}, \;\; \mathbb{P}(\sigma) \sim \frac{1}{\sigma}. \tag{4.5}$$

In the light of the Bayes' theorem **Eq.** 2.28, the posterior distribution over the parameter set $H, \beta, \lambda, \sigma$ can be written as:

$$\mathbb{P}(H, \beta, \lambda, \sigma|\mathcal{Y}) = C \cdot \mathcal{L}(H, \beta, \lambda, \sigma) \cdot \frac{1}{\lambda} \cdot \frac{1}{\sigma}, \tag{4.6}$$

where C is a normalization constant.

In order to receive the posterior distribution of the Hurst exponent it is necessary to marginalize out the remaining parameters and calculate the integral:

$$\int \frac{\left((\mathcal{Y} - F\beta^*(H,\sigma))^T \mathbf{\Omega}_{H,\sigma}^{-1}(\mathcal{Y} - F\beta^*(H,\sigma))\right)^{\frac{2-N}{2}}}{\sigma|\mathbf{\Omega}_{H,\sigma}|^{1/2}|F^T\mathbf{\Omega}_{H,\sigma}^{-1}F|^{1/2}} d\sigma. \tag{4.7}$$

Here, $\beta^*(H, \sigma)$ plays the role of the best linear unbiased predictor (BLUP) for the fixed effects:

$$\beta^*(H, \sigma) = (F^T\mathbf{\Omega}_{H,\sigma}^{-1}F)^{-1}F^T\mathbf{\Omega}_{H,\sigma}^{-1}\mathcal{Y}. \tag{4.8}$$

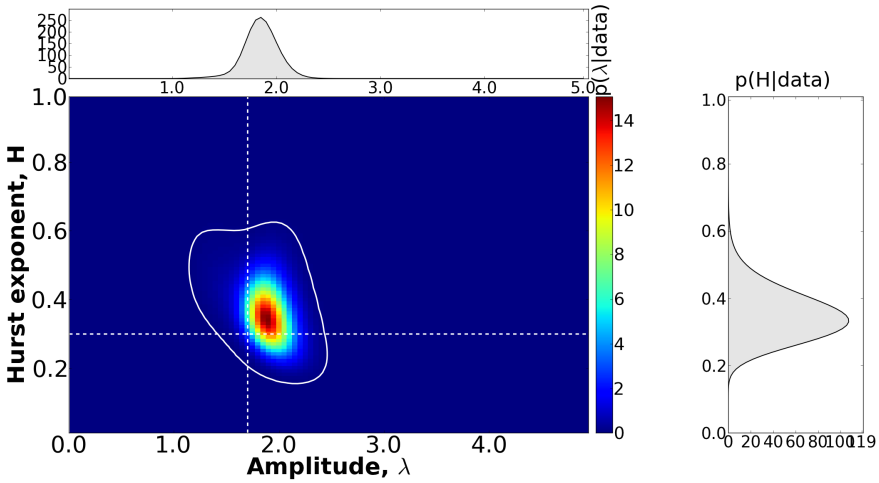

Figure 4.1: The Bayesian estimation of amplitude and Hurst exponent for the model $Y_t = 1.7B_t^{0.3} + \xi$ with $\xi \sim N(0, \lambda^2\sigma^2 I)$, $\sigma = 0.017$ and $\lambda = 1.7$.

Table 4.1: Estimation of the parameters for model (4.9).

Parameter	True value	Estimated value	90% confidence interval
H, Hurst exponent	0.3	0.325	$[0.225; 0.455]$
λ, Amplitude	1.7	1.825	$[1.625; 2.075]$

4.2 Estimation of the Hurst exponent from noisy data

In order to evaluate the proposed model, I generate next fBm data using the following model:

$$Y_t = 1.7B_t^{0.3} + \xi, \tag{4.9}$$

where $\xi \sim N(0, \lambda^2\sigma^2 I)$, with noise-amplitude ratio equal to 1%, $\sigma = 0.017, \lambda = 1.7$ and number of observation points fixed to 100.

Next, I estimate the value of the Hurst exponent from this data using the Bayesian method as discussed in *Section 4.1*. I present in **Fig.** 4.1 the obtained results in (H, λ)-plane, where the contour plots show the one dimensional projections of $p(H|data)$ and $p(\lambda|data)$. The intersection of the white dotted lines represents the true values of the parameters and the white solid curve 90% confidence interval. In **Table** 4.1, the estimated value of the Hurst exponent and amplitude with their confidence intervals of 90% are given. The analysis has shown that with 1% noise-amplitude ratio the received estimation values are in a good agreement with the true values of the parameters.

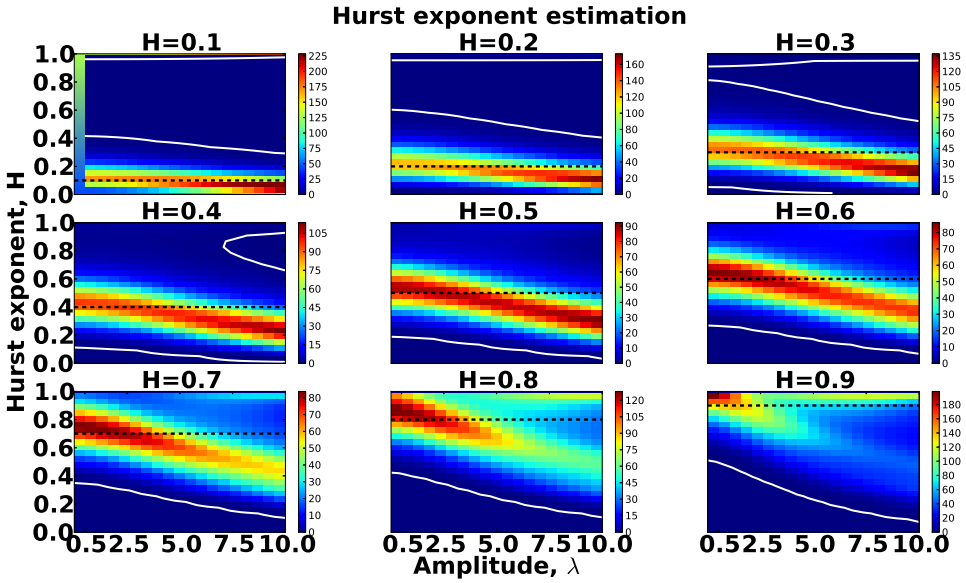

Figure 4.2: Estimation of the Hurst exponent for the model **Eq.** 4.1 with the Hurst exponent varying from 0.1 to 0.9, amplitude $\lambda \in [0.5, 10]$ and 10% noise-amplitude ratio. The number of the observation points is fixed to 100 and the number of different realizations is fixed to 50. The black dotted lines show the true values of the Hurst exponent and the white solid lines show 90% confidence intervals.

4.3 Noise-amplitude ratio

Next, I evaluate the robustness of the proposed Hurst exponent estimator with respect to the noise-amplitude level present in the data. The case studies for several different Hurst values are shown. The aim here is to determine the maximum percentage of the noise intensity according to the amplitude for which a correct estimation of the Hurst exponent from noisy data can be obtained using the proposed Bayesian method.

Thus, I use Monte Carlo simulations to generate 50 realizations with 100 point observations of the model **Eq.** 4.1, where H is varied in the interval $0.1, \ldots, 0.9$ with step 0.1 and $\xi \sim N(0, \lambda^2\sigma^2 I)$. The noise-amplitude ratio here is fixed to 1%, 5% and 10%. The amplitude λ assumes values form the interval $[0.5, 10]$ with 20 equidistant entries. Under this conditions, the estimation of the Hurst exponent depending on the 10% noise-amplitude ratio for the data is given in **Fig.** 4.2. Here, the black dotted lines represent the true value of the Hurst exponent and white solid lines - the corresponding 90% confidence intervals. The projection of the estimator is depicted on the (H, λ)-plane. Such representation of the results allow to emphasize visually the dependency between the noise-amplitude ratio and the value of the estimated Hurst exponent. As shown in the figure, with increasing amplitude value (*i.e.*

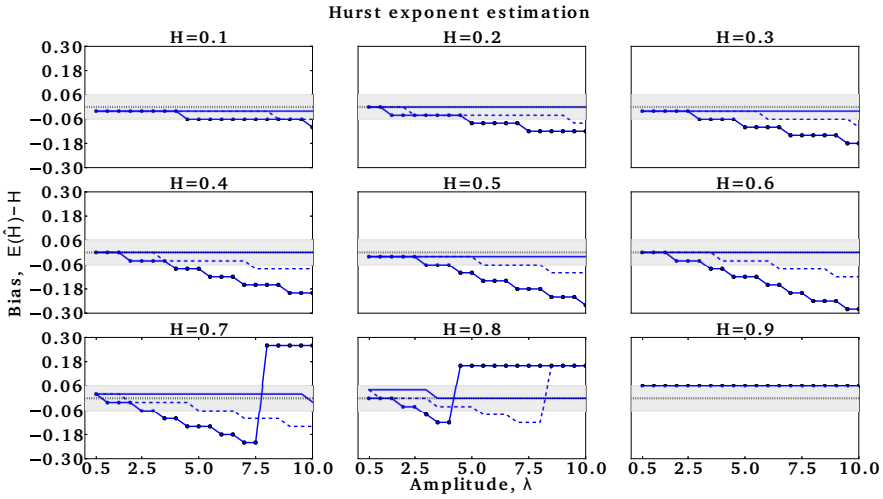

Figure 4.3: Bias of the maximal posterior estimator for the Hurst exponent, $\mathbb{E}(\hat{H}) - H$. The solid lines correspond to the 1% noise-amplitude ratio, dotted lines to 5% and dots represent the bias for 10% noise-amplitude ratio.

noise-amplitude ratio), the difference between the true value of the Hurst exponent and the estimated one increases, and the estimator of the Hurst exponent shifts to the short-range dependence area. Moreover, for highly correlated data ($H \sim 1$), even for small values of the amplitude, the correct estimation of the Hurst exponent is hard to achieve.

The bias of the estimator of the Hurst exponent, $\mathbb{E}(\hat{H})-H$, for all three cases is shown in **Fig.** 4.3. As depicted, the 1% noise-amplitude ratio allows better

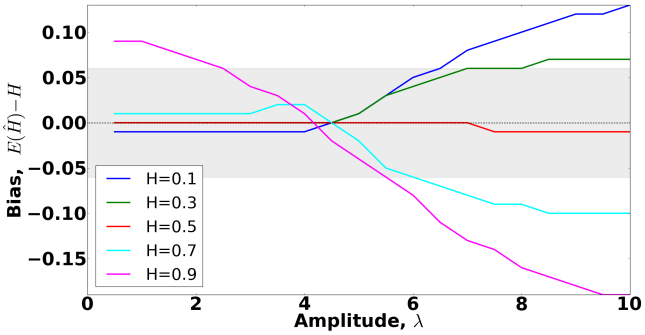

Figure 4.4: Bias of the maximal posterior estimator received with model **Eq.** 3.3 for the Hurst exponent, $\mathbb{E}(\hat{H}) - H$, for the 1% noise-amplitude ratio with the Hurst exponent $H \in [0.1, 0.9]$ with step 0.2, amplitude $\lambda \in [0.5, 10]$. Number of the observation points is fixed to 100 and the number of different realizations is fixed to 50.

Table 4.2: 6%-interval of goodness for the Hurst exponent according to the amplitude λ.

	1% ratio	5% ratio	10% ratio
$H = 0.1$	$[0.5, 10]$	$[0.5, 10]$	$[0.5, 9.5]$
$H = 0.2$	$[0.5, 10]$	$[0.5, 9]$	$[0.5, 4.5]$
$H = 0.3$	$[0.5, 10]$	$[0.5, 9.5]$	$[0.5, 4.5]$
$H = 0.4$	$[0.5, 10]$	$[0.5, 7]$	$[0.5, 3.5]$
$H = 0.5$	$[0.5, 10]$	$[0.5, 8]$	$[0.5, 4]$
$H = 0.6$	$[0.5, 10]$	$[0.5, 6]$	$[0.5, 3]$
$H = 0.7$	$[0.5, 10]$	$[0.5, 6.5]$	$[0.5, 3]$
$H = 0.8$	$[0.5, 10]$	$[0.5, 5]$	$[0.5, 2.5]$
$H = 0.9$	$[0.5, 10]$	$[0.5, 10]$	$[0.5, 10]$

estimation of the Hurst exponent. Moreover, **Fig.** 4.3 shows that for all Hurst exponent values, the model **Eq.** 4.1 can estimate correctly its value (within 6% which depicted with gray color in the figure) for 1% noise-amplitude ratio. Based on these results, I take a 6%-interval as an empirical goodness test for the bias of the Hurst exponent. Thus, I state that the estimator of the Hurst exponent performs well if the bias fulfills the condition: $|\mathbb{E}(\hat{H}) - H| \leq 0.06$.

For different Hurst exponent values, the range of noise-amplitude ratio that satisfy this condition varies. In **Table** 4.2, the 6%-intervals for all three cases of the noise-amplitude ratio are given.

Next, I show that better estimation of the Hurst exponent can be achieved by proposing a method which allows to consider the noise intensity present in the data. For this reason, I compare the estimation of the Hurst exponent when noise in the data is not taken into account (previous deterministic model **Eq.** 3.3) in contrast to the stochastic model **Eq.** 4.1 and show that explicit consideration of the noise intensity has significant advantages. For this purpose, I quantify the bias of the maximal posterior estimator for the Hurst exponent, $\mathbb{E}(\hat{H}) - H$, for the model **Eq.** 3.3 with 1% noise-amplitude ratio. The results are shown in **Fig.** 4.4. With stochastic data, the bias of the estimation of the Hurst exponent lies within 6% of the goodness test. In contrast to this, if the stochasticity of the data is not considered explicitly, *i.e.* in model **Eq.** 3.3, the bias of the estimation of the H is beyond the boundaries. This clearly favors the model **Eq.** 4.1 as more robust with respect to the noise-amplitude ratio.

4.4 Discussion

I have considered for a first time the estimation of the Hurst exponent from noisy data using the Bayesian approach. I show that this method has significant advantages in comparison to the case where noise intensity is not taken into consideration explicitly in the model. Moreover, I have identified

the noise-amplitude ratio for which correct estimation of the Hurst exponent using the Bayesian approach can be performed. Additionally, I also define a goodness test to identify the bias for the estimator. This allowed to show that the Bayesian method is quite sensitive to the noise influence and can detect noise-amplitude ratio even up to 5%.

The results obtained within this study are of significant advantage when working with real-world data, *e.g.* in the geophysical systems [98]. Taking into account that in the case of the deterministic model, the proposed Bayesian estimator outperforms the results obtained with the classical version of DFA, I expect that the same will also hold true in the case when stochasticity of the data is considered explicitly in the model. However, a more detailed comparison of the Bayesian method with advanced versions of DFA [25, 68] remain an open question to be solved.

Chapter 5

Bayesian estimation of the self-similarity exponent from fractional Gaussian noise

> Maybe there is an alternative viewpoint that no one has found because we have become so culturally attached to the house that Gauss built.
>
> Marcus du Sautoy

Many geophysical systems exhibit non-trivial multi-scale correlation structures. In particular, hydrological processes such as rainfall, water levels, runoff, *etc.* are often well-modeled as stationary stochastic processes in discrete time [130]. Therefore, in contrast to the model with fractional Brownian motion from *Chapter 3*, I consider in this chapter the stationary stochastic process of fractional Gaussian noise, which is actually the increment process of fBm. Here I implement the Bayesian approach which provides both estimation of the quantity of interest and a way to account for the uncertainty of its value with respect to the noisy signal, instead of considering its integrated version. The method is tested on synthetic realizations of fractional Gaussian noise processes.

5.1 Fractional Gaussian noise

Fractional Gaussian noise (fGn) is a Gaussian stochastic process $\{G_t^H, t \in \mathbb{R}\}$, that can formally be viewed as a derivative of a fractional Brownian motion $\{B_t^H, t \in \mathbb{R}\}$. It is the discrete process of increments for any fixed $s > 0$ [88, 115]:

$$G_t^H(s) = B_{t+s}^H - B_t^H, \;\; t \in \mathbb{R}. \tag{5.1}$$

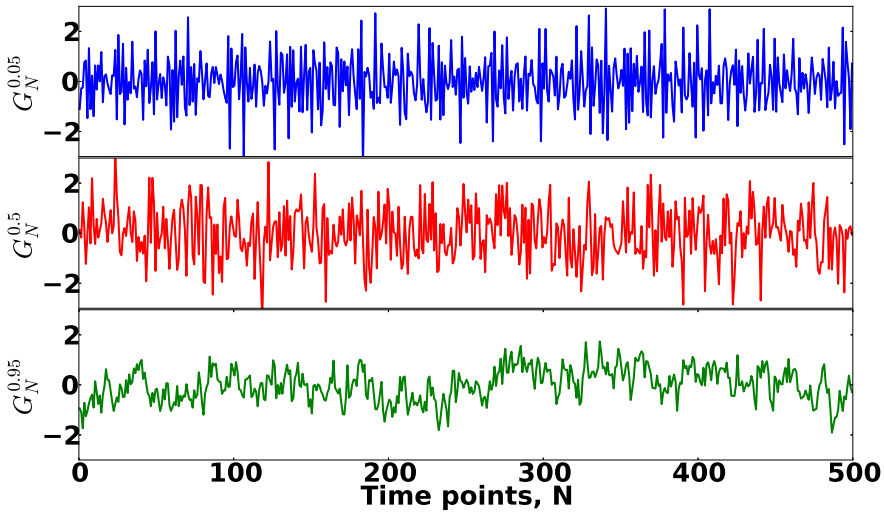

Figure 5.1: Simulation of fractional Gaussian noise of length $N = 500$ with Hurst exponent $H = 0.05$ (blue line), $H = 0.5$ (red line) and $H = 0.95$ (green line).

Any representation of $\{B_t^H, t \in \mathbb{R}\}$ induces a representation for $\{G_t^H, t \in \mathbb{R}\}$. The fGn is called a standard fractional Gaussian noise if $\sigma^2 = VarG_1^H = 1$. Moreover, the fractional Gaussian noise has some remarkable properties [115, 119]:

i) $\{G_t^H, t \in \mathbb{R}\}$ is a stationary Gaussian process;
ii) $\{G_t^H, t \in \mathbb{R}\}$ is a zero mean process: $\mathbb{E}G_t^H = 0$;
iii) $\mathbb{E}(G_t^H)^2 = \sigma^2 = \mathbb{E}(B_1^H)^2$.

The fractional Gaussian noise is characterized by its autocovariance function. Using the correlation of the fBm we see that:

$$\rho(k) = \mathbb{E}(G_{t+k}^H G_t^H) = \frac{\sigma^2}{2}(|k-1|^{2H} - 2|k|^{2H} + |k+1|^{2H}), \ k \in \mathbb{R}. \tag{5.2}$$

Similarly to the fBm, the fractional Gaussian noise can be classified into three subclasses according to the value of H: if $H = 1/2$ then $\rho(k) = 0$; if $0 < H < 1/2$ then $\rho(k) < 0$ and if $1/2 < H < 1$ then $\rho(k) > 0$.

The following consequences are important: for $H = 1/2$, G_t^H has independent identically distributed Gaussian variables and uncorrelated time series after one single time step; for $H \neq 1/2$ however, as the time lag k gets large the correlation function decays asymptotically like

$$\rho(k) \sim \sigma^2 H(2H-1)|k|^{2H-2}, \quad k \to \infty. \tag{5.3}$$

For $1/2 < H < 1$ the correlation function tends to zero slowly so that $\sum_{k=-\infty}^{\infty} \rho(k)$ diverges. This is the long-range dependencies case. On the other hand, if $0 <$

$H < 1/2$, $\sum_{k=-\infty}^{\infty} |\rho(k)| < \infty$ and $\sum_{k=-\infty}^{\infty} \rho(k) = 0$. The first relation follows from fact that $\sum_{k=1}^{\infty} k^{2H-2} < \infty$, and the second is due to the telescoping nature of $\rho(k)$. **Fig.** 5.1 depicts an example of fGn in the case of $H < \frac{1}{2}$, $H = \frac{1}{2}$ and $H > \frac{1}{2}$.

5.2 Definition of the model

The model I define here is similar to the model (**Eq.** 3.3) from *Chapter 3* but based on the fractional Gaussian noise:

$$\mathcal{Y} = \lambda G^H + \beta, \tag{5.4}$$

where G^H is the fractional Gaussian noise, $\lambda > 0$ is the amplitude, and $\beta \in \mathbb{R}$ is just an offset. For fixed parameters, the observations $\mathcal{Y}$ are multivariate Gaussian random variables with the mean value and covariance given by

$$\mathbb{E}(\mathcal{Y}) = F\,\beta, \quad \mathbb{E}(\mathcal{Y}\mathcal{Y}^t) = \mathbf{\Sigma}_H, \tag{5.5}$$

with $\mathbf{\Sigma}_H$ given by **Eq.** 5.2 and $F^t = [1, ..., 1]$ is the vector with N components, where $[.]^t$ denotes a transpose vector.

I apply the Bayes' theorem (**Eq.** 2.28) in order to obtain the following posterior density of the parameters:

$$\mathbb{P}(\lambda, \beta, H|\mathcal{Y}) = C\frac{1}{\lambda^{N+1}|\mathbf{\Sigma}_H|^{1/2}} e^{-\frac{(\mathcal{Y}-F\beta)^t \mathbf{\Sigma}_H^{-1}(\mathcal{Y}-F\beta)}{2\lambda^2}}. \tag{5.6}$$

I use in this case uninformative priors, $\mathbb{P}(\beta, \lambda, H) \simeq \lambda^{-1}$ (a Jeffreys-like prior for the scale parameter λ) since no apriori information about the parameters defining the process are known (*Section 2.2.1*).

The normalization constant C is chosen in order to normalize the posterior probability density (integral equals to 1). Thus, the expression under the exponent function $\left(\frac{(\mathcal{Y}-F\beta)^t \mathbf{\Sigma}_H^{-1}(\mathcal{Y}-F\beta)}{2\lambda^2}\right)$ can be written as the following quadratic polynomial in β

$$F^t \mathbf{\Sigma}_H^{-1} F\,\beta^2 - 2\beta F^t \mathbf{\Sigma}_H^{-1}\mathcal{Y} + \mathcal{Y}^t \mathbf{\Sigma}_H^{-1}\mathcal{Y}. \tag{5.7}$$

From this we see that the posterior can be written in the following "Gaussian" form:

$$\begin{aligned} \mathbb{P}(\lambda, \beta, H|\mathcal{Y}) &= C\frac{1}{\lambda^{N+1}|\mathbf{\Sigma}_H|^{1/2}} e^{-\frac{R^2(H)}{2\lambda^2}} e^{-\frac{\gamma^2(H)(\beta-\beta^*(H))^2}{2\lambda^2}} &(5.8)\\ &= C'\frac{e^{-\frac{R^2(H)}{2\lambda^2}}}{\gamma(H)\lambda^N|\mathbf{\Sigma}_H|^{1/2}}\, f_{\beta^*(H),\frac{\lambda^2}{\gamma^2(H)}}(\beta). &(5.9)\end{aligned}$$

Here f_{μ,σ^2} is the one dimensional Gaussian probability density function with mean μ and variance σ^2. Therefore, the residuum $R^2(H)$ and the mode $\beta^*(H)$ depend on H via

$$\gamma^2(H) = F^t \mathbf{\Sigma}_H^{-1} F \tag{5.10}$$

$$\beta^*(H) = \frac{F^t \mathbf{\Sigma}_H^{-1} \boldsymbol{Y}}{\gamma^2(H)}, \tag{5.11}$$

$$R^2(H) = \mathcal{Y}^t \mathbf{\Sigma}_H^{-1} \mathcal{Y} - \gamma^2(H)\, \beta^*(H)^2. \tag{5.12}$$

Eq. 5.8 describes the full posterior information about all the parameters jointly. In order to obtain the marginal distribution of the parameter we are interested in, it is necessary to treat the other parameters as completely unknown and to integrate over them. For example, integrating over λ, β and H gives correspondingly

$$\mathbb{P}(\beta, H|\mathcal{Y}) = \frac{C_1}{|\mathbf{\Sigma}_H|^{1/2}(R^2(H) + \gamma^2(H)(\beta - \beta^*(H))^2)^{N/2}}, \tag{5.13}$$

$$\mathbb{P}(\lambda, H|\mathcal{Y}) = \frac{C_2}{\gamma(H)\lambda^N |\mathbf{\Sigma}_H|^{1/2}} e^{-\frac{R(H)^2}{2\lambda^2}} \tag{5.14}$$

and

$$\mathbb{P}(\beta, \lambda|Y) = C_3 \lambda^{-N-1} \int_0^1 \frac{e^{-\frac{R^2(H)+\gamma^2(H)(\beta-\beta^*(H))^2}{2\lambda^2}}}{|\mathbf{\Sigma}_H|^{1/2}}\, dH.$$

In the same way, I have obtained the posterior distribution of λ, β and H as

$$\mathbb{P}(H|\mathcal{Y}) = \frac{C_1'}{\gamma(H)|\mathbf{\Sigma}_H|^{1/2}\, R^{N-1}(H)}, \tag{5.15}$$

$$\mathbb{P}(\lambda|\mathcal{Y}) = C_2'\, \lambda^{-N} \int_0^1 \frac{e^{-\frac{R^2(H)}{2\lambda^2}}}{\gamma(H)|\mathbf{\Sigma}_H|^{1/2}} dH \tag{5.16}$$

and

$$\mathbb{P}(\beta|\mathcal{Y}) = C_3' \int_0^1 \frac{e^{-\gamma^2(H)(\beta-\beta^*(H))^2}}{R^N(H)|\mathbf{\Sigma}_H|^{1/2}} dH. \tag{5.17}$$

Thus, the point estimator can be obtained as the location of the maximum of the posterior densities. For instance, I set

$$\hat{H} = \text{argmax}\, \mathbb{P}(H|\mathcal{Y}), \tag{5.18}$$

this in turn yields an estimator for the offset

$$\hat{\beta} = \beta^*(\hat{H}). \tag{5.19}$$

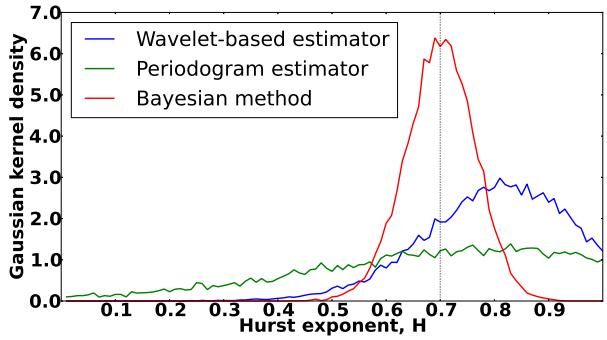

Figure 5.2: Gaussian kernel density of the estimator of the Hurst exponent for $H = 0.7, N = 128$ with $20,000$ realizations by periodogram method (green line), wavelet-based estimator (blue line) and Bayesian method (red line).

5.3 Application to synthetic data

To show how the proposed method works in comparison with other commonly used techniques, I performed $20,000$ realizations by Monte-Carlo simulations for the fixed Hurst exponent $H = 0.7$ and obtained series of length $N = 128$. Moreover, I used the wavelet-based joint estimator [6] and the periodogram estimator [110] to compare with the point estimator driven with Bayesian method (**Eq.** 5.18). In particular, Abry and Veitch in [6] for the first time proposed the explicit formula for the estimator of H based on the wavelet decomposition with wavelet coefficients $d_x(j,k)$ performing the weighted least squares fit between the scales j_1 and j_2 of the wavelet decomposition:

$$\hat{H}(j_1,j_2) = \frac{1}{2}\left[\frac{\sum\limits_{j=j_1}^{j_2} S_j j\eta_j - \sum\limits_{j=j_1}^{j_2} S_j j \sum\limits_{j=j_1}^{j_2} S_j\eta_j}{\sum\limits_{j=j_1}^{j_2} S_j \sum\limits_{j=j_1}^{j_2} S_j j^2 - \left(\sum\limits_{j=j_1}^{j_2} S_j j\right)^2} + 1\right],$$

Table 5.1: The interval for $\hat{H}$ with $p \geq 90\%$ for synthetic data with $H = 0.7$ and $N = 128$.

Method	$\hat{H}$	Interval for $\hat{H}$ with $p \geq 90\%$
Periodogram estimator	0.83	$[0.21, 1.31]$
Wavelet-based estimator	0.81	$[0.57, 1.05]$
Bayesian approach	0.71	$[0.59, 0.81]$
True value	0.7	

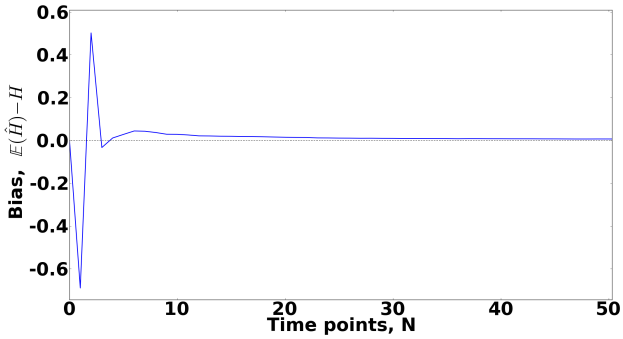

Figure 5.3: The validation test for Bayesian estimation with 500 realizations for the Hurst exponent $H = 0.7$.

where $\eta_j = \log_2(\frac{1}{\eta_j}\sum_k |d_x(j,k)|^2)$ and $S_j = (n\ln^2 2)/2^{j+1}$. This analysis has a low computational complexity, both with respect to time and memory, and allows to test large data sets. In contrast to this, Robinson [110] considered the estimator which regresses the logarithm of periodogram function, $\ln I(\lambda_i)$, on $\ln(2\sin(\lambda_i/2))^2$, where $i = l, l+1, \ldots, n^H$ with l being the lower truncation point which tends to infinity more slowly than n^H.

In the considered example, for each of the realization I calculate the maximum posterior estimator $\hat{H}$ of Bayesian method, wavelet-based joint estimator and periodogram method estimator. The wavelet-based joint estimator is implemented as proposed in [123]. **Fig.** 5.2 shows the comparison between all three methods. Moreover, in **Table** 5.1 the intervals containing $\geq 90\%$ of the distribution are presented. It is clearly visible that the Bayesian method outperforms the wavelet and periodogram methods, since the posterior information is more sharp.

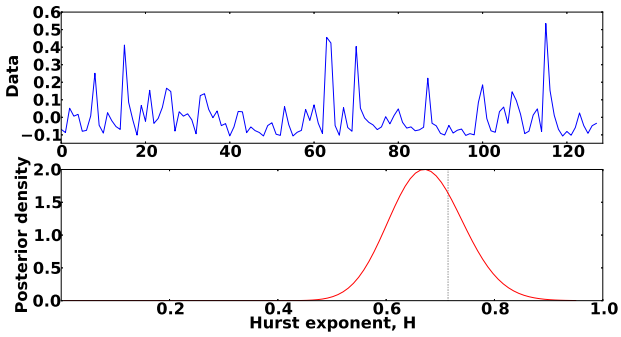

Figure 5.4: (*top*) Simulation of Rosenblatt process for $N = 129$ and $H = 0.7$. (*bottom*) The obtained averaged posterior densities for point estimator drives with Bayesian method for 1,000 realization for the Hurst exponent $H = 0.7$ of length $N = 129$.

Next, in order to make the validation test of the proposed method I quantify the bias of the maximal posterior estimator $\hat{H}$, $\mathbb{E}(\hat{H}) - H$, as a function of the number of data points. For that I generate 500 realizations of fractional Gaussian noise starting from one single observation point. Furthermore, **Fig.** 5.3 showed that even when starting from 20 data points, the bias quickly decays.

In addition, I also implement the Matlab code from [5] for the wavelet-based synthesis of the Rosenblatt process and drive its increments to test how the proposed estimator works on non-Gaussian data (definition of the Rosenblatt process is given in *Section 3.4*). **Fig.** 5.4 shows an example of the increment process for the Rosenblatt process with Hurst exponent $H = 0.7$ for data sets of the length $N = 129$ (*top*) and the obtained averaged posterior densities for the point estimator received with the Bayesian method for 1,000 realizations (*bottom*). The interval containing $\geq 90\%$ of the distribution here is $[0.596, 0.838]$.

5.4 Application to the Nile River data

Next, I apply the described estimator for the Hurst exponent from *Section 5.2* to time series of the annual minimum water level of the Nile River for the years $622-1284$ A.D (663 observations), measured at the Roda Gauge near Cairo. The data is publicly available at StatLib archive: http://lib.stat.cmu.edu/S/beran and depicted in **Fig.** 5.5 a) *(top)*. Moreover, **Fig.** 5.5 a) *(bottom)* shows the first discrete integral of the data, defined as

$$L(k) = \sum_{i=1}^{i=k} (Y_i - \bar{Y}), \quad k = 1, \ldots, N. \tag{5.20}$$

The analysis of this data, as discussed in *Section 5.2*, revealed the following results of the implementation of the Bayesian method as shown in **Fig.** 5.5b)-d). Here the marginalized posterior densities of H, λ and β are depicted. The posterior estimates (maximum of posterior) of the parameters together with their 90% confidence intervals are given in **Table** 5.2.

The Nile River data used in this work was also studied by several authors in order to estimate the Hurst exponent. Thus, I have used this study in order to compare the Bayesian method I propose here with other available methods for estimation of the Hurst exponent. For example, in [78], the modified

Table 5.2: Confidence intervals on the estimated parameters.

Parameter	Estimate	Confidence $\geq 90\%$
$\hat{H}$, Hurst exponent	0.83	$[0.78, 0.87]$
$\hat{\beta}$, Offset	11.51	$[10.9, 12.1]$
$\hat{\lambda}$, Amplitude	0.89	$[0.81, 1.04]$

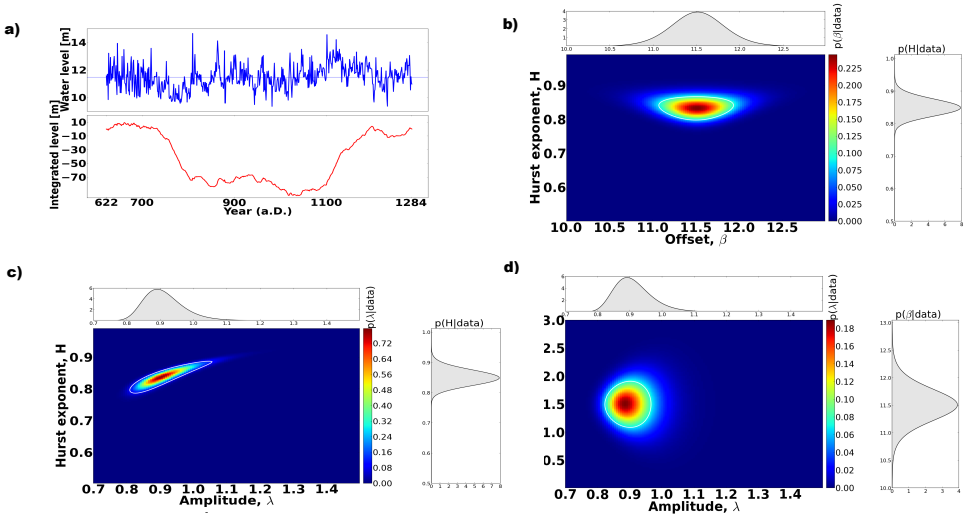

Figure 5.5: a) (*top*) The time series of minimal water levels of the Nile River near Cairo; (*bottom*) the integrated time series. In figures b)-d) are shown the normalized two dimensional marginal posterior densities where the maximum indicate the most likely estimates in: b) $H - \beta$ plain for the Nile River; c) $\lambda - \beta$ plain for the Nile River; d) $H - \lambda$ plain for the Nile River. On the axis, the one dimensional projections of the posterior densities are depicted. The white contour-line encloses 90% of posterior probability. It therefore quantifies the posterior uncertainty of the parameters together with their posterior dependency.

periodogram estimator is used and a comparison with the periodogram estimator, the Whittle estimator, the wavelet maximum likelihood estimator and estimator based on the associated fBm is presented. In **Table** 5.3, the summary of the known results from [78], additionally as the result obtained by Bayesian method are given. As shown, the values obtained by other methods are in the confidence interval of the result obtained here, except for the result obtained by the periodogram estimator. In addition, the Bayesian method provides esti-

Table 5.3: Estimation of the Hurst exponent for different methods.

Estimator	$\hat{H}$
Modified periodogram estimator	0.85
Periodogram estimator	0.90
Whittle estimator	0.84
Wavelet maximum likelihood estimator	0.82
Estimator based on the associated fBm	0.80
Bayesian Method	0.83

mations of all parameters involved in the model and allows also to obtain the error bars for each of them.

5.5 Discussion

In this chapter I proposed a Bayesian estimation technique of the Hurst parameter for the fractional Gaussian noise process. I considered a general model where in addition I took the offset and the amplitude as parameters. This technique, besides the point estimation of the most probable parameter's value, yields as well confidence intervals that enclose a given percentage of the posterior uncertainty. The idea of considering the increment process of fractional Brownian motion is based on the implementation of fractional Gaussian noise to hydrologic data. Moreover, instead of using the integrated process, the method I proposed here allows to estimate the Hurst exponent from the data directly.

In order to test the proposed method to real-world applications, I used it to estimate the Hurst exponent from the time series of the Nile River. The data was already studied with several different techniques to which I additional compared by findings. The Bayesian method I propose here provides similar values for the Hurst exponent as estimated with the contemporary models. However, in contrast to previous findings, the definition of the Bayesian method as given in *Section 5.2* allows also the estimation for the amplitude and the offset on one hand, and the error bars for all estimates, on the other hand. As Bayesian estimation naturally provides the confidence intervals of the estimated value, now it is possible to see how the previous results fulfill the confidence intervals of my estimating results. Except for the periodogram method, all other techniques show results within the confidence bounds. However, periodogram method usually underestimates the value of the Hurst exponent if its expected value is assumed to be less than 0.5 and overestimates it in the opposite case. Moreover, I showed here, that the Bayesian method outperforms the periodogram method and the wavelet-based estimator. Additionally, I checked the increments of the Rosenblatt process, and found that the method proposed in *Section 5.2* has a good performance even for non-Gaussian process and short data sets.

Chapter 6

Estimation of the Hurst exponent from fixational eye movements data

> Our visual system has a built-in paradox. We must fixate our gaze to inspect the minute details of our world; yet, if we were to fixate perfectly, the entire world would fade from view.
>
> Susana Martinez-Conde

Since the 17th century, researchers showed a particular interest to study different aspects of eye movements. Jurin [66] was the first who discovered that the eyes continue to perform miniature movements while human observers a static target. An example of such movements is depicted in **Fig.** 6.1.

Nowadays, video-based recordings of eye movements renewed the scientific interest in these fixational eye movements [41, 89, 90]. It is well-established that different components of fixational eye movements (FEM) are classified as tremor, drift, and microsaccades [35, 40, 91, 111] (see **Fig.** 6.2). The two most important components of fixational eye movements are *(i)* physiological drift, a slow random component that could be characterized as a diffusion process, and *(ii)* microsaccades, rapid small-amplitude movements for which it is unclear if they are generated by the same mechanism as physiological drift or by a separate generating process. Tremor is a small-amplitude, oscillatory component superimposed to drift that cannot be resolved with video-based recording devices used in the current study.

The detection [21, 41, 43], the characterization [1, 2, 21] or modeling of the underlying generating dynamic [22, 59, 103, 104] of FEM has been a topic of many studies.

Recently, Engbert and Kliegl [42] suggested that fixational eye movements might be interpreted as realizations of a fractional Brownian motion (see also [77]). The Hurst exponent, which parameterizes fBm, indicates its correlation behavior and allows the detection of the long-range dependencies inside the signal. Thus, FEM has been investigated so far using the approaches such as DFA and standard deviation analysis (SDA) [8] in work of Mergenthaler and Engbert [95], and a change of the scaling behavior from a persistent behavior on short time scales to anti-persistent behavior on long time scales was observed.

However, due to the advantages of the Bayesian method over DFA for the estimation of the Hurst exponent as discussed in previous sections, here, I aim to apply the Bayesian method for the study of FEM. In particular, I investigate the estimation problem for the Hurst exponent from a sequence of increments of a discrete process of FEM. This allows to examine the Hurst exponent value according to the scale factor. With this approach, the Hurst exponent keeps its value constant for a sampled sequence of fBm at all scale levels. In order to test these properties on the realizations of FEM I examine both, experimental and simulated FEM data. In the first case, the existence of the scale level starting from which the data acquires the anti-persistent behavior, turning the signal into short-memory process, is shown. Additionally, I demonstrate that the method I propose here allows for a reliable estimation of the Hurst exponent based on different trials from the same participant. I also validate the proposed method on simulated time series, mimicking the FEM based on a time-delayed random-walk model [95] and a model based on a self-avoiding walk [44] that generates both slow movements and microsaccades.

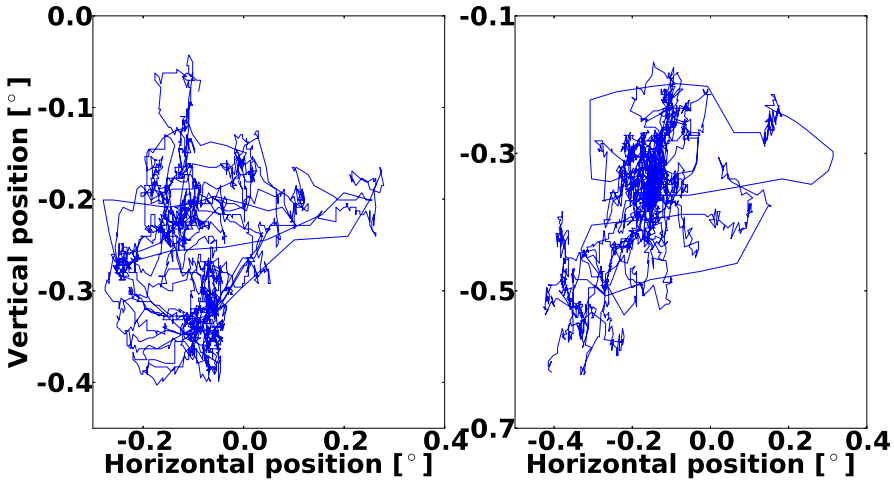

Figure 6.1: Example of fixational eye movements.

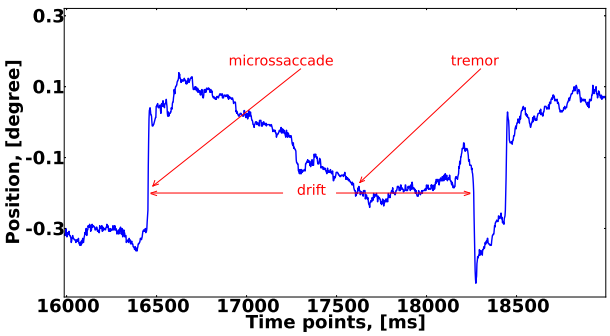

Figure 6.2: Example of fixational eye movements. The most important components drift and microsaccades are quite visible. Tremor is a very small-amplitude oscillatory component.

6.1 Model and method

The studies of the scaling behavior of fixational eye movements by DFA and SDA [42, 95] detected two different types of it on short and long time scales: persistent behavior on the short time scale and anti-persistent on the long one. I approach this estimation task in the following way: as self-similarity of fractional Brownian motion B_t^H implies for fractional Gaussian noise G_t^H that

$$G_{\alpha t}^H(\alpha s) = B_{\alpha t+\alpha s}^H - B_{\alpha t}^H = \alpha^H G_t^H(s). \tag{6.1}$$

Next, considering the discrete process

$$g_{s,n}^H = G_{ns}^H(s) = B_{[n+1]s}^H - B_{ns}^H, \tag{6.2}$$

it follows that $g_{\alpha s,n}^H = \alpha^H g_{s,n}^H$, and thus, all the increment processes are the same up to a multiplicative factor. I therefore take

$$g_n^H = B_{n+1}^H - B_n^H \tag{6.3}$$

as a reference discrete Gaussian increment process, and now apply the Bayes' theorem (as discussed in *Section 5.2*) to the model:

$$\mathcal{Y} = \lambda g^H + \beta, \tag{6.4}$$

where $H \in]0,1[$ is the Hurst exponent, $\lambda > 0$, $\lambda \in \mathbb{R}$ is the amplitude, and $\beta \in \mathbb{R}$ is the offset.

In order to analyze the scale dependency of the Hurst exponent of a sampled signal I propose the following: given a discrete signal d_n, for any discrete scale $a \in \{1, 2, \dots\}$ I consider the sequence of increments over the scale a:

$$D_{a,n} = d_{[n+1]a} - d_{na}. \tag{6.5}$$

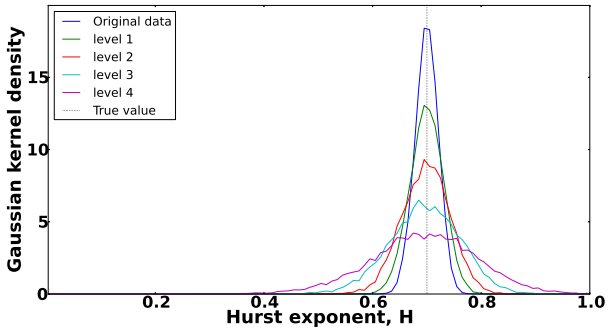

Figure 6.3: Gaussian kernel density of the maximal point estimator of the Hurst exponent obtained by the Bayesian approach for discrete fGn. The original value of the Hurst exponent is fixed to $H = 0.7$.

This can be regarded as a discrete wavelet transform with respect to the rather singular wavelet $\delta_1 - \delta_0$, which justifies the identification of a as a scale [37, 85]. I now propose to analyze the dependency of the Hurst exponent on the level j corresponding to the scale $a = 2^j$ by applying a discrete fGn based Bayesian analysis to each of these signals, as outlined in *Section* 5.2. Note that for a sampled sequence of Brownian motion, the current analysis would yield a constant Hurst exponent across all scales (**Fig.** 6.3). Here, I simulate data of fGn of length $N = 1,000$, and perform $10,000$ realizations using the Monte Carlo simulation method. For each realization I estimate the Hurst exponent value of the subsampled time series according to the **Eq.** 5.15, using a sliding window of length $N = 150$ when the length of the data is above 300, and N equals the data length elsewhere. Next, the data is decomposed up to scale level $j = 5$. Since increasing the level by one reduces the length of the time series to one half, the time series at scale level $j = 5$ corresponds to a length of $N = 32$ data samples. Additionally, the 95% confidence intervals for the estimated Hurst exponent values according to the scale levels are given in **Table** 6.1. Therefore, the analysis provides a mean of discriminate homogeneous, scale-invariant processes for which the correlation structure depends on the scale.

Table 6.1: The 95% confidence interval for $\hat{H}$ for artificial fGn with H=0.7.

Level	95% confidence interval
Original data	$[0.658, 0.741]$
level 1	$[0.639, 0.759]$
level 2	$[0.611, 0.784]$
level 3	$[0.568, 0.824]$
level 4	$[0.502, 0.88]$
level 5	$[0.386, 0.99]$
True value	0.7

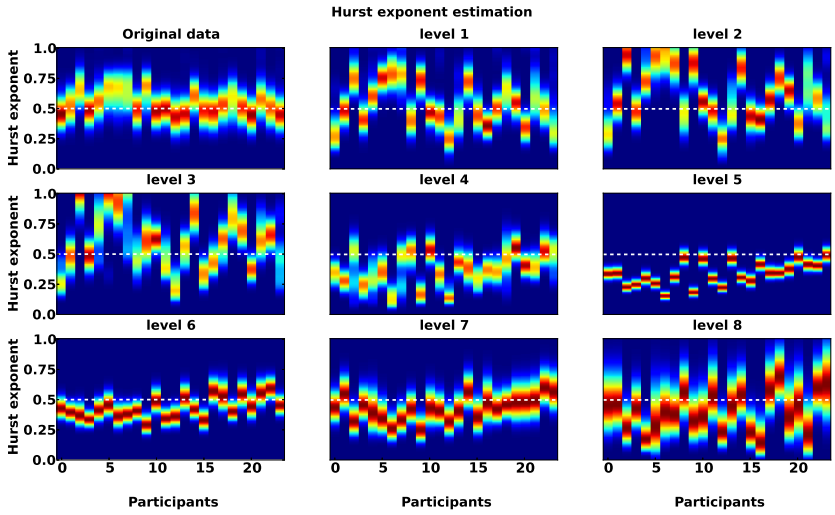

Figure 6.4: Bayesian estimation of the Hurst exponent for time series of fixational eye movements. The analysis applied on different scales corresponding to the subsampling decomposition. White dotted lines denote the value of the Hurst exponent equals 0.5.

6.2 Fixational eye movements data

The method proposed here is next applied to experimental FEM data (detailed information on experimental setup can be found in [21, 43, 95]) for which it has been shown that it can be described significantly well using fBm [42, 77]. The data used for my investigations was collected in one session for each of the 24 participants (average age of 22 years). Participants were asked to fixate a dot of black 3x3 pixels on white background on a computer display. Eye movements have been video-based recorded with an EyeLink II system. Each participant was required to perform 30 trials of 20 seconds each. Further, a check for blinks, which EyeLink II represents as missing data samples, was implemented to reduce the loss of data. As soon as participants performed a blink the trials were restarted. I used 622 trials after visual inspection for data loss based on eye blinking. Next, I removed the microsaccades by an amplitude scale-invariant detection method [21] and investigate only the drift movements.

Although Bayesian estimation can be used without data-preprocessing, in this case, a certain preprocessing step is necessary in order to use the increment series of FEM, since the underlying process in proposed model is fGn. Next, I perform the estimation of the Hurst exponent as described in the previous section, using a sliding window of $N = 150$ samples length (up to scale level 6). This analysis allows me to obtain time series on 8 scale levels in addition to the original time series.

The obtained results are presented in **Fig.** 6.4. Using the method proposed in *Section 6.1*, I have estimated the value of the Hurst exponent as $H \approx 0.5$ for the original, complete time series of FEM. Moreover, for the scale levels

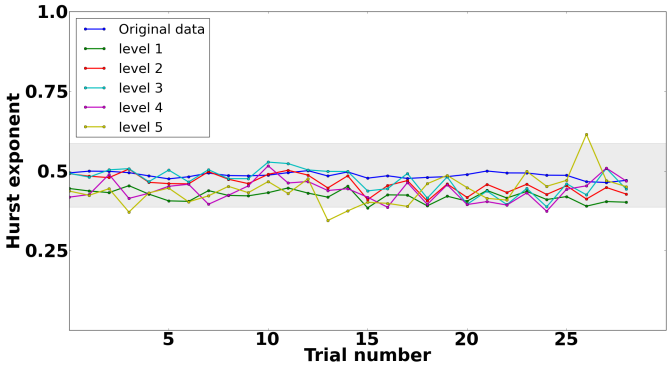

Figure 6.5: Hurst exponent estimation for one of the participants of *Group I*. The gray area denotes the 10% error bar.

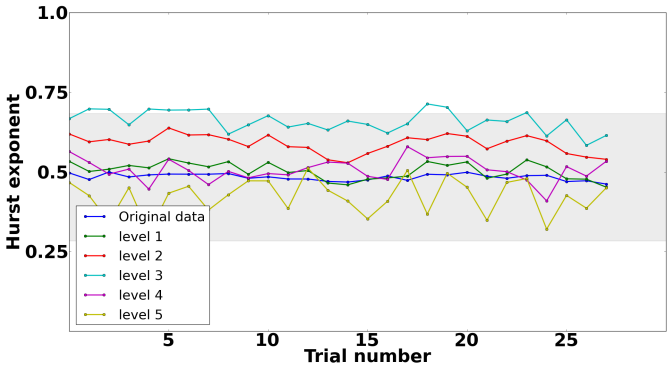

Figure 6.6: Hurst exponent estimation for one of the participants of *Group II*. The gray area denotes the 20% error bar.

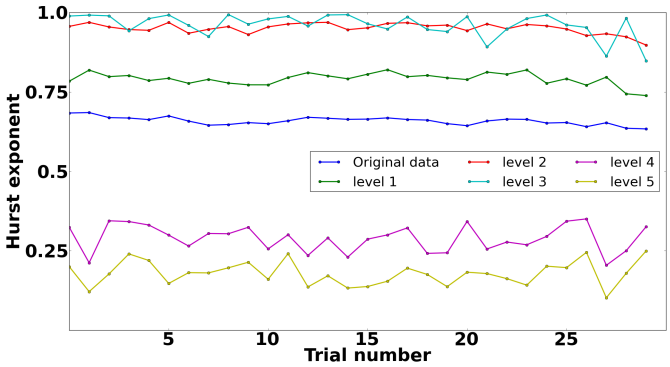

Figure 6.7: Hurst exponent estimation for one of the participants of *Group III*. The independence of the Hurst exponent from the trial is clearly visible.

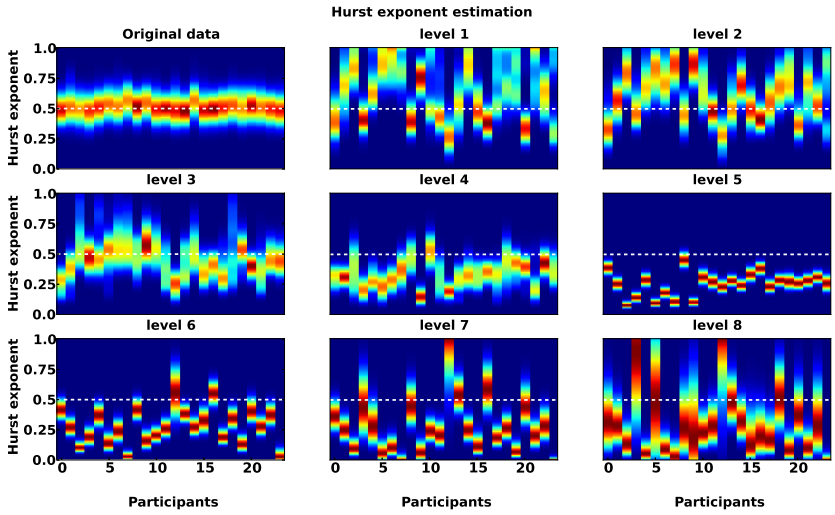

Figure 6.8: Bayesian estimation of the Hurst exponent for time series of fixational eye movements with microssacades. The analysis applied on different scales corresponding to the subsampling decomposition. White dotted lines denote the value of the Hurst exponent equals 0.5.

$j = 1$ to 3, I observe persistent behavior for more than half of the participants, since the Hurst exponent is $H > 0.5$. However, starting from scale level 4, the estimated values of the Hurst exponent makes a significant transition into anti-persistent area ($H < 0.5$). This is especially visible for the data at level 5. For larger scales, the estimator oscillates near the value of $H = 0.5$. However, the location of the maximum of the posterior values of the Hurst exponent is still in the anti-persistent region for the majority of the participants.

As shown previously, the Hurst exponent for fBm does not change its value for different scale levels. However, for FEM data, the observed results can be divided into 3 different groups: *Group I*: the estimation of the Hurst exponent keep its value constant for 3 participants within 10% error over the mean value (this shown in **Fig.** 6.5); *Group II*: the results are qualitatively comparable to the one in **Fig.** 6.5, except instead of 10%, a 20% error interval is necessary to describe changes of the Hurst exponent (this holds for 11 participants, shown in **Fig.** 6.6). In contrast to this the remaining 10 participants (*Group III*) show significantly different behavior, because values of the Hurst exponent for different scale levels are different (see **Fig.** 6.7). However, for participants from each of the 3 different groups, the estimation of the Hurst exponent indicates persistence over different trials. We therefore speculate that the FEM data does not fully resemble the properties of fBm.

Considering the data with microsaccades Mergenthaler in his work [94] found that on short time scales (up to 200 ms), the eye movements are persistent and that for timescales up to 60 ms, the microsaccades generate persistent behavior, where in the time lag between 60 ms and 200 ms, the persistent behavior is caused by drift. On long time scales, he also observed anti-persistent

behavior between 300 ms and 900 ms mainly derived from the drift. In [94] it was shown that there is a strong influence of microsaccades at time scales larger than 900 ms, due to the fact that the anti-persistence in this region weakens when the microsaccades are removed.

The analysis of FEM data with microsaccades by Bayesian approach is presented in **Fig.** 6.8. As it shown, the existence of the microsaccades only slightly influence on the estimated value of the Hurst exponent. From the other hand, that differences do not change the memory of the signal, and the transition into anti-persistent area is still observed at scale level 4. However, for some scale levels (*e.g.* scales 6 and 7) the estimated Hurst exponent values are more dispersed comparing to the case where the microsaccades were removed.

6.3 Hurst exponent estimation from simulated eye movements data

Next, I apply the proposed method **Eq.** 6.5 to analyze simulated time series, generated from the random-walk model of FEM with neurophysiological delays [95] and from the integrated model of slow FEM and microsaccades [44], which is based on a self-avoiding random walk.

6.3.1 Neurophysiological delay model

The neurophysiological delay model is implemented as a discrete map. It is based on the idea to summarize three terms: an autoregressive term, a noise term and a term with negative feedback in order to stabilize fixational eye movements:

$$w_{i+1} = (1-\gamma)w_i + \xi_i - \lambda \ \tan(\epsilon w_{i-\tau}). \tag{6.6}$$

The autoregressive term $(1-\gamma)w_i$ here generates the persistent correlations at the short time scales; the noise term ξ_i is Gaussian noise with $\langle \xi_i \rangle = 0$ and $\langle \xi_i \xi_j \rangle = \sigma^2 \delta_{ij}$; and the term $\lambda \ \tan(\epsilon w_{i-\tau})$ generates the anti-persistent behavior on the long time scale. The parameters in the model are: the feedback strength λ, the parameter for variation of the stepness of the control function ϵ, and the physiological delay τ. The positions of the eye are represented in the following form:

$$x_{i+1} = x_i + w_{i+1} + \eta_i, \tag{6.7}$$

where η_i is an additive noise term with $< \eta_i >= 0$ and $< \eta_i \eta_j >= \rho^2 \delta_{ij}$ and standard deviation ρ. Note that all parameters of the model have direct physiological interpretations. Thus, tonic units activity can be approximated by an additional noise source $< \eta_i >$ added to eye position. The noise term ξ_i represents excitatory burst neurons baseline activity.

For simulation purposes, I use parameters as proposed in [95] ($\gamma = 0.25$, $\lambda = 0.15$, $\sigma = 0.075$, $\rho = 0.35$, $\epsilon = 1.1$, and $\tau = 70$) to produce $1,000$ realizations of length $N = 200,000$ (one iteration step corresponds to 1 ms). Only the

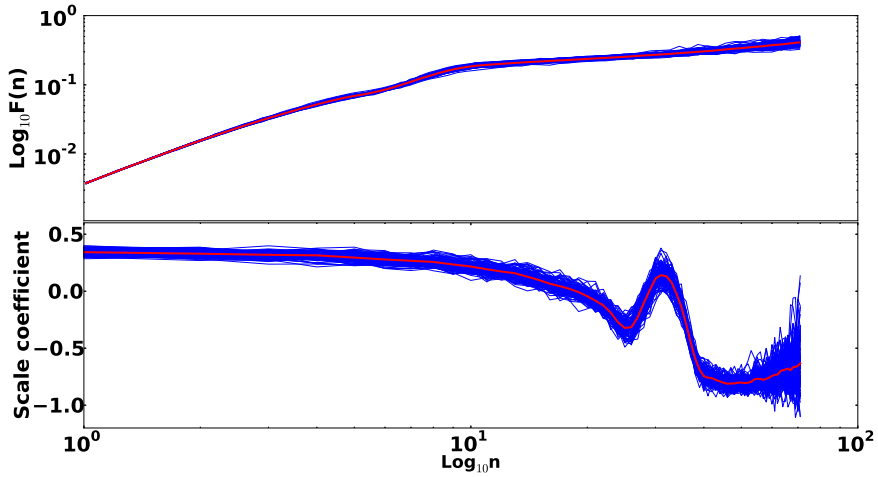

Figure 6.9: *(top)* Scaling behavior estimated via detrended fluctuation analysis. *(bottom)* Slope analysis.

last $10,000$ values were taken into consideration to remove possible transient effects. In **Fig.** 6.9 I present the results for the detrended fluctuation analysis of the simulated data with fix parameters to show the good agreement with one from work of Mergenthaler and Engbert [95].

6.3.2 Integrated model of FEM

The integrated model of FEM as described in [44], is based on the results obtained from neurophysiological delay model. In particular, it is based on the observed persistent behavior on short time scales [95]. Here, the model of FEM is described by a self-avoiding random walk in a quadratic potential due to the fact that FEM are confined to a small area:

$$u(i,j) = \lambda L \left(\left(\frac{i - i_0}{i_0} \right)^2 + \left(\frac{j - j_0}{j_0} \right)^2 \right), \tag{6.8}$$

where the site (i_0, j_0) is the center of the LxL lattice and can be interpreted as the rostral pole of the superior colliculus related to movements with vanishing amplitudes and λ is a slope parameter. In the proposed model the walker always moves to the minimum of the sum of the activation and the potential across the four neighboring sites with position:

$$(i\prime j\prime) = \arg\min_{(k,l)\in N(i,j)}\{h_{kl} + u(k,l)\}, \tag{6.9}$$

where the value h_{kl} is interpreted as the activation of a neuron (k,l) in the motor map of the superior colliculus and $N(i,j) = \{(k,l)|\ k = i \pm 1 \text{ and } l = j \pm 1\}$.

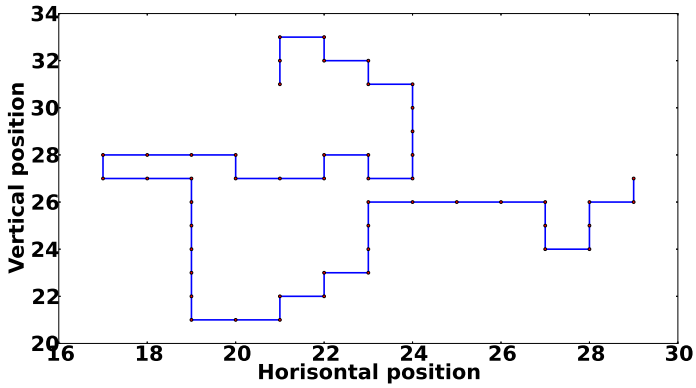

Figure 6.10: A sample trajectory of simulation of the self-avoiding walk in a quadratic potential.

In **Fig.** 6.10 I show an example of the trajectory of simulation of the self-avoiding walk in a quadratic potential on the lattice for the first 50 simulated data points, which do not intersect themself.

I use here as well the parameters as proposed in [44]: the data are simulated on a lattice of 51 cells, the center is at $(25, 25)$ with slope parameter equal to 1. I produce $1,000$ realizations of length $N = 10,000$ each.

In order to apply the Bayesian method for the Hurst exponent estimation in both cases, I calculate the Hurst exponent from subsampled time series of the increment processes as described in *Section 6.1*. The estimation was implemented by a sliding window of width $N = 150$ along the time series. In **Fig.** 6.11, the estimated values for the Hurst exponent at the different decomposition levels are shown. Here, the Hurst exponent values of the original simulated data are the same as for real data of FEM ($H \approx 0.5$). Moreover, for the neurophysiological model, the Hurst exponent keeps this value for all scale levels and do not show any changes in memory. However, this is not fully observed with the integrated model. Here, the transition to the anti-persistent area at level 2 is detected. Such kind of behavior is observed with real data of FEM. Thus, my analysis suggests that the integrated model is compatible with experimental data, since the obtained results show qualitatively similar behavior.

6.4 Discussion

In this chapter I analyzed the scaling exponent from realization of the fixational eye movements data using a Bayesian approach as proposed in *Chapter 5* of this thesis. I applied the Bayesian method to subsampled data of fixational eye movements. This approach allowed me to conclude that the FEM data do not fully resemble the properties of fBm. In particular, the Hurst exponent value of the fBm within the proposed approach keeps its value unchangeable

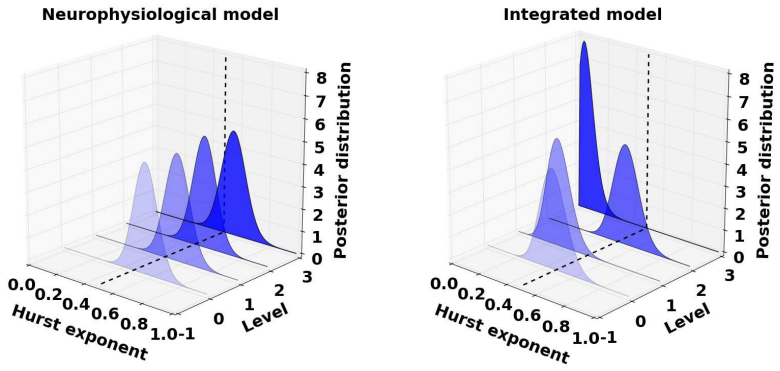

Figure 6.11: Estimation of the Hurst exponent for two simulating models of fixational eye movements: *(left)* model of fixational eye movements data with neurophysiological delays, *(right)* an integrated model of fixational eye movements.

for all the scale levels. However, I observed the same feature only for the half (from 24) participants within the 20% error bar over the mean value.

Additionally, given the results, I found that starting from specific scale level, the Hurst exponent changes its value and the signal turns from persistent to anti-persistent area. These results are observable for both data containing microsaccades and data where microsaccades were removed. My findings are in good agreement with the results discussed in [94], even though the author used there the detrended fluctuation analysis. Moreover, I showed that the values of the Hurst exponent are unique for all participants.

I provide also the validation of the method on simulated models of fixational eye movements data. Here, I analyzed the neurophysiological delay model which is based on the combination of tree components responsible for short-memory on long scales, long-memory on short time scales and noise; and the integrated model based on the self-avoiding walk in potential. The observed results showed that the integrated model demonstrates the behavior of the scaling coefficient in a good agreement with the scalings observed in the real data of fixational eye movements.

Chapter 7

Summary and outlook

> Whenever a new finding is reported to the world people say: It is probably not true. Later on, when the reliability of a new finding has been fully confirmed, people say: OK, it may be true but it has no real significance. At last, when even the significance of the finding is obvious to everybody, people say: Well, it might have some significance, but the idea is not new.
>
> Michel de Montaigne

This work has been focused on the investigation of a novel technique for the estimation of the self-similarity exponent from Gaussian processes, in particular, on the Bayesian approach for estimation of the Hurst exponent. The intention here is to define and analyze an estimation method that provides as outcome a robust and accurate estimator, characterized by an increased efficiency in comparison to other modern methods. In addition, the proposed technique is also compared to several well-established methods, such as the detrended fluctuation analysis, the periodogram method and the wavelet-based estimator. The obtained results showed that the Bayesian method proposed in this thesis yields more precise solutions than the classical estimation methods. The novelties reported in thesis can be therefore summarized as follows:

First, I propose a Bayesian-based method for estimation of the Hurst exponent from fractional Brownian motion in a form of a linear mixed effects model. This allows not only to estimate the Hurst exponent value, but in contrast to previous methods, it allows for simultaneous estimation of the confidence intervals, as well as inference of the remaining parameter values. To that extent, I have shown that:

• the proposed method can be applied successfully to short data sets as well as unevenly sampled data;

• the method outperforms the standard measure of detrended fluctuation analysis even in the case when the full distribution of the posterior probabilities for the possible parameters values received from the detrended fluctuation analysis were used;

• the Bayesian method was additionally evaluated for non-Gaussian data, *i.e* the Rosenblatt process. I have shown that the proposed method provides a good estimation of the Hurst exponent in this case as well. However, similarly to DFA, the Bayesian approach fails to estimate correctly the Hurst exponent from α-stable Lévy motion. This is mainly due to the fact that Lévy motion is a jump process, whereas the Bayesian method is defined for the processes with a Gaussian nature. Moreover, the type of data to which the proposed estimator could be generally applied was discussed in details.

• Bayesian method was also applied to the real-world data of the Dow-Jones Industrial Average values. The results were compared with previous findings. In contrast to former studies, I have shown that a shift of the Hurst exponent to persistent area occurs in the DJIA data.

Second, I have extended the Bayesian estimator of the Hurst exponent in order to take noisy data in consideration as well. To that extent, I have defined a linear mixed effects model with noise term as a basic model to investigate stochastic data. I have shown that when the noise is explicitly considered in the model, the Bayesian estimator provides better inference of the parameters of interest.

Third, I proposed a Bayesian estimation of the Hurst exponent for Gaussian data when using the noise signal directly, instead of taking its integrated process as standard in previous studies. For this purpose, I defined a model for the Hurst exponent estimation based on a fractional Gaussian noise. I further applied this method to study historical data of the Nile River, identifying the Hurst exponent, the amplitude and the offset of the signal within confidence bounds.

Fourth, using the described Bayesian analysis as a basis, I developed an estimation approach which can be regarded as a discrete transform with respect to the singular wavelet decomposition and applied it to analyze data of fixational eye movements. The findings showed several significant properties of the data not identified in the previous studies:

• for each participant of the experiment there exists own value of the Hurst exponent, which remains constant independently on the trial;

• the scale level which changes the memory of the signal was shown constant within reasonable error bounds for all participants;

• previous studies have suggested that the data of fixational eye movements has the same scaling characteristics as the fractional Brownian motion. However, here it was shown that this property is not true in all cases.

• the simulated data of the fixational eye movements based on two different techniques, generated from the random-walk model of FEM with neurophysiological delays and from an integrated model of slow FEM and microsaccades which is based on a self-avoiding random walk, were tested to distinguish the models' properties reliably. The results showed that the integrated model is compatible with the experimental data, displaying qualitatively a similar behavior.

To summarize, this work was focused on the estimation and analysis of the scaling structure of self-similar processes. For this purpose, I proposed an estimation method based on the Bayesian analysis. However, although this work allowed to define the basic properties of a Bayesian-based approach for the estimation of the Hurst exponent in several different cases, *i.e.* considering deterministic case or taking explicitly the stochastic nature of the data in the model, sufficient amount of work still lies ahead. For *e.g.* it is necessary to provide mathematical restrictions to the data-types to which proposed methods could be applied or, more general, to provide a framework to identify which of the proposed methods could be applied when dealing with particular data set. Also, a more comprehensive understanding of the stochastic case requires a detailed comparison between the robustness of the proposed method with the classical or modified versions of detrended fluctuation analysis. Additionally, in the case of fractal processes, it is necessary to have a scale-invariant model based on the wavelet decomposition as a basis for the estimation, rather than the translation-invariant models as standard in the contemporary research.

List of Figures

List of Tables

Appendix

Publications and conference presentations

Publications

N. Makarava, S. Benmehdi and M. Holschneider (2011) *Phys. Rev. E*, 84, 021109

In this study we propose a Bayesian approach to the estimation of the Hurst exponent in terms of linear mixed models. Even for unevenly sampled signals and signals with gaps, our method is applicable. We test our method by using artificial fractional Brownian motion of different length and compare it with the detrended fluctuation analysis technique. The estimation of the Hurst exponent of a Rosenblatt process is shown as an example of a H-self-similar process with non-Gaussian dimensional distribution. Additionally, we perform an analysis with real data, the Dow-Jones Industrial Average closing values, and analyze its temporal variation of the Hurst exponent.

S. Benmehdi, N. Makarava, N. Benhamidouche and M. Holschneider (2011) *Nonlinear Processes in Geophysics*, 18, 441

The aim of this paper is to estimate the Hurst parameter of Fractional Gaussian Noise (FGN) using Bayesian inference. We propose an estimation technique that takes into account the full correlation structure of this process. Instead of using the integrated time series and then applying an estimator for its Hurst exponent, we propose to use the noise signal directly. As an application we analyze the time series of the Nile River, where we find a posterior distribution which is compatible with previous findings. In addition, our technique provides natural error bars for the Hurst exponent.

N. Makarava and M. Holschneider (2012) *European Physical Journal B*, 85(8), 272,

We consider a model based on the fractional Brownian motion under the in-

fluence of noise. We implement the Bayesian approach to estimate the Hurst exponent of the model. The robustness of the method to the noise intensity is tested using artificial data from fractional Brownian motion. We show that estimation of the parameters achieved when noise is considered explicitly in the model. Moreover, we identify the corresponding noise-amplitude level that allow to receive the correct estimation of the Hurst exponents in various cases.

N. Makarava, M. Bettenbühl, R. Engbert and M. Holschneider (2012) *Europhysics Letters*, submitted

In this study we reevaluate the estimation of the self-similarity exponent of fixational eye movements using Bayesian theory. Our analysis is based on a subsampling decomposition, which permits an analysis of the signal up to some scale factor. We demonstrate that our approach can be applied to simulated data from mathematical models of fixational eye movements to distinguish the models' properties reliably.

Conference presentations

N. Makarava and M. Holschneider
Oral presentation at 7^{th} *Conference on Statistical Computation and Complex Systems*, September 2011, Padua (Italy).

N. Makarava, N. Schütz and M. Holschneider
Poster presentation at 45^{th} *American Geophysical Union annual Fall Meeting*, December 2011, San Francisco USA.

Bibliography

[1] R. V. Abadi and E. Gowen. Characteristics of saccadic intrusions. *Vision Research*, 44:2675–2690, 2004.

[2] R. V. Abadi, C. J. Scallan, and R. A. Clement. The characteristics of dynamic overshoots in square-wave jerks, and in congenital and manifest latent nystagmus. *Vision Research*, 40(20):2813–2829, 2000.

[3] R. P. Abelson. On the surprising longevity of flogged horses: Why there is a case for the significance test. *Psychological Science*, 8:12–15, 1997.

[4] P. Abry, P. Flandrin, M. S. Taqqu, and D. Veitch. Self-similarity and long-range dependence through the wavelet lens. In P. Doukhan, G. Oppenheim, and M.S. Taqqu, editors, *Theory and Applications of Long-Range Dependence*, pages 527–556. Birkhäuser, 2002.

[5] P. Abry and V. Pipiras. Wavelet-based synthesis of the Rosenblatt process. *Signal Processing*, 86:2326–2339, September 2006.

[6] P. Abry and D. Veitch. Wavelet analysis of long-range-dependent traffic. *IEEE Transactions on Information Theory*, 44(1):2–15, 1998.

[7] R. J. Adler, R. E. Feldman, and C. Gallagher. Analysing stable time series. In R. J. Adler, R. E. Feldman, and M.S. Taqqu, editors, *A Practical Guide to Heavy Tails: Statistical Techniques and Applications*, pages 133–159. Birkhäuser, 1998.

[8] P. Allegrini, M. Barbi, P. Grigolini, and B. J. West. Dynamical model for DNA sequences. *Phys. Rev. E*, 52:5281–5296, Nov 1995.

[9] S. Arianos and A. Carbone. Detrending moving average algorithm: A closed-form approximation of the scaling law. *Physica A*, 382:9–15, 2007.

[10] A. Arneodo, B. Audit, E. Bacry, S. Manneville, J. F. Muzy, and S. G. Roux. Thermodynamics of fractal signals based on wavelet analysis: application to fully developed turbulence data and DNA sequences. *Physica A: Statistical Mechanics and its Applications*, 254(1-2):24–45, May 1998.

[11] D. Bakan. The test of significance in psychological research. *Psychological Bulletin*, 66(6):423–437, 1966.

[12] J. M. Bardet and P. Bertrand. Definition, properties and wavelet analysis of fractional Brownian motion. *Fractals*, 15:73–87, 2007.

[13] J. M. Bardet and I. Kammoun. Asymptotic properties of the detrended fluctuation analysis of long range dependent process. *IEEE Transactions on information theory*, 54(5):2041 – 2052, 2007.

[14] J. M. Bardet, A. Philippe, G. Oppenheim, M. Taqqu, S. Stoev, and G. Lang. Semi-parametric estimation of the long-range dependence parameter: A survey. In P. Doukhan, G. Oppenheim, and M.S. Taqqu, editors, *Theory and Applications of Long-Range Dependence*, pages 557–577. Birkhäuser, 2003.

[15] M. Bartolozzi, C. Mellen, T. Di Matteo, and T. Aste. Multi-scale correlations in different futures markets. *The European Physical Journal B*, 58:207–220, 2007.

[16] T. Bayes and R. Price. An essay towards solving a problem in the doctrine of chance. *Philosophical Transactions of the Royal Society of London*, 53:370–418, 1763.

[17] I. Belov, A. Kabašinskas, and L. Sakalauskas. A study of stable models of stock markets. *Information technology and control*, 35(1):34–56, 2006.

[18] S. Benmehdi, N. Makarava, N. Benhamidouche, and M. Holschneider. Bayesian estimation of the self-similarity exponent of the Nile River fluctuation. *Nonlinear Processes in Geophysics*, 18(3):441–446, 2011.

[19] J. O. Berger and R. L. Wolpert. *The Likelihood Principle*. Institute of Mathematical Statistics, Hayward (CA), 1988.

[20] Z. D. Berger. Bayesian and frequentist models: legitimate choices for different purposes. *Annals of Epidemiology*, 13:583, 2003.

[21] M. Bettenbühl, C. Paladini, K. Mergenthaler, R. Kliegl, R. Engbert, and M. Holschneider. Microsaccade characterization using the continuous wavelet transform and principal component analysis. *Journal of Eye Movement Research*, 3(5)(1):1–14, 2010.

[22] M. Bettenbühl, M. Rusconi, R. Engbert, and M. Holschneider. Markov models for sequences of microsaccades. Conference Abstract BC11: Computational Neuroscience and Neurotechnology Bernstein Conference and Neurex Annual Meeting 2011. Frontiers in Computational Neuroscience, 2011.

[23] J. C. Breton and J. F. Coeurjolly. Confidence intervals for the Hurst parameter of a fractional Brownian motion based on finite sample size. *Statistical Inference for Stochastic Processes*, 15(1):1–26, 2010.

[24] J. C. Breton, I. Nourdin, and G. Peccati. Exact confidence intervals for the Hurst parameter of a fractional Brownian motion. *Electronic Journal of Statistics*, 3:416–425, 2009.

[25] A. Bunde, S. Havlin, J. W. Kantelhardt, T. Penzel, J.-H. Peter, and K. Voigt. Correlated and uncorrelated regions in heart-rate fluctuations during sleep. *Phys. Rev. Lett.*, 85:3736–3739, 2000.

[26] K. Burnecki and A. Weron. Fractional Lévy stable motion can model subdiffusive dynamics. *Phys. Rev. E*, 82(2):021130, 2010.

[27] B. P. Carlin and T. A. Louis. *Bayes and Empirical Bayes Methods for Data Analysis*. Chapman and Hall/CRC, 2000.

[28] A. M. Churilla, W. A. Gottschalke, L. S. Liebovitch, L. Y. Selector, A. T. Todorov, and S. Yeandle. Membrane potential fluctuations of human t-lymphocytes have fractal characteristics of fractional Brownian motion. *Annal of Biomedical Engineering*, 24(1):99–108, 1996.

[29] K. A. Clarke. Testing nonnested models of international relations: Reevaluating realism. *American Journal of Political Science*, 45(3):724–744, 2001.

[30] E. A. Codling, M. J. Plank, and S. Benhamou. Random walk models in biology. *Journal of the Royal Society Interface*, 5(25):813–834, 2008.

[31] J. J. Collins and C. J. De Luca. Open-loop and closed-loop control of posture: a random-walk analysis of center-of-pressure trajectories. *Experimental Brain Research*, 95(23):308–318, 1993.

[32] P. L. Conti, L. De Giovanni, and M. Naldi. Blind maximum likelihood estimation of traffic matrices under long-range dependent traffic. *Computer Networks*, 54:2626–2639, 2010.

[33] P. L. Conti, L. De Giovanni, S. A. Stoev, and M. S. Taqqu. Confidence intervals for the long memory parameter based on wavelets and resampling. *Statistica Sinica*, 18(2):559–579, 2008.

[34] P. L. Conti, A. Lijoi, and F. Ruggeri. Long-range dependence and performance in telecom networks: Research articles. *Applied Stochastic Models Business and Industry*, 20:305–321, 2004.

[35] T. N. Cornsweet. Determination of the stimuli for involuntary drifts and saccadic eye movements. *Journal of the Optical Society of America*, 46:987–988, 1956.

[36] G. D'Agostini. *Bayesian Reasoning in Data Analysis. A Critical Introduction*. World Scientific Publishing, 2003.

[37] I. Daubechies. *Ten Lectures on Wavelets*. CBMS-NSF Series in Applied Mathematics. SIAM, 1992.

[38] E. Demidenko. *Mixed Models: Theory and Applications*. Wiley Series in Probability and Statistics, New Jersey, 2004.

[39] T. Dieker. *Simulation of Fractional Brownian motion*. PhD thesis, University of Twente, 2004.

[40] R. W. Ditchburn. The function of small saccades. *Vision Research*, 20(3):271–272, 1980.

[41] R. Engbert and R. Kliegl. Microsaccades uncover the orientation of covert attention. *Vision Research*, 43(9):1035–1045, 2003.

[42] R. Engbert and R. Kliegl. Microsaccades keep the eyes' balance. *Psychological Science*, 15:431–436, 2004.

[43] R. Engbert and K. Mergenthaler. Microsaccades are triggered by low retinal image slip. *Proceedings of the National Academy of Sciences of the United States of America*, 103:7192–7197, 2006.

[44] R. Engbert, K. Mergenthaler, P. Sinn, and A. Pikovsky. An integrated model of fixational eye movements and microsaccades. *Proceedings of the National Academy of Sciences*, 108:E765–E770, 2011.

[45] F. Esposti, M. Ferrario, and M. G. Signorini. A blind method for the estimation of the Hurst exponent in time series: theory and application. *Chaos: An Interdisciplinary Journal of Nonlinear Science*, 18(3):033126, 2008.

[46] K. Falconer. *Fractal Geometry: Mathematical Foundations and Applications*. J. Wiley & Sons, New York, 1993.

[47] R. A. Fisher. *Statistical Methods for Research Workers*. Oliver and Boyd, Edinburgh, 1925.

[48] P. Flandrin. Wavelet analysis and synthesis of fractional Brownian motion. *IEEE Transactions on information theory*, 38(2):910–917, 1992.

[49] A. Gelman, J. B. Carlin, H. S. Stern, and D. B. Rubin. *Bayesian Data Analysis*. Chapman & Hall/CRC, 2004. Second edition.

[50] J. Geweke and S. Porter-Hudak. The estimation and application of long memory time series models. *Journal of Time Series Analysis*, 4:221–238, 1983.

[51] M. Gilmore, C. X. Yu, T. L. Rhodes, and W. A. Peebles. Investigation of rescaled range analysis, the Hurst exponent, and long-time correlations in plasma turbulence. *Physics of Plasmas*, 9(4):1312–1317, 2002.

[52] T. Gneiting and M. Schlather. Stochastic models that separate fractal dimension and Hurst effect. *SIAM Review*, 46:269–282, 2003.

[53] A. L. Goldberger, L.A.N. Amaral, J. M. Hausdorff, P. C. Ivanov, C. K. Peng, and H. E. Stanley. Fractal dynamics in physiology: Alterations with disease and aging. *Proceedings of the National Academy of Sciences of the United States of America*, 46(1):2466–2472, 2002.

[54] A.L. Goldberger, L.A.N. Amaral, L. Glass, J.M. Hausdorff, P.C. Ivanov, R.G. Mark, J.E. Mietus, G.B. Moody, C.K. Peng, and H.E. Stanley. Components of a new research resource for complex physiologic signals. *PhysioBank, PhysioToolkit, and Physionet*, Circulation 101(23):e215-e220, 2000.

[55] M. P. Golombek, R. E. Arvidson, J. F. Bell III, P. R. Christensen, J. A. Crisp, L. S. Crumpler, B. L. Ehlmann, R. L. Fergason, J. A. Grant, R. Greeley, A. F. C. Haldemann, D. M. Kass, T. J. Parker, J. T. Schofield, S. W. Squyres, and R. W. Zurek. Assessment of mars exploration rover landing site predictions. *Nature*, 436(7047):44–8, 2005.

[56] G. H. Golub and C. F. Van Loan. *Matrix Computations*. Johns Hopkins University Press, Baltimore, 1983.

[57] D. Grech and Z. Mazur. Statistical properties of old and new techniques in detrended analysis of time series. *Acta Physica Polonica B*, 36(8):2403–2413, 2005.

[58] C. Grimm and G. Schlüchtermann. *IP Traffic Theory and Performance*. Springer, 2008.

[59] Z. M. Hafed. Mechanisms for generating and compensating for the smallest possible saccades. *European Journal of Neuroscience*, 33(11):2101–2113, 2011.

[60] J. M. Hausdorff. Gait dynamics, fractals and falls: Finding meaning in the stride-to-stride fluctuations of human walking. *Human Movement Science*, 26(4):555–589, 2007.

[61] K. Hu, P. C. Ivanov, Z. Chen, P. Carpena, and H. E. Stanley. Effect of trends on detrended fluctuation analysis. *Phys. Rev. E*, 64:011114, Jun 2001.

[62] H. E. Hurst. Long-term storage capacity of reservoirs. *Transactions of the American Society of Civil Engineers*, 116:770–799, 1951.

[63] R. C. Hwa, C. B. Yang, S. Bershadskii, J. J. Niemela, and K. R. Sreenivasan. Critical fluctuation of wind reversals in convective turbulence. *Phys. Rev. E*, 72:066308, 2005.

[64] P. C. Ivanov, L.A.N. Amaral, A. L. Goldberger, S. Havlin, M. G. Rosenblum, Z. R. Struzik, and H. E. Stanley. Multifractality in human heartbeat dynamics. *Nature*, 399:461–465, 1999.

[65] H. Jeffreys. An invariant form for the prior probability in estimation problems. *Proc. R. Soc. Lond. A*, (186):453–461, 1946.

[66] J. Jurin. An essay upon distinct and indistinct vision. *In R.Smith (Ed.): Compleat system of opticks in four books*, 2:115–171, 1738.

[67] S. D. Kang. *Simulation of the fractional Brownian motion*. PhD thesis, J.W.Goethe University, 2008.

[68] J. W. Kantelhardt, E. Koscielny-Bunde, H. H. A. Rego, S. Havlin, and A. Bunde. Detecting long-range correlations with detrended fluctuation analysis. *PHYSICA A*, 295:441, 2001.

[69] J. W. Kantelhardt, S. A. Zschiegner, E. Koscielny-Bunde, S. Havlin, A. Bunde, and H. E. Stanley. Multifractal detrended fluctuation analysis of nonstationary time series. *Physica A: Statistical Mechanics and its Applications*, 316:87–114, 2002.

[70] H. Kettani and J. A. Gubner. A novel approach to the estimation of the Hurst parameter in self-similar traffic. In *Proceedings of the 27th Annual IEEE Conference on Local Computer Networks*, LCN '02, page 0160, Washington, DC, USA, 2002. IEEE Computer Society.

[71] A. N. Kolmogorov. Wienersche Spiralen und einige andere interessante Kurven im Hilbertschen Raum. *Doklady Akad. Nauk SSSR*, 26:115–118, 1940.

[72] L. Lacasa, B. Luque, J. Luque, and J. C. Nuño. The visibility graph: A new method for estimating the Hurst exponent of fractional Brownian motion. *Europhysics Letters*, 86(3):30001, 2009.

[73] J. W. Lamperti. Semi-stable stochastic processes. *Transactions of the American Mathematical Society*, 104:62–78, 1962.

[74] G. F. Lawler. *Introduction to Stochastic Processes*. Chapman & Hall/CRC Probability Series, 2006. Second edition.

[75] W. E. Leland, M. S. Taqqu, W. Willinger, and D. V. Wilson. On the self-similar nature of ethernet traffic (extended version). *IEEE/ACM Trans. Netw.*, 2:1–15, 1994.

[76] S. Lennartz and A. Bunde. Eliminating finite-size effects and detecting the amount of white noise in short records with long-term memory. *Phys. Rev. E*, 79:066101, 2009.

[77] J.-R. Liang, S. Moshel, A. Z. Zivotofsky, A. Caspi, R. Engbert, R. Kliegl, and S. Havlin. Scaling of horizontal and vertical fixational eye movements. *Phys. Rev. E*, 71:031909, 2005.

[78] Y. Liu, Y. Liu, K. Wang, T. Jiang, and L. Yang. Modified periodogram method for estimating the Hurst exponent of fractional Gaussian noise. *Phys. Rev. E*, 80(6):066207, 2009.

[79] A. W. Lo. Long-term memory in stock market prices. *Econometrica*, 59(5):1279–1313, 1991.

[80] M. Maejima and C. A. Tudor. Wiener integrals with respect to the Hermite process and a non-central limit theorem. *Stochastic Analysis and Applications*, 25:1043–1056, 2007.

[81] P. C. Mahalanobis. On the generalized distance in statistics. *In Proceedings National Institute of Science*, 2(1):49–55, 1936.

[82] N. Makarava, S. Benmehdi, and M. Holschneider. Bayesian estimation of self-similarity exponent. *Phys. Rev. E*, 84(2):021109, 2011.

[83] N. Makarava, M. Bettenbühl, R. Engbert, and M. Holschneider. Bayesian estimation of the scaling parameter of fixational eye movements. *Europhysics Letters*, 100:40003, 2012.

[84] N. Makarava and M. Holschneider. Estimation of the Hurst exponent from noisy data: a Bayesian approach. *European Physical Journal B*, 85(8):272, 2012.

[85] S. Mallat. *A Wavelet Tour of Signal Processing*. Academic Press, New York, USA, 1998.

[86] B. B. Mandelbrot. *The Fractal Geometry of Nature*. Freeman, New York, 1983.

[87] B. B. Mandelbrot and J. W. Van Ness. Fractional Brownian motions, fractional noises and applications. *SIAM Review*, 10(4):422–437, 1968.

[88] B .B. Mandelbrot and J. R. Wallis. Computer experiments with fractional Gaussian noises. *Water Resour, Res*, 5:242, 1969.

[89] S. Martinez-Conde, S. L. Macknik, and D. H. Hubel. Microsaccadic eye movements and firing of single cells in the striate cortex of macaque monkeys. *Nature Neuroscience*, 3:251–258, 2000.

[90] S. Martinez-Conde, S. L. Macknik, and D. H. Hubel. The role of fixational eye movements in visual perception. *Nature Reviews Neuroscience*, 5:229–240, 2004.

[91] S. Martinez-Conde, S. L. Macknik, X. G. Troncoso, and D. H. Hubel. Microsaccades: A neurophysiological analysis. *Trends in Neuroscience*, 32:463–475, 2009.

[92] J. A. Matos, S. Gama, H. J. Ruskin, A. Sharkasi, and M. Crane. Time and scale Hurst exponent analysis for financial markets. *Physica A*, 387(15):3910–3915, 2008.

[93] E. J. McCoy and A. T. Walden. Wavelet analysis and synthesis of stationary long-memory processes. *Journal of Computational and Graphical Statistics*, 5:26–56, 1996.

[94] K. Mergenthaler. *The control of fixational eye movements*. PhD thesis, University of Potsdam, 2007.

[95] K. Mergenthaler and R. Engbert. Modeling the control of fixational eye movements with neurophysiological delays. *Physical Review Letters*, 98(13):138104, 2007.

[96] R. Metzler and J. Klafter. The random walk's guide to anomalous diffusion: a fractional dynamics approach. *Physics Reports*, 339(1):1 – 77, 2000.

[97] S. Millen and R. Beard. Estimation of the Hurst exponent for the Burdekin River using the Hurst-Mandelbrot rescaled range statistic. In *Proceedings of the First Queensland Statistics Conference*, 2003.

[98] C. Narteau, J. L. Le Mouël, J. P. Poirier, E. Sepúlveda, and M. G. Shnirman. On a small scale roughness of the core-mantle boundary. *Phys. Earth Planet. Int.*, 191:49–61, 2001.

[99] J. Neyman and E. S. Pearson. On the problem of the most efficient tests of statistical hypotheses. *Philosophical Transactions of the Royal Society A*, 231:289–337, 1933.

[100] J. P. Nolan. *Stable Distributions - Models for Heavy Tailed Data*. Birkhäuser, Boston, 2011. In progress, Chapter 1 online at academic2.american.edu/∼jpnolan.

[101] I. Norros. On the use of fractional Brownian motion in the theory of connectionless networks. *IEEE Journal on Selected Areas in Communications*, 13(6):953–962, 1995.

[102] A. O'Hagan and J. Forster. *Kendalls Advanced Theory of Statistic 2B: Bayesian Inference*. Edward Arnold, London, 2004. Second edition.

[103] J. Otero-Millan, S.L. Macknik, A. Serra, R.J. Leigh, and S. Martinez-Conde. Triggering mechanisms in microsaccade and saccade generation: a novel proposal. *Annals of the New York Academy of Sciences*, 1233(1):107–116, 2011.

[104] J. Otero-Millan, A. Serra, R. J. Leigh, X. G. Troncoso, S. L. Macknik, and S. Martinez-Conde. Distinctive features of saccadic intrusions and microsaccades in progressive supranuclear palsy. *The Journal of Neuroscience*, 31(12):4379, 2011.

[105] C. K. Peng, S. V. Buldyrev, S. Havlin, M. Simons, H. E. Stanley, and A. L. Goldberger. Mosaic organization of DNA nucleotides. *Phys Rev E*, 49:1685–1689, 1994.

[106] C. K. Peng, S. Havlin, H. E. Stanley, and A. L. Goldberger. Quantification of scaling exponents and crossover phenomena in nonstationary heartbeat time series. *Chaos*, 5:82–87, 1995.

[107] B. Pilgram and D. T. Kaplan. A comparison of estimators for 1/f noise. *Physica D: Nonlinear Phenomena*, 114(1-2):108–122, 1998.

[108] H. Raiffa and R. Schlaifer. *Applied Statistical Decision Theory*. Division of research, Graduate school of business administration, Harvard University, Boston, 1961.

[109] N. Ravishanker and Z. Qiou. Bayesian inference for time series with infinite variance stable innovations. In R. J. Adler, R. E. Feldman, and M.S. Taqqu, editors, *A Practical Guide to Heavy Tails: Statistical Techniques and Applications*, pages 259–283. Birkhäuser, 1998.

[110] P.M. Robinson. Log-periodogram regression of time series with long range dependence. *The Annals of Statistics*, 23:1048–1072, 1995.

[111] M. Rolfs. Microsaccades: small steps on a long way. *Vision Research*, 49:2415–2441, 2009.

[112] M. Rosenblatt. Independence and dependence. *In Proceedings of the Fourth Berkeley Symposium on Mathematical Statistics and Probability*, 2:431–443, 1961.

[113] R. M. Royall. The effect of sample size on the meaning of significance tests. *The American Statistician*, 40(4):313–315, 1986.

[114] G. Samorodnitsky. Long range dependence. *Foundations and Trends in Stochastic Systems*, 1(3):163–257, 2006.

[115] G. Samorodnitsky and M. S. Taqqu. *Stable Non-Gaussian Random Processes*. Chapman and Hall, New York, 1994.

[116] S. Stoev, M. S. Taqqu, C. Park, G. Michailidis, and J. S. Marrond. Lass: a tool for the local analysis of self-similarity. *Computational Statistics and Data Analysis*, 50:2447–2471, 2006.

[117] M. Taqqu, V. Teverovsky, and W. Willinger. Estimators for long range dependence: an empirical study. *Fractals*, 3:785–798, 1995.

[118] M. S. Taqqu. Weak convergence to fractional Brownian motion and to the Rosenblatt process. *Z. Wahrscheinlichkeitstheorie verw. Gebiete*, 31:287–302, 1975.

[119] M. S. Taqqu. Fractional Brownian motion and long-range dependence. In P. Doukhan, G. Oppenheim, and M.S. Taqqu, editors, *Theory and Applications of Long-Range Dependence*, pages 5–39. Birkhäuser, 2003.

[120] C. A. Tudor. Analysis of the Rosenblatt process. *ESAIM Probability and Statistics*, 12, 2008.

[121] D. L. Turcotte and L. Greene. A scale-invariant approach to flood-frequency analysis. *Stochastic Hydrology and Hydraulics*, 7(1):33–40, 1993.

[122] J. M. Van Noortwijk, H. J. Kalk, and E. H. Chbab. Bayesian estimation of design loads. *Heron*, 49:189–205, 2004.

[123] D. Veitch and P. Abry. Ldestimate, available at: http://www.cubinlab.ee.unimelb.edu.au/˜ darryl/secondorder_code.html, 2002.

[124] E.-J. Wagenmakers. A practical solution to the pervasive problems of *p*-values. *Psychonomic Bulletin and Review*, 14(5):779–804, 2007.

[125] E.-J. Wagenmakers, M. D. Lee, T. Lodewyckx, , and G. Iverson. Bayesian versus frequentist inference. In H. Hoijtink, I. Klugkist, and P. A. Boelen, editors, *Bayesian Evaluation of Informative Hypotheses*, pages 181–207. Springer: New York, 2008.

[126] A. Weron, K. Burnecki, S. Mercik, and K. Weron. Complete description of all self-similar models driven by Lévy stable noise. *Phys. Rev. E*, 71(1):016113, 2005.

[127] B. J. West, P. Grigolini, R. Metzler, and T. F. Nonnenmacher. Fractional diffusion and Lévy stable processes. *Phys. Rev. E*, 55(1):99, 1997.

[128] P. Whittle. The analysis of multiple stationary time series. *Journal of the Royal Statistical Society. Series B (Methodological)*, 15(1):125 –139, 1953.

[129] L. Xu, P. C. Ivanov, K. Hu, Z. Chen, A. Carbone, and H. E. Stanley. Quantifying signals with power-law correlations: A comparative study of detrended fluctuation analysis and detrended moving average techniques. *Phys. Rev. E*, 71:051101, 2005.

[130] V. M. Yevjevich. *Stochastic Processes in Hydrology*. Water Resources Publications, 1972.

Acknowledgements

I would like to express my appreciation and gratefulness to all the people who have contributed to make this work possible. My deepest gratitude to Prof. Dr. Matthias Holschneider for giving me the opportunity to work in his group, providing constantly his endless support and confidence. Without his guidance and fruitful scientific discussions, this work would not have been possible.

Let me thank the Cognitive Science Program, of the University of Potsdam, especially Prof. Dr. Ralf Engbert for the interdisciplinary discussions, through which I have increased my knowledge in cognition.

I am indebt to the DAAD and the Graduate School NADI, University of Potsdam providing the financial support which made possible my work at the Nonlinear Dynamics group at the University of Potsdam.

I have greatly appreciated the many scientific discussions and fruitful cooperation with my colleagues from the group, providing an inspiring scientific working environment. I am sincerely thankful to Dr. rer. nat. habil. Aneta Koseska for enlightened discussions and scientific advices, leading to a significant improvement of my work. Special thanks to all Faculty staff especially to S. Neiße and V. Gustavs for all the technical assistance I was given throughout this project, for their kindness and friendliness. I would like to thank Prof. Dr. Patrice Abry and Prof. Dr. Vladas Pipiras for providing the Matlab synthesis routines necessary for simulating some data. I am very grateful to my parents and my sister Elena, whose constant loving and unconditional support has encouraged me to move always forward, and for teaching me that kindness and faith are most valuable assets in life.