

Statistical Analysis and Stochastic Modelling of

Atmospheric Boundary Layer Wind

**Dissertation**

zur Erlangung des akademischen Grades  
*Doctor rerum naturalium* (Dr. rer. nat.)  
im Fachbereich C – Mathematik und Naturwissenschaften  
(Fachgruppe Physik)  
an der Bergischen Universität Wuppertal

WUB-DIS 2009-07

vorgelegt von  
**Thomas Laubrich**  
geboren am 07. August 1980 in Karl-Marx-Stadt

**Juni 2009**

erstellt am  
MAX-PLANCK-INSTITUT FÜR PHYSIK KOMPLEXER SYSTEME  
DRESDEN

Diese Dissertation kann wie folgt zitiert werden:

urn:nbn:de:hbz:468-20090797

[<http://nbn-resolving.de/urn/resolver.pl?urn=urn%3Anbn%3Ade%3Ahbz%3A468-20090797>]

Eingereicht am:

2. Juni 2009

Wiss. Betreuer:

Prof. Dr. H. Kantz

2. Gutachter und Prüfer:

Prof. Dr. A. Klümper

Prüfer:

Prof. Dr. R. Koppmann

# Abstract

The present thesis deals with the time series analysis of atmospheric wind velocity recordings by taking into account that the recordings do not reflect ideal turbulence and with the modelling of the wind velocity by stochastic processes.

As a consequence of the poor comprehension of atmospheric turbulence, time series analysis is an appropriate tool to study atmospheric wind. Based on fluctuation statistics and superstatistics, a variety of statistical tools which allow to analyse wind velocity recordings are developed and validated. It is verified that atmospheric boundary layer wind speed data represent in good approximation a succession of ideal turbulence periods with different parameters. The algorithms are able estimate the time evolution of the turbulence parameters. However, the results also indicate the existence of periods in which the wind velocity recording does not reflect turbulence. The focus is not only on the velocity increment but additionally on the wind speed variation around the mean wind speed (which itself is treated as a random variable). It is confirmed that the wind speed fluctuation is roughly normally distributed and that there is a proportionality between the standard deviation of the fluctuation and the mean wind speed.

Motivated by this last result, the finding of a stochastic process which has the same fluctuation statistics as atmospheric boundary layer wind speed is another goal of this thesis. It is shown that the mentioned fluctuation property is not trivial. Nevertheless, the first order geometric auto regressive process has the desired property under certain conditions regarding its parameters. The extent to which this stochastic process is a suitable model for wind velocity simulation is analysed.

# Kurzfassung

Die vorgelegte Dissertation behandelt zum einen die Zeitreihenanalyse atmosphärischer Windgeschwindigkeiten unter der Berücksichtigung, dass atmosphärische Turbulenz nicht idealer Turbulenz entspricht, und zum anderen die Simulation der Windgeschwindigkeit mittels stochastischer Prozesse.

Zeitreihenanalyse ist ein geeignetes Mittel atmosphärische Turbulenz zu studieren. Basierend auf der Fluktuations- und Superstatistik werden statische Verfahren entwickelt und validiert, mit Hilfe dessen Windgeschwindigkeits-Daten analysiert werden können. Es wird gezeigt, dass eine Zeitreihe bodennaher Windgeschwindigkeiten in guter Näherung aus aufeinander folgenden Zeitintervallen besteht, in denen ideale Turbulenz mit unterschiedlichen Parametern wiedergespiegelt wird. Die Algorithmen sind in der Lage, die zeitliche Entwicklung dieser Turbulenzparameter zu erfassen. Aber die Ergebnisse deuten auch auf die Existenz von Intervallen hin, in denen keine Turbulenz wiedergespiegelt wird. Dabei werden nicht nur die Windgeschwindigkeits-Inkremente untersucht, sondern auch die Windgeschwindigkeits-Fluktuationen um den laufenden Mittelwert herum, welcher ebenfalls als Zufallsvariable behandelt wird. Es wird bestätigt, dass die Fluktuationen annäherungsweise normal verteilt sind und dass es eine Proportionalität zwischen der Standardabweichung der Fluktuation und der mittleren Windgeschwindigkeit gibt.

Das letzte Ergebnis ist die Motivation dafür, geeignete stochastische Prozesse zu finden, die die gleiche Fluktuationsstatistik wie atmosphärischer Grenzschichtenwind besitzen. Dieses ist ein weiteres Ziel der vorliegenden Arbeit. Es wird gezeigt, dass die eben genannte Eigenschaft der Fluktuation nicht trivial ist. Der autoregressive Prozess erster Ordnung besitzt hingegen unter bestimmten Bedingungen bezüglich der gewählten Parameter diese Eigenschaft. Das Ausmaß, inwieweit dieser Prozess dazu geeignet ist, bodennahen Wind zu simulieren, wird analysiert.

# Nomenclature

App. ....	appendix
AR ....	auto regressive
ARMA .....	auto regressive moving average
cdf .....	cumulative distribution function
cf. ....	confer
Chap. ....	chapter
e.g. ....	for example, from Latin <i>exempli gratia</i>
Eq. ....	equation
Ex. ....	example
Fig. ....	figure
i.e. ....	that is, from Latin <i>id est</i>
iid .....	independent identically distributed
K41 .....	Kolmogorov 1941
K62 .....	Kolmogorov 1962
MA .....	moving average
mgf .....	moment generating function
NSE .....	Navier-Stokes equation
pdf .....	probability density function
pmf .....	probability mass function
rv .....	random variable
Sec. ....	section
Tab. ....	table
TI .....	turbulence intensity

# Contents

<b>1</b>	<b>Motivation</b>	<b>1</b>
<b>I</b>	<b>Theoretical Background</b>	<b>5</b>
<b>2</b>	<b>Statistics</b>	<b>6</b>
2.1	Probability and random variable . . . . .	6
2.2	Distribution . . . . .	8
2.3	Expected value and other characterisations . . . . .	10
2.4	Transformation . . . . .	12
2.5	Joint distribution and correlation . . . . .	14
2.6	Multivariate transformation . . . . .	17
2.7	Sum of continuous random variables and stable distributions . . . . .	18
2.8	Random sample estimation . . . . .	22
2.9	Kolmogorov-Smirnov test . . . . .	24
<b>3</b>	<b>Stochastic Processes</b>	<b>28</b>
3.1	Stationary processes . . . . .	28
3.2	Autocorrelation function . . . . .	28
3.3	$\delta$ -correlated processes . . . . .	29
3.4	Gaussian processes . . . . .	30
3.5	Transformed Gaussian processes . . . . .	31
3.6	Discretisation of continuous processes . . . . .	33
3.7	ARMA processes . . . . .	34
<b>4</b>	<b>Turbulence Theory</b>	<b>37</b>
4.1	The Navier-Stokes equation . . . . .	37
4.2	Periodic boundary condition and non-locality . . . . .	39
4.3	Conservation laws . . . . .	40
4.4	Reynolds number and turbulence . . . . .	42
4.5	Kolmogorov 1941 theory for infinite Reynolds number . . . . .	44
4.6	Kolmogorov 1962 theory for infinite Reynolds number . . . . .	45
4.7	Increment distribution . . . . .	48

<b>5</b>	<b>Superstatistics</b>	<b>50</b>
5.1	The concept . . . . .	50
5.2	Typical weight functions . . . . .	51
<b>II</b>	<b>Application and Discussion</b>	<b>53</b>
<b>6</b>	<b>Wind Speed Statistics</b>	<b>54</b>
6.1	Taylor hypothesis . . . . .	54
6.2	Non-stationarity and fluctuation statistics . . . . .	55
6.3	Increment statistics . . . . .	62
6.4	Superstatistics . . . . .	68
6.5	Summary . . . . .	77
<b>7</b>	<b>Discussion of the Superstatistical Algorithm</b>	<b>80</b>
7.1	The algorithm in a nutshell . . . . .	80
7.2	Application to an ideal series . . . . .	81
7.3	Application to a series which is not ideal . . . . .	84
7.4	Wind speed increments and conclusion . . . . .	86
<b>8</b>	<b>Fluctuation Statistics of Stochastic Processes and Wind Speed Modelling</b>	<b>87</b>
8.1	Conditioned fluctuation distribution . . . . .	88
8.2	Stationary Gaussian processes . . . . .	89
8.3	$\chi^2$ -distributed white noise . . . . .	93
8.4	Geometric AR(1) process . . . . .	95
8.5	Increment distribution of the geometric AR(1) process . . . . .	99
8.6	Fitting the parameters of the geometric AR(1) process to wind speed data . . . . .	102
<b>9</b>	<b>Summary and Outlook</b>	<b>106</b>
9.1	Summary . . . . .	106
9.2	Outlook . . . . .	109
<b>III</b>	<b>Appendix</b>	<b>112</b>
<b>A</b>	<b>Frequently used Distributions and their Properties</b>	<b>113</b>
A.1	Binomial distribution . . . . .	113
A.2	Gaussian distribution . . . . .	114
A.3	$\chi^2$ -distribution . . . . .	115
A.4	Log-normal distribution . . . . .	115
A.5	Cauchy distribution . . . . .	117
A.6	Lévy distribution . . . . .	117
A.7	Uniform distribution . . . . .	117
A.8	Exponential distribution . . . . .	118

A.9 Interpretation of skewness and kurtosis . . . . .	118
A.9.1 Skewness . . . . .	119
A.9.2 Kurtosis . . . . .	120
<b>B Influence of the <math>T_s</math>-estimation</b>	<b>122</b>
<b>C Derivation of Fluctuation Statistics</b>	<b>129</b>
C.1 Stationary Gaussian processes . . . . .	129
C.2 $\chi^2$ -distributed white noise . . . . .	130
C.3 Geometric AR(1) process . . . . .	132
C.4 Symmetry of the stationary AR(1) process . . . . .	133
<b>Bibliography</b>	<b>141</b>
<b>Acknowledgements</b>	<b>142</b>
<b>Versicherung</b>	<b>143</b>
<b>Index</b>	<b>144</b>



# List of Examples

2.1	Fair coin game (introduction).	7
2.2	Fair coin game (probability).	8
2.3	Fair coin game (pmf/cdf).	9
2.4	Linear transformation of a normally distributed rv.	12
2.5	Linear transformation of a $\chi^2$ -distributed rv.	12
2.6	Mgf of a normal distribution.	13
2.7	Central moments of a normal distribution.	14
2.8	Log-normal distribution.	14
2.9	Correlation between $X \sim N_{0,1}$ and $X^2$ .	16
2.10	Three fully anti-correlated rv's.	17
2.11	Alternative way to derive the univariate transformation rule.	18
2.12	Sum of normally distributed random variables.	20
2.13	Sum of $\chi^2$ -distributed rv's.	20
2.14	Sum of two uniformly distributed rv's.	20
2.15	Sum of iid normally distributed rv's.	21
2.16	Sum of iid $\chi^2$ -distributed rv's.	21
2.17	Special case $S_{2,0,\gamma,\delta}$ (normal distribution).	22
2.18	Special case $S_{1,0,\gamma,\delta}$ (Cauchy distribution).	22
2.19	Special case $S_{1/2,1,\gamma,\delta}$ (Lévy distribution).	22
2.20	Exemplary Kolmogorov-Smirnov test.	26
3.1	Four-point correlation of a Gaussian process.	30
3.2	Standard Gaussian white noise.	30
3.3	Standard Wiener process.	31
3.4	General Wiener process.	31
3.5	Ornstein-Uhlenbeck process.	31
3.6	$\chi^2_1$ -process.	32
3.7	Geometric Ornstein-Uhlenbeck process.	33
3.8	Geometric Gaussian white noise.	33
3.9	Geometric Brownian motion.	33
3.10	Discrete Wiener process.	33
3.11	Discrete Ornstein-Uhlenbeck process.	34
3.12	MA(1) process.	35
3.13	AR(1) process.	35

5.1	Dynamical realisation of superstatistics. . . . .	50
5.2	$\chi^2$ -distributed $\beta$ . . . . .	51
5.3	$\chi^{-2}$ -distributed $\beta$ . . . . .	51
5.4	Log-normally distributed $\beta$ . . . . .	52

# List of Figures

1.1	Global cumulative installed wind power capacity. . . . .	1
1.2	Estimated operating time of wind turbine components. . . . .	2
2.1	Fair coin game's pmf and cdf. . . . .	9
2.2	Mnemonic for the terms platy- and leptokurtic. . . . .	11
2.3	Exemplary Kolmogorov-Smirnov test. . . . .	27
4.1	The periodicity box $B_L$ . . . . .	40
4.2	Pipe flow experiment by Reynolds (1883). . . . .	42
4.3	Intermittency model by constructing a fluctuating mean energy dissipation rate per unit mass. . . . .	46
6.1	The Taylor (1938) hypothesis. . . . .	55
6.2	The distribution of the air speed measured in the free jet experiment by Renner <i>et al.</i> (2001). . . . .	56
6.3	Three 24 h wind speed measurements. . . . .	57
6.4	Conditioned fluctuation histograms. . . . .	59
6.5	Kolmogorov-Smirnov test for the conditioned fluctuation. . . . .	60
6.6	Estimated standard deviation of the fluctuation conditioned on $V$ . . . . .	60
6.7	Estimated standard deviation of the fluctuation conditioned on $V$ . . . . .	62
6.8	Proportionality between the standard deviation of the fluctuation and the mean wind speed. . . . .	63
6.9	The estimated TI as a function of time. . . . .	64
6.10	The estimated standard deviation of the fluctuation as a function of the mean wind speed. . . . .	65
6.11	The TI as a function of time scale $m$ . . . . .	65
6.12	Wind velocity increment histograms. . . . .	67
6.13	The increment distribution parameters as a function of increment length $s$ . . . . .	68
6.14	Schematic sketch of the superstatistical approach. . . . .	69
6.15	Estimation of $\kappa_s^{(m)}$ as a function of $m$ . . . . .	71
6.16	Estimated time scales. . . . .	72
6.17	Superstatistical analysis of a increment series with $s = 8$ . . . . .	73
6.18	Superstatistical analysis of a 2 h resolved diurnal cycle. . . . .	74
6.19	Superstatistical analysis of a increment series with $s = 8$ . . . . .	75
6.20	Superstatistical analysis of a 2 h resolved diurnal cycle. . . . .	76

6.21	Superstatistical analysis of an increment series with $s = 8$ . . . . .	77
6.22	Superstatistical analysis of a 2 h resolved diurnal cycle. . . . .	78
7.1	Estimation of $\kappa^{(m)}$ as a function of $m$ for three ideal processes. . . . .	82
7.2	Value of the Kolmogorov-Smirnov test variable as a function of $T$ . . . . .	82
7.3	Estimation of $\mu_{\ln\beta}$ and $\sigma_{\ln\beta}^2$ as a function of $T$ . . . . .	83
7.4	Estimation of $\kappa^{(m)}$ as a function of $m$ for three processes which are not ideal. . . . .	84
7.5	Value of the Kolmogorov-Smirnov test variable as a function of $T$ . . . . .	85
7.6	Estimation of $\mu_{\ln\beta}$ and $\sigma_{\ln\beta}^2$ as a function of $T$ . . . . .	85
8.1	Estimated mean and standard deviation of the fluctuation conditioned on $\bar{X}_n^{(m)} = \bar{x}_n$ for a generated AR(1) series. . . . .	91
8.2	The $m$ -dependence of $\phi_m/\theta_m^2 - 1$ and $\sigma\sqrt{\theta_m^2 - \phi_m^2}/\theta_m$ for a stationary AR(1) process. . . . .	92
8.3	Estimated mean and standard deviation of the fluctuation conditioned on $\bar{X}_n^{(m)} = \bar{x}_n$ for generated $\chi^2$ -distributed white noise. . . . .	94
8.4	The $m$ -dependence of $\alpha(m, k)$ . . . . .	95
8.5	Estimated mean and standard deviation of the fluctuation conditioned on $\bar{X}_n^{(m)} = \bar{x}_n$ for a generated geometric AR(1) series. . . . .	98
8.6	The $m$ -dependence of $\alpha(m, b)$ . . . . .	99
8.7	Estimation of the fluctuation distributions conditioned on mean $\bar{x}$ . . . . .	104
8.8	Estimation of the increment distributions with increment length $\tau = 1$ s. . . . .	104
8.9	Estimation of the turbulence intensity as a function of time scale $T$ . . . . .	105
8.10	Estimation of $\beta_\tau$ and $\lambda_\tau^2$ as a function of $\tau$ . . . . .	105
A.1	Skew normal distribution. . . . .	119
A.2	Generalised normal distribution. . . . .	120
B.1	The test value of the Kolmogorov-Smirnov normality test for the $\Lambda$ -series as a function of $T_s$ . . . . .	123
B.2	The test value of the Kolmogorov-Smirnov normality test for the $\Lambda$ -series as a function of $T_s$ . . . . .	125
B.3	Estimation of the parameters $\Lambda_s$ and $\lambda_s^2$ as a function of the value of $T_s$ . . . . .	126
B.4	Comparison between the parameters estimated by the Queiros (2007) method and those estimated by the Castaing <i>et al.</i> (1990) method. . . . .	126
B.5	Comparison between the parameters estimated by the Queiros (2007) method and those estimated by the Castaing <i>et al.</i> (1990) method. . . . .	127
B.6	Comparison between the parameters estimated by the Queiros (2007) method and those estimated by the Castaing <i>et al.</i> (1990) method. . . . .	127

# List of Tables

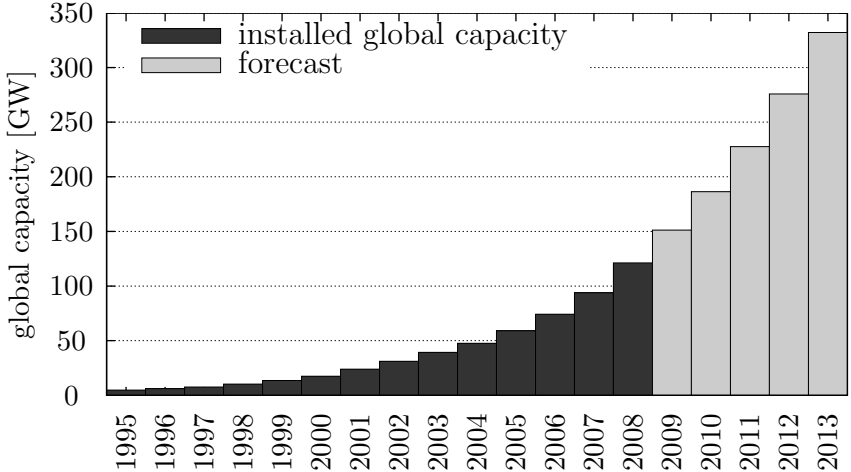
2.1	Properties of the probability function. . . . .	7
2.2	Error types in hypothesis testing. . . . .	25
4.1	Dimension of the quantities in the NSE. . . . .	39
4.2	Properties of the average of periodic functions. . . . .	41
4.3	Summary of cascade models. . . . .	47

# Chapter 1

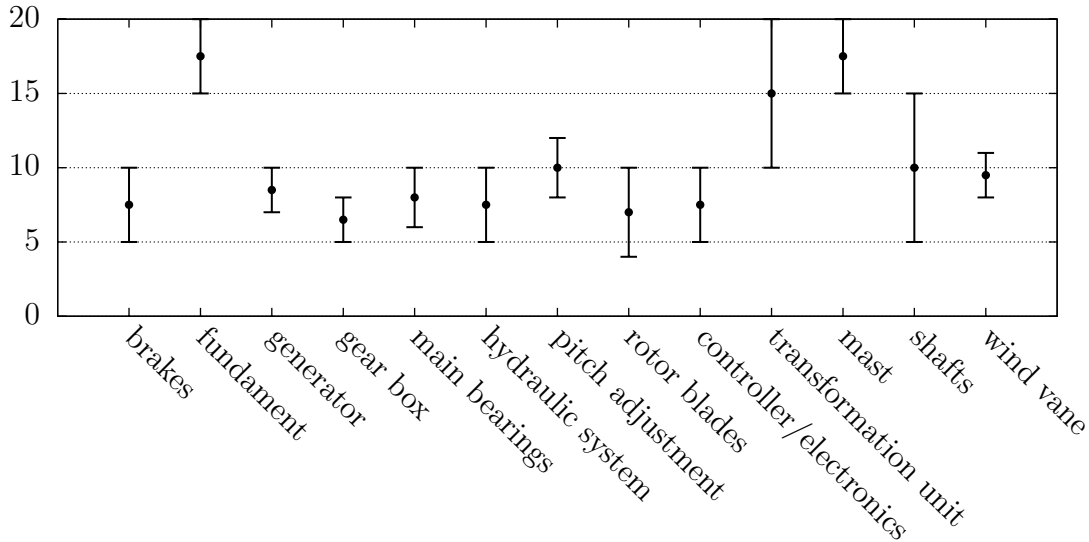
## Motivation

The global demand for energy is increasing at breathtaking pace. In particular, developing countries such as China and India show a sharp increase. A serious investment in new power sources and grid infrastructure is essential because recovering fossil fuels becomes gradually more difficult and thus more expensive. Pullen *et al.* (2008b) claim that wind power has the advantage that it can be deployed faster than any other energy supply technology. Even large offshore wind farms, which require a great level of infrastructure and an elaborate grid network connection, can be installed from start to finish in less than two years. This compares favourably with the much longer time scale for conventional power stations such as nuclear reactors. Fig. 1.1 visualises that the global cumulative installed capacity of wind power increased from 5 GW in 1995 to 120 GW in 2008. Pullen *et al.* (2008a) forecast 330 GW in 2013.

The cost efficient use of wind power is not only a challenging task for engineers but also for natural scientists. Neumann *et al.* (2002) analysed the breakdown reasons of 1500 wind power plants between 1999–2002. The wind turbines were designed to work for twenty years, but Fig. 1.2 shows that most components need to be replaced or need intensive maintaining after five to ten years of operation. It appears that the energy resource, i.e., boundary



**Figure 1.1:** Global cumulative installed capacity and forecast by Pullen *et al.* (2008a).



**Figure 1.2:** Estimated operating time after which an intensive maintaining is required. The study is published by [Neumann \*et al.\* \(2002\)](#).

layer wind, is responsible for the attrition. Apparently, the dynamic loads caused by wind field fluctuations were underestimated because the load estimation procedures for the design of wind turbines often do not accurately treat the statistical nature of loads (see [Veers & Butterfield, 2001](#)). A better understanding of the statistical properties of boundary layer turbulence under realistic conditions is indispensable because realistic models of input wind fields are needed for numerical simulations estimating the loads on structures and their expected lifetimes. This includes the need for good statistical evidence of extreme wind gusts, their relative frequency, their spatial extension, and their temporal correlations.

The fundamental equation of turbulence, the Navier-Stokes equation, is known since [Navier \(1823\)](#). Still today, more than 180 years later, an understanding of solutions to this equation remains incomplete. In 2000, the Clay Mathematics Institute declared the proof of the existence and smoothness of the Navier-Stokes solutions in  $\mathbb{R}^3$  (see [Fefferman, 2000](#)) as an unsolved problem with US\$ 1,000,000 being awarded for a correct solution. Starting with the pioneering experiments by [Reynolds \(1883\)](#), studying the phenomenology of a turbulent flow has by now a long tradition. For instance, the “Album of Fluid Motion” by [Van Dyke \(1982\)](#) displays photographs of the flow around an obstacle and the wake behind it.

Back in the early 20<sup>th</sup> century, [Richardson \(1920\)](#) studied atmospheric turbulence. “Does the wind possess a velocity?” seems to be a foolish question at first sight. Yet, it was posed by [Richardson \(1926\)](#) because the limit of  $\Delta x/\Delta t$ , where  $\Delta x$  labels the displacement of an air parcel in  $x$ -direction over time  $\Delta t$ , might not exist for  $\Delta t \rightarrow 0$ . Hence, a statistical interpretation of turbulence was required. The foundation was laid by the works of [Kolmogorov \(1941a,b,c, 1962\)](#). Since then, studies about the dissipation rate by e.g. [Batchelor & Townsend \(1949\)](#), the spatial correlations by e.g. [Anselmet \*et al.\* \(1984\)](#); [Meneveau & Sreenivasan \(1991\)](#); [Vincent & Meneguzzi \(1991\)](#), the increment statistics by e.g. [Castaing](#)

*et al.* (1990), the Markov properties by e.g. Renner *et al.* (2001), or the life time statistics of coherent structures by e.g. Schneider & Eckhardt (2008) revealed a lot of insight into the properties of turbulence (see also the overview by Sreenivasan & Antonia, 1997). Naturally, this list is far away from being complete.

However, the transition regimes between isotropic turbulence as one idealisation and laminar flow as another idealisation is still poorly understood. 125 years after Reynolds (1883) made his famous pipe flow experiment, Eckhardt & Schneider (2008) addressed the issue: “How does a flow in a pipe become turbulent?”. Recent studies by Eckhardt *et al.* (2008) and Schneider *et al.* (2008) make a progress to answer this question but the final solution is still due.

Similarly, studying and understanding the turbulence outside the laboratory, more specifically, in the atmospheric boundary layer (ABL), is a demanding task for scientists. The air flow in the ABL, i.e., in the lowest few 1–2 km of the atmosphere (see e.g. Stull, 1988; Wallace & Hobbs, 2006), is strongly influenced by surface roughness and hence by orography and land use, but even more by thermal effects through heating from the ground. The intensity of solar radiation and the prevailing surface wind direction change much faster than large scale pressure differences which generate the overall wind conditions. Summarising, the turbulence generating mechanism changes with time. Even though studies about the statistical analysis of ABL wind velocity can be found in the literature, for instance the works by Beck *et al.* (2005a), and Kholmyansky *et al.* (2007), it is the work by Boettcher (2005); Boettcher *et al.* (2007) which explicitly takes into account that ABL turbulence is not ideal turbulence. Nonetheless, the authors used the approach of a Weibull distributed mean wind speed. This distribution is frequently found empirically and is hence an accepted fact (see Burton *et al.*, 2004) but there is no theoretical justification for it.

As a consequence of the poor comprehension of atmospheric turbulence, time series analysis is an appropriate tool to study atmospheric wind. This rises the question: “How to deal with a time series which was recorded in the ABL and which does not reflect ideal turbulence?”. The present thesis is intended to give (a few) answers to this problem. After reviewing the theoretical background of statistics, stochastic processes, ideal turbulence, and superstatistics in Chaps. 2, 3, 4, and 5, respectively, Chaps. 6 and 7 develop and validate a variety of statistical tools which allow to analyse wind velocity recordings. It is verified that ABL wind speed data represent in good approximation a succession of ideal turbulence periods with different parameters. Additionally, the algorithms are able to indicate the existence of periods in which the wind velocity recording does not reflect turbulence at all. The focus is not only on the velocity increment as in Boettcher (2005); Boettcher *et al.* (2007) but also on the wind speed variation around the mean wind speed (which itself is treated as a random variable). It is confirmed that the wind speed fluctuation is roughly Gaussian and that there is a proportionality between the standard deviation of the fluctuation and the mean wind speed.

Motivated by this last result, Chap. 8 is dedicated to finding a stochastic process which has the same fluctuation statistics as ABL wind speed. After formulating a recipe which computes the fluctuation distribution conditioned on the value of the moving average for any



given stochastic process, this technique is applied to a diversity of processes showing that the mentioned fluctuation property is not trivial. Nevertheless, it is shown that the first order geometric auto regressive process has the desired property under certain conditions regarding its parameters. Moreover, it is shown that this process has increment statistics similar to ABL wind speed so that this process is a suitable candidate to simulate wind velocity during a period in which its properties are close to ideal turbulence. Finally, the problem of how to fit the parameters to real data is discussed.

Chap. 9 summarises the results of the presented thesis. Additionally, it proposes further studies.

**Part I**

**Theoretical Background**

# Chapter 2

## Statistics

A first question which comes to mind when dealing with statistics is: What is probability? A mathematical definition of a belief in the chance of an event occurring is quite difficult. [Casella & Berger \(2002\)](#) writes

Unfortunately, matters are not that simple. There are some technical difficulties to overcome. We will not dwell on these technicalities; although they are of importance, they are usually of more interest to probabilists than to statisticians.

Guided by [Casella & Berger \(2002\)](#), which if not stated otherwise serves as reference, this chapter is intended to explain the basics of probability theory without going into the details of these technicalities. An at least weak familiarity with probability theory is required for a firm understanding of statistics.

### 2.1 Probability and random variable

Before defining the meaning of probability, some basics are introduced.

A *random experiment* is an experiment whose outcome cannot be determined in advance. The set of all possible outcomes is called the *sample space*  $S$ . An *event*  $E \subset S$  is said to occur if the outcome of the experiment is an element of  $S$ . Working with subsets of  $S$  is simplified by defining the term *sigma algebra* (or *Borel field*)  $\mathcal{B}(S)$ . It denotes a collection of subsets of  $S$  and has the following properties. It

- contains the empty set, i.e.,  $\emptyset \in \mathcal{B}(S)$ ,
- is closed under complementation, i.e.,  $E \in \mathcal{B}(S) \Rightarrow E^c \in \mathcal{B}(S)$ , and
- is closed under countable unions, i.e.,  $E_1, E_2, \dots \in \mathcal{B}(S) \Rightarrow \bigcup_{i=1}^{\infty} E_i \in \mathcal{B}(S)$ .

The *probability function* is a function  $P(\cdot)$  with domain  $\mathcal{B}(S)$  satisfying the following three axioms, which are known as the *axioms of probability* or *Kolmogorov's axioms*.

1.  $P(E) \geq 0$  for all  $E \in \mathcal{B}(S)$ .

$P(\emptyset) = 0$
$P(E) \leq 1$
$P(E^c) = 1 - P(E)$
$P(E_1 \cup E_2) = P(E_1) + P(E_2) - P(E_1 \cap E_2)$
$E_1 \subset E_2 \Rightarrow P(E_1) \leq P(E_2)$

**Table 2.1:** Properties of the probability function.

2.  $P(S) = 1$ .

3. If  $E_1, E_2, \dots, E_n \in \mathcal{B}(S)$  are pairwise disjoint, then  $P(\bigcup_{i=1}^{\infty} E_i) = \sum_{i=1}^n P(E_i)$ .

Consequently,  $P(E)$  is a real number between zero and unity for any event  $E \in \mathcal{B}(S)$ . This number is interpreted as the *probability* of the event. Tab. 2.1 lists some useful properties of the probability function when applied to a single event. These rules are convenient for calculating more complicated probabilities. Most of them are fairly self-evident so that they are stated here without proofs.

A general method to define a probability function for an experiment with countable sample space without having to check whether the axioms of probability are satisfied is as follows. If  $S = \{s_1, s_2, \dots\}$  is a countable set and  $p_1, p_2, \dots$  are non-negative numbers that sum to unity, the function

$$P : E \in \mathcal{B}(S) \mapsto P(E) = \sum_{i: s_i \in E} p_i \tag{2.1}$$

is a probability function. Clearly,  $p_i$  reflects the probability of the event  $\{s_i\}$ , i.e.,  $P(\{s_i\}) = p_i$ .

The following example demonstrates that for certain random experiments it is easier to deal with a summary variable rather than with the original probability structure.

**Example 2.1:** A simple game where the player tosses a fair coin and wins one Euro<sup>1</sup> when getting head otherwise loosing one Euro. The sample space is therefore  $S = \{-1, 1\}$  and the probability function is given by  $P(\{+1\}) = P(\{-1\}) = 1/2$ . The sample space of tossing the coin  $n$  times has  $2^n$  elements, where each of which is an ordered string of  $\pm 1$ 's of length  $n$ . Of course, the player is only interested in how often the coin shows head. This is merely an integer between zero and  $n$  and much easier to deal with than the original sample space. Mathematically speaking, the player defines a map from the original sample space to the set  $\{0, 1, \dots, n\}$ : a so-called random variable. ■

---

<sup>1</sup>The unit Euro can be exchanged by any other unit.

Generally, a *random variable* (rv)  $X$  is a function from the sample space of the experiment into the real numbers. It is said to be discrete (continuous) if its range  $\mathcal{X} \subseteq \mathbb{R}$  is countable (uncountable). For a discrete rv with range  $\mathcal{X} = \{x_1, \dots, x_m\}$ , the probability of  $X$  taking the value  $x_i$  is denoted by  $\mathbb{P}[X = x_i]$  and given by

$$\mathbb{P}[X = x_i] = P(\{s_j \in S : X(s_j) = x_i\}). \quad (2.2)$$

Similarly, for a continuous rv with range  $\mathcal{X}$  a legitimate probability function reads

$$\mathbb{P}[X \in A] = P(\{s \in S : X(s) \in A\}), \quad (2.3)$$

where  $A$  is any subset of  $\mathcal{X}$ , i.e.,  $A \in \mathcal{B}(\mathcal{X})$ .

**Example 2.2:** Fair coin game mentioned in Ex. 2.1. The range  $\mathcal{X}$  corresponds to the set  $\{0, 1, \dots, n\}$ . The probability of winning  $k$  Euro (and consequently losing  $n - k$  Euro) corresponds to  $\mathbb{P}[X = k]$ . Each of the strings  $s$  of size  $n$  is equally likely with  $P(\{s\}) = 2^{-n}$  so that is the probability of  $X = k$  is equal to  $2^{-n}$  times the number of string containing exactly  $k$  +1's. That is

$$\mathbb{P}[X = k] = \frac{1}{2^n} \binom{n}{k}, \quad (2.4)$$

where the so-called *binomial coefficients* are defined by

$$\binom{n}{k} = \frac{n!}{k!(n-k)!}. \quad (2.5)$$

■

## 2.2 Distribution

Regarding a discrete rv  $X$ , Eq. (2.2) with range  $\mathcal{X}$  can be interpreted as a function mapping from  $\mathcal{X}$  to  $[0, 1]$ . It contains all information about the probability of  $X$  taking any value  $x$  out of  $\mathcal{X}$  and is therefore called *probability mass function* (pmf)

$$f_X(x) = \mathbb{P}[X = x] \quad (2.6)$$

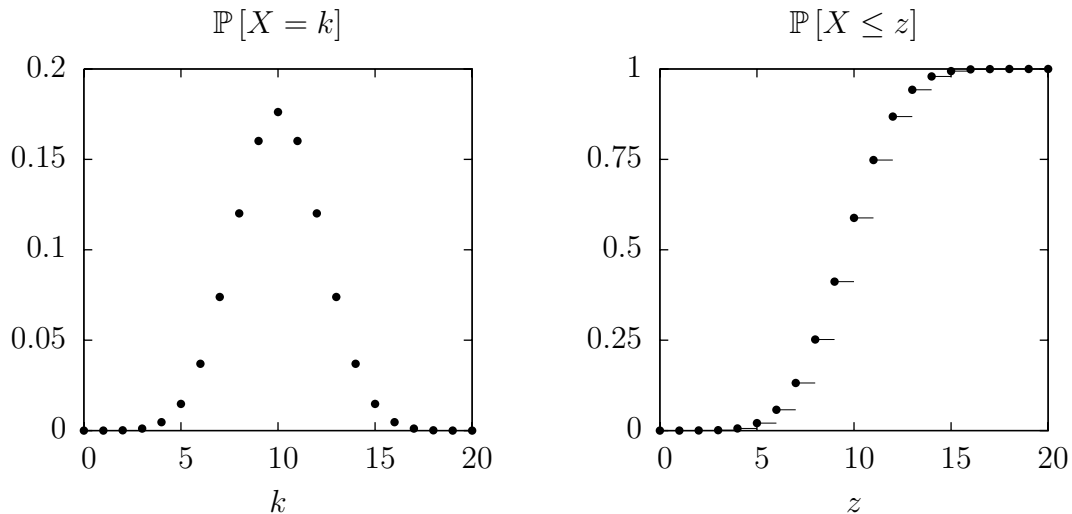
with  $\sum_{x \in \mathcal{X}} f_X(x) = 1$ . For a continuous rv the *probability density function* (pdf)  $f_X : \mathcal{X} \mapsto \mathbb{R}_{\geq 0}$  is defined such that

$$\int_A dx f_X(x) = \mathbb{P}[X \in A] \quad (2.7)$$

for any  $A \in \mathcal{B}(\mathcal{X})$ . Hence, it has to satisfy  $\int_{\mathcal{X}} dx f_X(x) = 1$ . Another function used to characterise a rv is the *cumulative distribution function* (cdf). It is defined by

$$F_X(z) = \mathbb{P}[X \leq z] = \begin{cases} \sum_{x \leq z} \mathbb{P}[X = x] & \text{for discrete rv's and} \\ \int_{x \leq z} dx f_X(x) & \text{for continuous rv's.} \end{cases} \quad (2.8)$$

so that its domain is not restricted to  $\mathcal{X}$ , i.e.,  $F_X : \mathbb{R} \mapsto [0, 1]$ . It has the following three properties:



**Figure 2.1:** Fair coin game's pmf (2.4) (left) and cdf (right) with  $n = 20$ .

- $\lim_{z \rightarrow -\infty} F_X(z) = 0$  and  $\lim_{z \rightarrow \infty} F_X(z) = 1$ .
- $F_X(z)$  is a non-decreasing function of  $z$ .
- $F_X(z)$  is right-continuous, i.e.,  $\lim_{z \searrow z_0} F_X(z) = F_X(z_0)$ .

Two rv's  $X$  and  $Y$  are said to be *identically distributed* if

$$F_X(z) = F_Y(z) \tag{2.9}$$

for every  $z \in \mathbb{R}$ .

**Example 2.3:** The fair coin game's pmf  $k \mapsto \mathbb{P}[X = k]$  and cdf  $z \mapsto \mathbb{P}[X \leq z]$  are depicted in Fig. 2.1 for  $n = 20$ . It is obvious that the number of getting tails is identically distributed to the number of getting heads. ■

In order to settle a frequently used notation in this text, the following symbols are introduced. If the rv  $X$  is  $f_X$ -distributed, it is denoted by

$$X \sim f_X. \tag{2.10}$$

If the two rv's  $X$  and  $Y$  are identically distributed, it is written as

$$X \sim Y. \tag{2.11}$$

## 2.3 Expected value and other characterisations

A rv is fully described by its pmf/pdf or its cdf, but capturing important information in a few single quantities sometimes suffices. These could be measures of location (e.g. expected value, median, mode, etc.), measures of statistical dispersion (e.g. standard deviation, variance), and measures of the shape of the distribution (e.g. skewness, kurtosis).

The *expected value* or *expectation value*  $\mathbb{E}[X]$  of a rv  $X$  is defined by

$$\mathbb{E}[X] = \begin{cases} \sum_{x \in \mathcal{X}} x f_X(x) & \text{if } X \text{ is discrete,} \\ \int_{\mathcal{X}} dx x f_X(x) & \text{if } X \text{ is continuous,} \end{cases} \quad (2.12)$$

where  $\mathcal{X}$  denotes the range of  $X$  and  $f_X(\cdot)$  denotes its pmf/pdf. The expected value is not to be confused with the most probable outcome. The *modes* are the values  $x$  where  $f_X(x)$  takes its maximum. If a pmf/pdf has multiple modes, it is said to be multimodal. Otherwise it is a so-called unimodal distribution. The third location measure which shall be introduced is the *median*, i.e., the value  $z$  where  $\mathbb{P}[X \leq z] \geq 1/2$  and  $\mathbb{P}[X \geq z] \geq 1/2$ . In other words, the median separates the half containing the lower values of  $X$  from the half containing the higher values of  $X$ .

Eq. (2.12) can be generalised to obtain the expected value of a rv  $g(X)$ , where  $g : \mathcal{X} \mapsto \mathbb{R}$  and  $\mathcal{X}$  denotes the range of  $X$ :

$$\mathbb{E}[g(X)] = \begin{cases} \sum_{x \in \mathcal{X}} g(x) f_X(x) & \text{if } X \text{ is discrete,} \\ \int_{\mathcal{X}} dx g(x) f_X(x) & \text{if } X \text{ is continuous.} \end{cases} \quad (2.13)$$

If  $g(x)$  is convex, i.e.,  $g(\lambda x_1 + (1 - \lambda)x_2) \leq \lambda g(x_1) + (1 - \lambda)g(x_2)$  for any  $x_{1,2} \in \mathcal{X}$  and  $\lambda \in [0, 1]$  such that  $\lambda x_1 + (1 - \lambda)x_2 \in \mathcal{X}$ , the *Jensen's inequality*

$$\mathbb{E}[g(X)] \geq g(\mathbb{E}[X]) \quad (2.14)$$

holds true. It should be mentioned that this inequality relies on properties of real-valued functions rather than on any statistical property. Eq. (2.13) is used to define further characteristics of a rv. The  $n^{\text{th}}$  *moment* or *raw moment* of a rv is given by  $\mathbb{E}[X^n]$  if it exists. The  $n^{\text{th}}$  *central moment* is defined by  $\mathbb{E}[(X - \mathbb{E}[X])^n]$ . A special case is the *variance*

$$\text{Var}[X] = \mathbb{E}[(X - \mathbb{E}[X])^2] = \mathbb{E}[X^2] - (\mathbb{E}[X])^2, \quad (2.15)$$

which according to Jensen's inequality (2.14) is always positive. Its square root, the *standard deviation*, is a measure of the rv's dispersion. If the expected value is positive, it might be convenient to express the standard deviation in units of the expected value as

$$\text{CV}[X] = \frac{\sqrt{\text{Var}[X]}}{\mathbb{E}[X]}, \quad (2.16)$$

which is known as the *coefficient of variation*.

\* In case any of my readers may be unfamiliar with the term “kurtosis” we may define mesokurtic as “having  $\beta_2$  equal to 3,” while platykurtic curves have  $\beta_2 < 3$  and leptokurtic  $> 3$ . The important property which follows from this is that platykurtic curves have shorter “tails” than the



normal curve of error and leptokurtic longer “tails.” I myself bear in mind the meaning of the words by the above *memoria technica*, where the first figure represents platypus, and the second kangaroos, noted for “lepping,” though, perhaps, with equal reason they should be hares!

**Figure 2.2:** Mnemonic for the terms platy- and leptokurtic provided by Student (1927).  $\beta_2$  stands for the kurtosis.

In addition to expected value and variance, two dimensionless quantities, which are of importance to unimodal distributions, are introduced. The *skewness*

$$\text{Skew}[X] = \frac{\mathbb{E}[(X - \mathbb{E}[X])^3]}{\text{Var}[X]^{3/2}} \quad (2.17)$$

is a measure of the lack of symmetry. A pmf/pdf is said to be *symmetric* about the point  $a$  if for all  $\Delta x > 0$   $f_X(a + \Delta x) = f_X(a - \Delta x)$ . Its expectation value and median is  $a$ . Furthermore, it has zero skewness because all odd central moments vanish. App. A.9.1 illustrates the meaning of skewness exemplified by a class of pdf’s with zero mean and unit variance. The *kurtosis*

$$\mathbb{K}\text{urt}[X] = \frac{\mathbb{E}[(X - \mathbb{E}[X])^4]}{\text{Var}[X]^2}, \quad (2.18)$$

is a measure of the peakedness or flatness of the pmf/pdf relative to the normal distribution. That is, a pdf with high kurtosis has a distinct peak near the mean, decline rather rapidly, and has heavy tails. Conversely, a pdf with low kurtosis is flat around the mean. The normal distribution, see App. A.2, has always kurtosis three. A pdf which has same kurtosis as the normal distribution is called *mesokurtic*. A pdf with larger (lower) kurtosis is called *leptokurtic* (*platykurtic*)—a mnemonic provided by Student (1927) is shown in Fig. 2.2. Jensen’s inequality in form of  $\mathbb{E}[(X - \mathbb{E}[X])^4] \geq (\mathbb{E}[(X - \mathbb{E}[X])^2])^2$  implies that the kurtosis is bounded from below by unity. App. A.9.2 illustrates the meaning of kurtosis exemplified by a class of pdf’s with zero mean, unit variance, and zero skewness.



## 2.4 Transformation

It is often the case, that the pmf/pdf of the rv  $X$  is known, but the pmf/pdf of the rv  $Y = g(X)$ , where  $g : \mathcal{X} \mapsto \mathcal{Y}$ , is of interest. For the sake of simplicity, only continuous rv's are discussed in this section. The reader is referred to [Casella & Berger \(2002\)](#) for discrete rv's. The most common case is when  $g(\cdot)$  is a monotone function, whose inverse  $g^{-1}(\cdot)$  has a continuous derivative in  $\mathcal{Y}$  and when the pdf  $f_X(\cdot)$  is continuous on  $\mathcal{X}$ . The pdf of  $Y$  is then given by

$$f_Y(y) = \begin{cases} f_X(g^{-1}(y)) \left| \frac{d}{dy} g^{-1}(y) \right| & \text{for } y \in \mathcal{Y}, \\ 0 & \text{otherwise.} \end{cases} \quad (2.19)$$

A very convenient, but maybe a bit careless, notation is

$$f_X(x) dx = f_Y(y) dy, \quad (2.20)$$

which is frequently used when substituting variables under an integral containing pdf's.

A frequently used transformation is the *linear transformation*  $Y = aX + b$ , where  $a \neq 0$  and  $b$  are constants. The inverse of the function  $g(x) = ax + b$  has the derivative  $1/a$  so that

$$f_Y(y) = \frac{1}{|a|} f_X\left(\frac{y-b}{a}\right). \quad (2.21)$$

Additionally, the following formulas are useful:

$$\mathbb{E}[aX + b] = a\mathbb{E}[X] + b, \quad (2.22a)$$

$$\text{Var}[aX + b] = a^2 \text{Var}[X], \quad (2.22b)$$

$$\text{Skew}[aX + b] = \text{Skew}[X], \quad \text{and} \quad (2.22c)$$

$$\mathbb{Kurt}[aX + b] = \mathbb{Kurt}[X]. \quad (2.22d)$$

The last two equations demonstrate that skewness and kurtosis are invariant under linear transformation.

**Example 2.4:** Linear transformation of a normally distributed rv  $X \sim N_{\mu, \sigma^2}$  (see App. [A.2](#) for notation). The pdf of  $Y = aX + b$  with  $a \neq 0$  is given by

$$f_Y(y) = \frac{1}{\sqrt{2\pi a^2 \sigma^2}} \exp\left[-\frac{(y - (a\mu + b))^2}{2a^2 \sigma^2}\right] \quad (2.23)$$

so that  $aX + b$  is of a normal distribution with mean  $a\mu + b$  and variance  $a^2\sigma^2$ . ■

**Example 2.5:** Linear transformation of a  $\chi^2$ -distributed rv  $X \sim \chi_k^2$  (see App. [A.3](#) for notation). The pdf of  $Y = aX + b$  with  $a \neq 0$  is given by

$$f_Y(y) = \begin{cases} \frac{1}{2^{|a|} \Gamma(k/2)} \left(\frac{y-b}{2a}\right)^{k/2-1} \exp\left[-\frac{y-b}{2a}\right] & \text{for } \frac{y-b}{a} > 0, \\ 0 & \text{otherwise} \end{cases} \quad (2.24)$$

and there is no general function  $l = l(k, a, b)$  such that the expression can be written as

$$f_Y(y) = \chi_l^2(y) = \begin{cases} \frac{1}{2\Gamma(l/2)} \left(\frac{y}{2}\right)^{l/2-1} \exp\left[-\frac{y}{2}\right] & \text{for } \frac{y-b}{a} > 0, \\ 0 & \text{otherwise.} \end{cases} \quad (2.25)$$

The only solution is the trivial solution  $a = 1$  and  $b = 0$ , i.e.,  $Y = X$ . Hence, a rv obtained by a non-trivial linear transformation of a  $\chi^2$ -distributed rv is not of a  $\chi^2$ -distribution. ■

Another transformation, which plays an important role in statistics, is the *exponential transformation*  $Y = e^{tX}$ , where  $t \neq 0$  is a constant and  $Y$  can only take positive values. The inverse of the function  $g(x) = e^{tx}$  has the derivative  $(ty)^{-1}$  so that

$$f_Y(y) = \begin{cases} \frac{1}{y|t|} f_X\left(\frac{\ln y}{t}\right) & \text{for } y > 0 \text{ and} \\ 0 & \text{for } y \leq 0. \end{cases} \quad (2.26)$$

The expectation value of  $Y$  is not directly related to the expected value of  $X$  because  $\mathbb{E}[Y]$  depends on all higher moments of  $X$  (if they exist) due to

$$\mathbb{E}[e^{tX}] = 1 + \sum_{n=1}^{\infty} \frac{t^n}{n!} \mathbb{E}[X^n]. \quad (2.27)$$

Nevertheless, the function  $M_X : t \mapsto \mathbb{E}[e^{tX}]$  which is defined by

$$M_X(t) = \mathbb{E}[e^{tX}] = \int_{-\infty}^{\infty} dx e^{tx} f_X(x), \quad (2.28)$$

and thus is related to the Laplace transformation of the distribution (if it exists) contains full information about the higher moments of  $X$  due to

$$\mathbb{E}[X^n] = \left. \frac{d^n}{dt^n} M_X(t) \right|_{t=0}. \quad (2.29)$$

It is therefore named *moment generating function* (mgf). Higher moments of  $Y = e^{tX}$  are given by

$$\mathbb{E}[Y^n] = M_X(nt). \quad (2.30)$$

**Example 2.6:** Mgf of normal distribution  $N_{\mu, \sigma^2}$ . If  $X$  is normally distributed with mean  $\mu$  and variance  $\sigma^2$  its mgf is given by

$$\begin{aligned} \int_{-\infty}^{\infty} dx e^{tx} N_{\mu, \sigma^2}(x) &= \frac{1}{\sqrt{2\pi\sigma^2}} \int_{-\infty}^{\infty} dx \exp\left[tx - \frac{(x-\mu)^2}{2\sigma^2}\right] \\ &= \frac{1}{\sqrt{2\pi\sigma^2}} \int_{-\infty}^{\infty} dx \exp\left[-\frac{x^2 - 2x(\mu + \sigma^2 t) + \mu^2}{2\sigma^2}\right] \\ &= \exp\left[-\frac{\mu^2 - (\mu + \sigma^2 t)^2}{2\sigma^2}\right] \int_{-\infty}^{\infty} dx N_{\mu + \sigma^2 t, \sigma^2}(x) \\ &= \exp\left[\mu t + \frac{\sigma^2}{2} t^2\right]. \end{aligned} \quad (2.31)$$

This result can be used to obtain e.g. the first two moments of  $X$  by  $\mathbb{E}[X] = M'_X(0) = \mu$  and  $\mathbb{E}[X^2] = M''_X(0) = \sigma^2 + \mu^2$ . ■

**Example 2.7:** Central moments of a normal distribution. If  $X$  is normally distributed with mean  $\mu$  and variance  $\sigma^2$  the shifted rv  $X - \mu$  is according to (2.21) normally distributed with zero mean and variance  $\sigma^2$ . Its mgf is given by  $M_{X-\mu}(t) = \exp(\sigma^2 t^2/2)$  which can be expanded to

$$M_{X-\mu}(t) = 1 + \sum_{q=1}^{\infty} \frac{\sigma^{2q}}{2^q q!} t^{2q} \quad (2.32)$$

so that its derivatives evaluated at  $t = 0$  can be read off directly. Hence, the odd central moments are zero and the even central moments are given by

$$\mathbb{E}[(X - \mu)^{2q}] = \sigma^{2q} \frac{(2q)!}{2^q q!} = \sigma^{2q} (2q - 1)!!, \quad (2.33)$$

where  $(\cdot)!!$  denotes the *double factorial* (see Weisstein, 2003) which is defined by the recursion formula

$$k!! = \begin{cases} 1 & \text{for } k = 0 \text{ or } k = 1, \\ k(k-2)!! & \text{for } k \geq 2. \end{cases} \quad (2.34)$$

**Example 2.8:** Log-normal distribution (see App. A.4 for notation). If  $X \sim N_{\mu, \sigma^2}$  with mgf  $M_X(\cdot)$ , the rv  $Y = e^X$  is log-normally distributed with density

$$f_Y(y) = \frac{1}{y\sqrt{2\pi\sigma^2}} \exp\left[-\frac{(\ln y - \mu)^2}{2\sigma^2}\right] \quad (2.35)$$

for  $y > 0$ . The higher moments are given by

$$\mathbb{E}[Y^n] = M_X(n) = \exp\left[n\mu + \frac{n^2\sigma^2}{2}\right]. \quad (2.36)$$

## 2.5 Joint distribution and correlation

The previous sections considered only one rv. This section is intended to generalise the concept of a distribution to more than one rv.

Consider  $n$  rv's  $X_1, X_2, \dots, X_n$  with sample set  $\mathcal{X} = \mathcal{X}_1 \times \mathcal{X}_2 \times \dots \times \mathcal{X}_n$ , their *joint probability mass function* (discrete rv's) or *joint probability density function* (continuous rv's) is a function  $\rho : \mathcal{X} \mapsto \mathbb{R}_{\geq 0}$  with

$$\rho_{(n)}(x_1, \dots, x_n) = \mathbb{P}[X_1 = x_1 \wedge \dots \wedge X_n = x_n] \quad (2.37a)$$

or

$$\int_A dx_1 \dots dx_n \rho_{(n)}(x_1, \dots, x_n) = \mathbb{P}[(X_1, \dots, X_n) \in A] \quad (2.37b)$$

for any  $A \in \mathcal{B}(\mathcal{X})$ , respectively. The corresponding marginal pmf's and pdf's are given by

$$f_i(x_i) = \sum_{x'_1} \dots \sum_{x'_{i-1}} \sum_{x'_{i+1}} \dots \sum_{x'_n} \rho_{(n)}(x'_1, \dots, x'_{i-1}, x_i, x'_{i+1}, \dots, x'_n) \quad (2.38a)$$

and

$$f_i(x_i) = \int dx'_1 \dots dx'_{i-1} dx'_{i+1} \dots dx'_n \rho_{(n)}(x'_1, \dots, x'_{i-1}, x_i, x'_{i+1}, \dots, x'_n) \quad (2.38b)$$

for  $i = 1, \dots, n$ , respectively. Additionally, the joint distribution gives information about the distribution of the  $n - 1$  rv's  $X_1, \dots, X_{i-1}, X_{i+1}, \dots, X_n$  if the  $i^{\text{th}}$  rv  $X_i$  is fixed  $X_i = x_i$ . The *conditioned* pmf or pdf of  $X_i$  is given by

$$\rho_{(n-1)}(x_1, \dots, x_{i-1}, x_{i+1}, \dots, x_n | X_i = x_i) = \frac{\rho_{(n)}(x_1, \dots, x_{i-1}, x_i, x_{i+1}, \dots, x_n)}{f_i(x_i)}. \quad (2.39)$$

Iterating this equation is frequently used to compute the joint distribution of  $n$  rv's when the conditioned and marginal distributions are known. A rv  $X_i$  is said to be *independent* of the remaining  $n - 1$  rv's if the joint pmf/pdf factorises such that

$$\rho_{(n)}(x_1, \dots, x_n) = \rho_{(n-1)}(x_1, \dots, x_{i-1}, x_{i+1}, \dots, x_n) \times f_i(x_i), \quad (2.40)$$

i.e., the left hand side of (2.39) does not depend on  $x_i$ . As a consequence, if the  $n$  rv's  $X_1, \dots, X_n$  are mutually independent, the joint distribution completely factorises such that

$$\rho_{(n)}(x_1, \dots, x_n) = \prod_{i=1}^n f_i(x_i). \quad (2.41)$$

If furthermore  $f_i(\cdot) = f(\cdot)$  for all  $i = 1, \dots, n$  so that

$$\rho_{(n)}(x_1, \dots, x_n) = \prod_{i=1}^n f(x_i), \quad (2.42)$$

the  $n$  rv's are called *independent identically distributed* (iid).

Therefore, two independent rv's  $X$  and  $Y$  with distribution  $f_X$  and  $f_Y$ , respectively, have

$$\begin{aligned} \mathbb{E}[g(X)h(Y)] &= \begin{cases} \sum_x \sum_y g(x)h(y)\rho_{(2)}(x, y) & \text{(discrete)} \\ \int dx dy g(x)h(y)\rho_{(2)}(x, y) & \text{(continuous)} \end{cases} \\ &= \begin{cases} \sum_x g(x)f_X(x) \sum_y h(y)f_Y(y) & \text{(discrete)} \\ \int dx g(x)f_X(x) \int dy h(y)f_Y(y) & \text{(continuous)} \end{cases} \\ &= \mathbb{E}[g(X)] \mathbb{E}[h(Y)] \end{aligned} \quad (2.43)$$

for any functions  $g(\cdot)$  and  $h(\cdot)$ . Especially,  $\mathbb{E}[XY] = \mathbb{E}[X]\mathbb{E}[Y]$ .

If the two rv's  $X$  and  $Y$  are not independent, the expected value of  $XY$  might not be the product of the individual expected values. The difference between  $\mathbb{E}[XY]$  and  $\mathbb{E}[X]\mathbb{E}[Y]$  is referred to as the *covariance* between  $X$  and  $Y$ , i.e.,

$$\text{Cov}[X, Y] = \mathbb{E}[(X - \mathbb{E}[X])(Y - \mathbb{E}[Y])] = \mathbb{E}[XY] - \mathbb{E}[X]\mathbb{E}[Y]. \quad (2.44)$$

The covariance is used to determine the variance of  $aX + bY$ , where  $X$  and  $Y$  are any rv's whereas  $a$  and  $b$  are constants:

$$\begin{aligned} \text{Var}[aX + bY] &= \mathbb{E}[(aX + bY - \mathbb{E}[aX + bY])^2] \\ &= \mathbb{E}[(a(X - \mathbb{E}[X]) + b(Y - \mathbb{E}[Y]))^2] \\ &= a^2 \text{Var}[X] + b^2 \text{Var}[Y] + 2ab \text{Cov}[X, Y]. \end{aligned} \quad (2.45)$$

It is proven in (Casella & Berger, 2002, Thm. 4.5.7) that

$$|\text{Cov}[X, Y]| \leq \sqrt{\text{Var}[X] \text{Var}[Y]}, \quad (2.46)$$

so that it is convenient to normalise the covariance and define

$$\text{Corr}[X, Y] = \frac{\text{Cov}[X, Y]}{\sqrt{\text{Var}[X] \text{Var}[Y]}}, \quad (2.47)$$

which is known as the *linear correlation* or simply *correlation*. It is a real number between  $-1$  and  $+1$ . The theorem also proves that  $\text{Corr}[X, Y] = 1$  ( $\text{Corr}[X, Y] = -1$ ) if and only if there exist numbers  $a > 0$  ( $a < 0$ ) and  $b$  such that  $Y = aX + b$ . Two rv's  $X$  and  $Y$  are said to be *correlated* if  $\text{Corr}[X, Y] > 0$  and *anti-correlated* if  $\text{Corr}[X, Y] < 0$ . On the other hand, if the correlation between  $X$  and  $Y$  vanishes,  $X$  and  $Y$  are called *uncorrelated random variables*, which is not to be mistaken with independent rv's. Independent rv's are always uncorrelated, but uncorrelated rv's are not necessarily independent.

**Example 2.9:** Correlation between  $X \sim N_{0,1}$  and  $X^2 \sim \chi_1^2$ . The rv's  $X$  and  $X^2$  are clearly not independent, but their linear correlation vanishes because

$$\text{Corr}[X, X^2] = \frac{\mathbb{E}[X^3] - \mathbb{E}[X]\mathbb{E}[X^2]}{\sqrt{\text{Var}[X] \text{Var}[X^2]}} = \frac{\mathbb{E}[X^3]}{\sqrt{2}} = 0. \quad (2.48)$$

■

The concept of correlation can be generalised to  $n \in \mathbb{N}$  rv's  $X_1, \dots, X_n$  with finite variances. The *correlation matrix* is an  $n \times n$  matrix defined by

$$(\mathbf{\Gamma})_{ij} = \text{Corr}[X_i, X_j]. \quad (2.49)$$

It is symmetric because  $\text{Corr}[X_i, X_j] = \text{Corr}[X_j, X_i]$  and its diagonal entries are unity. The *average correlation*  $\rho_n$  is obtained by averaging over the off-diagonal elements, i.e.,

$$\rho_n = \frac{1}{n(n-1)} \sum_{i=1}^n \sum_{\substack{j=1 \\ j \neq i}}^n (\mathbf{\Gamma})_{ij}, \quad (2.50)$$

and plays an essential role for the mean and variance estimation of a random sample, cf. Sec. 2.8. The average correlation is bounded from above by +1 which only occurs if  $\text{Corr}[X_i, X_j] = 1$  for any combination  $i, j$ . As for the lower bound on the other hand, it is not possible to have  $\text{Corr}[X_i, X_j] = -1$  for any combination  $i, j$  with  $j \neq i$ .

**Example 2.10:** Three fully anti-correlated rv's. Consider three rv's  $X, Y, Z$  with finite variance. If  $\text{Corr}[X, Y] = -1$  and  $\text{Corr}[X, Z] = -1$  it follows that  $\text{Corr}[Y, Z] = +1$ . This is because there are constants  $a, b, c, d$  such that  $Y = aX + b$  and  $X = cZ + d$  with  $a, c < 0$ . Consequently,  $Y = acZ + ad + b$  with  $ac > 0$  so that  $\text{Corr}[Y, Z] = +1$ . ■

Rasch (1989) shows that  $\mathbf{\Gamma}$  is positive semi-definite, i.e., for any  $\vec{c} \in \mathbb{R}^n$  the product  $\vec{c}^\top \mathbf{\Gamma} \vec{c}$  is non-negative. Hence, the sum of all entries must be non-negative. Splitting this sum into the trace of  $\mathbf{\Gamma}$  plus the sum over the off-diagonal entries results in

$$\sum_{i,k=1}^n (\mathbf{\Gamma})_{ik} = n + \sum_{i=1}^n \sum_{\substack{j=1 \\ j \neq i}}^n (\mathbf{\Gamma})_{ij} \geq 0. \quad (2.51)$$

As a consequence,

$$-\frac{1}{n-1} \leq \rho_n \leq 1, \quad (2.52)$$

so that for  $n \rightarrow \infty$  the average correlation is non-negative, i.e.,  $0 \leq \rho_\infty \leq 1$ .

## 2.6 Multivariate transformation

This section derives the distribution of the rv

$$Z = g(X_1, \dots, X_n), \quad (2.53)$$

where  $g : \mathcal{X}_1 \times \dots \times \mathcal{X}_n \mapsto \mathcal{Z}$  is a function of  $n$  rv's  $X_1, \dots, X_n$  with joint pmf/pdf  $\rho_{(n)}(x_1, \dots, x_n)$ . Note that all  $X_i$ 's have to be either discrete or continuous.

The distribution of  $Z$  is a function  $f_z : \mathcal{Z} \mapsto \mathbb{R}_{\geq 0}$  with

$$f_Z(z) = \mathbb{P}[Z = z] = \sum_{\{x_1, \dots, x_n : g(x_1, \dots, x_n) = z\}} \rho_{(n)}(x_1, \dots, x_n) \quad (2.54)$$

for discrete rv's because the probability for  $Z = z$  is given by sum over the probabilities of all possible states with  $g(x_1, \dots, x_n) = z$ . For continuous rv's it is given by

$$f_Z(z) = \int dx_1 \dots dx_n \delta(z - g(x_1, \dots, x_n)) \rho_{(n)}(x_1, \dots, x_n) \quad (2.55)$$

because

$$\begin{aligned}
\int_A dz f_Z(z) &= \int dx_1 \dots dx_n \rho_{(n)}(x_1, \dots, x_n) \int_A dz \delta(z - g(x_1, \dots, x_n)), \\
&= \int_{g(x_1, \dots, x_n) \in A} dx_1 \dots dx_n \rho_{(n)}(x_1, \dots, x_n) \\
&= \mathbb{P}[g(X_1, \dots, X_n) \in A],
\end{aligned} \tag{2.56}$$

i.e.,

$$\mathbb{P}[Z \in A] = \int_A dz f_Z(z) \quad \forall A \in \mathcal{B}(\mathcal{Z}). \tag{2.57}$$

Eqs. (2.54) and (2.55) can be written as

$$f_Z(z) = \mathbb{E}[\delta(z, g(X_1, \dots, X_n))], \tag{2.58}$$

where

$$\delta(x, y) = \begin{cases} \delta_{x,y} & \text{(Kronecker } \delta) \text{ for the discrete case and} \\ \delta(x - y) & \text{(Dirac } \delta) \text{ for the continuous case.} \end{cases} \tag{2.59}$$

**Example 2.11:** Derivation of the univariate transformation  $Y = g(X)$  (continuous) rule (2.19). The distribution of  $Y$  is according to (2.55) given by  $f_Y(y) = \int dx f_X(x) \delta(y - g(x))$ . The  $\delta$ -function can be expressed as  $\delta(x - g^{-1}(y))/|g'(g^{-1}(y))|$ , where it is assumed that  $g^{-1}(\cdot)$  and  $g'(\cdot)$  exist. Integrating over  $x$  yields  $f_Y(y) = f_X(g^{-1}(y)) \left| \frac{d}{dy} g^{-1}(y) \right|$ , which corresponds to (2.19). ■

## 2.7 Sum of continuous random variables and stable distributions

Consider two continuous rv's  $X$  and  $Y$  with their joint distribution  $\rho_{(2)}(x, y)$ . The question is: what is the distribution of the rv  $Z = X + Y$ ? Of course, this is just a special case of a multivariate transformation. However, it is worth discussing this special case thoroughly because it allows a clear and accessible introduction to concepts like convolutions, characteristic functions and stable distributions. These ideas are commonly used in statistics.

Consider the rv  $Z = \sum_{i=1}^n X_i$  to be the sum of  $n$  rv's with joint distribution  $\rho_{(n)}(x_1, \dots, x_n)$ . Applying (2.55) leads to

$$f_Z(z) = \int dx_1 \dots dx_n \delta\left(z - \sum_{i=1}^n x_i\right) \rho_{(n)}(x_1, \dots, x_n). \tag{2.60}$$

The  $\delta$ -function can be used to reduce the  $n$  dimensional integral down to a  $n - 1$  dimensional integral. The calculation procedure strongly depends on the structure of  $\rho_{(n)}(x_1, \dots, x_n)$  and therefore on the correlations between the rv's.

An important special case can be derived for independent rv's  $X_i$ 's, where the integral

$$f_Z(z) = \int dx_1 \dots dx_n \delta(z - \sum_{i=1}^n x_i) f_1(x_1) \dots f_n(x_n) \quad (2.61)$$

can be computed using the convolutions. The *convolution* of two pdf's  $f : \mathbb{R} \mapsto \mathbb{R}$  and  $g : \mathbb{R} \mapsto \mathbb{R}$  is defined by

$$f \star g : x \mapsto \int_{-\infty}^{\infty} dy f(y)g(x - y). \quad (2.62)$$

It is commutative, i.e.,

$$f \star g = g \star f, \quad (2.63)$$

and [Yosida \(1980\)](#) extends the definition of the convolution in a unique way such that the associative law

$$f \star (g \star h) = (f \star g) \star h \quad (2.64)$$

holds true. That is,

$$(f \star g \star h)(z) = \int_{-\infty}^{\infty} dx_1 dx_2 dx_3 f(x_1)g(x_2)h(x_3) \delta(z - x_1 - x_2 - x_3). \quad (2.65)$$

More generally, the convolution of  $n$  distributions  $f_i(\cdot)$  ( $i = 1, \dots, n$ ) is written as

$$\left( \bigstar_{i=1}^n f_i \right) (z) = \int_{\infty} \delta \left( z - \sum_{i=1}^n x_i \right) \prod_{i=1}^n f_i(x_i) dx_i \quad (2.66)$$

so that (2.61) becomes

$$f_Z(z) = \left( \bigstar_{i=1}^n f_i \right) (z). \quad (2.67)$$

The computation of this convolution is sometimes quite difficult, but applying the Fourier transformation

$$\mathcal{F}[f] : t \mapsto \int_{-\infty}^{\infty} dx f(x)e^{itx} \quad (2.68)$$

to the above equation and using the *convolution theorem*

$$\mathcal{F} \left[ \bigstar_{i=1}^n f_i \right] = \prod_{i=1}^n \mathcal{F}[f_i] \quad (2.69)$$

leads to

$$f_Z(z) = \mathcal{F}^{-1} \left[ \prod_{i=1}^n \mathcal{F}[f_i] \right] (z), \quad (2.70)$$

where the inverse Fourier transformation in this convention reads

$$\mathcal{F}^{-1}[F] : x \mapsto \frac{1}{2\pi} \int_{-\infty}^{\infty} dt F(t)e^{-itx}. \quad (2.71)$$



It is often convenient to work directly in the Fourier space using characteristic functions. The *characteristic function* of a rv  $X$  which is  $f_X$ -distributed is defined by

$$\varphi_X(t) = \mathbb{E} [e^{itX}] = \mathcal{F} [f_X] (t). \quad (2.72)$$

Hence, the characteristic function of  $Z = \sum_{i=1}^n X_i$  is given by

$$\varphi_Z(t) = \prod_{i=1}^n \varphi_{X_i}(t) \quad (2.73)$$

with  $\varphi_{X_i}(\cdot)$  denoting the characteristic function of  $X_i$ .

**Example 2.12:** Sum of normally distributed random variables. Consider  $X_1, \dots, X_n$  to be  $n$  independent normally distributed random variables with  $\mathbb{E}[X_j] = \mu_j$  and  $\text{Var}[X_j] = \sigma_j^2$  for  $j = 1, \dots, n$ . The characteristic functions are given by  $\varphi_{X_j}(t) = \exp [i\mu_j t - \sigma_j^2 t^2 / 2]$  so that the characteristic function of  $Z = \sum_{j=1}^n X_j$  reads

$$\varphi_Z(t) = \prod_{j=1}^n \exp \left[ i\mu_j t - \frac{\sigma_j^2 t^2}{2} \right] = \exp \left[ it \sum_{j=1}^n \mu_j - \frac{t^2}{2} \sum_{j=1}^n \sigma_j^2 \right]. \quad (2.74)$$

Taking the inverse Fourier transformation yields that  $Z$  is normally distributed with mean  $\sum_{j=1}^n \mu_j$  and variance  $\sum_{j=1}^n \sigma_j^2$ . ■

**Example 2.13:** Sum of  $\chi^2$ -distributed rv's (see App. A.3 for notation). Consider  $Z = \sum_{j=1}^n X_j$ , where  $X_j \sim \chi_{k_j}^2$  with  $k_j$  denoting the degrees of freedom. The characteristic functions are given by  $\varphi_j(t) = (1 - 2it)^{-k_j/2}$  so that the characteristic function of  $Z = \sum_{i=1}^n X_i$  reads

$$\varphi_Z = \prod_{j=1}^n (1 - 2it)^{-k_j/2} = (1 - 2it)^{-\sum_{j=1}^n k_j/2}. \quad (2.75)$$

Taking the inverse Fourier transformation yields that  $Z$  is also  $\chi^2$ -distributed with  $\sum_{j=1}^n k_j$  degrees of freedom. That is, the number of degrees of freedom simply add. ■

**Example 2.14:** Sum of two uniformly distributed rv's (see App. A.7 for notation). Consider  $Z = X + Y$ , where both  $X$  and  $Y$  are uniformly distributed between  $a$  and  $b$ , i.e.,  $X, Y \sim U_{a,b}$ . This problem is best solved without taking the detour with characteristic functions and solve (2.61) directly by

$$f_Z(z) = \int_{-\infty}^{\infty} dx U_{a,b}(x) U_{a,b}(z - x) \quad (2.76)$$

There is only a contribution to the integral if  $a \leq x \leq b$  and  $z - b \leq x \leq z - a$  so that

$$\begin{aligned}
 f_Z(z) &= \frac{1}{(b-a)^2} \int_{[a,b] \cap [z-b, z-a]} dx \\
 &= \frac{1}{(b-a)^2} \begin{cases} 0 & \text{for } z \leq 2a, \\ z - 2a & \text{for } 2a < z < a + b, \\ b - a & \text{for } z = a + b, \\ 2b - z & \text{for } a + b < z < 2b, \\ 0 & \text{for } z \geq 2b, \end{cases} \quad (2.77)
 \end{aligned}$$

which is a member of the triangular distribution family. ■

The examples showed that the sum of independent normally distributed random variables is again normally distributed with added mean and variance and that the sum of  $\chi^2$ -distributed rv's is also of a  $\chi^2$ -distribution with added degrees of freedom. However, the sum of two uniformly distributed rv's is not uniformly distributed. That is, there are some distributions, where the distribution of the sum of iid rv's belongs to the same class as the distribution of the individual summands.

In order to discuss this problem systematically, the sum of  $n$  rv's is considered, which are not only independent but in addition identically distributed. This problem was initially studied by P. Lévy in the 1920's and subject of a recent work by Nolan (2009). A pdf  $f_X$  is said to be *stable* if there are constants  $a > 0$  and  $b$  such that the sum of  $n$  iid rv's  $X_1, \dots, X_n$  with  $X_j \sim f_X$  for  $j = 1, \dots, n$  satisfies

$$a \sum_{j=1}^n X_j + b \sim f_X. \quad (2.78)$$

**Example 2.15:** Sum of iid normally distributed rv's. As shown in Ex. 2.12, the rv  $Z = \sum_{j=1}^n X_j$ , where  $X_j \sim N_{\mu, \sigma^2}$  for  $j = 1, \dots, n$ , is normally distributed with mean  $n\mu$  and variance  $n\sigma^2$ . The linearly transformed rv  $Z/\sqrt{n} - \mu(\sqrt{n} - 1)$  is  $N_{\mu, \sigma}$ -distributed, cf. Ex. 2.4 or Eq. (A.9). Hence, the normal distribution is stable. ■

**Example 2.16:** Sum of iid  $\chi^2$ -distributed rv's. As shown in Ex. 2.13, the rv  $Z = \sum_{j=1}^n X_j$  with  $X_j \sim \chi_k^2$  for  $j = 1, \dots, n$  is  $\chi^2$ -distributed with  $nk$  degrees of freedom. However, it is not possible to shift and scale  $Z$  such that  $aZ + b$  is  $\chi^2$ -distributed with  $k$  degrees of freedom (it is shown in Ex. 2.5 that a non-trivial linear transformation of a  $\chi^2$ -distributed rv is not  $\chi^2$ -distributed). Hence, the  $\chi^2$ -distribution is not stable. ■

The class of stable distributions can be obtained by using characteristic functions. Denoting  $\varphi_X(t) = \mathbb{E}[e^{itX_j}]$  for all  $j = 1, \dots, n$ , the characteristic function of  $Z = \sum_{j=1}^n X_j$  reads  $\varphi_Z(t) = \varphi_X(t)^n$ . The linear transformation rule (2.21) implies that  $\varphi_{aZ+b}(t) = e^{itb} \varphi_X(at)^n$  for  $a > 0$ , which has to be equal to  $\varphi_X(t)$  if  $f_X$  is a stable distribution, i.e.,

$$\varphi_X(t) = e^{itb} \varphi_X(at)^n. \quad (2.79)$$

Lévy (1925) and Khintchine & Lévy (1936) solved the general problem of determining the entire class of stable distributions. The *Lévy skew  $\alpha$ -stable distributions* are given by characteristic functions

$$\begin{aligned}\varphi_X(t) &= \mathcal{F}[S_{\alpha,\beta,\gamma,\delta}](t) \\ &= \begin{cases} \exp\{-\gamma^\alpha |t|^\alpha [1 - i\beta \operatorname{sgn}(t) \tan(\pi\alpha/2)] + it\delta\} & \text{for } \alpha \neq 1, \\ \exp\{-\gamma |t| [1 + 2i\beta/\pi \operatorname{sgn}(t) \ln |t|] + it\delta\} & \text{for } \alpha = 1. \end{cases} \end{aligned} \quad (2.80)$$

Consequently, a general stable distribution requires four parameters to describe: an index of stability  $0 < \alpha \leq 2$ , a skewness parameter  $-1 \leq \beta \leq 1$ , a scale parameter  $\gamma > 0$ , and a location parameter  $\delta \in \mathbb{R}$ . Inserting this expression into (2.79) yields

$$a = \frac{1}{n^{1/\alpha}} \quad \text{and} \quad b = \begin{cases} \delta [1 - n^{1-1/\alpha}] & \text{for } \alpha \neq 1, \\ -2\beta\gamma/\pi \ln n & \text{for } \alpha = 1. \end{cases} \quad (2.81)$$

There are three special cases of  $S_{\alpha,\beta,\gamma,\delta}$  which are discussed in the following three examples.

**Example 2.17:** Normal distribution ( $\alpha = 2, \beta = 0$ ). The rv  $X \sim S_{2,0,\gamma,\delta}$  has the characteristic function  $\varphi_X(t) = \exp[-\gamma^2 t^2 + it\delta]$  and is therefore normally distributed with mean  $\delta$  and variance  $2\gamma^2$ :  $S_{2,0,\gamma,\delta}(x) = N_{\delta,2\gamma^2}(x)$ . The solution for the constants  $a$  and  $b$  for the sum of  $n$  independent  $N_{\mu,\sigma^2}$ -distributed rv's is given by  $a = 1/\sqrt{n}$  and  $b = -\mu(\sqrt{n} - 1)$ . This result agrees with the result found in Ex. 2.15. ■

**Example 2.18:** Cauchy distribution ( $\alpha = 1, \beta = 0$ ). The rv  $X \sim S_{1,0,\gamma,\delta}$  has the characteristic function  $\varphi_X(t) = \exp[-\gamma |t| + it\delta]$  and is therefore of a Cauchy distribution  $\text{Cy}_{\delta,\gamma}$  (see App. A.5 for notation). Eq. (2.79) for the sum of  $n$  iid rv's is solved by  $a = 1/n$  and  $b = 0$ . That is, if  $X_j \sim \text{Cy}_{\delta,\gamma}$  for  $j = 1, \dots, n$  are iid rv's, its average  $1/n \sum_{j=1}^n X_j$  is of a  $\text{Cy}_{\delta,\gamma}$ -distribution, too. ■

**Example 2.19:** Lévy distribution ( $\alpha = 1/2, \beta = 1$ ). The rv  $X \sim S_{1/2,1,\gamma,\delta}$  has the characteristic function  $\varphi_X(t) = \exp\{-[1 - i\operatorname{sgn}(t)]\sqrt{\gamma |t|} + it\delta\}$  and is therefore of a Lévy distribution  $\text{Lv}_{\delta,\gamma}$  (see App. A.6 for notation). Eq. (2.79) for the sum of  $n$  iid rv's is solved by  $a = 1/n^2$  and  $b = \delta(1 - 1/n)$ . ■

## 2.8 Random sample estimation

Consider a random sample of finite size  $n$  consisting of identically distributed rv's  $X_j$  with  $j = 1, \dots, n$ . Each rv is drawn from an unknown distribution  $f_X(\cdot)$ . The goal is to determine the pmf/pdf and its parameterisation.

Firstly, the case of iid rv's is discussed. Denoting the values of the rv's by  $x_j$  ( $j = 1, \dots, n$ ), the estimated distribution is formally written as

$$\tilde{f}_X(x) = \frac{1}{n} \sum_{j=1}^n \delta(x - x_j) \quad (2.82)$$

and is difficult to present graphically. A (normalised) *histogram* is used instead. The range of  $x \in [a, b]$  is split into  $m$  bins of width  $\Delta x = (b - a)/m$ . The  $k^{\text{th}}$  bin covers  $B_k = [a + (k - 1)\Delta x, a + k\Delta x[$  around the bin's centre  $\xi_k = a + (k - 1/2)\Delta x$  for  $k = 1, \dots, m$ . The estimated histogram  $\tilde{h}_X : \{\xi_1, \dots, \xi_m\} \mapsto \mathbb{R}$  is determined by

$$\tilde{h}_X(\xi_k) = \int_{B_k} dx \tilde{f}_X(x) = \frac{1}{n} \sum_{\{j: x_j \in B_k\}} 1 \quad (2.83)$$

for  $k = 1, \dots, m$ , i.e., by counting the number of rv's whose value lies in bin  $k$ . The estimated cdf however can be plotted directly and does not require any binning. It is given by

$$\tilde{F}_X(x) = \int_{-\infty}^x dx' \tilde{f}_X(x') = \frac{1}{n} \sum_{\{j: x_j \leq x\}} 1. \quad (2.84)$$

Often it is not the estimation of the pmf/pdf/cdf which is of interest, but rather the mean  $\mu_X$  and variance  $\sigma_X^2$  need to be estimated. The arithmetic average

$$\bar{X} = \frac{1}{n} \sum_{j=1}^n X_j \quad (2.85)$$

is itself a rv with expected value  $\mathbb{E}[\bar{X}] = \mu_X$  and variance  $\text{Var}[\bar{X}] = \sigma_X^2/n$ . That is, if the population size  $n$  is sufficiently large, the standard deviation of  $\bar{X}$ , which is a measure for the estimation uncertainty, is negligible so that the expectation value of the set  $\{x_1, \dots, x_n\}$

$$\mathbb{E}\{x_1, \dots, x_n\} = \frac{1}{n} \sum_{j=1}^n x_j \quad (2.86)$$

is a reasonable *estimated expectation value* of  $X_j$ . As for the variance, the rv

$$S^2 = \frac{1}{n-1} \sum_{j=1}^n [X_j - \bar{X}]^2 = \frac{\sum_{j=1}^n X_j^2 - n\bar{X}^2}{n-1} \quad (2.87)$$

has expectation

$$\begin{aligned} \mathbb{E}[S^2] &= \frac{1}{n-1} \mathbb{E} \left[ \sum_{j=1}^n [(X_j - \mu_X) - (\bar{X} - \mu_X)]^2 \right] \\ &= \frac{1}{n-1} \left[ n\sigma_X^2 + n\text{Var}[\bar{X}] - 2 \sum_{j=1}^n \mathbb{E}[(X_j - \mu_X)(\bar{X} - \mu_X)] \right] \\ &= \frac{1}{n-1} [n\sigma_X^2 + \sigma_X^2 - 2\sigma_X^2] \\ &= \sigma_X^2. \end{aligned} \quad (2.88)$$

As a consequence, defining the variance of the set  $\{x_1, \dots, x_n\}$  by

$$\text{Var} \{x_1, \dots, x_n\} = \frac{1}{n-1} \sum_{j=1}^n [x_j - \bar{x}]^2 = \frac{\sum_{j=1}^n x_j^2 - n\bar{x}^2}{n-1} \quad (2.89)$$

with  $\bar{x} = \mathbb{E} \{x_1, \dots, x_n\}$  is a reasonable *estimated variance* of  $X_j$ . Note that there is the factor  $n-1$  instead of  $n$  in the denominator of the definition.

Secondly, the case of correlated rv's  $X_j$  is discussed, where the estimated mean and variance might not reflect the actual mean and variance of the underlying distribution. The expected value of  $\bar{X} = \sum_{j=1}^n X_j/n$  is still  $\mu_X$ , but its variance reads

$$\begin{aligned} \text{Var} [\bar{X}] &= \frac{\sigma_X^2}{n^2} \sum_{j,j'=1}^n \text{Corr} [X_j, X_{j'}] \\ &= \frac{\sigma_X^2}{n} + \rho_n \sigma_X^2 \frac{n-1}{n} \xrightarrow{n \rightarrow \infty} \sigma_X^2 \rho_\infty, \end{aligned} \quad (2.90)$$

where  $\rho_n$  denotes the average correlation (2.50) and is a real number between  $-1/(n-1)$  (as many anti-correlated rv's as possible) and 1 (fully perfectly correlated). Consequently, even for an infinite population size the estimation uncertainty is positive depending on the value of  $\rho_\infty \in [0, 1]$ . Regarding the variance estimation, the expected value of  $S^2$  is given by

$$\begin{aligned} \mathbb{E} [S^2] &= \frac{1}{n-1} \mathbb{E} \left[ \sum_{j=1}^n [(X_j - \mu_X) - (\bar{X} - \mu_X)]^2 \right] \\ &= \frac{1}{n-1} \left[ n\sigma_X^2 + n\text{Var} [\bar{X}] - \frac{2}{n} \sum_{j,j'=1}^n \text{Cov} [X_j, X_{j'}] \right] \\ &= \sigma_X^2 - \rho_n \sigma_X^2 \xrightarrow{n \rightarrow \infty} \sigma_X^2 (1 - \rho_\infty), \end{aligned} \quad (2.91)$$

which potentially underestimates the true variance.

## 2.9 Kolmogorov-Smirnov test

In applied research it is common to speculate the population distribution (or at least any of its parameters) and try to verify this speculation based on available data. This is known as hypothesis testing. This section however will only cover a specific test: the Kolmogorov-Smirnov test, which is described in standard inferential statistics books such as [Sheskin \(1997\)](#) and [Hartung \(2005\)](#).

Some basic vocabulary needs to be introduced first. A *hypothesis* is a statement about the population. The two complementary hypothesis are called null hypothesis  $H_0$  and the alternative hypothesis  $H_1 = \neg H_0$ . There is of course no 100 % warranty that the decision to accept/reject the null hypothesis is correct. A *type I error* refers to rejecting the null hypothesis when in fact it is correct. The probability of making the type I error is often

	$H_0$ is true	$H_0$ is false
test accepts $H_0$	no error	type II
test rejects $H_0$	type I	no error

**Table 2.2:** Error types in hypothesis testing.

denoted by  $\alpha$  and named *level of significance*. Of course, this error can be minimised by never rejecting the null hypothesis, i.e.,  $\alpha = 0$ . However, this rule makes it also impossible to discover situations in which the null hypothesis is indeed wrong. This type of error is called *type II error*, i.e., erroneously keeping the null hypothesis when it is false. Its probability is often denoted by  $\beta$ , see Tab. 2.2.

The idea of the *Kolmogorov-Smirnov test* is to test whether  $n$  rv's, which can be treated as being mutually independent, are drawn from the hypothesized distribution  $p_X$ , i.e.,

$$H_0 : X_j \sim p_X. \quad (2.92)$$

It is convenient to work with the cdf's as it is mentioned in the previous section that the cdf of the random sample  $\{X_1, \dots, X_n\}$  can be estimated directly from data without any binning. The “distance” between the estimated cdf  $\tilde{F}_X$  and hypothesized cdf  $\Phi_X$  is given by

$$d_X^{\text{KS}} = \sup_z |\tilde{F}_X(z) - \Phi(z)|. \quad (2.93)$$

Calculating the supremum numerically is a bit tricky. However, using that the estimated cdf has the form of a step function, the supremum can be computed by finding the maximum of a finite set. That is, if the values of the rv's are denoted by  $x_1, \dots, x_n$ , respectively, they can be rearranged into  $\{z_1, \dots, z_n\} = \{x_1, \dots, x_n\}$  such that  $z_1 \leq z_2 \leq \dots \leq z_n$ . Additionally,  $z_0$  is defined to be  $-\infty$  so that  $\tilde{F}_X(z_0) = 0$ . The estimated cdf reads

$$\tilde{F}_X(z_l) = \frac{1}{n} \max_{k=0, \dots, n} \{k : z_k = z_l\} \quad (2.94)$$

for  $l = 0, 1, \dots, n$ . Using that  $\Phi_X(\cdot)$  is monotonically increasing, it suffices to compute the maximum value

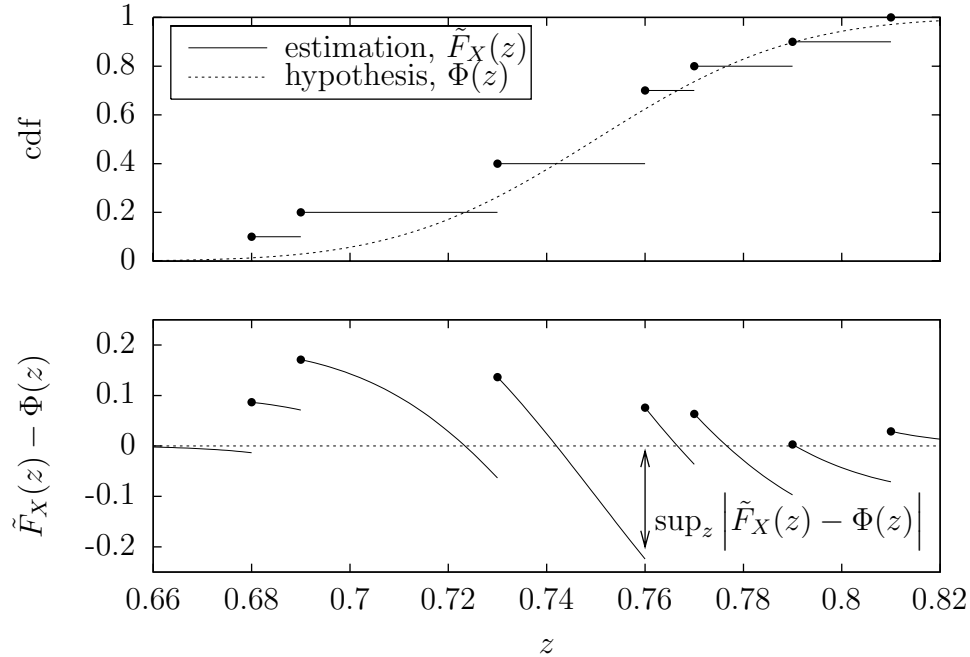
$$d_X^{\text{KS}} = \max_{l=1, \dots, n} \left\{ \left| \tilde{F}_X(z_l^*) - \Phi(z_l) \right|, \left| \tilde{F}_X(z_l) - \Phi(z_l) \right| \right\}, \quad (2.95)$$

where  $z_l^* = \max\{z_k : z_k < z_l\}$ . The larger  $d_X^{\text{KS}}$  is, the less likely it is that  $H_0$  is true. The null hypothesis is rejected if  $d_X^{\text{KS}}$  is larger than a given critical value. The crucial question is: how large should this critical value be such that making a type I error when rejecting  $H_0$  has probability  $\alpha < 1$ . Note that  $\alpha = 0$  means “never reject  $H_0$ ” so that  $\alpha$  is usually chosen to be 1 %, 5 %, or 10 %. Hence, the critical value should decrease with increasing  $\alpha$ . Moreover, the critical value should depend on the sample size as well and decrease with increasing sample size. That is, the null hypothesis is rejected with significance level  $\alpha$  if

$$\sqrt{n} d_X^{\text{KS}} \geq D^{\text{KS}}(\alpha, n). \quad (2.96)$$

Note that the right hand side of this equation still depends on  $n$ . However, this dependence is neglected for  $n > 40$ . The critical values are tabled in works like [Miller \(1956\)](#); [Daniel \(1990\)](#); [Hartung \(2005\)](#).

**Example 2.20:** Exemplary Kolmogorov-Smirnov test for a random sample of size  $n = 10$  to illustrate the concept. Consider ten rv's with values  $x_1 = 0.79$ ,  $x_2 = 0.68$ ,  $x_3 = 0.76$ ,  $x_4 = 0.73$ ,  $x_5 = 0.69$ ,  $x_6 = 0.77$ ,  $x_7 = 0.76$ ,  $x_8 = 0.76$ ,  $x_9 = 0.73$ , and  $x_{10} = 0.81$  and test them for being drawn from a normal distribution with mean 0.75 and variance 0.001, i.e.,  $\Phi(x) = \Phi_{0.75,0.001}(x)$  (see App. A.2 for notation). Fig. 2.3 illustrates the procedure to find  $d_X^{\text{KS}} \approx 0.22$ . The critical value reads  $D^{\text{KS}}(\alpha = 1\%, n = 10) = 1.55$  so that with significance level of 1 % the null hypothesis cannot be rejected. ■



$l$	0	1	2	3	4	5	6	7	8	9	10
$z_l$	$-\infty$	0.68	0.69	0.73		0.76	0.77	0.79	0.81		
$z_l^*$	-	$-\infty$	0.68	0.69		0.73	0.76	0.77	0.79		
$\tilde{F}_X(z_l)$	0.00	0.10	0.20	0.40		0.70	0.80	0.90	1.00		
$\tilde{F}_X(z_l^*)$	-	0.00	0.10	0.20		0.40	0.70	0.80	0.90		
$\Phi(z_l)$	0	0.01	0.03	0.26		0.62	0.74	0.90	0.97		
$\tilde{F}_X(z_l) - \Phi(z_l)$	-	0.09	0.17	0.14		0.08	0.06	0.00	0.03		
$\tilde{F}_X(z_l^*) - \Phi(z_l)$	-	-0.01	0.07	-0.06		-0.22	-0.04	-0.10	-0.07		

$$\therefore d_X^{\text{KS}} \approx 0.22$$

**Figure 2.3:** Exemplary Kolmogorov-Smirnov test from Ex. 2.20. The first graph shows the estimated and hypothetical cdf, whereas the second graph displays the difference. The table illustrates the algorithm to find the supremum (2.93) by computing the maximum (2.95). All values are rounded to two decimal digits.



# Chapter 3

## Stochastic Processes

Guided by [Sobczyk \(1991\)](#) and [Ross \(1996\)](#), which if not stated otherwise serve as references, this chapter deals with stochastic processes, which play an essential role in time series analysis and of course in wind velocity modelling.

A *stochastic process*  $\mathbf{X} = \{X(t) : t \in \mathcal{T}\}$  is a collection of rv's. The set  $\mathcal{T}$  is called the *index set* and for each  $t \in \mathcal{T}$  the variable  $X(t)$  is a rv, which can be either discrete or continuous. The index  $t$  is often interpreted as time and thus  $X(t)$  is often interpreted as the state of the process at time  $t$ . If  $\mathcal{T}$  is a countable set, the process  $\mathbf{X}$  is called (time) discrete, whereas if  $\mathcal{T}$  is uncountable, the process  $\mathbf{X}$  is said to be (time) continuous. Any realisation of  $\mathbf{X}$  is called a *sample path* and denoted by  $x(t)$  with  $t \in \mathcal{T}$ .

The variety of stochastic processes is very broad. There are for example processes, where the rv's  $X(s)$  and  $X(t)$  with  $s, t \in \mathcal{T}$  are independent/not independent or of the same distribution/not of the same distribution etc.

The following sections discuss some classifications and give examples of discrete and continuous stochastic processes.

### 3.1 Stationary processes

A stochastic process is said to be *stationary of order*  $p \in \mathbb{N}$  if the joint distribution of any  $k \leq p$  rv's of this process  $X(t_1), X(t_2), \dots, X(t_k)$  with  $t_1, t_2, \dots, t_k \in \mathcal{T}$  is identical to the joint distribution of the rv's  $X(t_1 + s), X(t_2 + s), \dots, X(t_k + s)$  for any real  $s$  such that  $t_1 + s, t_2 + s, \dots, t_k + s \in \mathcal{T}$ . As a special case, the distribution of the rv  $X(t)$  with  $t \in \mathcal{T}$  of a stochastic process which is stationary of at least first order does not depend on  $t \in \mathcal{T}$ , i.e.,  $X(s) \sim X(t)$  for any  $s, t \in \mathcal{T}$ . If a process is stationary of all orders, it is called *strictly stationary*.

### 3.2 Autocorrelation function

In general, the correlation between the two rv's  $X(s)$  and  $X(t)$  of a stochastic process  $\mathbf{X}$  with  $s, t \in \mathcal{T}$  and  $\text{Var}[X(s)], \text{Var}[X(t)] < \infty$  is a function of both times  $s$  and  $t$ . This

function is named *correlation function*  $\Gamma_{\mathbf{X}} : \mathcal{T} \times \mathcal{T} \mapsto [-1, +1]$  with

$$\Gamma_{\mathbf{X}}(s, t) = \text{Corr}[X(s), X(t)]. \quad (3.1)$$

If the process is at least second order stationary,  $\Gamma_{\mathbf{X}}(s, t) = \Gamma_{\mathbf{X}}(s + \tau, t + \tau)$  for any  $\tau$  such that  $s + \tau$  and  $t + \tau$  are elements of  $\mathcal{T}$ . Choosing  $\tau = -s$  leads to  $\Gamma_{\mathbf{X}}(s, t) = \Gamma_{\mathbf{X}}(0, t - s)$  which is a function solely of  $t - s$ . Choosing  $\tau = -t$  leads to  $\Gamma_{\mathbf{X}}(s, t) = \Gamma_{\mathbf{X}}(s - t, 0)$  which is a function solely of  $s - t$ . Hence, the correlation between the rv's  $X(s)$  and  $X(t)$  of a second order stationary process is a function of  $|s - t|$  and called *auto correlation function*

$$\gamma_{\mathbf{X}}(\tau) = \text{Corr}[X(t), X(t + \tau)]. \quad (3.2)$$

This function is symmetric, i.e.,  $\gamma_{\mathbf{X}}(\tau) = \gamma_{\mathbf{X}}(-\tau)$ , bounded by  $|\gamma_{\mathbf{X}}(\tau)| \leq 1$ , and takes the value  $\gamma_{\mathbf{X}}(0) = 1$ .

For the estimation of the mean and variance of a discrete stochastic process with  $\mathcal{T} = \{1, 2, \dots, n\}$  the average correlation  $\rho_n$  defined in Eq. (2.50) plays a crucial role. If the process is at least second order stationary this quantity is given by

$$\begin{aligned} \rho_n &= \frac{1}{n(n-1)} \sum_{t=1}^n \sum_{\substack{s=1 \\ s \neq t}}^n \gamma_{\mathbf{X}}(s-t) \\ &= \frac{2}{n-1} \sum_{\tau=1}^{n-1} \left(1 - \frac{\tau}{n}\right) \gamma_{\mathbf{X}}(\tau). \end{aligned} \quad (3.3)$$

A common case for an auto correlation function is that above a certain threshold  $\tau > \tau_0$  the contribution to the sum can be neglected and therefore  $\rho_{\infty}$  vanishes. That is, in order to estimate the mean and variance of a time series with  $\gamma_{\mathbf{X}}(\tau) \neq 0$  for  $\tau \geq 1$ , it is required that the time series is long enough.

### 3.3 $\delta$ -correlated processes

If the two rv's  $X(s)$  and  $X(t)$  of a stochastic process for any  $s \neq t$  ( $s, t \in \mathcal{T}$ ) are uncorrelated, the process is called  *$\delta$ -correlated* because

$$\text{Corr}[X(s), X(t)] = \delta_{s,t} = \begin{cases} 1 & \text{for } s = t \text{ and} \\ 0 & \text{for } s \neq t. \end{cases} \quad (3.4)$$

Note that even though  $\delta_{s,t} = \delta_{s-t,0}$  is a function of  $s - t$ , the process might not be second order stationary because  $X(s)$  and  $X(t)$  can have totally different distributions. It only says that  $X(s)$  and  $X(t)$  are uncorrelated. But if the process is second order stationary, its auto correlation function is  $\gamma_{\mathbf{X}}(\tau) = \delta_{\tau,0}$ .

### 3.4 Gaussian processes

A stochastic process  $\mathbf{X}$  with index set  $\mathcal{T}$  is said to be a *Gaussian process* or *normal process* if for any integer  $n$  and any subset  $\{t_1, \dots, t_n\} \subset \mathcal{T}$  the joint distribution of the rv's  $X(t_1), \dots, X(t_n)$  is a (multivariate) normal distribution. That is, the multivariate characteristic function reads

$$\begin{aligned} \varphi_{X(t_1), \dots, X(t_n)}(\lambda_1, \dots, \lambda_n) &= \mathbb{E} \left[ \exp \left( i \sum_{k=1}^n \lambda_k X(t_k) \right) \right] \\ &= \exp \left[ i \sum_{k=1}^n \mu_k s_k - \frac{1}{2} \sum_{k,l=1}^n \lambda_k \sigma_{kl}^2 \lambda_l \right], \end{aligned} \quad (3.5)$$

where  $\mu_k = \mathbb{E}[X(t_k)]$  and  $\sigma_{kl}^2 = \text{Cov}[X(t_k), X(t_l)]$  for  $k, l = 1, \dots, n$ . A Gaussian process  $\mathbf{X}$  has the following basic properties (see [Sobczyk, 1991](#)):

1. The process is entirely determined by its mean function  $\mu_{\mathbf{X}}(t) = \mathbb{E}[X(t)]$  and *covariance function*  $R_{\mathbf{X}}(s, t) = \text{Cov}[X(s), X(t)]$ .
2. All processes obtained by linear transformation of  $\mathbf{X}$  are also normal.
3. The central moments of odd order are equal to zero, i.e., for any finite  $q \in \mathbb{N}$

$$\mathbb{E}[(X(t_1) - \mu(t_1)) \cdots (X(t_{2q-1}) - \mu(t_{2q-1}))] = 0. \quad (3.6)$$

The even central moments can be expressed in terms of the covariance function by

$$\begin{aligned} &\mathbb{E}[(X(t_1) - \mu(t_1)) \cdots (X(t_{2q}) - \mu(t_{2q}))] \\ &= \sum_{\text{pairs}} R_{\mathbf{X}}(t_{i_1}, t_{i_2}) \cdots R_{\mathbf{X}}(t_{i_{2q-1}}, t_{i_{2q}}), \end{aligned} \quad (3.7)$$

where summation is over the  $(2q - 1)!!$  possible ways of dividing the set  $\{1, \dots, 2q\}$  into  $q$  combination of pairs  $(i_1, i_2), (i_3, i_4), \dots, (i_{2q-1}, i_{2q})$ .

It immediately follows that a second order stationary Gaussian process is strictly stationary

**Example 3.1:** Four-point correlation of a Gaussian process. A special case of (3.7) is  $\mathbb{E}[\hat{X}(t_1)\hat{X}(t_2)\hat{X}(t_3)\hat{X}(t_4)] = R_{12}R_{34} + R_{13}R_{24} + R_{14}R_{23}$ , where  $\hat{X}(t_i) = X(t_i) - \mu(t_i)$  and  $R_{ij} = R_{\mathbf{X}}(t_i, t_j)$  for  $i, j \in \{1, 2, 3, 4\}$ . ■

The following examples describe some Gaussian processes which are needed in this thesis.

**Example 3.2:** The *standard Gaussian white noise*  $\{\xi(t) : t \geq 0\}$  consists of normally distributed rv's  $\xi(t)$  with zero mean and unit variance. Any two rv's  $\xi(s)$  and  $\xi(t)$  with  $s \neq t$  ( $s, t \in \mathcal{T}$ ) are defined to be independent. The process can be discrete or continuous. ■

**Example 3.3:** The *standard Wiener process* or *standard Brownian motion*  $\mathbf{W} = \{W(t) : t \geq 0\}$  is characterised by  $W(0) = 0$  and has independent increments

$$W(s) - W(t) \sim N_{0,|s-t|}, \quad (3.8)$$

i.e., the increments  $W(t_1) - W(t_0)$ ,  $W(t_3) - W(t_2)$ ,  $\dots$ ,  $W(t_k) - W(t_{k-1})$  are mutually independent for arbitrary  $0 < t_0 < t_1 < t_2 < \dots < t_k$  ( $k \in \mathbb{N}$ ). Hence,  $W(t) \sim N_{0,t}$  so that the process is not stationary. The mean function reads  $\mu_{\mathbf{W}}(t) = 0$  and the covariance function is given by  $R_{\mathbf{W}}(s, t) = \min\{s, t\}$ . ■

**Example 3.4:** The *general Wiener process*  $\mathbf{X} = \{X(t) : t \geq 0\}$  is defined by  $X(t) = \alpha t + \sqrt{\beta}W(t)$  ( $\alpha \in \mathbb{R}$ ,  $\beta \in \mathbb{R}_{\geq 0}$ ), where  $\{W(t) : t \geq 0\}$  denotes a standard Wiener process. It is again a Gaussian process with  $X(t) \sim N_{\alpha t, \beta^2 t}$  and has mutually independent increments with  $X(s) - X(t) \sim N_{\alpha(s-t), \beta|s-t|}$ . The mean and covariance function read  $\mu_{\mathbf{X}}(t) = \alpha t$  and  $R_{\mathbf{X}}(s, t) = \beta \min\{s, t\}$ , respectively. ■

**Example 3.5:** *Ornstein-Uhlenbeck process.* The processes in the previous examples are non-stationary, whereas the Ornstein-Uhlenbeck process  $\mathbf{U} = \{U(t) : t \geq 0\}$  is stationary. It is defined by

$$U(t) = \lambda e^{-\beta t} W(e^{2\beta t}) \quad (3.9)$$

with  $\{W(t) : t \geq 0\}$  denoting a standard Wiener process. Being a linear transformation of a normal process,  $\mathbf{U}$  is a Gaussian process, too. It suffices to compute the mean and covariance function reading  $\mu_{\mathbf{U}}(t) = 0$  and

$$R_{\mathbf{U}}(s, t) = \lambda^2 e^{-\beta(s+t)+2\beta \min\{s,t\}} = \lambda^2 e^{-\beta|s-t|}, \quad (3.10)$$

respectively. Hence, the process is stationary and has an exponentially decaying auto correlation function given by  $\gamma_{\mathbf{U}}(\tau) = e^{-\beta|\tau|}$ . ■

## 3.5 Transformed Gaussian processes

Non-Gaussian stochastic processes can be obtained by taking a nonlinear transformation of a Gaussian process. That is, if  $\mathbf{X} = \{X(t) : t \in \mathcal{T}\}$  is a Gaussian process and  $f : \mathcal{T} \times \mathbb{R} \mapsto \mathbb{R}$ , the set  $\mathbf{Y} = \{Y(t) : t \in \mathcal{T}\}$  with

$$Y(t) = f(t, X(t)) \quad (3.11)$$

is again a stochastic process.

It is convenient to expand (3.11) in terms of Hermite polynomials  $H_n(\cdot)$  with  $n \in \mathbb{N}_0$  (see [Bronstein et al., 1999](#), Eq. (9.63) for the definition), i.e.,

$$Y(t) = f(t, X(t)) = \sum_{k=0}^{\infty} \alpha_k(t) H_k \left( \frac{X(t) - \mu_{\mathbf{X}}(t)}{\sqrt{R_{\mathbf{X}}(t, t)}} \right), \quad (3.12)$$

because these polynomials constitute an orthogonal system with respect to the standard normal pdf. [Granger & Newbold \(1976\)](#) show that

$$\mathbb{E} \left[ H_n \left( \frac{X(s) - \mu_{\mathbf{X}}(s)}{\sqrt{R_{\mathbf{X}}(s, s)}} \right) H_m \left( \frac{X(t) - \mu_{\mathbf{X}}(t)}{\sqrt{R_{\mathbf{X}}(t, t)}} \right) \right] = \Gamma_{\mathbf{X}}(s, t)^n n! \delta_{n, m}, \quad (3.13)$$

where  $\Gamma_{\mathbf{X}}(s, t) = \mathbb{C} \text{orr} [X(s), X(t)]$  denotes the correlation function of  $\mathbf{X}$ . Due to  $H_0(\hat{x}) = 1$ , the mean function of  $\mathbf{Y}$  is given by

$$\mu_{\mathbf{Y}}(t) = \mathbb{E} [Y(t)] = \alpha_0(t). \quad (3.14)$$

Therefore, the covariance function reads

$$R_{\mathbf{Y}}(s, t) = \mathbb{E} [Y(s)Y(t)] - \alpha_0(s)\alpha_0(t) = \sum_{k=1}^{\infty} k! \alpha_k(s)\alpha_k(t) \gamma_{\mathbf{X}}(s, t)^k. \quad (3.15)$$

**Example 3.6:** The  $\chi_1^2$ -process is obtained by squaring a Gaussian process with zero mean and unit variance, i.e.,  $Y(t) = X(t)^2 \sim \chi_1^2$  with  $X(t) \sim N_{0,1}$ . Therefore,  $\alpha_0 = \alpha_2 = 1$  while any other  $\alpha_l$  is set to zero so that  $\mu_{\mathbf{Y}}(t) = 1$  and  $R_{\mathbf{Y}}(s, t) = 2\gamma_{\mathbf{X}}(s, t)^2$  and hence  $\text{Var} [Y(t)] = 2$ . ■

A special case of (3.11), the exponential distribution, leads to the class of *geometric Gaussian processes* which is characterised by

$$Y(t) = e^{X(t)} \quad (3.16)$$

so that the rv  $Y(t)$  is log-normally distributed. It has expectation

$$\mathbb{E} [Y(t)] = \exp \left[ \mu_{\mathbf{X}}(t) + \frac{R_{\mathbf{X}}(t, t)}{2} \right] \quad (3.17)$$

and coefficient of variation

$$\mathbb{C} \text{V} [Y(t)] = \sqrt{e^{R_{\mathbf{X}}(t, t)} - 1}. \quad (3.18)$$

In order to compute the correlation function  $\Gamma_{\mathbf{Y}}(s, t) = R_{\mathbf{Y}}(s, t) / \sqrt{R_{\mathbf{Y}}(s, s)R_{\mathbf{Y}}(t, t)}$  of  $\mathbf{Y}$ , the exponential function needs to be expanded in terms of Hermite polynomials. ([Granger & Newbold, 1976](#), Eq. (A.3)) states that  $\exp[ax - a^2/2] = \sum_{n=0}^{\infty} H_n(x) a^n / n!$  leading to

$$Y(t) = e^{\mu_{\mathbf{X}}(t) + R_{\mathbf{X}}(t, t)/2} \sum_{k=0}^{\infty} \frac{R_{\mathbf{X}}(t, t)^k}{k!} H_k \left[ \frac{X(t) - \mu_{\mathbf{X}}(t)}{\sqrt{R_{\mathbf{X}}(t, t)}} \right] \quad (3.19)$$

so that

$$\Gamma_{\mathbf{Y}}(s, t) = \frac{e^{R_{\mathbf{X}}(s, t)} - 1}{\sqrt{(e^{R_{\mathbf{X}}(s, s)} - 1)(e^{R_{\mathbf{X}}(t, t)} - 1)}}. \quad (3.20)$$

The following example describes a geometric Gaussian process which will play an essential role in this thesis when it comes to the simulation of wind speed.

**Example 3.7:** The *geometric Ornstein-Uhlenbeck process*  $\mathbf{G}$  is defined by  $\{G(t) : t \geq 0\}$  with  $G(t) = e^{U(t)}$ , where  $U(t)$  denotes an Ornstein-Uhlenbeck process (3.9). As  $R_{\mathbf{U}}(t, t) = \lambda^2 = \text{const}$ , the coefficient of variation is constant, i.e.,

$$\mathbb{C}\mathbb{V}[G(t)] = \sqrt{e^{\lambda^2} - 1}. \quad (3.21)$$

As the process is stationary, its auto correlation function is given by

$$\gamma_{\mathbf{G}}(\tau) = \frac{\exp[\lambda^2 e^{-\beta|\tau|}] - 1}{\exp[\lambda^2] - 1} \quad (3.22)$$

by employing (3.20). ■

The following two examples describe some geometric Gaussian processes which have applications beyond the scope of this thesis.

**Example 3.8:** The *geometric Gaussian white noise* is simply  $\{e^{\alpha+\beta\xi(t)} : t \in \mathcal{T}\}$  with  $\{\xi(t) : t \in \mathcal{T}\}$  denoting standard Gaussian white noise. The process (or a sum of several processes) is used to model non-Gaussian background noise for developing e.g. sea clutter radars (see Kassam & Thomas, 1976; Al-Hussaini & Turner, 1979) or wireless communication devices (see Beaulieu *et al.*, 1993; Cardieri & Rappaport, 2000; Renzo & Graziosi, 2009). ■

**Example 3.9:** The *geometric Brownian motion* or *geometric Wiener process* is defined by  $\{e^{\alpha t + \beta W(t)} : t \geq 0\}$ , where  $\{W(t) : t \geq 0\}$  denotes a standard Wiener process. It is characterised by  $\mu(t) = \alpha t$  and  $\sigma^2(t) = \beta^2 t$ . The process is used for instance in financial mathematics as a model for the stock price (see Steele, 2001; Baaquie, 2004). ■

## 3.6 Discretisation of continuous processes

The previous two sections introduced a couple of continuous processes. There is a convenient way to obtain a discrete stochastic process by discretising a continuous stochastic process.

If  $\mathbf{X} = \{X(t) : t \geq 0\}$  is a continuous stochastic process, its *discretised process*  $\{X_n : n \in \mathbb{N}_0\}$  is defined by

$$X_n = X(n\Delta t), \quad (3.23)$$

where  $\Delta t$  is an intrinsic parameter, the *discretisation time*. The following examples derive very prominent discretised processes.

**Example 3.10:** *Discrete Wiener process.* The increments of a Wiener process  $\{W(t) : t \geq 0\}$  described in Ex. 3.3 and Eq. (3.8) are normally distributed so that

$$W(t + \Delta t) = W(t) + \sqrt{\Delta t} \xi(t) \quad (3.24)$$

with  $\{\xi(t) : t \geq 0\}$  denoting standard Gaussian white noise. That is, its discretised process  $\{W_n : n \in \mathbb{N}\}$  is generated by

$$W_{k+1} = W_k + \xi_k \sqrt{\Delta t} \quad (3.25)$$

with  $k \in \mathbb{N}_0$ ,  $W_0 = 0$  and  $\{\xi_k : k \in \mathbb{N}_0\}$  denoting discrete standard Gaussian white noise. Iterating (3.25)  $n$  times yields a non-recursive description of the discrete Wiener process

$$W_n = \sqrt{\Delta t} \sum_{k=0}^{n-1} \xi_k \sim N_{0, n\Delta t}. \quad (3.26)$$

■

**Example 3.11:** *Discrete Ornstein-Uhlenbeck process.* Discretising the Ornstein-Uhlenbeck process (3.9) leads to a stationary auto regressive process of first order, see Sec. 3.7. Its recursion formula can be derived from

$$U(t + \Delta t) = \lambda e^{-\beta(t+\Delta t)} W(e^{2\beta(t+\Delta t)}) \quad (3.27)$$

by using that according to (3.24) the rv  $W(e^{2\beta(t+\Delta t)})$  can be written as

$$W(e^{2\beta t} e^{2\beta\Delta t}) = W(e^{2\beta t}) + \xi(t) e^{\beta t} \sqrt{e^{2\beta\Delta t} - 1} \quad (3.28)$$

with  $\{\xi(t) : t \geq 0\}$  denoting standard Gaussian white noise. Consequently, Eq. (3.27) becomes

$$U(t + \Delta t) = e^{-\beta\Delta t} U(t) + \lambda \xi(t) \sqrt{1 - e^{-2\beta\Delta t}} \quad (3.29)$$

so that the discretised process  $\{U_n : n \in \mathbb{N}\}$  is given by

$$U_{k+1} = aU_k + \lambda \xi_k \sqrt{1 - a^2} \quad (3.30)$$

with  $k \in \mathbb{N}_0$ ,  $U_0 = 0$ ,  $a = e^{-\beta\Delta t} \in [0, 1]$  and  $\{\xi_k : k \in \mathbb{N}_0\}$  denoting discrete standard Gaussian white noise. As a consequence of the stationarity of the Ornstein-Uhlenbeck process, its discretised process is stationary as well. Moreover,  $U_n \sim U(n\Delta t) \sim N_{0, \lambda^2}$  and the auto correlation function reads  $\gamma(k) = a^{|k|}$  for  $k \in \mathbb{Z}$ . ■

## 3.7 ARMA processes

The Ex. 3.11 in the previous section derived a representative of an auto regressive process of first order. This process belongs to the class of auto regressive moving average (ARMA) processes.

Following [Montgomery et al. \(2008\)](#) the concept of a *linear filter* is introduced first. It is a linear operation from one discrete stochastic process  $\mathbf{Z} = \{Z_n : n \in \mathbb{Z}\}$  to another process  $\mathbf{X} = \{X_n : n \in \mathbb{Z}\}$  and is defined by

$$X_n = \sum_{k=-\infty}^{\infty} \psi_k Z_{n-k}. \quad (3.31)$$

Linear filters are said to be

- time-invariant if the coefficients  $\psi_k$  do not depend on time,
- physically realisable if  $\psi_k = 0$  for  $k < 0$ , and
- stable if  $\sum_{k=-\infty}^{\infty} |\psi_k| < \infty$ .

Applying a time-invariant, stable linear filter to a stationary process  $\mathbf{Z}$  with mean  $\mu_{\mathbf{Z}}$ , variance  $\sigma_{\mathbf{Z}}^2$ , and auto correlation function  $\gamma_{\mathbf{Z}}(\cdot)$  leads to a stationary process  $\mathbf{X}$  with mean  $\mu_{\mathbf{X}} = \mu_{\mathbf{Z}} \sum_{k=-\infty}^{\infty} \psi_k$ , variance  $\sigma_{\mathbf{X}}^2 = \sigma_{\mathbf{Z}}^2 \sum_{k,l=-\infty}^{\infty} \psi_k \psi_l \gamma_{\mathbf{Z}}(l-k)$ , and auto correlation function

$$\gamma_{\mathbf{X}}(\tau) = \frac{\sum_{k,l=-\infty}^{\infty} \psi_k \psi_l \gamma_{\mathbf{Z}}(\tau + l - k)}{\sum_{k,l=-\infty}^{\infty} \psi_k \psi_l \gamma_{\mathbf{Z}}(l - k)}. \quad (3.32)$$

The main application of linear filters is to model or to approximate a relationship between the input process and the output process while the true relationship is not known. This is done by estimating the parameters  $\psi_i$ , which unless there is a cut-off can be infinitely many. These models are based on observations in form of a time series and can be used to forecast the corresponding observable (see [Kantz & Schreiber, 2004](#); [Montgomery et al., 2008](#)).

A special case of these processes is obtained by applying a time-invariant, physically realisable, and stable linear filter on standard Gaussian white noise  $\mathbf{Z} = \{\xi_n : n \in \mathbb{Z}\}$ , i.e.,

$$X_n = \sum_{k=0}^{\infty} \psi_k \xi_{n-k}. \quad (3.33)$$

This process is again a Gaussian process and called *moving average process*. It has zero mean and its variance is equal to  $\sum_{k=0}^{\infty} \psi_k^2$ . Its auto correlation function reads

$$\gamma_{\mathbf{X}}(\tau) = \frac{\sum_{k=0}^{\infty} \psi_k \psi_{k+|\tau|}}{\sum_{k=0}^{\infty} \psi_k^2}. \quad (3.34)$$

If furthermore  $\psi_k = 0$  for  $k \geq q \in \mathbb{N}$ , the auto correlation function vanishes for  $\tau \geq q$ . The process is said to be of order  $q$  and is abbreviated by  $\text{MA}(q)$ .

**Example 3.12:** The  $\text{MA}(1)$  process is characterised by  $\psi_0 \neq 0$  and  $\psi_k = 0$  for  $k > 0$ . Its variance reads  $\psi_0^2$ . The auto correlation function is given by  $\gamma_{\mathbf{X}}(0) = 1$  and  $\gamma_{\mathbf{X}}(\tau) = 0$  for  $|\tau| > 0$ . ■

The  $\text{MA}(\infty)$  process at time  $n$  is characterised by the sum of infinite disturbances from the past. However, for adequately modelling such a process it is required to estimate only a finite number of weights. The solution to the problem is to assume that the infinitely many weights follow a distinct pattern and are represented by only a few parameters.

**Example 3.13:** An  $\text{MA}(\infty)$  process described by only two parameters is e.g. given by  $\psi_k = ba^k$  with  $a \in \mathbb{R}$  and  $b > 0$ . The process is stationary if (and only if)  $\sum_{k=0}^{\infty} b|a|^k < \infty$ , i.e.,  $|a| < 1$ . The variance of the process reads  $b^2/(1-a^2)$  and the auto correlation function



is given by  $\gamma_{\mathbf{X}}(\tau) = a^{|\tau|}$ . For general  $a \in \mathbb{R}$ , the rv  $X_n = b \sum_{k=0}^{\infty} a^k \xi_{n-k}$  can be written as  $X_n = b\xi_n + ab \sum_{k=0}^{\infty} a^k \xi_{n-1-k} = b\xi_n + aX_{n-1}$ . Equivalently,

$$X_{n+1} = aX_n + b\xi_n, \quad (3.35)$$

where the indices of the white noise term were re-labelled. If  $a = 1$ , this process corresponds to a discrete Wiener process (3.25) with discretisation time  $\Delta t = b^2$ . If  $0 \leq a < 1$ , this process corresponds to a discrete Ornstein-Uhlenbeck process (3.30) with discretisation time  $\Delta t = -\ln(a)/\beta$  and variance  $\lambda^2 = b^2/(1 - a^2)$ . ■

Eq. (3.35) can be understood as a regression of  $X_{n+1}$  and  $X_n$  and can be generalised to describe the class of *auto regressive processes* of order  $p$ , in short AR( $p$ ), by

$$X_{n+1} = \sum_{k=0}^{p-1} a_k X_{n-k} + b\xi_n, \quad (3.36)$$

where  $a_0, \dots, a_{p-1} \in \mathbb{R}$  and  $b > 0$  are  $p + 1$  parameters. Exemplified by the AR(1) process, the previous example demonstrated that (3.36) does not necessarily describe a stationary process.

More generally, the class of *auto regressive moving average processes* ARMA( $p, q$ ) with  $p, q \in \mathbb{N}_0$  is described by allowing  $q + 1$  past noise terms to contribute, i.e.,

$$X_{n+1} = \sum_{k=0}^{p-1} a_k X_{n-k} + \sum_{j=0}^q b_j \xi_{n-j}. \quad (3.37)$$

It is characterised by  $p + q + 1$  parameters:  $a_0, \dots, a_{p-1} \in \mathbb{R}$ ,  $b_0, \dots, b_q \in \mathbb{R}_{\geq 0}$ . It should be noted that an ARMA( $p, q$ ) process is Gaussian with mean function  $\mu(n) = 0$  and is consequently characterised by the covariance function  $R(n, m)$ . In particular, ARMA( $p, 0$ ) = AR( $p$ ) and ARMA( $0, q$ ) = MA( $q$ ).

# Chapter 4

## Turbulence Theory

Guided by [Frisch \(1995\)](#); [Pope \(2000\)](#); [Sagaut & Cambon \(2008\)](#), which if not stated otherwise serve as references, this chapter introduces the theory of turbulence.

### 4.1 The Navier-Stokes equation

The Navier-Stokes equation (NSE), which is known since [Navier \(1823\)](#), is the fundamental equation describing a flow of particles under the *continuum hypothesis*, i.e., the discrete molecular nature of fluids can be treated as continuum. From a physical point of view, it is a momentum balance equation for an incompressible fluid.

Before deriving the NSE, the *law of mass conservation* is reviewed. It is a balance equation connecting the variation of mass with its flux and production within a volume  $V$ . Denoting an infinitesimal volume element by  $d^3\vec{r}$ , the surface of  $V$  by  $\partial V$ , and its infinitesimal normal vector by  $d\vec{s}$ , the law of mass conservation reads

$$\underbrace{\frac{d}{dt} \int_V \rho(\vec{r}, t) d^3\vec{r}}_{\text{variation}} = - \underbrace{\oint_{\partial V} \rho(\vec{r}, t) \vec{u}(\vec{r}, t) \cdot d\vec{s}}_{\text{flux}} + \underbrace{\int_V q(\vec{r}, t) d^3\vec{r}}_{\text{production}}, \quad (4.1)$$

where  $\rho$ ,  $\vec{u}$ , and  $q$  are the density, the velocity, and the rate of mass production, respectively. These fields are assumed to be continuous in position  $\vec{r}$  and time  $t$ . Writing the surface integral, i.e., the flux term, as a volume integral over the divergence yields

$$\int_V \left[ \frac{\partial \rho}{\partial t} + \nabla \cdot (\rho \vec{u}) - q \right] d^3\vec{r} = 0, \quad (4.2)$$

where the notation for the  $\vec{r}$  and  $t$  dependence is omitted for the sake of clarity. This relation has to hold true for any volume  $V$  so that

$$\frac{\partial \rho}{\partial t} + \nabla \cdot (\rho \vec{u}) = q \quad (4.3)$$

which is known as the *continuity equation*.

In the same way as for the mass, the conservation of the  $i^{\text{th}}$  component of the momentum  $\rho(\vec{r}, t)u_i(\vec{r}, t)$  per unit volume reads

$$\frac{\partial \rho u_i}{\partial t} + \nabla \cdot (\rho u_i \vec{u}) = b_i + f_i \quad (4.4)$$

with  $\vec{b}_i(\vec{r}, t)$  and  $f_i(\vec{r}, t)$  denoting the *body force* and *surface force*, respectively. Both forces are sources and sinks of momentum. The body force of interest is gravity  $\vec{b}(\vec{r}, t) = -\rho(\vec{r}, t)\nabla\Psi(\vec{r}, t)$  due to gravitational potential  $\Psi(\vec{r}, t)$ . More contributions to the body force, such as buoyancy or magnetic forces (in magnetohydrodynamics), can be regarded as well but are not considered here. The surface forces are of molecular nature and described by the symmetric *stress tensor*  $\tau_{ij}(\vec{r}, t)$  by  $f_i(\vec{r}, t) = \sum_{k=1}^3 \partial\tau_{ki}(\vec{r}, t)/\partial x_k$  with

$$\tau_{ij}(\vec{r}, t) = -P(\vec{r}, t)\delta_{ij} + \mu \left( \frac{\partial u_i(\vec{r}, t)}{\partial x_j} + \frac{\partial u_j(\vec{r}, t)}{\partial x_i} - \frac{2\delta_{ij}}{3} \nabla \cdot \vec{u}(\vec{r}, t) \right). \quad (4.5)$$

Here,  $P = P(\vec{r}, t)$  and  $\mu = \text{const}$  stand for pressure and coefficient of viscosity, respectively. Applying the chain rule, Eq. (4.4) becomes

$$u_i \frac{\partial \rho}{\partial t} + \rho \frac{\partial u_i}{\partial t} + \sum_{k=1}^3 \left[ u_i \frac{\partial}{\partial r_k} \rho u_k + \rho u_k \frac{\partial u_i}{\partial r_k} \right] = b_i + f_i \quad (4.6)$$

so that

$$u_i \underbrace{\left[ \frac{\partial \rho}{\partial t} + \nabla \cdot (\rho \vec{u}) \right]}_{=q} + \rho \left[ \frac{\partial u_i}{\partial t} + (\vec{u} \cdot \nabla) u_i \right] = b_i + f_i. \quad (4.7)$$

A system with no mass production, i.e.,  $q(\vec{r}, t) = 0$ , has consequently

$$\rho \left[ \frac{\partial u_i}{\partial t} + (\vec{u} \cdot \nabla) u_i \right] = b_i + f_i \quad (4.8)$$

for  $i = 1, 2, 3$ .

An incompressible fluid has constant density and no mass production. Inserting  $\rho(\vec{r}, t) = \rho_0 = \text{const}$  and  $q(\vec{r}, t) = 0$  into (4.3) leads to

$$\nabla \cdot \vec{u} = 0. \quad (4.9)$$

That is, the velocity field of an incompressible fluid flow is divergence-free or in other words *solenoidal*. The stress tensor becomes

$$\tau_{ij} = -P\delta_{ij} + \mu \left( \frac{\partial u_i}{\partial x_j} + \frac{\partial u_j}{\partial x_i} \right) \quad (4.10)$$

so that the surface force is given by

$$f_i = \sum_{k=1}^3 \frac{\partial \tau_{ki}}{\partial x_i} = -\frac{\partial P}{\partial x_i} + \mu \nabla^2 u_i. \quad (4.11)$$

quantity	dimension	quantity	dimension
$\vec{u}(\vec{r}, t)$	$\text{LT}^{-1}$	$\nu$	$\text{L}^2\text{T}^{-1}$
$p(\vec{r}, t)$	$\text{MT}^{-2}\text{L}^{-1}$	$\rho_0$	$\text{ML}^{-3}$

**Table 4.1:** Dimension of the quantities in the NSE (L...length, T...time, M...mass).

Eq. (4.8) with  $\rho(\vec{r}, t) = \rho_0 = \text{const}$  becomes

$$\frac{\partial \vec{u}}{\partial t} + (\vec{u} \cdot \nabla) \vec{u} = -\frac{1}{\rho_0} \nabla p + \nu \nabla^2 \vec{u} \quad (4.12)$$

with *modified pressure*  $p(\vec{r}, t) = P(\vec{r}, t) + \rho_0 \Psi(\vec{r}, t)$  and *kinematic viscosity*  $\nu = \mu/\rho_0$ . Eq. (4.12) is known as the *Navier-Stokes equation*.

As a summary, the flow of an incompressible fluid with constant density and kinematic viscosity is described by the NSE together with the solenoidal condition (4.9). The problem is self-consistent and well posed: there are four parameters,  $\vec{u}$  and  $p$ , and four equations, (4.9) and (4.12). The dimensions of the quantities are listed in Tab. 4.1.

## 4.2 Periodic boundary condition and non-locality

The NSE is nonlinear due to the *advection term*  $(\vec{u} \cdot \nabla) \vec{u}$ . However, it is not this nonlinearity which makes it difficult to solve the NSE but rather its non-locality, which is shown in this section.

It is easy to eliminate the (modified) pressure from the NSE by taking the divergence of (4.12) and using the solenoidal condition so that

$$\nabla^2 p = -\rho_0 \sum_{i,j=1}^3 \frac{\partial^2}{\partial x_i \partial x_j} u_i u_j. \quad (4.13)$$

This equation is an instance of the Poisson equation  $\nabla^2 p(\vec{r}, t) = \sigma(\vec{r}, t)$  with  $\sigma(\vec{r}, t) = -\rho_0 \sum_{i,j=1}^3 \partial_{ij}^2 u_i(\vec{r}, t) u_j(\vec{r}, t)$ .

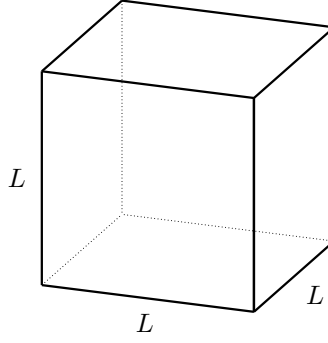
If the fluid fills all of the space  $\mathbb{R}^3$ , its unboundness leads to some mathematical difficulties. Hence, a periodic boundary condition for the space variable  $\vec{r}$  is assumed, i.e.,

$$\vec{u}(\vec{r} + \vec{s}L, t) = \vec{u}(\vec{r}, t) \quad \text{and} \quad p(\vec{r} + \vec{s}L, t) = p(\vec{r}, t) \quad (4.14)$$

with  $\vec{s} \in \mathbb{Z}^3$ . Consequently, it suffices to consider the flow in the *periodicity box*  $B_L$  visualised in Fig. 4.1.

Going from physical space into the Fourier space with

$$p(\vec{r}, t) = \sum_{\vec{k} \in \mathbb{Z}^3} e^{2\pi i \vec{k} \cdot \vec{r} / L} \hat{p}(\vec{k}, t) \quad \text{and} \quad \sigma(\vec{r}, t) = \sum_{\vec{k} \in \mathbb{Z}^3} e^{2\pi i \vec{k} \cdot \vec{r} / L} \hat{\sigma}(\vec{k}, t) \quad (4.15)$$



**Figure 4.1:** The periodicity box  $B_L$  as shown in Frisch (1995).

results in

$$\hat{p}(\vec{k}, t) = -\frac{L^2}{4\pi^2} \frac{\hat{\sigma}(\vec{k}, t)}{k^2}, \quad (4.16)$$

where  $\vec{k} \in \mathbb{Z}^3 \setminus \{\vec{0}\}$  and  $k^2 = \vec{k} \cdot \vec{k}$ . This equation does not yield a value for the coefficient  $\hat{p}(\vec{0}, t)$ , which reflects that the solution of the Poisson equation is defined up to an additive constant. Still, the NSE is invariant under adding a constant term to the pressure so that the value of  $\hat{p}(\vec{0}, t)$  is irrelevant. Going back into physical space yields

$$p(\vec{r}, t) = \frac{\rho_0}{4\pi} \int_{B_L} \frac{d^3 \vec{s}}{|\vec{r} - \vec{s}|} \sum_{i,j=1}^3 \frac{\partial^2}{\partial x_i \partial x_j} u_i(\vec{s}, t) u_j(\vec{s}, t) + \text{const} \quad (4.17)$$

(see Pope, 2000, Eq. (2.49)). This equation demonstrates the non-locality problem, i.e., in order to compute the pressure at point  $\vec{r}$  the velocity at all other points  $\vec{s} \in B_L$  needs to be known.

An alternative way to eliminate the pressure from (4.12) is to take the curl of the NSE because  $\nabla \times \nabla p$  is equal to zero. It is convenient to work with the *vorticity* defined by

$$\vec{\omega}(\vec{r}, t) = \nabla \times \vec{u}(\vec{r}, t). \quad (4.18)$$

Applying the curl operator to the NSE and writing its advection term  $(\vec{u} \cdot \nabla) \vec{u}$  as  $\frac{1}{2} \nabla |\vec{u}|^2 - \vec{u} \times (\nabla \times \vec{u})$  leads to the vorticity equation

$$\frac{\partial}{\partial t} \vec{\omega} = \nabla \times (\vec{u} \times \vec{\omega}) + \nu \nabla^2 \omega. \quad (4.19)$$

Here, it is used that  $\nabla \times \nabla |\vec{u}|^2$  vanishes. Solving the vorticity equation by inverting (4.18) leads however to the same non-locality problem.

### 4.3 Conservation laws

This section deals with global conservation laws, i.e., the conservation of quantities averaged over the whole volume. The reader is referred to Lamb (1932) for some local conservation laws.

$\langle \nabla f \rangle = \vec{0}$ $\langle (\nabla g) f \rangle = - \langle f \nabla g \rangle$ $\langle (\nabla^2 g) f \rangle = - \langle (\nabla f) \cdot (\nabla g) \rangle$ $\langle \vec{u} \cdot (\nabla \times \vec{v}) \rangle = \langle \nabla \times \vec{u} \cdot \vec{v} \rangle$ $\langle \vec{u} \cdot (\nabla^2 \vec{v}) \rangle = - \langle \nabla \times \vec{u} \cdot (\nabla \times \vec{v}) \rangle \text{ if } \nabla \cdot \vec{v} = 0$
---

**Table 4.2:** Properties of the average of periodic functions (Frisch, 1995, Eqs. (2.17) to (2.21)). The average is taken over the periodicity box.

As in the previous section, periodic boundary conditions are assumed and the spatial mean of any quantity  $f(\vec{r}, t)$  is denoted by

$$\langle f(\vec{r}, t) \rangle_{\vec{r}} = \frac{1}{L^3} \int_{B_L} f(\vec{r}, t) d^3 \vec{r}. \quad (4.20)$$

Properties of the average of a periodic function (scalar/vector) are listed in Tab. 4.2.

**Conservation of momentum.** Applying the mean operator to the NSE (4.12) yields  $\frac{d}{dt} \langle \vec{u} \rangle = 0$ . This equation states that the mean momentum per unit volume  $\rho_0 \vec{u}$  is conserved. Clearly, this is a consequence of the NSE representing a momentum balance equation.

**Conservation of energy.** Applying the average operator to  $\vec{u} \cdot (4.12)$  leads to  $\frac{d}{dt} \langle \frac{1}{2} |\vec{u}|^2 \rangle = -\nu \langle |\nabla \times \vec{u}|^2 \rangle$ . Defining the mean energy per unit mass by

$$E(t) = \frac{1}{2} \langle |\vec{u}(\vec{r}, t)|^2 \rangle_{\vec{r}} \quad (4.21)$$

and the mean *enstrophy* per unit mass by

$$\Omega(t) = \frac{1}{2} \langle |\vec{\omega}(\vec{r}, t)|^2 \rangle_{\vec{r}}, \quad (4.22)$$

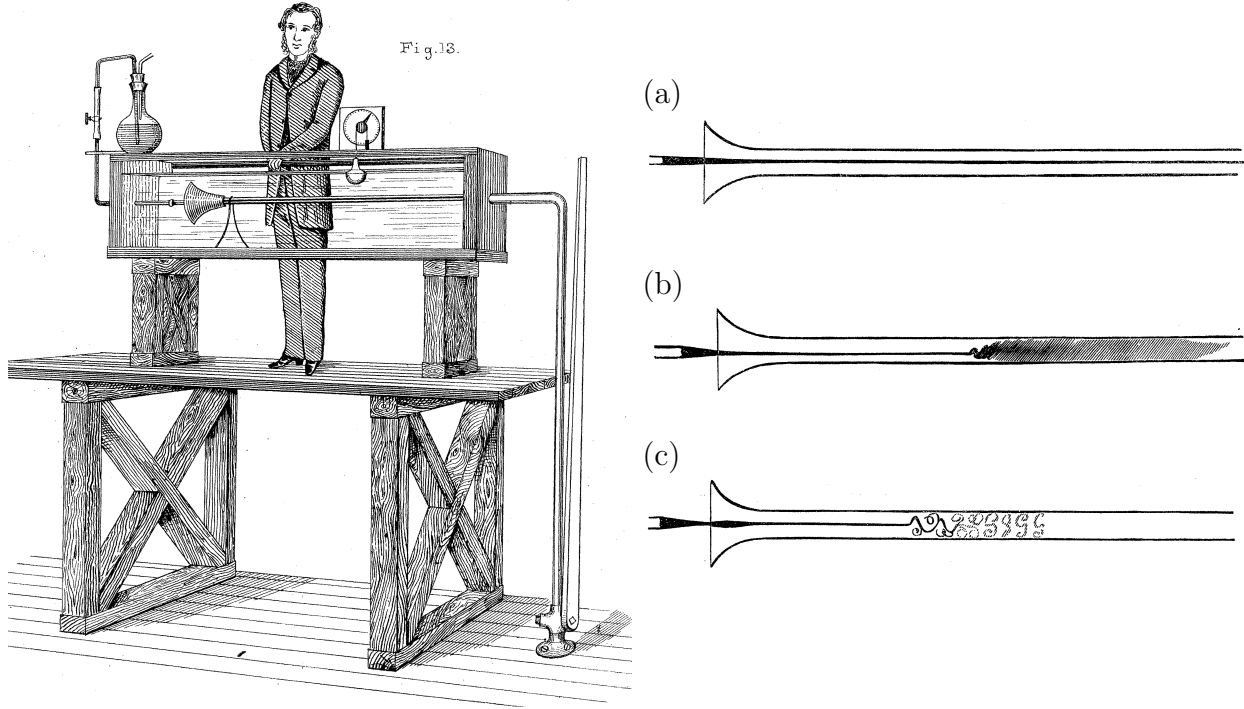
the equation can be interpreted as an energy balance stating that the *energy dissipation rate* per unit mass is

$$\varepsilon(t) = \frac{dE(t)}{dt} = -2\nu \Omega(t) \quad (4.23)$$

and thus proportional to the kinematic viscosity times the mean enstrophy per unit mass.

**Conservation of helicity.** In a similar way to the conservation of energy, the conservation of mean *helicity* per unit mass

$$H(t) = \frac{1}{2} \langle \vec{u}(\vec{r}, t) \cdot \vec{\omega}(\vec{r}, t) \rangle_{\vec{r}} \quad (4.24)$$



**Figure 4.2:** Pipe flow experiment by Reynolds (1883). (a) Picture for low velocity. (b) Picture for larger velocity. (c) Picture for larger velocity as seen by the light of electric sparks.

can be derived. It is given by

$$\frac{dH(t)}{dt} = -2\nu H_\omega(t) \quad (4.25)$$

with

$$H_\omega(t) = \frac{1}{2} \langle \vec{\omega}(\vec{r}, t) \cdot \nabla \times \omega(\vec{r}, t) \rangle_{\vec{r}}. \quad (4.26)$$

## 4.4 Reynolds number and turbulence

Reynolds (1883) made his famous pipe flow experiment by injecting coloured water into a flow of clear water through a tube, see Fig. 4.2. For sufficiently low velocity, the streak of colour extended in a straight line through the tube, Fig. 4.2 (a). With increasing velocity, the colour band mixed up at some point in the tube and filled the rest of the tube, Fig. 4.2 (b). Viewing the tube by the light of electric sparks, distinct curls were observed, Fig. 4.2 (c).

For an experimental setup such as Fig. 4.2, it is convenient to express the NSE in terms of dimensionless quantities. That is, a length and velocity scale  $L_0$  and  $U_0$ , respectively, need to be defined. The dimensionless independent variables are then  $\hat{r} = \vec{r}/L_0$  and  $\hat{t} = t U_0/L_0$

so that the dimensionless dependent variables are

$$\hat{\vec{u}}(\hat{r}, \hat{t}) = \frac{\vec{u}\left(\hat{r} L_0, \hat{t} \frac{L_0}{U_0}\right)}{U_0} \quad \text{and} \quad \hat{p}(\hat{r}, \hat{t}) = \frac{p\left(\hat{r} L_0, \hat{t} \frac{L_0}{U_0}\right)}{\rho_0 U_0^2}. \quad (4.27)$$

Defining the dimensionless Nabla operator  $\hat{\nabla} = \partial/\partial\hat{r}$ , the solenoidal condition becomes  $\hat{\nabla} \cdot \hat{\vec{u}} = 0$ . Multiplying the NSE (4.12) by  $L_0/U_0^2$  results in

$$\frac{\partial \hat{\vec{u}}}{\partial \hat{t}} + (\hat{\vec{u}} \cdot \hat{\nabla}) \hat{\vec{u}} = -\hat{\nabla} \hat{p} + \frac{1}{Re} \hat{\nabla}^2 \hat{\vec{u}}. \quad (4.28)$$

Evidently, the only parameter of the dimensionless NSE is the *Reynolds number*

$$Re = \frac{U_0 L_0}{\nu}. \quad (4.29)$$

Changing the scales from  $L_0$  and  $U_0$  to  $L'_0$  and  $U'_0$ , respectively, keeps the form of the dimensionless NSE invariant, but the Reynolds number changes from  $Re$  to  $Re' = Re U'_0/U_0 L'_0/L_0$ . This is known as *Reynolds number similarity*.

Concerning the [Reynolds \(1883\)](#) experiment, increasing the velocity in fact increased  $Re$  because the geometry of the setup and the kinematic viscosity of the water remained constant. As a conclusion, if the flow in an experiment has Reynolds number smaller than a specific critical value, which clearly depend on the setup, the flow is found to be *laminar*, i.e., there is only few molecule exchange perpendicular to the mean flow direction. If on the other hand the Reynolds number is larger than the critical value, the flow is *turbulent*. That is, there are apparently random transverse motions superposed to the laminar flow.

[Richardson \(1922\)](#) considered fully developed turbulence to be composed of eddies of different sizes and summarised the matter by the poem

Big whorls have little whorls  
That feed on their velocity,  
And little whorls have lesser whorls  
And so on to viscosity.

It should be noted that below a certain scale, known as the *Kolmogorov dissipation length*  $\eta$ , no eddies are observed. On the other hand, the length scale on which the turbulent motion is generated is called the *integral length*  $\ell_0$ .

It should be noted that there are efficient ways to reach a high Reynolds number in an experiment. Usually the geometry is fixed, but using a fluid with low kinematic viscosity, such as liquid Helium near its critical point (see [Donnelly, 1999](#)) or sulfur hexafluoride at 10 to 20 bar (see [Nobach et al., 2008](#)), leads to high Reynolds number.



## 4.5 Kolmogorov 1941 theory for infinite Reynolds number

There is no fully deductive theory for turbulence. A turbulent motion appears to be random, but yet some statistics can be found. The observable of interest is the velocity difference

$$\vec{u}_{\vec{\ell}}(\vec{r}, t) = \vec{u}(\vec{r} + \vec{\ell}, t) - \vec{u}(\vec{r}, t) \quad (4.30)$$

which is a vector. A suitable scalar quantity is the *longitudinal velocity increment*

$$u_{\parallel\vec{\ell}}(\vec{r}, t) = \vec{u}_{\vec{\ell}}(\vec{r}, t) \cdot \frac{\vec{\ell}}{|\vec{\ell}|} \quad (4.31)$$

which can be treated as a random variable (rv). [Kolmogorov \(1941a,b,c\)](#) states three basic hypotheses which make it possible to derive a hand-full of laws for the statistics of  $u_{\parallel\vec{\ell}}(\vec{r}, t)$  for infinite Reynolds number directly from the NSE. The hypotheses are assumed for  $|\vec{\ell}| \ll \ell_0$ , i.e., for small scale turbulence. This celebrated theory is named the *Kolmogorov 1941 theory*, in short K41 theory.

**Symmetry.** The first assumption is that for infinite Reynolds number and far away from any boundaries the statistics of longitudinal velocity increments (4.31) is isotropic, i.e., the statistics of  $u_{\parallel\vec{\ell}}(\vec{r}, t)$  does not depend on the direction of  $\vec{\ell}$  but only on its magnitude  $\ell$ . Furthermore, the statistics shall not depend on  $\vec{r}$ , i.e., it is translational invariant, and if the turbulence generating mechanism, i.e., outer parameters, remain constant in time, the statistics of (4.31) does not depend on  $t$ , either. Consequently, the  $p^{\text{th}}$  moment of  $u_{\parallel\vec{\ell}}(\vec{r}, t)$  is just a function of  $\ell$  and gives rise to the *longitudinal structure function*

$$S_p(\ell) = \mathbb{E} \left[ \left( u_{\parallel\vec{\ell}}(\vec{r}, t) \right)^p \right] \quad (4.32)$$

with  $p \in \mathbb{N}$ .

**Scaling.** The second assumption is that there is a scaling exponent  $h \in \mathbb{R}$  such that the statistics of  $u_{\parallel\lambda\vec{\ell}}(\vec{r}, t)$  is the same as the statistics of  $\lambda^h u_{\parallel\vec{\ell}}(\vec{r}, t)$  for any  $\lambda \in \mathbb{R}_{\geq 0}$ . Consequently,

$$S_p(\lambda\ell) = \lambda^{ph} S_p(\ell). \quad (4.33)$$

**Universality.** The third assumption is that the turbulent flow has a finite, non-vanishing, and constant mean energy dissipation rate  $\varepsilon$  per unit mass. Together with scale  $\ell$ , it uniquely and universally determine the statistical properties of the flow. By dimensional arguments it is possible to relate the longitudinal structure function to the mean energy dissipation rate per unit mass by noting that  $\dim S_p(\ell, t) = \text{length}^p / \text{time}^p$  and  $\dim \varepsilon(t) = \text{length}^2 / \text{time}^3$ . Thus,

$$S_p(\ell) \propto [\varepsilon\ell]^{p/3} \quad (4.34)$$

for any  $p \in \mathbb{N}$ . The proportionality factor for a given  $p$  must not depend on the Reynolds number because the latter is assumed to be infinite. Hence, it must be a universal constant.

Based on these assumptions [Kolmogorov \(1941a\)](#) derived an exact and non-trivial relation from the NSE, which is known as the *four-fifths law*

$$S_3(\ell) = -\frac{4}{5}\varepsilon\ell. \quad (4.35)$$

A consequence of (4.33) and (4.34) is that the scaling exponent is equal to  $h = 1/3$ . [Obukhov \(1941a,b\)](#) argued that the *two-thirds law*

$$S_2(\ell) \propto (\varepsilon\ell)^{2/3}, \quad (4.36)$$

which is a special case of (4.34), leads to the energy spectrum

$$E(k) \propto \varepsilon^{2/3}k^{-5/3}, \quad (4.37)$$

where  $k$  denotes the length of the wave vector. As the Kolmogorov's hypothesis are only assumed for small scale  $\ell$ , this energy spectrum only holds for large wave numbers  $k$ .

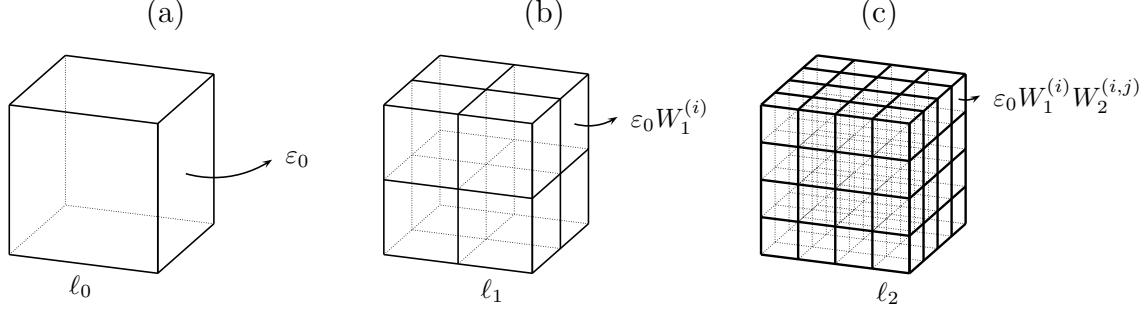
## 4.6 Kolmogorov 1962 theory for infinite Reynolds number

[Batchelor & Townsend \(1949\)](#) gave experimental evidence that the energy dissipation rate on small spatial scales is very unevenly distributed in space. Consequently, small scale turbulent eddies are intermittent, i.e., occurring at irregular intervals. This spatial inhomogeneity becomes more marked with decreasing scale. Therefore, [Kolmogorov \(1962\)](#) refined the previously made hypothesis by allowing the energy dissipation rate to fluctuate spatially. With decreasing scale this fluctuation becomes stronger. This theory is called the *Kolmogorov 1962 theory*, in short K62 theory and can be visualised as follows (see [Frisch, 1995](#)).

Consider a cube of side  $\ell_0$  (the integral length) in which the motion of the fluid is turbulent. It is characterised by the mean energy dissipation rate<sup>2</sup>  $\varepsilon_0$ , see Fig. 4.3 (a). Going into smaller scales is done systematically by splitting this cube into eight sub-cubes of side  $\ell_1 = \ell_0/2$  which are labelled by “ $(i)$ ” with  $1 \leq i \leq 8$ . The energy dissipation rate averaged over sub-cube  $(i)$  is hypothesized to be  $\varepsilon_0$  times  $W_1^{(i)}$ , where  $W_1^{(i)}$  with  $1 \leq i \leq 8$  are positive iid rv's with unit expectation, see Fig. 4.3 (b). Zooming into smaller scales means splitting each sub-cube  $(i)$  again into eight sub-sub-cubes of side  $\ell_2 = \ell_0/4$ . They are labelled by  $(i, j)$  with  $1 \leq i, j \leq 8$ . The energy dissipation rate averaged over sub-sub-cube  $(i, j)$  is hypothesized to be  $\varepsilon_0 W_1^{(i)} W_2^{(i,j)}$ , where  $W_1^{(i)}$  and  $W_2^{(i,j)}$  with  $1 \leq i, j \leq 8$  are positive iid rv's with unit expectation. This is visualised in Fig. 4.3 (c). That is, after  $n \in \mathbb{N}$

---

<sup>2</sup>per unit mass



**Figure 4.3:** Intermittency model by constructing a fluctuating mean energy dissipation rate per unit mass.

iterations, the mean energy dissipation rate averaged over the cube  $(i_1, i_2, \dots, i_n)$ , which is of side  $\ell_n = \ell_0/2^n$ , is given by  $\varepsilon_0 \prod_{k=1}^n W_k^{(i_1, \dots, i_k)}$ , i.e.,  $\varepsilon_0$  times a product of  $n$  iid rv's with unit expectation value. So, with increasing  $n$  and thus on smaller and smaller scales  $\ell_n$ , the (locally averaged) energy dissipation rate per unit mass becomes more and more fluctuating. Its  $q^{\text{th}}$  moment at scale  $\ell = 2^{-n}\ell_0$  does not depend on the exact location of  $(i_1, \dots, i_n)$  and is therefore

$$\mathbb{E}[\varepsilon(\ell)^q] = \varepsilon_0^q \left(\frac{\ell}{\ell_0}\right)^{\tau_q} \quad (4.38)$$

with

$$\tau_q = -\log_2 \mathbb{E}[W^q]. \quad (4.39)$$

Here,  $W$  denotes a rv which is of the same distribution as the  $W_k$ 's. Due to  $\mathbb{E}[W] = 1$ ,  $\tau_1$  is restricted to be zero.

As for the longitudinal structure functions  $S_p(\ell)$ , Eq. (4.34) needs to be generalised because the right hand side is no longer a constant but rather a rv. Replacing the right hand side by its expectation value leads to

$$S_p(\ell) \propto \mathbb{E}\left[(\varepsilon(\ell)\ell)^{p/3}\right] = (\varepsilon_0\ell)^{p/3} \left(\frac{\ell}{\ell_0}\right)^{\tau_{p/3}} \quad (4.40)$$

which is proportional to  $\ell^{p/3+\tau_{p/3}}$ . Thus, the exponent  $\xi_p$  defined by  $S_p(\ell) \propto \ell^{\xi_p}$  reads

$$\xi_p = \frac{p}{3} + \tau_{p/3} \quad (4.41)$$

and is restricted to  $\xi_3 = 1$  by the four-fifths law. The relation (4.41) is experimentally verified by e.g. [Meneveau & Sreenivasan \(1991\)](#). So far, nothing has been said about the distribution of the  $W_k$ 's making this relation quite general.

[Obukhov \(1962\)](#) and [Kolmogorov \(1962\)](#) argued that the logarithm of the (small scale) energy dissipation rate corresponds to a sum of (many) iid rv's. The central limit theorem implies that the logarithm of the energy dissipation rate is normally distributed so that the energy dissipation rate is log-normally distributed. Additionally, the distribution of

K41	exponent: $\xi_p = p/3$ intermittency correction: $2 - \xi_6 = 0$ distribution of $W$ : $W = 1$ with unit probability
log-normal model	exponent: $\xi_p = p/3 - \mu p(p-3)/18$ intermittency correction: $2 - \xi_6 = \mu > 0$ distribution of $W$ : log-normally distributed $W$
$\beta$ -model	exponent: $\xi_p = p/3 + (p/3 - 1) \log_2 \beta$ intermittency correction: $2 - \xi_6 = -\log_2 \beta > 0$ distribution of $W$ : $W = 1/\beta$ with probability $\beta$ $W = 0$ with probability $1 - \beta$
log-Poisson model	exponent: $\xi_p = p/9 + 2 - 2(2/3)^{p/3}$ intermittency correction: $2 - \xi_6 = 2/9 \approx 0.22$ distribution of $W$ : log-Poisson distributed $W$

**Table 4.3:** Summary of cascade models. The log-normal model corresponds to [Kolmogorov \(1962\)](#). The  $\beta$  model is stated in the way of [Novikov & Stewart \(1964\)](#). The log-Poisson is stated in the form of [She & Waymire \(1995\)](#).

$W$  is also assumed to be of a log-normal distribution  $L_{W_0, \lambda^2}(W)$  with position parameter  $W_0 = \exp(-\frac{\mu}{2} \ln 2)$  and shape parameter  $\lambda^2 = \mu \ln 2$ , where  $\mu > 0$  is a dimensionless constant. This leads to

$$\tau_q = -\frac{\mu}{2} q(q-1) \quad (4.42)$$

so that

$$\xi_p = \frac{p}{3} - \frac{\mu}{18} p(p-3). \quad (4.43)$$

This is known as the *log-normal model*. The parameter  $\mu$  reflects the *intermittency correction* because it corresponds to the difference between the K41 sixth order exponent, which is 2, and the intermittent sixth order exponent, which is  $\xi_6 = 2 - \mu$ . It should be noted that the K41 theory is a special case of K62 for  $\mu = 0$ . The experimentally measured value for the intermittency correction is  $0.2 \leq \mu \leq 0.25$  (see e.g. [Anselmet \*et al.\*, 1984](#); [Vincent & Meneguzzi, 1991](#)). Regarding the energy dissipation rate at scale  $\ell$ , Eq. (4.38) reveals that

$$\mathbb{E}[\varepsilon(\ell)^q] = \varepsilon_0^q \left(\frac{\ell}{\ell_0}\right)^{-\frac{\mu}{2} q(q-1)} = \left[ \varepsilon_0 \left(\frac{\ell}{\ell_0}\right)^{\mu/2} \right]^q \exp \left[ \frac{q^2}{2} \mu \ln \frac{\ell_0}{\ell} \right]. \quad (4.44)$$

It is therefore log-normally distributed with position parameter  $\varepsilon_0(\ell/\ell_0)^{\mu/2}$  and position parameter  $\mu \ln(\ell_0/\ell)$ .

Albeit [Novikov \(1971\)](#) and [Mandelbrot \(1972\)](#) clarified inherent problems of the log-normal model, it is widely used in turbulence (see e.g. [Castaing \*et al.\*, 1990](#); [Boettcher,](#)

2005). In fact, experiments show that Eq. (4.43) yields good results for  $p \leq 8$  (see e.g. Frisch, 1995, Fig. 8.8). Renner *et al.* (2001) points out that measurements for  $p > 8$  are very difficult and contain large errors. This explains the coexistence of further models, such as

- the  $\beta$ -model introduced by Novikov & Stewart (1964) and generalised by Novikov (1969, 1971, 1990) or
- the log-Poisson model introduced by She & L ev eque (1994), whose work is originally based on energy filaments, and formulated in a statistical framework by She & Waymire (1995).

These models are summarised in Tab. 4.3.

## 4.7 Increment distribution

This section derives a formula for the distribution of the longitudinal velocity increment (4.31) with spatial scale  $\ell$  by treating it as a rv  $U_\ell$ . It should be mentioned that in this section capital letters are used to label a rv while small letters are used to denote its value.

Castaing *et al.* (1990) argues that the simplest assumption for the distribution of  $U_\ell$  within a cell with mean energy dissipation rate  $\varepsilon(\ell)$  is a symmetric Gaussian. Its standard deviation  $\Sigma_\ell$  is proportional to  $(\varepsilon(\ell)\ell)^{1/3}$  and therefore also a rv. Its distribution is denoted by  $f_\ell(\sigma_\ell)$  so that the marginal distribution for the longitudinal velocity increment corresponds to the superposition

$$p_\ell(u_\ell) = \int_0^\infty d\sigma_\ell f_\ell(\sigma_\ell) N_{0,\sigma_\ell^2}(u_\ell). \quad (4.45)$$

This distribution is called the *symmetric Castaing distribution*. This model generates a longitudinal structure function given by

$$\begin{aligned} S_p(\ell) &= \mathbb{E}[U_\ell^p] = \int_0^\infty d\sigma_\ell f_\ell(\sigma_\ell) \int_{-\infty}^\infty du_\ell u_\ell^p N_{0,\sigma_\ell^2}(u_\ell) \\ &= \begin{cases} 0 & \text{for odd } p \text{ and} \\ (p-1)!! \mathbb{E}[\Sigma_\ell^p] & \text{for even } p. \end{cases} \end{aligned} \quad (4.46)$$

Due to  $\Sigma_\ell \propto (\varepsilon(\ell)\ell)^{1/3}$ , the structure function of even order is  $S_p(\ell) \propto \mathbb{E}[(\varepsilon(\ell)\ell)^{p/3}]$  and in agreement with (4.40). However, the even order structure function vanishes. Hence, the assumption of a (symmetric) normally distributed longitudinal velocity increment within a cell with given mean energy dissipation rate needs to be refined. If this distribution is denoted by  $p(u_\ell | \sigma_\ell)$  with standard deviation  $\sigma_\ell$  being the value of the rv  $\Sigma_\ell \propto (\varepsilon(\ell)\ell)^{1/3}$ , Eq. (4.45) becomes

$$p_\ell(u_\ell) = \int_0^\infty d\sigma_\ell f_\ell(\sigma_\ell) p(u_\ell | \sigma_\ell). \quad (4.47)$$

If  $p(u_\ell | \sigma_\ell)$  can be written as

$$p(u_\ell | \sigma_\ell) = \frac{1}{\sigma_\ell} q\left(\frac{u_\ell}{\sigma_\ell}\right) \quad (4.48)$$

with  $\ell$ -independent distribution  $q(z)$ , the longitudinal structure function reads

$$S_p(\ell) = \mathbb{E}[U_\ell^p] = \int_0^\infty d\sigma_\ell f_\ell(\sigma_\ell) \int_{-\infty}^\infty \frac{du_\ell}{\sigma_\ell} u_\ell^p q\left(\frac{u_\ell}{\sigma_\ell}\right). \quad (4.49)$$

Substituting  $u_\ell/\sigma_\ell$  by  $z$ , this equation turns into

$$S_p(\ell) = \mathbb{E}[\Sigma_\ell^p] \mathbb{E}[Z^p] \quad (4.50)$$

with  $\Sigma_\ell$  and  $Z$  being  $f_\ell$ - and  $q$ -distributed rv's, respectively. Eq. (4.45) is recovered for  $q(z) = N_{0,1}(z)$ . [Castaing \*et al.\* \(1990\)](#) shows that the distribution

$$q(z) \propto \exp\left[-\frac{z^2}{2} \left(1 + a_s \frac{z}{\sqrt{1+z^2}}\right)\right] \quad (4.51)$$

with universal factor  $a_s > 0$  is a suitable candidate. Note that the so-defined  $q$ -distribution has negative skewness. The corresponding left skewed increment distribution (4.47) is thus called the *asymmetric Castaing distribution*. Experimental measurements of laboratory turbulent data, such as done by [Renner \*et al.\* \(2001\)](#), verify the (negative) skewness of the increment distribution.

# Chapter 5

## Superstatistics

There are many complex systems whose dynamics can be understood by a superposition of several dynamics on different time scales, e.g., a Brownian motion in an environment with (slowly) changing temperature. The formalism of superposing two statistics was introduced by [Beck & Cohen \(2003\)](#); [Beck \(2004c\)](#) and called *superstatistics*. It is a general idea which has been applied to a variety of complex systems, such as fluid/air turbulence by [Reynolds \(2003\)](#); [Beck \(2004b\)](#); [Rizzo & Rapisardia \(2005\)](#); [Jung & Swinney \(2005\)](#); [Beck et al. \(2005a\)](#); [Beck \(2007\)](#), pattern forming systems by [Daniels et al. \(2004\)](#), cosmic rays by [Beck \(2004a\)](#), solar flares by [Baiesi et al. \(2006\)](#), mathematical finance by [Ausloos & Ivanova \(2003\)](#), random matrix theory by [Abul-Magd \(2006\)](#), networks by [Abe & Thurner \(2005\)](#), hydro-climatic fluctuations by [Porporato et al. \(2006\)](#), delay statistics in traffic models by [Briggs & Beck \(2007\)](#), and astrophysics by [Chavanis \(2006\)](#). The superstatistical concept is mainly used to interpret observables with non-Gaussian distribution.

### 5.1 The concept

Consider a complex dynamical system in a stationary non-equilibrium state which is driven by external forces. Generally, a complex system may be inhomogeneous, i.e., the phase space is partitioned into small spatio-temporal cells, each of which is described by a different relevant system parameter  $\beta > 0$ . The system in each cell is described by ordinary Maxwell-Boltzmann statistics while  $\beta$  varies from cell to cell according to a probability density  $f(\beta)$ , or alternatively  $\beta$  is uniform throughout the whole phase space but changes with time. The stationary distribution arises as a superposition of *Boltzmann factors*  $e^{-\beta E}$ , where  $E$  denotes the relevant thermodynamical variable of the system. That is,

$$p(E) = \int_0^\infty d\beta f(\beta) \frac{e^{-\beta E}}{Z(\beta)} \quad (5.1)$$

with  $Z(\beta) = \int dE e^{-\beta E}$ .

**Example 5.1:** Dynamical realisation of superstatistics. Consider the Langevin equation (see [Sobczyk, 1991](#), for the mathematical definition) for a variable  $u$  with friction constant  $\gamma > 0$  and noise strength  $\sigma > 0$  given by

$$\dot{u} = \gamma F(u) + \sigma \xi(t), \quad (5.2)$$

where  $\xi(t)$  denotes standard Gaussian white noise and  $F(u) = -V'(u)$  denotes the drift force due to potential  $V(u)$ . For constant friction and noise strength, the stationary probability density of  $u$  is  $p(u|\beta) = e^{-\beta V(u)}/Z(\beta)$  with  $\beta = \gamma/\sigma^2$  and  $Z(\beta) = \int du e^{-\beta V(u)}$  (see [Beck, 2004c](#)). If however  $\beta$ , which is  $f$ -distributed, varies slowly with time such that the system temporarily reaches equilibrium, the marginal distribution of  $u$  reads  $p(u) = \int_0^\infty d\beta f(\beta) p(u|\beta)$ . ■

## 5.2 Typical weight functions

The distribution  $f(\beta)$  is determined by the spatio-temporal fluctuation of  $\beta$ . Three typical situations and the corresponding tails of  $p(E)$  are discussed in the next examples. For a more general discussion the reader is referred to [Touchette & Beck \(2005\)](#).

**Example 5.2:**  $\chi^2$ -distributed  $\beta$ . If there are many (nearly) independent microscopic rv's contributing to  $\beta$  in an additive way, their sum can be rescaled such that it approaches a standard normally distributed rv  $X \sim N_{0,1}$  due to the central limit theorem. In total, there can be  $n$  such subsystems and hence  $n$  Gaussian rv's denoted by  $X_1, \dots, X_n$ . If  $\beta = \sum_{i=1}^n X_n^2 > 0$ , it is  $\chi^2$ -distributed with  $n$  degrees of freedom so that

$$f(\beta) = \chi_n^2(\beta) = \frac{1}{2\Gamma(n/2)} \left(\frac{\beta}{2}\right)^{n/2-1} e^{-\beta/2}. \quad (5.3)$$

Hence, Eq. (5.1) becomes

$$p(E) = \frac{1}{\Gamma(n/2)2^{n/2}} \int_0^\infty \beta^{n/2+1} e^{-\beta(E+1/2)}, \quad (5.4)$$

which can be solved by using the Laplace transformation technique. That is,

$$p(E) = \frac{n}{(1+2E)^{n/2+1}} \quad (5.5)$$

having power law tails for large energy  $E$ . Note the specific distribution of the microscopic rv's is irrelevant making this superstatistics universal. ■

**Example 5.3:**  $\chi^{-2}$ -distributed  $\beta$ . The same consideration as in the previous example can be applied to the “temperature”  $\beta^{-1}$  so that

$$f(\beta) = \chi_n^{-2}(\beta) = \frac{2}{\Gamma(n/2)} (2\beta)^{-n/2-1} e^{-1/(2\beta)}. \quad (5.6)$$



Hence, Eq. (5.1) becomes

$$p(E) = \frac{(2E)^{n/4}}{2^{n/2-1}\Gamma(n/2)} K_{n/2}(\sqrt{2E}) \quad (5.7)$$

which is derived in [Sattin & Salasnich \(2002\)](#). Here,  $K_{n/2}(\cdot)$  denotes the modified Bessel function of order  $n/2$  (see [Bronstein \*et al.\*, 1999](#), Sec. 9.1.2.6.2). Consequently, this superstatistics produces exponential decays in  $\sqrt{E}$ , i.e.,  $p(E) \propto e^{-c\sqrt{E}}$  for large energy  $E$ . Again, the specific distribution of the microscopic rv's is irrelevant making this superstatistics universal. ■

**Example 5.4:** Log-normally distributed  $\beta$ . The previous two examples described  $\beta$  and  $\beta^{-1}$ , respectively, as a sum of many contributions. In this example it is generated by a multiplicative process. That is, if there are many positive microscopic rv's, the sum of their logarithms approaches a normally distributed rv  $\ln X$  so that  $X$  is log-normally distributed. If there are  $n$  multiplicative contributions to  $\beta$ , i.e.,  $\beta = \prod_{i=1}^n X_i$  with log-normally distributed  $X_1, \dots, X_n$ , the distribution of  $\beta$  is the log-normal distribution

$$f(\beta) = L_{\beta_0, \lambda^2}(\beta) = \frac{1}{\beta\sqrt{2\pi\lambda^2}} \exp\left[-\frac{1}{2\lambda^2} \ln^2 \frac{\beta}{\beta_0}\right] \quad (5.8)$$

with position and shape parameter  $\beta_0$  and  $\lambda^2$ , respectively. ■

**Part II**

**Application and Discussion**

# Chapter 6

## Wind Speed Statistics

This chapter is intended to discuss the statistical analysis of atmospheric boundary layer (ABL) wind velocity and more specifically its horizontal component. Ideal turbulence as described in Chap. 4 is used as a reference for ABL turbulence.

Time series recordings of the horizontal component are obtained from measurements made with cup anemometers at the [Lammefjord \(1987\)](#) site. Located at  $55^\circ 47' 41''$  N and  $11^\circ 26' 52''$  E, the Lammefjord is a flat terrain at the base of the Odsherred peninsula in Denmark. The wind speed is recorded with frequency  $\nu = 8$  Hz over twelve successive days. This produces about  $6.9 \times 10^5$  data points per day. For further references the time series is denoted by

$$u_n = \sqrt{u_x^2(t = n/\nu) + u_y^2(t = n/\nu)} \quad (6.1)$$

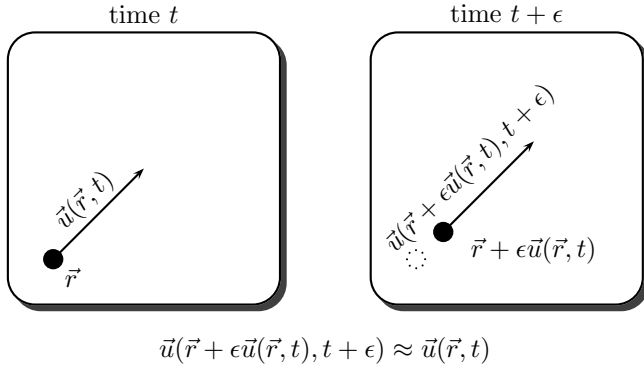
for  $n \in \mathbb{N}_0$ . The anemometers are installed at 10 m, 20 m, and 30 m altitude. The data gathered at 10 m above ground are used, which clearly exemplify ABL turbulence.

### 6.1 Taylor hypothesis

The problem with wind velocity measurements in general is that it only reconstructs the velocity field  $\vec{u}(\vec{r}, t)$ , or any component, at fixed  $\vec{r}$  as a function of time  $t$ . Contrarily, the turbulence theory introduced in Chap. 4 is build upon  $\vec{u}(\vec{r}, t)$  as being a function of location and time. It is impractical though to measure the wind velocity at all locations at the same time. Nonetheless, the wind velocity  $\vec{u}(\vec{r}, t)$  is connected to air masses traveling with velocity  $\vec{u}(\vec{r}, t)$  so that after an infinitesimal time interval  $\epsilon$ , the same, or nearly the same, velocity is expected to be measured at the location  $\vec{r} + \epsilon \vec{u}(\vec{r}, t)$ , i.e.,  $\vec{u}(\vec{r} + \epsilon \vec{u}(\vec{r}, t), t + \epsilon) \approx \vec{u}(\vec{r}, t)$ , see Fig. 6.1. In the limit  $\epsilon \rightarrow 0$  this becomes  $\frac{d}{dt} \vec{u}(\vec{r}, t) = 0$ . Of course, the concept can be generalised to any observable  $\xi$  which is “tied” to the moving air parcels, such as velocity components, temperature, humidity, etc, (see e.g. [Stull, 1988](#)). That is,

$$\xi(\vec{r} + \epsilon \vec{u}(\vec{r}, t), t + \epsilon) \approx \xi(\vec{r}, t), \quad (6.2)$$

or equivalently  $\frac{d}{dt} \xi(\vec{r}, t) = 0$ .



**Figure 6.1:** The [Taylor \(1938\)](#) hypothesis. The left panel shows a parcel which is at time  $t$  located at  $\vec{r}$  and moving with velocity  $\vec{u}(\vec{r}, t)$ . The right panel depicts that the same parcel after infinitesimal time  $\epsilon$  is located at  $\vec{r} + \epsilon\vec{u}(\vec{r}, t)$  and has nearly the same velocity.

The idea was introduced by [Taylor \(1938\)](#) and generalised to macroscopic time scales. It is thus called *Taylor hypothesis*. It assumes that the fluctuation of the velocity is small compared to the mean velocity  $\vec{U}$  measured at fixed location  $\vec{r}_0$  over time period  $T$ . Additionally, the fluctuation of the mean velocity direction is assumed to be negligible making the hypothesis be very useful for a (mainly) unidirectional turbulent flow. In that case, Eq. (6.2) can be extrapolated to relate  $\xi(\vec{r}, t_0)$  with fixed time  $t_0$  and variable  $\vec{r}$  to  $\xi(\vec{r}_0, t)$  and  $\vec{u}(\vec{r}_0, t)$  measured at fixed position  $\vec{r}_0$  as a function of time:  $\xi(\vec{r}, t_0) \approx \xi(\vec{r}_0, t)$  with  $\vec{r} = \vec{r}_0 - \vec{U}\Delta t$ ,  $t = t_0 + \Delta t$ , and  $0 \leq \Delta t \leq T$ . From a statistical point of view, the velocity fluctuation in units of the mean velocity corresponds to the coefficient of variation and is in this context called *turbulence intensity* (TI) over period  $T$ , (see e.g. [Burton et al., 2004](#)). It states the percentage of the mean flow which is represented by the fluctuation. For the Taylor hypothesis to be satisfactory, the TI should be very small. Though, [Willis & Deardorff \(1976\)](#) suggested that a TI up to 50% suffices. In any case, [Lumley \(1965\)](#); [Pinton & Labbé \(1994\)](#) developed corrections in case the TI is not small enough.

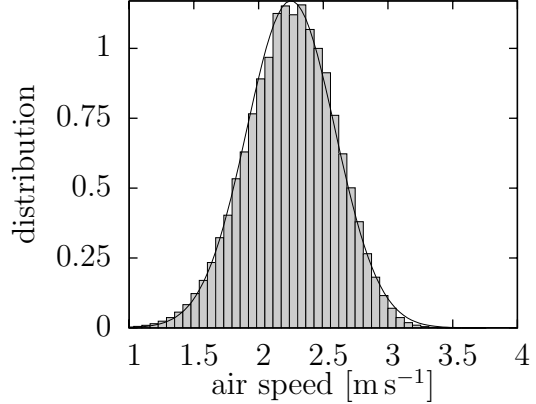
## 6.2 Non-stationarity and fluctuation statistics

It is customary in turbulence research to decompose the wind velocity, or its horizontal component in the case considered here, during a period  $T$  as

$$\begin{array}{ccccccc} u(t) & = & U & + & u'(t) & & \\ \text{(signal)} & & \text{(mean)} & & \text{(fluctuation)} & & \end{array} \quad (6.3)$$

where  $U$  is the average value of  $u(t)$  over this period. As explained in the previous section, the TI is defined as the root mean square of  $u'(t)$  in units of  $U$ . It describes the strength of the instantaneous turbulence so that it depends on the thermal behaviour of the atmosphere (see [Burton et al., 2004](#)). For instance if the air near the ground warms up on a sunny day, it may become buoyant enough to rise up through the atmosphere. This causes a pattern of convection cells which are experienced as large-scale turbulent eddies. It is therefore not

**Figure 6.2:** The distribution of the air speed measured in the free jet experiment by Renner *et al.* (2001). The solid line corresponds to a Gaussian curve with mean  $2.25 \text{ m s}^{-1}$  and standard deviation  $0.341 \text{ m s}^{-1}$ .



surprising if the TI is different at two successive days. Additionally, the value of the TI depends on the considered period length  $T$ . It is however difficult to fix a suitable time scale because Boettcher (2005) states that the atmospheric winds exhibit variations on any time scale, ranging from seconds (and less) up to centuries. It is often distinguished between large scale variations such as diurnal, weekly, and seasonal and small scale variations which mirror ordinary turbulence as described in Chap. 4. The existence of a mesoscale gap is suggested by van der Hoven (1957), but it is also strongly criticised by Lovejoy *et al.* (2001) and its references. Consequently, no matter how long  $T$  is, the mean  $U$  fluctuates.

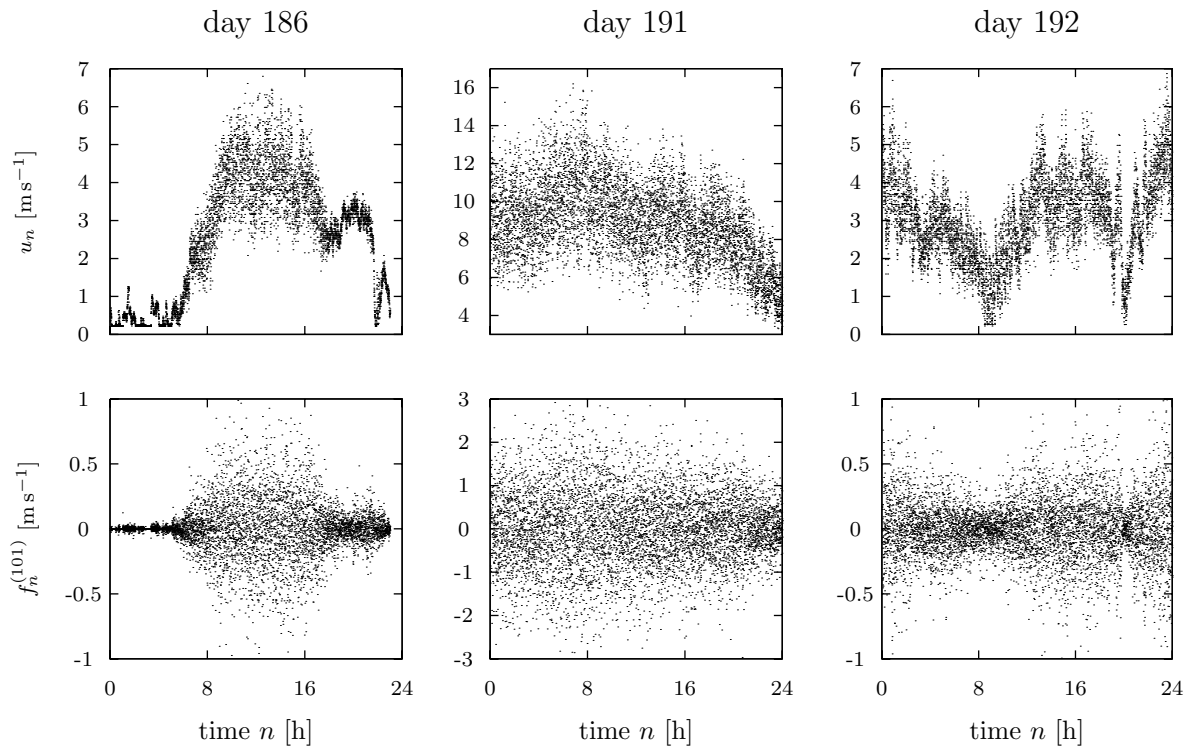
For fixed  $T$ , Burton *et al.* (2004) states that turbulent wind speed variations are roughly normally distributed about the mean wind speed  $U$  with standard deviation  $U \times \text{TI}$ . Laboratory experiments such as the air into air free jet experiment by Renner *et al.* (2001), which produce a turbulent flow with fixed mean air speed  $U$ , verify the picture of normally distributed  $u'(t)$ , see Fig. 6.2. Approximating the distribution of the wind speed fluctuation by a Gaussian distribution is only justified if the TI is small. Otherwise, the probability for  $u(t) < 0$  forecast by the Gaussian model with zero mean and standard deviation  $U \times \text{TI}$  might be too large. It is given by

$$\mathbb{P}[u(t) < 0] = \mathbb{P}[u'(t) < -U] = \frac{1}{2} \operatorname{erfc} \left( \frac{1}{\text{TI}\sqrt{2}} \right), \quad (6.4)$$

see App. A.2. For small TI the argument in the complementary error function becomes large. Using its asymptotic expansion, (see Abramowitz & Stegun, 1964, Eq. (7.1.23)), the equation above becomes

$$\begin{aligned} \mathbb{P}[u(t) < 0] &= \frac{\text{TI}}{\sqrt{2\pi}} \exp \left[ -\frac{1}{2\text{TI}^2} \right] \sum_{n=0}^{\infty} (-1)^n \frac{(2n)!}{n!} \left( \frac{\text{TI}^2}{2} \right)^n \\ &\approx \frac{\text{TI}(1 - \text{TI}^2)}{\sqrt{2\pi}} \exp \left[ -\frac{1}{2\text{TI}^2} \right], \end{aligned} \quad (6.5)$$

where the series is aborted after two terms. In order to get a feeling for the numbers, if the TI is below 20%  $\mathbb{P}[u(t) < 0]$  is smaller than  $3 \times 10^{-7}$ .



**Figure 6.3:** The top row shows the wind speed data for three different days of the [Lammefjord \(1987\)](#) measurement. The second row displays the fluctuation (6.6) for  $m = 101$  which corresponds to a time window of 12.5 s.

Wind speed measurements at the [Lammefjord \(1987\)](#) site are very non-stationary. The top row of Fig. 6.3 shows three 24 h recordings illustrating three time series:

- The time series recorded at day 186 shows low wind activity during the night hours whereas strong winds occur during daytime.
- The time series recorded at day 191 looks comparably stationary, whereas
- the time series recorded at day 192 has many up and downs.

Thus, a derivation of direct statistics, e.g., the estimation of probability density, mean, variance, etc. over 24 h, becomes difficult to interpret. Of course, it is not uninteresting to estimate the 24 h mean and variance, but they are different from the 24 one-hour means and variances, respectively, estimated every successive hour. That is, the statistics depends on the time and length of the considered period.

In spirit of (6.3), the signal  $u_n$  is transformed into a fluctuation series

$$f_n^{(m)} = u_n - \bar{u}_n^{(m)}, \quad (6.6)$$

where the value of the moving average  $\bar{u}_n^{(m)}$  is taken over  $m = 2\tilde{m} + 1$  ( $\tilde{m} \in \mathbb{N}_0$ ) time steps, i.e.,

$$\bar{u}_n^{(m)} = \frac{1}{m} \sum_{k=-\tilde{m}}^{\tilde{m}} u_{n+k}. \quad (6.7)$$

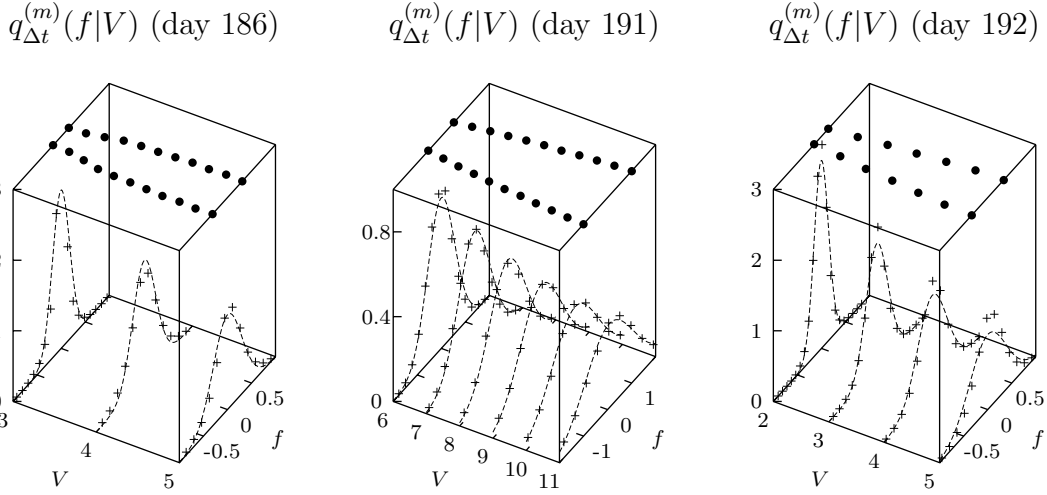
If  $m$  is chosen to be small enough such that the series  $\bar{u}_n^{(m)}$  still contains the main characteristics of the series  $u_n$ , the fluctuation  $f_n^{(m)}$  is a measure for the volatility of the wind speed at time  $n$ . If on the other hand  $m$  is chosen too large, large temporal trends, e.g., the above mentioned up and downs in  $u_n$ , are not reflected by  $\bar{u}_n^{(m)}$  so that the fluctuation at time  $n$  treated as rv has non-vanishing expectation value. The second row of Fig. 6.3 displays  $f_n^{(m)}$  for  $m$  exemplarily chosen to be 101, which for a 8 Hz measurement covers the time window of 12.5 s. It can be seen that the fluctuation series has zero expectation and that its volatility becomes larger as the wind speed  $u_n$ , and hence the mean wind speed  $\bar{u}_n^{(m)}$ , increases. It is therefore useful to investigate the statistics of the fluctuation conditioned on the value  $V$  of the moving average.

It is assumed that the time series  $f_n^{(m)}$  for  $n$  such that  $|\bar{u}_n^{(m)} - V| \leq \Delta V/2$  for comparably small  $\Delta V$  is (first order) stationary over a diurnal cycle. Then, it is possible to verify that  $f_n^{(m)} \in \{f_n^{(m)} : n \in \mathcal{N}_{\Delta V}^{(m)}(V)\}$  with

$$\mathcal{N}_{\Delta V}^{(m)}(V) = \{n : \bar{u}_n^{(m)} \in [V - \Delta V/2, V + \Delta V/2]\} \quad (6.8)$$

is roughly normally distributed and that there is a proportionality between the standard deviation of the set  $\{f_n^{(m)} : n \in \mathcal{N}_{\Delta V}^{(m)}(V)\}$  and  $V$  if the TI, which represents the proportionality factor, remains approximately constant over 24 h. As described in Sec. 2.8, the conditioned distribution  $q_{\Delta V}^{(m)}(f|V)$  for a variety of  $V$  can be estimated by

$$q_{\Delta V}^{(m)}(f|V) = \langle \delta(f - f_n^{(m)}) \rangle_{n \in \mathcal{N}_{\Delta V}^{(m)}(V)} \quad (6.9)$$



**Figure 6.4:** Estimated distribution of the fluctuation conditioned on mean wind speed  $V$  over  $m = 101$  time steps for the [Lammefjord \(1987\)](#) wind data (day 186, day 191, and day 192 with  $\Delta V = 0.2 \text{ ms}^{-1}$ ,  $\Delta V = 0.5 \text{ ms}^{-1}$ , and  $\Delta V = 0.5 \text{ ms}^{-1}$ , respectively). Note that only for some selected values of  $V$  the conditioned distributions are shown. The crosses symbolise the estimation (6.9) whereas the dashed lines represent Gaussian distributions. The standard deviation of the fluctuation is again plotted on top of the boxes with black dots.

and plotted as histograms. These are drawn in the boxes of Fig. 6.4. The dashed lines correspond to symmetric normal distributions with the same standard deviation as estimated from the set  $\{f_n^{(m)} : n \in \mathcal{N}_{\Delta V}^{(m)}(V)\}$ . These standard deviations are again plotted on top of the boxes, indicating that there might be a proportionality between the fluctuation standard deviation and  $V$ . For further reference, the proportionality factor is denoted by  $\alpha(m)$ .

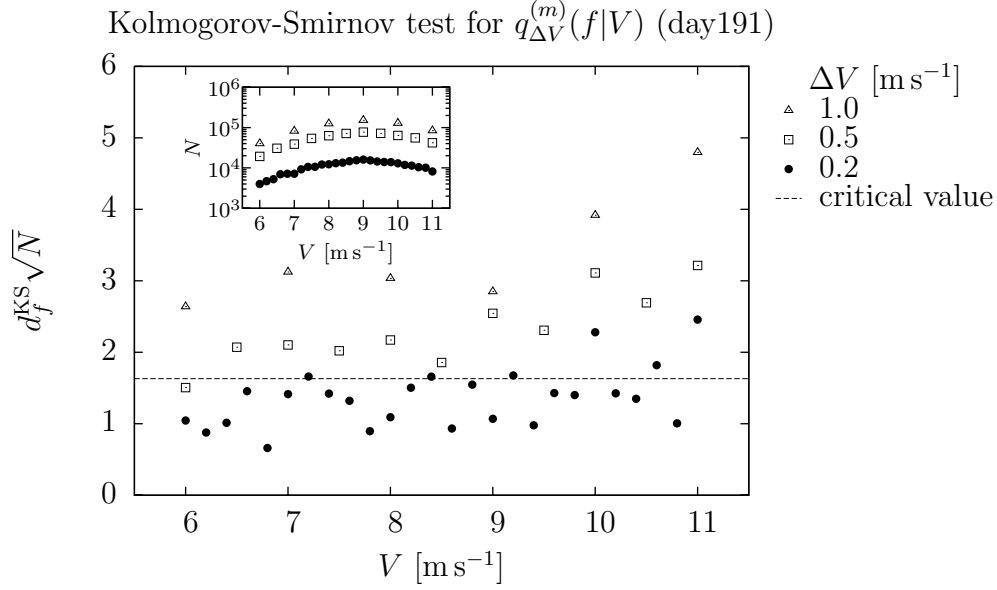
Regarding the Gaussian shape of the distributions, it is difficult to verify because the shape depends on the binning of  $V$ . Due to  $\Delta V > 0$ , it might correspond to a superposition of normal distributions with standard deviations ranging from  $\alpha(m) \times (V - \Delta V/2)$  to  $\alpha(m) \times (V + \Delta V/2)$ . A Kolmogorov-Smirnov test which is performed exemplarily on the data gathered at day 191 according to Sec. 2.9 is shown in Fig. 6.5. It yields that for sufficiently small  $\Delta V$  the hypothesis “ $q_{\Delta V}^{(m)}(f|V)$  is a normal distribution” cannot be rejected with significance level 1%.

The exact shape of  $q_{\Delta V}^{(m)}(f|V)$  is not crucial, the important question is whether the estimated standard deviation  $\sigma_{f; \Delta V}^{(m)}(V)$  of the set  $\{f_n^{(m)} : n \in \mathcal{N}_{\Delta V}^{(m)}(V)\}$  is proportional to  $V$  in the limit  $\Delta V \rightarrow 0$ , i.e.,

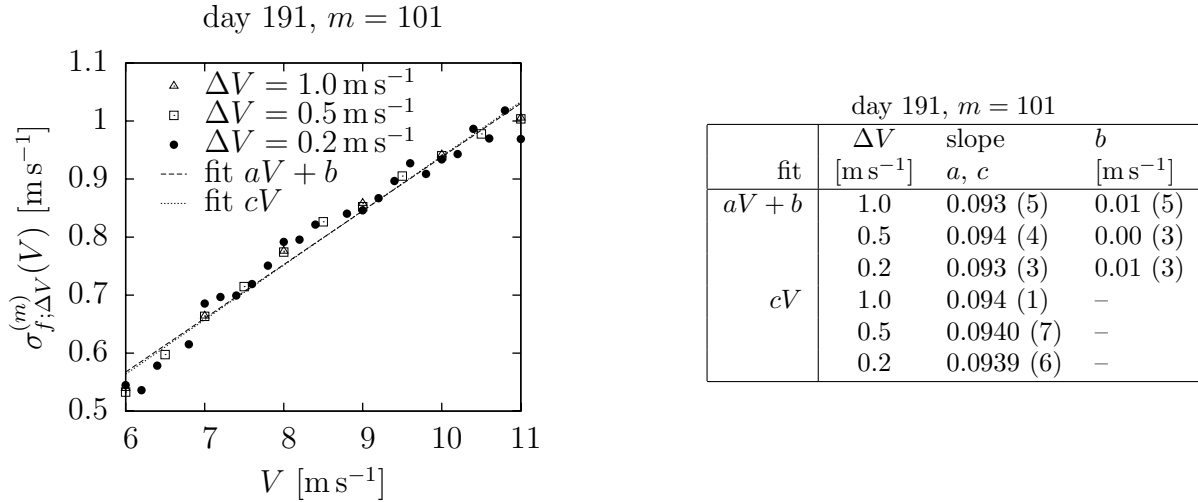
$$\lim_{\Delta V \rightarrow 0} \sigma_{f; \Delta V}^{(m)}(V) \propto V. \quad (6.10)$$

Fig. 6.6 displays the standard deviation as a function of  $V$  for a variety of  $\Delta V$ . It can be seen that there is a proportionality between  $\sigma_{f; \Delta V}^{(m)}(V)$  and  $V$ . This is confirmed by fitting the two functions  $f(V) = aV + b$  and  $g(V) = cV$  with  $a, b, c \in \mathbb{R}$  to the data points using





**Figure 6.5:** Kolmogorov-Smirnov tests for  $q_{\Delta V}^{(m)}(f|V)$  with  $m = 101$  for the [Lammeffjord \(1987\)](#) data gathered at day 191, see Sec. 2.9. Three binnings of  $V$  are on display:  $\Delta V = 0.2, 0.5, 1.0 \text{ m s}^{-1}$ . The dashed line represents the critical value  $D^{\text{KS}}(1\%, N > 40) = 1.63$  for the 1% significant level. (inner panel) Number of data points  $N = |\mathcal{N}_{\Delta V}^{(m)}(V)|$ .



**Figure 6.6:** Estimated standard deviation of the fluctuation gathered at day 191 at the [Lammeffjord \(1987\)](#) site conditioned on mean wind speed  $V$  over  $m = 101$  time steps. The figure shows the raw data for a variety of  $\Delta V$  together with two fits ( $aV + b$  and  $cV$ ) obtained by linear regression. The results are summarised in the right table. The number in brackets denotes the fit error.

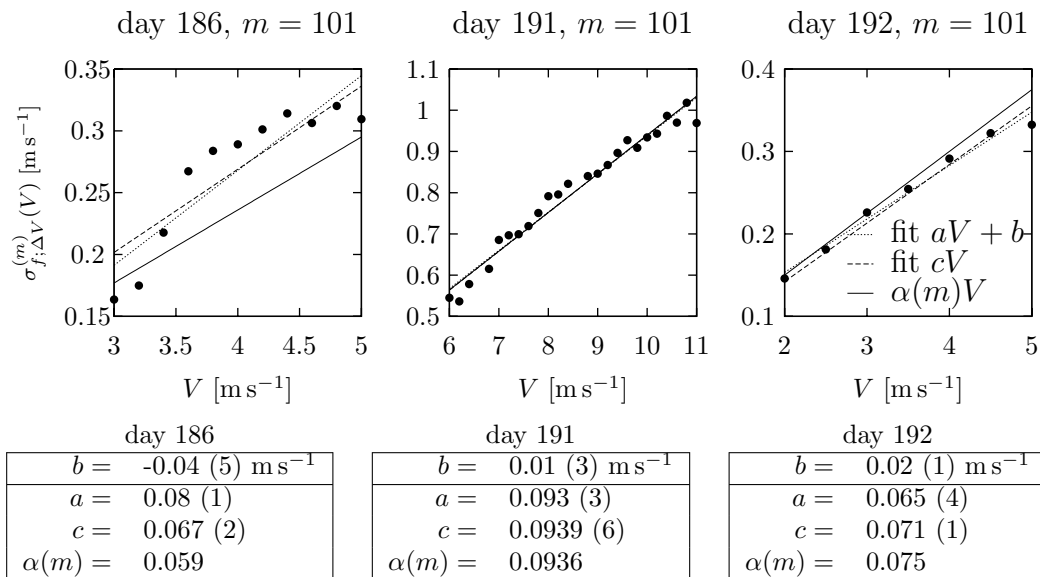
linear regression, (see e.g. [Witte & Witte, 2001](#)). The result is summarised in the right table of Fig. 6.6. It is evident that the data are in good agreement with (6.10). In order to obtain a value for the proportionality factor, the method of linear regression is unsuitable because its outcome depends on the  $V$ -binning, i.e.,  $\Delta V$ . The assumption that there is the proportionality  $\sigma_{f; \Delta V \rightarrow 0}^{(m)}(V) = \alpha(m)V$  implies that  $\text{Var} \left[ F_n^{(m)} \mid \bar{U}_n^{(m)} = V \right] = \alpha(m)^2 V^2$ , where  $F_n^{(m)}$  and  $\bar{U}_n^{(m)}$  are rv's reflecting the statistics of  $f_n^{(m)}$  and  $\bar{u}_n^{(m)}$ . Consequently,  $\text{Var} \left[ F_n^{(m)} / V \mid \bar{U}_n^{(m)} = V \right] = \alpha(m)^2$ , which can be written as  $\text{Var} \left[ F_n^{(m)} / \bar{U}_n^{(m)} \right] = \alpha(m)^2$ . Hence, the proportionality factor, which reflects the TI, can be estimated by estimating the standard deviation of the normalised fluctuation  $F_n^{(m)} / \bar{U}_n^{(m)}$ , i.e.,

$$\alpha(m) = \sqrt{\text{Var} \left\{ f_n^{(m)} / \bar{u}_n^{(m)} \right\}}, \quad (6.11)$$

where  $\text{Var} \{ \dots \}$  stands for the estimation of the set  $\{ \dots \}$ . Note that this method does not require any binning and takes a whole 24 h series into consideration. The proportionality factor is estimated for three 24 h recordings (day 186, 191, and 192) at the [Lammefjord \(1987\)](#) site with  $m = 101$ . They are given by 0.059, 0.0936, and 0.075, respectively, and visualised in Fig. 6.7. Eq. (6.11) is a good estimator of the proportionality factor for the data acquired at day 191, but there is a slight deviation from the proportionality for the data acquired at day 192. Furthermore, the data acquired at day 186 even show bad proportionality.

This can be due to the following reasons. Fig. 6.8 (a) suggests the tendency of the proportionality (6.10) for each of the twelve 24 h recordings gathered at the [Lammefjord \(1987\)](#) site. However, panel (b) reveals that the ratio  $\sigma_{f; \Delta V}^{(m)}(V)/V$  is approximately constant for mean wind speeds of above five metres per second. This could explain why the data gathered at the days 186 and 192 are not in agreement with (6.10) in Fig. 6.7 because the plot only considers mean wind speeds smaller than five metres per second. Another possibility can be a fluctuating TI over the diurnal cycle, i.e., even the time series  $f_n^{(m)}$  for  $n \in \mathcal{N}_{\Delta V}^{(m)}(V)$  is not (first order) stationary within 24 h. This possibility is checked by splitting a 24 h recording into 3, 12, and 48 sub-series, each of which containing 8 h, 2 h, and 30 min of data, respectively. That is, the TI is treated as being constant over periods of length  $\Delta\vartheta = 1/2, 2, 8$  h and its values, estimated via (6.11) in subsequent periods, are compared to each other. The result is shown in Fig. 6.9 and gives evidence that the TI on day 191 hardly fluctuates whereas the days 186 and 192 show larger TI fluctuations. More specifically, day 186 has hardly fluctuating TI between 8 h and 16 h. Its TI is (apparently) strongly fluctuating during the morning and evening hours in which, see Fig. 6.3, there is only little wind activity. The same behaviour is seen in the data gathered at day 192: in the time region with low wind activity (around 8 h and 20 h, see Fig. 6.3) the TI is comparably strongly fluctuating.

Concluding, this analysis shows that the data gathered at day 191 represent turbulent data over 24 h. Up to now, only  $m = 101$  is used. The day 191 data are therefore again analysed for a variety of  $m$  to check whether the proportionality (6.10) and the usefulness of (6.11) holds true for larger  $m$ . Fig. 6.10 shows the standard deviation of the fluctuation as a



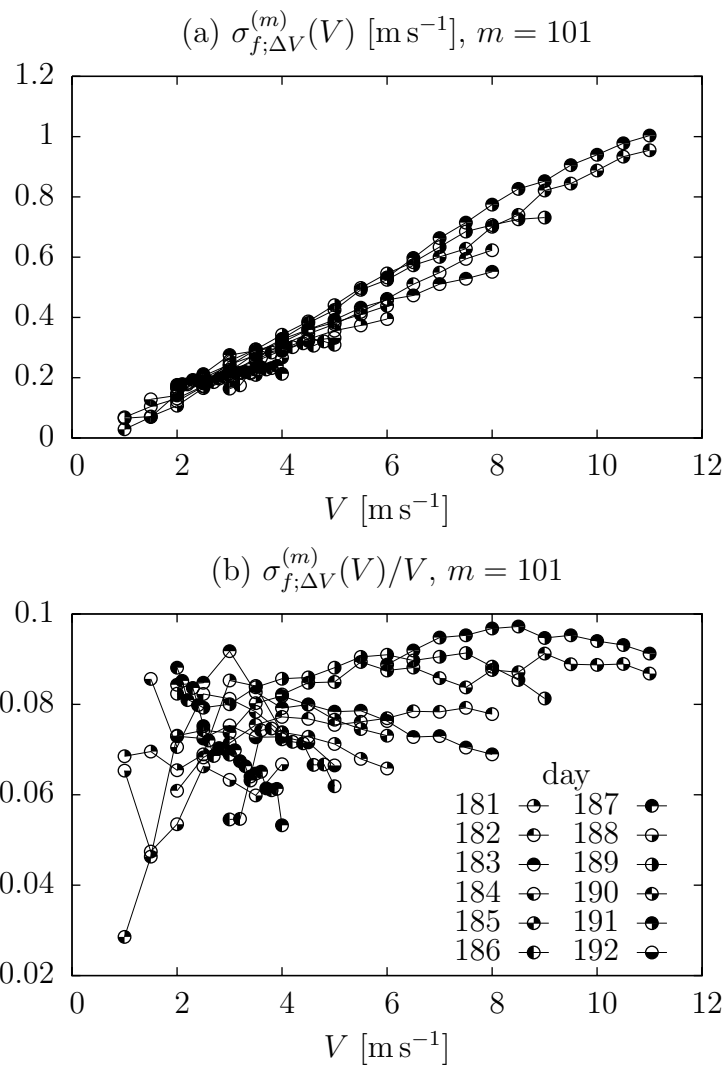
**Figure 6.7:** Estimated standard deviation of the fluctuation gathered at three days at the [Lammefjord \(1987\)](#) site conditioned on mean wind speed  $V$  over  $m = 101$  time steps. The figures show the raw data together with two fits ( $aV + b$  and  $cV$ ) obtained by linear regression and together with  $\alpha(m)V$ , where the slope  $\alpha(m)$  is obtained by (6.11). The results are summarised in the tables below. The number in brackets denotes the fit error.

function of the mean wind speed over 2.5 s, 12.5 s, and 125 s, i.e., the parameter  $m$  is equal to 21, 101, and 1001, respectively. It can be seen that the linearity (6.10) even holds true for much larger  $m$ . Note that the points for small and large  $V$ , which do not lie on the line, are obtained by estimating the standard deviation of a very small set  $\mathcal{N}_{\Delta V}^{(m)}(V)$ . It can also be concluded that Eq. (6.11) yields reasonable results for the proportionality factor  $\alpha(m)$ . Regarding the  $m$ -dependence of the TI, Fig. 6.11 depicts  $\alpha(m)$  as a function of  $m$ . It can be concluded that the TI increases with increasing  $m$ .

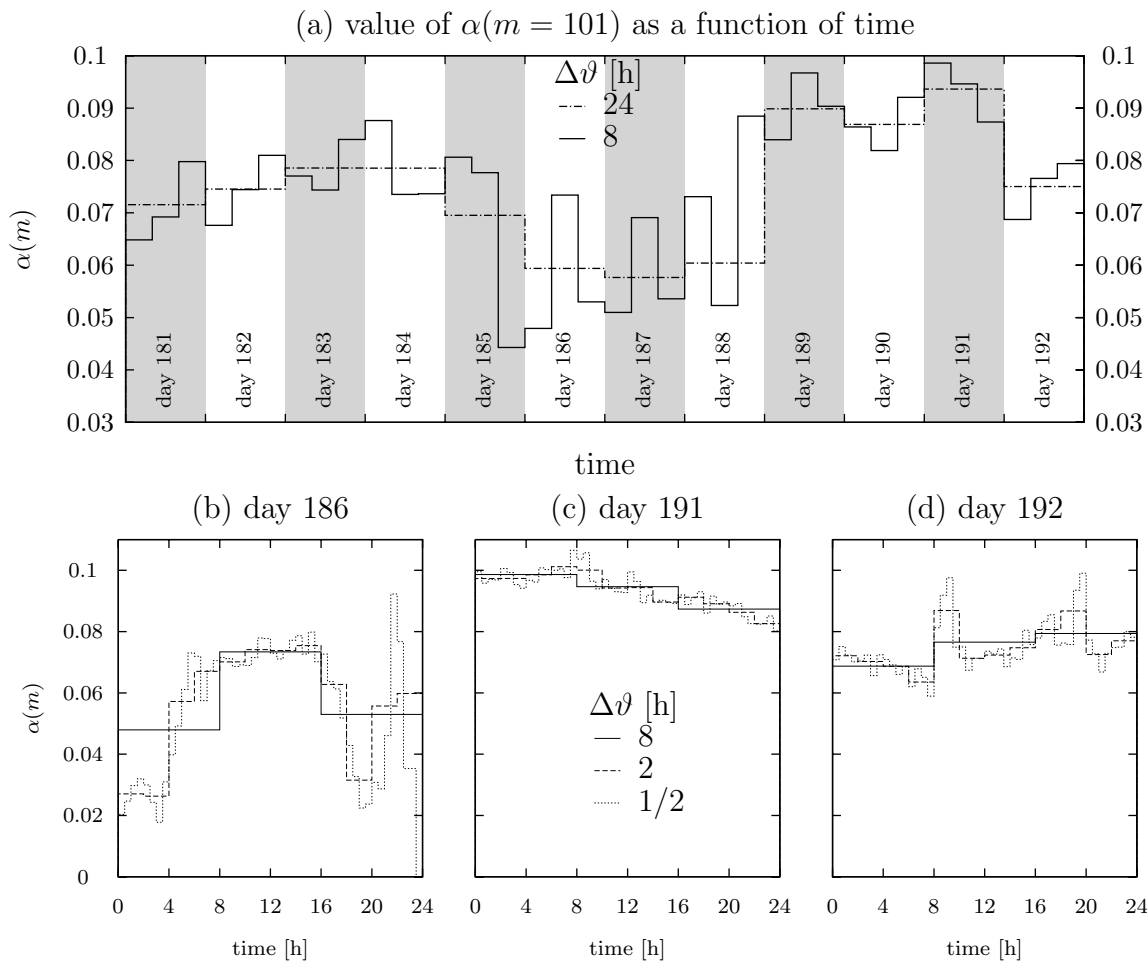
It should be mentioned that it is a non-trivial property of a time series that its fluctuation conditioned on the value  $V$  of the moving average  $V$  is normally distributed with zero mean and standard deviation proportional to  $V$ . This will be further studied in Chap. 8.

### 6.3 Increment statistics

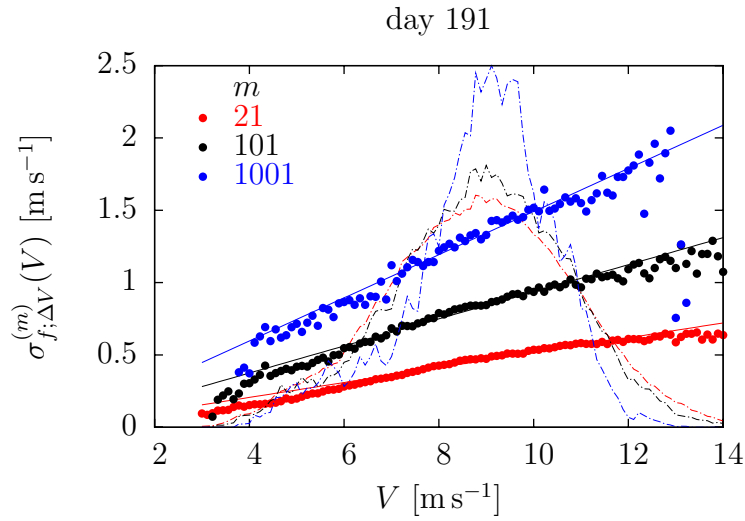
As described in Chap. 4 a turbulent flow is characterised by its spatial correlations. The longitudinal velocity increment (4.31) is  $[\vec{u}(\vec{r} + \vec{\ell}, t) - \vec{u}(\vec{r}, t)] \cdot \vec{\ell} / \ell$ . The [Taylor \(1938\)](#) hypothesis states that spatial correlations can be translated into temporal correlations if the flow is mainly unidirectional with small TI. Denoting the main flow direction by the unit vector  $\vec{e}_{\text{flow}}$  and choosing  $\vec{\ell}$  to be anti-parallel to  $\vec{e}_{\text{flow}}$ , i.e.,  $\vec{\ell} = -\ell \vec{e}_{\text{flow}}$ , the longitudinal increment is approximately the negative of the temporal increment  $u_{\text{flow}}(\vec{r}, t + \ell/U) - u_{\text{flow}}(\vec{r}, t)$ , where  $u_{\text{flow}} = \vec{u} \cdot \vec{e}_{\text{flow}}$  and  $U$  denotes the mean wind speed over a large time scale. However, a



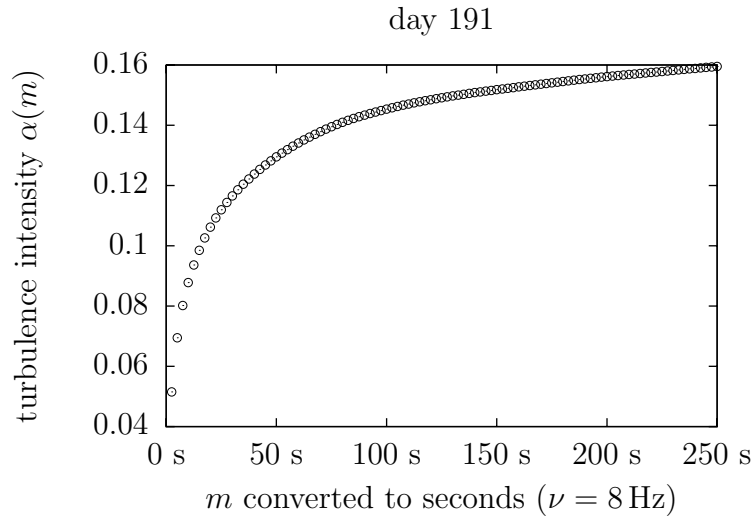
**Figure 6.8:** (a) The estimated standard deviation of the fluctuation as a function of the mean wind speed  $m = 101$  at the [Lammefjord \(1987\)](#) site for twelve 24 h recordings. (b) The estimated standard deviation of the fluctuation divided by the mean wind speed as a function of the mean wind speed.



**Figure 6.9:** The estimated proportionality factor (6.11) for the  $m = 101$  Lammefjord (1987) fluctuation in successive periods of lengths  $\Delta\vartheta$ . (a) “Long-term” behaviour over twelve days of measurement ( $\Delta\vartheta = 8, 24$  h). (b), (c), (d) Three 24 h samples individually with higher resolution ( $\Delta\vartheta = 1/2, 2, 8$  h).



**Figure 6.10:** The estimated standard deviation of the Lammefjord (1987) day 191 fluctuation as a function of the 2.5 s, 12.5 s, and 125 s mean wind speed by choosing  $m$  to be 21, 101, and 1001, respectively. The straight lines denotes  $\alpha(m)V$  with  $\alpha(m)$  being estimated by (6.11). The dash-dot line represents the distribution of  $V$  in arbitrary units and therefore the number of data points in  $\mathcal{N}_{\Delta V}^{(m)}(V)$ .



**Figure 6.11:** The TI of the Lammefjord (1987) day 191 data as a function of time scale  $m$  estimated by (6.11).

quantity which is much easier to determine and which is expected to have similar properties to  $u_{\text{flow}}(\vec{r}, t + \tau) - u_{\text{flow}}(\vec{r}, t)$  is the increment of the horizontal component

$$u_\tau(t) = u(t + \tau) - u(t). \quad (6.12)$$

The parameter  $\tau$  is called the *increment length* and plays the same role as the spatial scale  $\ell$ . Since  $u_\tau(t)$  is not exactly identical to a longitudinal velocity increment, its distribution might differ from the asymmetric Castaing distribution (4.47).

ABL wind speed measurements such as performed by Boettcher *et al.* (2003, 2007) and also by Beck *et al.* (2005a); Rizzo & Rapisardia (2005) showed that wind velocity increments are well described by the symmetric Castaing distribution. As mentioned in Sec. 4.7 it assumes that the velocity increment in a (temporal) cell with mean energy dissipation rate  $\varepsilon(\tau)$ <sup>3</sup> is given by the symmetric normal distribution

$$p(u_\tau | \beta) = \sqrt{\frac{\beta}{2\pi}} e^{-\beta u_\tau^2/2}, \quad (6.13)$$

where the variance is denoted by  $\beta^{-1}$ . The reason why this notation is used will become clear in the next section dealing with superstatistics. The inverse variance  $\beta$  of  $p(u_\tau | \beta)$  is proportional to  $\varepsilon(\tau)^{-2/3}$  and consequently a rv. The log-normal model for small scale turbulence with intermittency correction  $\mu$  assumes that  $\varepsilon(\tau)$  is for (small)  $\tau$  log-normally distributed with position parameter  $\propto (\tau/\tau_0)^{\mu/2}$  and shape parameter  $\mu \ln(\tau_0/\tau)$ , where  $\tau_0$  is a constant connected to the integral length. According to (A.20),  $\beta$  is also log-normally distributed with position parameter  $\beta_\tau \propto (\tau/\tau_0)^{-\mu/3}$  and shape parameter  $\lambda_\tau^2 = 4\mu/9 \ln(\tau_0/\tau)$ .

This is checked for the (time discrete) ABL wind speed time series  $u_n$  recorded at the Lammefjord (1987) site with frequency  $\nu = 8$  Hz. The observable of interest is the increment

$$u_{s;n} = u_{n+s} - u_n \quad (6.14)$$

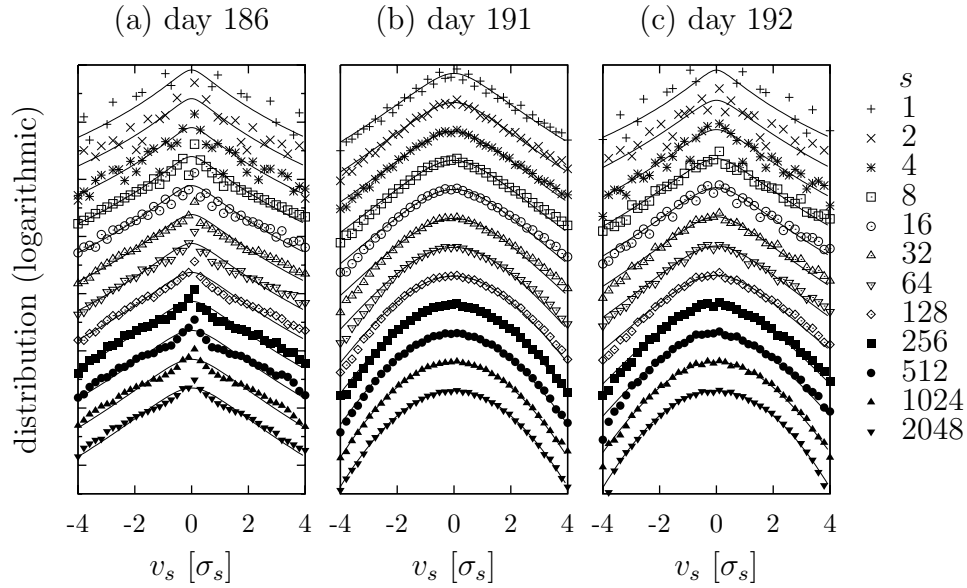
with increment length  $s = \nu\tau \in \mathbb{N}$ . It is expected that the increment  $v_s$ <sup>4</sup> is of the symmetric Castaing distribution

$$\begin{aligned} \text{Ca}_s(v_s) &= \int_0^\infty d\beta L_{\beta_s, \lambda_s^2}(\beta) N_{0, \beta^{-1}}(v_s) \\ &= \frac{1}{2\pi\lambda_s} \int_0^\infty \frac{d\beta}{\sqrt{\beta}} \exp \left[ -\frac{1}{2} \left( \frac{1}{\lambda_s} \ln \frac{\beta}{\beta_s} \right)^2 - \frac{1}{2} \beta v_s^2 \right] \end{aligned} \quad (6.15)$$

with position parameter  $\beta_s \propto s^{-\mu/3}$  and shape parameter  $\lambda_s^2$  which decreases linearly with  $\ln s$  and reaches zero for very large  $s$ , see Castaing *et al.* (1990) for a description of large

<sup>3</sup>The spatial scale  $\ell$  is replaced by the increment length  $\tau$

<sup>4</sup>The notation  $v_s$  is used instead of  $u_s$  to avoid confusion with the wind speed recorded at time  $s$ , which is labeled by  $u_s$ .



**Figure 6.12:** Increment histograms  $p_s(v_s)$  for three 24 h data recorded at the [Lammefjord \(1987\)](#) site. The x-axis represents  $v_s$  in units of the (estimated) standard deviation  $\sigma_s$ . The solid lines represent the corresponding (symmetric) Castaing distribution. The histograms are shifted and drawn in a semi-logarithmic plot for better visibility. **(a)** day 186. **(b)** day 191. **(c)** day 192.

scale turbulence which the log-normal model is not justified for. The two parameters can be estimated from the time series by estimating its variance  $\sigma_s^2$  and kurtosis  $k_s$ , i.e.,

$$\beta_s = \frac{1}{\sigma_s^2} \sqrt{\frac{k_s}{3}} \quad \text{and} \quad \lambda_s^2 = \ln \frac{k_s}{3}. \quad (6.16)$$

If the increment length  $s$  is sufficiently large such that  $\lambda_s^2$  is expected to be zero, the increment distribution (6.15) approaches a Gaussian distribution with kurtosis three and variance  $\beta_s^{-1}$  because it can be shown that  $\lim_{\lambda_s^2 \rightarrow 0} \text{Ca}_s(v_s) = N_{0, \beta_s^{-1}}(v_s)$  by using property (A.18) of the log-normal distribution.

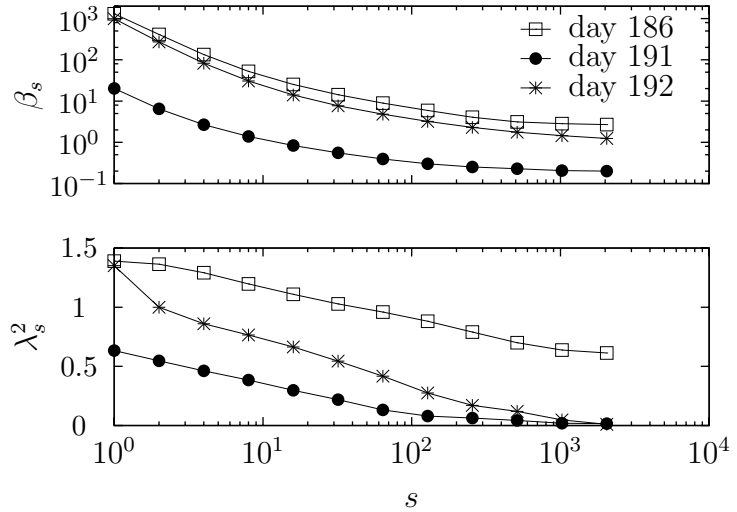
The increment distribution is estimated by

$$p_s(v_s) = \langle \delta(v_s - u_{s;n}) \rangle_n. \quad (6.17)$$

Its shape is compared to the shape of the Castaing distribution with position and shape parameter calculated according to (6.16) by plotting (6.17) semi-logarithmically for a variety of  $s$  ranging from 1 to 2048, see Fig. 6.12. That is in physical units with  $\nu = 8$  Hz, the increment length varies between 1/8 s and 4:16 min. It can be seen that the shape of the increment distribution is in good agreement with the symmetric Castaing distribution (6.15). Some distributions are slightly positively skewed.<sup>5</sup> It can also be noted that the increment

<sup>5</sup>The increment (6.14) is expected to have similar properties to the negative of the longitudinal velocity increment in flow direction. As the distribution of the latter is negatively skewed, the increment distribution  $p_s(v_s)$  might (but do not have to be) positively skewed.





**Figure 6.13:** The position and shape parameter  $\beta_s$  and  $\lambda_s^2$ , respectively, as a function of  $s$  for the data acquired at days 186, 191, and 192 at the [Lammefjord \(1987\)](#) site.

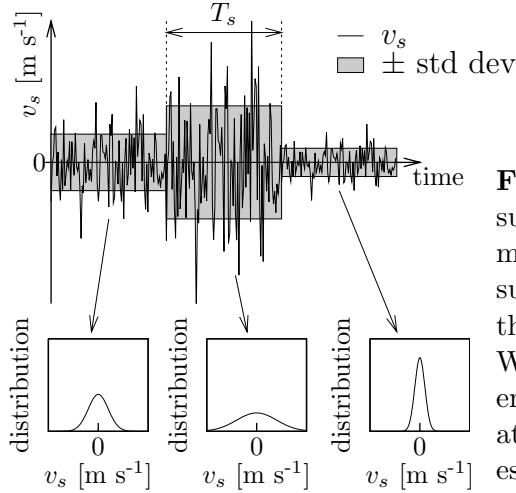
data gathered at day 186 do not tend towards a Gaussian distribution as  $s$  becomes larger. This is visualised in Fig. 6.13 depicting the  $s$ -dependency of the position and shape parameter estimated by (6.16). The plot shows evidence that  $\lambda_s^2$  decreases linearly with the logarithm of  $s$  for small  $s$ . For larger  $s$ ,  $\lambda_s^2$  approaches zero for the data gathered at day 191 and 192. For the day 186 data, there is no such settlement up to  $s = 2048$  which corresponds to 4:16 min in physical units. That is, the increment distribution for large increment length is still intermittent for the day 186 data.

[Boettcher et al. \(2007\)](#) study this large scale intermittency by treating the mean wind speed as a Weibull distributed rv. The authors investigate the increment distribution  $p_s(v_s)$  for large  $s$  as a function of the spread of the Weibull distribution. The work gives evidence that for days with highly fluctuating mean wind speed the increment distribution is not expected to become a Gaussian. The drawback with this approach is that a Weibull distributed mean wind speed is frequently found empirically and hence an accepted fact (see [Burton et al., 2004](#)) but that there is no theoretical justification for that.

## 6.4 Superstatistics

Sec. 6.2 concludes that the TI, which is a parameter characterising the turbulence, might vary in a diurnal cycle. The same question arises for the position and shape parameter of the increment distribution. These are also parameters describing the turbulence. This section develops a statistical test which quantifies the time dependence of the increment distribution parameters and checks whether an ABL wind speed recording, or just a part of it, reflects turbulence.

The results of the previous section about the increment distribution with increment length



**Figure 6.14:** Schematic sketch of the superstatistical approach. The increment series is considered to consist of successive cells of length  $T_s$ , in which the increment is normally distributed. Within such a cell the mean is considered to be zero and the standard deviation (std dev) is approximated by its estimation over the cell interval.

$s$  can be reformulated as hypothesis

$$H_0 : \begin{cases} v_s \sim \text{Ca}_s & \text{for all } s \text{ and} \\ \lambda_s^2 \rightarrow 0 & \text{for large } s. \end{cases} \quad (6.18)$$

Here,  $\lambda_s^2$  stands for the shape parameter of the increment distribution, which has hypothetically the shape of a symmetric Castaing distribution (6.15). There is no standard statistical test which directly checks (6.18), but it can be used that the symmetric Castaing distribution (6.15) corresponds to the superstatistics (5.1) with system parameter  $\beta$  and  $E = v_s^2/2$ . Due to the universality hypothesis, the system parameter  $\beta$  is related to the mean energy dissipation rate  $\varepsilon$  by  $\beta \propto \varepsilon^{-2/3}$ . According to Kolmogorov (1962) the latter can be thought as a product of positive iid rv's, cf. Fig. 4.3, making the (intermittent) statistics of the velocity increment  $v_s$  be an instance of the log-normal superstatistics derived in Ex. 5.4. As a consequence, Castaing's hypothesis (6.18) can be checked by using the superstatistics approach in the following way. Assuming that in a temporal cell with system parameter  $\beta$  the wind speed increment is normally distributed with vanishing mean and variance  $1/\beta$ , the hypothesis (6.18) is equivalent to the null hypothesis

$$H_0 : \begin{cases} \beta \text{ is for all } s \text{ log-normally distributed and} \\ \text{its shape parameter } \lambda_s^2 \rightarrow 0 \text{ for large } s. \end{cases} \quad (6.19)$$

That is, the time series  $u_{s;n}$  needs to be converted into a  $\beta$ -series so that  $H_0$  can be checked with standard statistical tests.

Beck *et al.* (2005b) and Queiros (2007) proposed an algorithm to compute the  $\beta$ -series. The assumption is that the increment time series  $u_{s;n}$  consists of cells in which the increment is of a symmetric Gaussian distribution. These cells are treated as being of equal length  $T_s$ . Fig. 6.14 sketches schematically the assumption. The  $\beta$ -series is obtained by estimating the variance in each cell and taking its inverse. As explained in Secs. 2.8 and 3.2, in order to

estimate the variance in each cell, their length  $T_s$  must be much larger than the decay time

$$\tau_s = \min_{t \in \mathbb{N}} \{t : \gamma_s(t) \leq \exp(-1)\}, \quad (6.20)$$

where the auto correlation function  $\gamma_s(\cdot)$  is estimated from the whole increment series  $(u_{s;n})_{n=1}^N$ . The essential step of the algorithm is to find a suitable time scale  $T_s$ . It can be identified with the scale on which the increment series is of a normal distribution. In order to estimate  $T_s$ , the increment series  $(u_{s;n})_{n=1}^N$  which is of size  $N$  is split into the sub-series

$$\underbrace{(u_{s;1}, \dots, u_{s;m})}_{\text{size } m}, \underbrace{(u_{s;m+1}, \dots, u_{s;2m})}_{\text{size } m}, \dots, \underbrace{(u_{s;(K-1)m+1}, \dots, u_{s;N})}_{\text{size } \leq m} \quad (6.21)$$

with  $m \in \mathbb{N}$  and  $K = N/m$  if  $N/m \in \mathbb{N}$  or  $K = \text{int}[N/m] + 1$  otherwise. Being a measure of Gaussianity, the sample kurtosis, which for a set  $\mathcal{A}$  of size  $|\mathcal{A}|$  and mean  $\bar{a} = \frac{1}{|\mathcal{A}|} \sum_{a \in \mathcal{A}} a$  is defined by

$$\mathbb{Kurt}^* \mathcal{A} = |\mathcal{A}| \times \frac{\sum_{a \in \mathcal{A}} (a - \bar{a})^4}{(\sum_{a \in \mathcal{A}} (a - \bar{a})^2)^2}, \quad (6.22)$$

is used to determine the Gaussianity in each sub-series:

$$\underbrace{(u_{s;1}, \dots, u_{s;m})}_{\mathbb{Kurt}^*\{\dots\} = \kappa_{s;1}^{(m)}}, \underbrace{(u_{s;m+1}, \dots, u_{s;2m})}_{\mathbb{Kurt}^*\{\dots\} = \kappa_{s;2}^{(m)}}, \dots, \underbrace{(u_{s;(K-1)m+1}, \dots, u_{s;N})}_{\mathbb{Kurt}^*\{\dots\} = \kappa_{s;K}^{(m)}}. \quad (6.23)$$

If the sub-series is normally distributed and reasonably large<sup>6</sup>, its estimated sample kurtosis is equal to three. Taking the mean of the sub-series sample kurtosis leads to

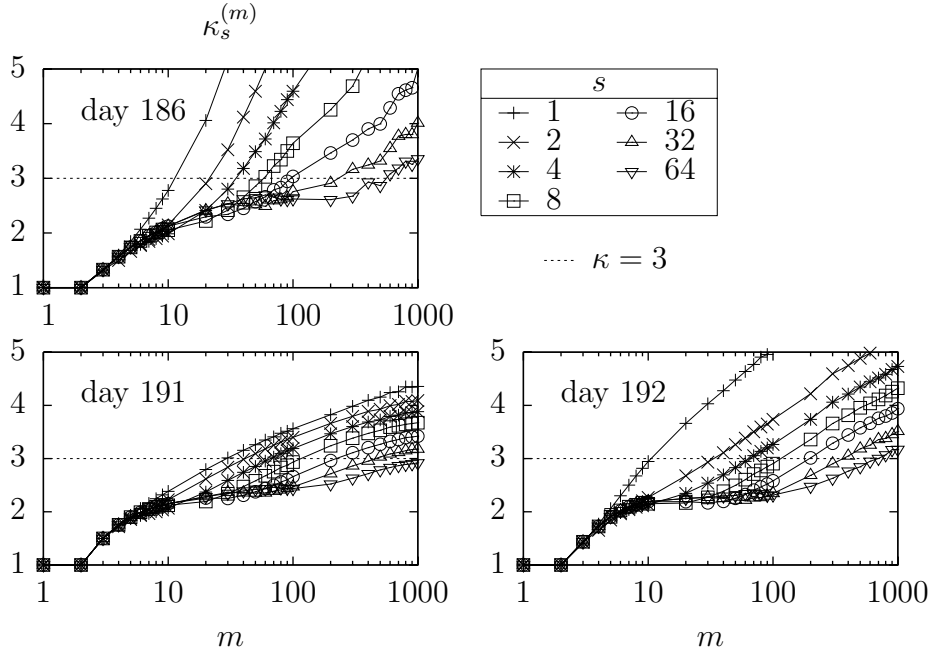
$$\kappa_s^{(m)} = \left\langle \kappa_{s;k}^{(m)} \right\rangle_{k=1, \dots, K} \quad (6.24)$$

being a function of  $m$ . In order to find the large time scale,  $m$  is varied between 1 and  $N$ . Due to  $\kappa_s^{(1)} = 1$ , see footnote 6, and  $\kappa_s^{(N)} = k_s \geq 3$  with  $k_s$  denoting the kurtosis of the whole increment series, there is a value  $m$  between 1 and  $N$  such that  $\kappa_s^{(m)}$  is closest to three. According to Beck *et al.* (2005b) and Queiros (2007), this value is taken as an estimation for the large time scale  $T_s$ .<sup>7</sup> Note that App. B discusses the influence of the value of  $T_s$  on the outcome of the analysis. As mentioned before, the  $\beta$ -series is simply

$$\underbrace{(u_{s;1}, \dots, u_{s;T_s})}_{1/\text{Var}\{\dots\} = \beta_1}, \underbrace{(u_{s;T_s+1}, \dots, u_{s;2T_s})}_{1/\text{Var}\{\dots\} = \beta_2}, \dots, \underbrace{(u_{s;(K-1)T_s+1}, \dots, u_{s;N})}_{1/\text{Var}\{\dots\} = \beta_K}. \quad (6.25)$$

<sup>6</sup>The star in the definition (6.22) symbolises that the sample kurtosis is biased with respect to the sample size. If for instance  $\mathcal{A}$  consists of only one or two elements, i.e.  $\mathcal{A} = \{a\}$  or  $\mathcal{A} = \{a_1, a_2\}$ , the sample kurtosis is always unity.

<sup>7</sup>As  $T_s$  is supposed to be large there should be no problem with using the biased kurtosis estimator (6.22). The advantage of this approach is that it always yields an estimation for  $T_s$ .



**Figure 6.15:** Estimation of  $\kappa_s^{(m)}$  as a function of  $m$  for three 24 h increment series recorded at the [Lammefjord \(1987\)](#) site (day 186, 191, and 192). The dashed line corresponds to  $\kappa = 3$ . The intersection  $\kappa_s^{(m)} = 3$  is an estimation for the time scale  $T_s$  shown in Fig. 6.16.

Its distribution is estimated by  $f_s(\beta) = \langle \delta(\beta - \beta_k) \rangle_{k=1, \dots, K}$  which, hypothetically, has the shape of a log-normal distribution. Thus, it is more convenient to consider the variables<sup>8</sup>

$$\Lambda_k = \ln \beta_k \quad (6.26)$$

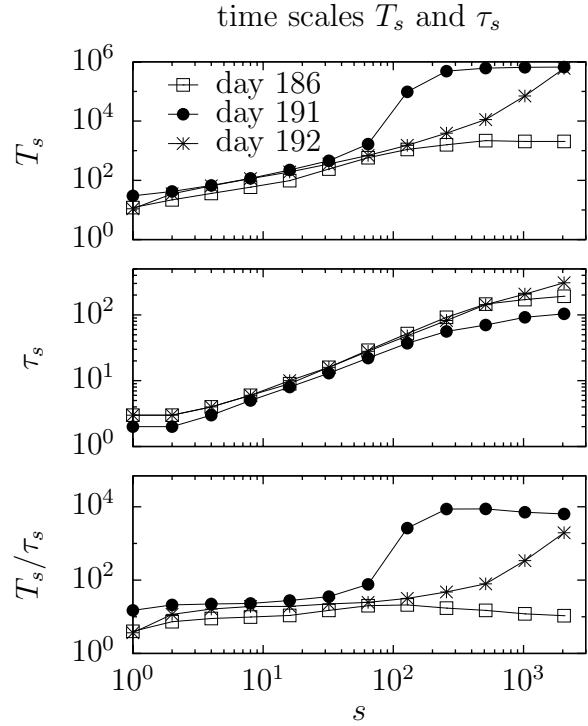
and estimate the distribution  $h_s(\Lambda) = \langle \delta(\Lambda - \Lambda_k) \rangle_{k=1, \dots, K}$ . The null hypothesis reads

$$H_0 : \begin{cases} \Lambda \sim N_{\Lambda_s, \lambda_s^2} & \text{for all } s \text{ and} \\ \lambda_s^2 \rightarrow 0 & \text{for large } s \end{cases} \quad (6.27)$$

with  $\Lambda_s = \mathbb{M}\text{ean} \{ \Lambda_1, \dots, \Lambda_K \}$  and  $\lambda_s^2 = \mathbb{V}\text{ar} \{ \Lambda_1, \dots, \Lambda_K \}$ . The hypothesis  $H_0$  is checked using the Kolmogorov-Smirnov test explained in Sec. 2.9. If it is satisfied, the increment series is of a Castaing distribution with position parameter  $\beta_s = e^{\Lambda_s}$  and shape parameter  $\lambda_s^2$ .

This algorithm is applied to three 24 h recorded increment series obtained from the [Lammefjord \(1987\)](#) measurement (day 186, 191, and 192) and discussed step by step. Fig. 6.15 displays the mean sub-series sample kurtosis (6.24) as a function of window size  $m$  for different increment lengths  $s$ . It is evident that  $\kappa_s^{(m)}$  increases with increasing  $m$  and has an intersection with three, which is used for the definition of the time scale  $T_s$ . It can be

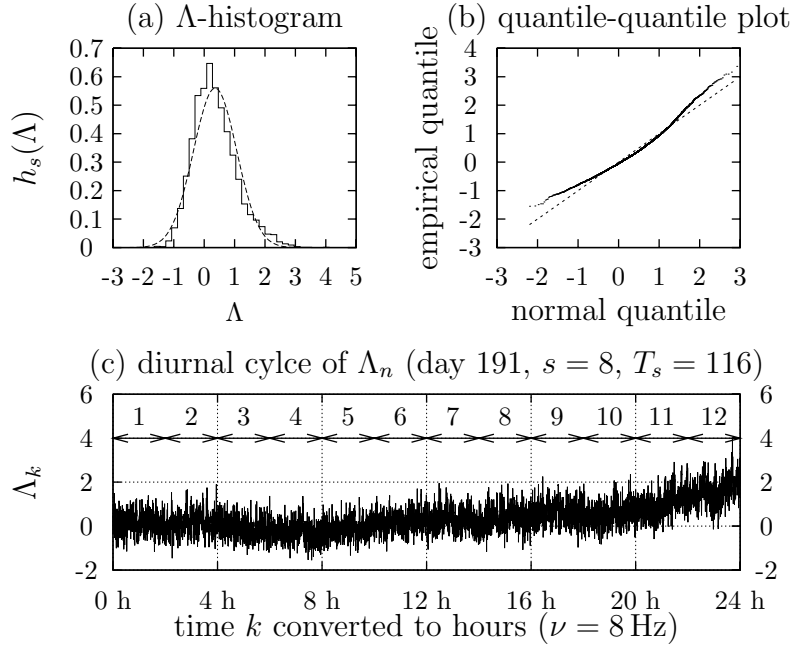
<sup>8</sup>For simplicity, the data are treated as being dimensionless.



**Figure 6.16:** Time scale estimation for three increment series (24 h recordings) obtained at the [Lammefjord \(1987\)](#) site (day 186, 191, and 192). Each series consists of roughly  $7 \times 10^5$  data points being the upper limit for the estimation of  $T_s$ . **(top)**  $T_s$  as a function of  $s$ . **(central)**  $\tau_s$  as a function of  $s$ . **(bottom)** Ratio  $T_s/\tau_s$ .

seen that the latter becomes larger with increasing  $s$ . This is visualised in the top panel of Fig. 6.16 depicting the  $s$ -dependence of  $T_s$ . As it describes the scale on which the increment process is of a normal distribution, an increasing  $T_s$  is in full agreement with the approach to a Gaussian increment distribution with increasing increment length. Its estimation is bounded from above by the length  $N$  of the increment series, i.e.,  $T_s \leq N$ . The day 191 and 192 data yield an estimation  $T_s \approx N$  for very large  $s$ . This means that the whole increment series can be regarded as being normally distributed, which is in agreement with Castaing's hypothesis (6.18) and with the histograms plotted in Fig. 6.12. In addition, Fig. 6.16 reveals that the day 186 increment series does not reach a Gaussian distribution up to  $s = 2048$ , which is also in accordance with Fig. 6.12. Last, but not least, the bottom panel of Fig. 6.16 verifies the existence of two separated time scales  $T_s \gg \tau_s$ . It pictures the ratio  $T_s/\tau_s$  being of the order of magnitude of 10 or larger.

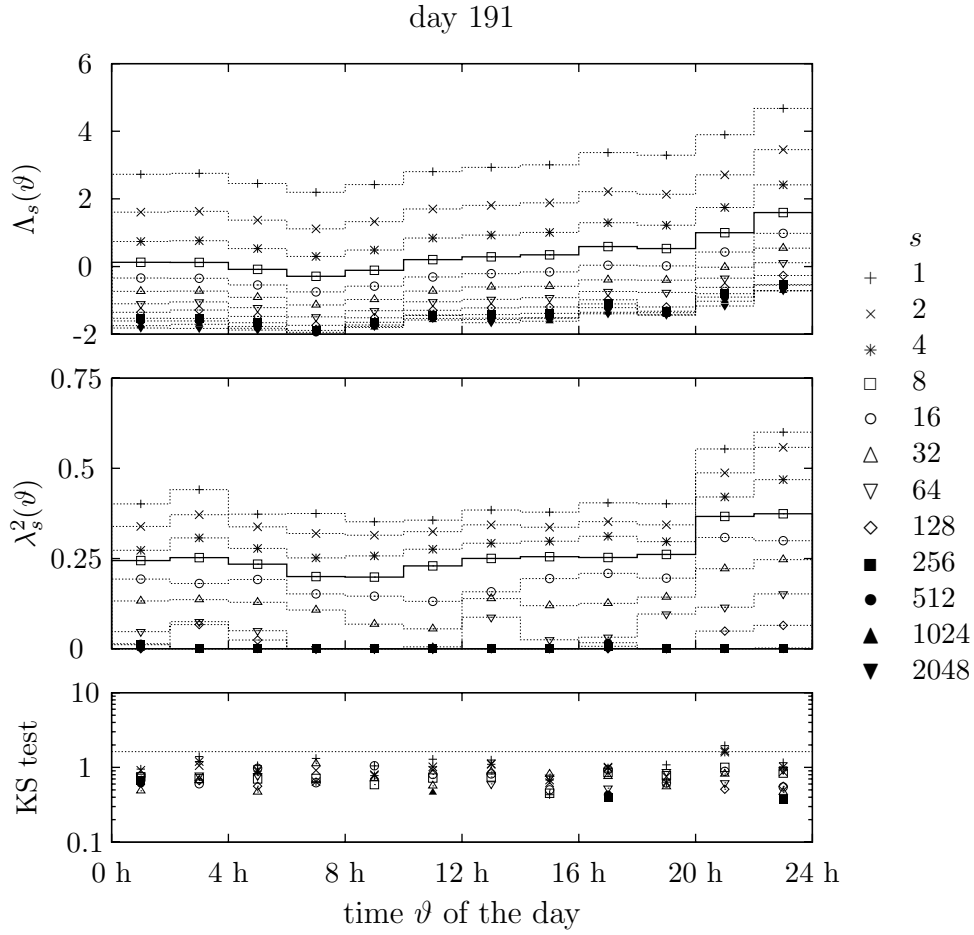
As an example, Fig. 6.17 (a) shows the estimated  $\Lambda$ -distribution  $h_s(\Lambda)$  for  $s = 8$  of the 24 h recording marked with day 191. The time scale  $T_s$  is estimated to be 116, which in physical units ( $\nu = 8$  Hz) corresponds to 14.5 s. It can be seen that  $h_s(\Lambda)$  is close to a normal distribution, but it has a systematic and statistically significant deviation. In fact, it has a positive skewness. The non-Gaussianity is underlined by the quantile-quantile plot in panel (b) of this figure. That is, on a scale of 24 h the turbulence hypothesis (6.27) is not satisfied. Nevertheless, panel (c), which shows the diurnal cycle of  $\Lambda$  over 24 h, indicates that the hypothesis might be fulfilled on a smaller period. Hence, the increment series is divided into twelve 2 h sub-samples, each of which represents a time  $\vartheta$  of the day. Each sub-sample is analysed with respect to superstatistics individually. This also includes computing



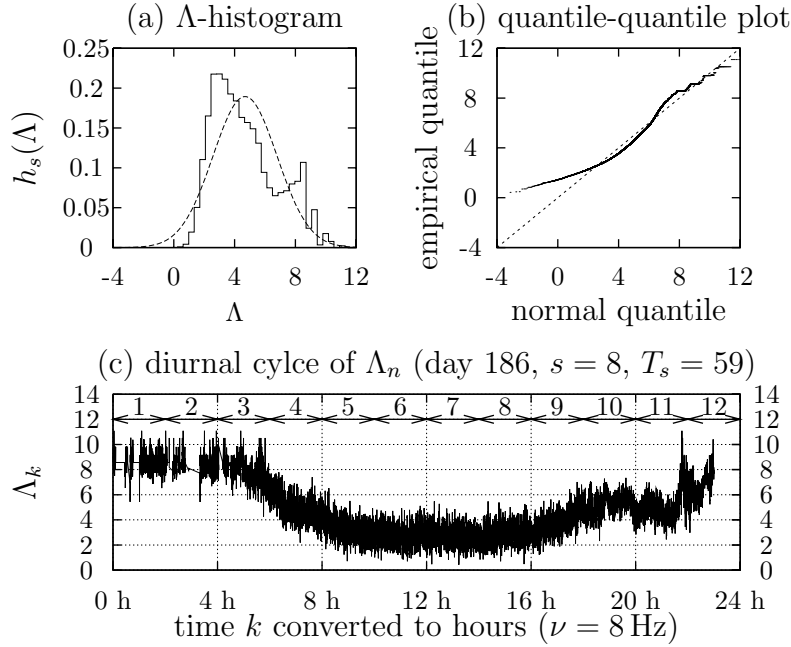
**Figure 6.17:** Superstatistical analysis of an increment series with  $s = 8$  (24 h recording) obtained at the [Lammefjord \(1987\)](#) site (day 191 with estimated  $T_s = 116$ ). **(a)**  $\Lambda$ -distribution  $h_s(\Lambda)$ . The dashed line corresponds to a Gaussian distribution with same mean and variance as  $h_s(\Lambda)$ . **(b)** Quantile-quantile plot for the  $\Lambda$ -distribution with respect to a normal distribution with same mean and variance. **(c)** Diurnal cycle of the  $\Lambda$ -series. The numbers 1 to 12 illustrate the twelve 2 h sub-samples which are individually investigated with respect to superstatistics.

and comparing the time scales  $T_s(\vartheta)$  and  $\tau_s(\vartheta)$  for each sub-sample. The  $\Lambda$ -series for each sub-sample is calculated and tested for Gaussianity (6.27) using a Kolmogorov-Smirnov test. This is done not only for  $s = 8$  but for a variety of increment lengths ranging from  $s = 1$  to  $s = 2048$ . Fig. 6.18 shows the result for each 2 h sub-sample. The top and central panel depict the mean and variance of  $\Lambda$  for each sub-sample, respectively. The bottom panel shows the test value of the Kolmogorov-Smirnov test and the critical value for the significance level 1%. The graph allows the conclusion that the hypothesis “ $\Lambda$  is normally distributed” cannot be rejected on a significance level 1%. That is, we can interpret  $e^{\Lambda_s}$  and  $\lambda_s^2$  as the position and shape parameter, respectively, of the corresponding Castaing distribution. Moreover, the  $\lambda_s^2$ -plot in the central panel reveals that for large  $s$  the shape parameter of each sub-sample approaches zero satisfying the hypothesis “ $\lambda_s^2 \rightarrow 0$  as  $s \gg 1$ ”. Therefore, it can be concluded that Castaing’s hypothesis (6.18) is fulfilled during time intervals of 2 h, but it is (slightly) violated on much larger time scales, such as 24 h, due to the time dependence of  $\Lambda_s(\vartheta)$  and  $\lambda_s^2(\vartheta)$ . In other words, the distribution shown in Fig. 6.17 is a superposition of Gaussians with different means  $\Lambda_s(\vartheta)$  and variances  $\lambda_s^2(\vartheta)$  for  $s = 8$  and is thus not exactly a normally shaped distribution.

The same analysis is done with the [Lammefjord \(1987\)](#) day 186 data which do not show a clear cross-over behaviour between an intermittent and Gaussian increment distribution, cf. Figs. 6.12, 6.13, 6.16. Firstly, the superstatistical algorithm is used to extract the  $\Lambda$ -series from the 24 h time series  $(u_{s;n})_{n=0}^{N-1}$  with  $s = 8$  and tested for hypothesis  $H_0$  in (6.27). The time scale  $T_s$  is estimated to be 58, which in physical units ( $\nu = 8$  Hz) corresponds to 7.4 s. Fig. 6.19 (a) shows the histogram  $h_s(\Lambda)$  for  $s = 8$ . It is clearly non-Gaussian shaped. Secondly, the increment series of the 24 h recording is also divided into twelve 2 h sub-samples. Each of which is analysed with respect to superstatistics individually. Fig. 6.20



**Figure 6.18:** Superstatistical analysis of a 2 h resolved diurnal cycle (day 191 at the [Lammefjord, 1987](#), site). The increment length is varied between  $s = 1$  and  $s = 2048$ . The solid line corresponds to  $s = 8$  and is therefore related to the graphs in Fig. 6.17. **(top)** Mean of  $\Lambda$  in sub-sample  $\vartheta$ . **(central)** Variance of  $\Lambda$  in sub-sample  $\vartheta$ . **(bottom)** Kolmogorov-Smirnov test variable  $d_{\Lambda}^{\text{KS}} \sqrt{N_{\Lambda}}$  for each sub-sample  $\vartheta$  ( $N_{\Lambda}$  denotes the number of  $\Lambda$ 's and depends on the  $T_s$  estimated in sub-sample  $\vartheta$ ). The dashed line corresponds to the critical value 1.63 for the significance level 1%.



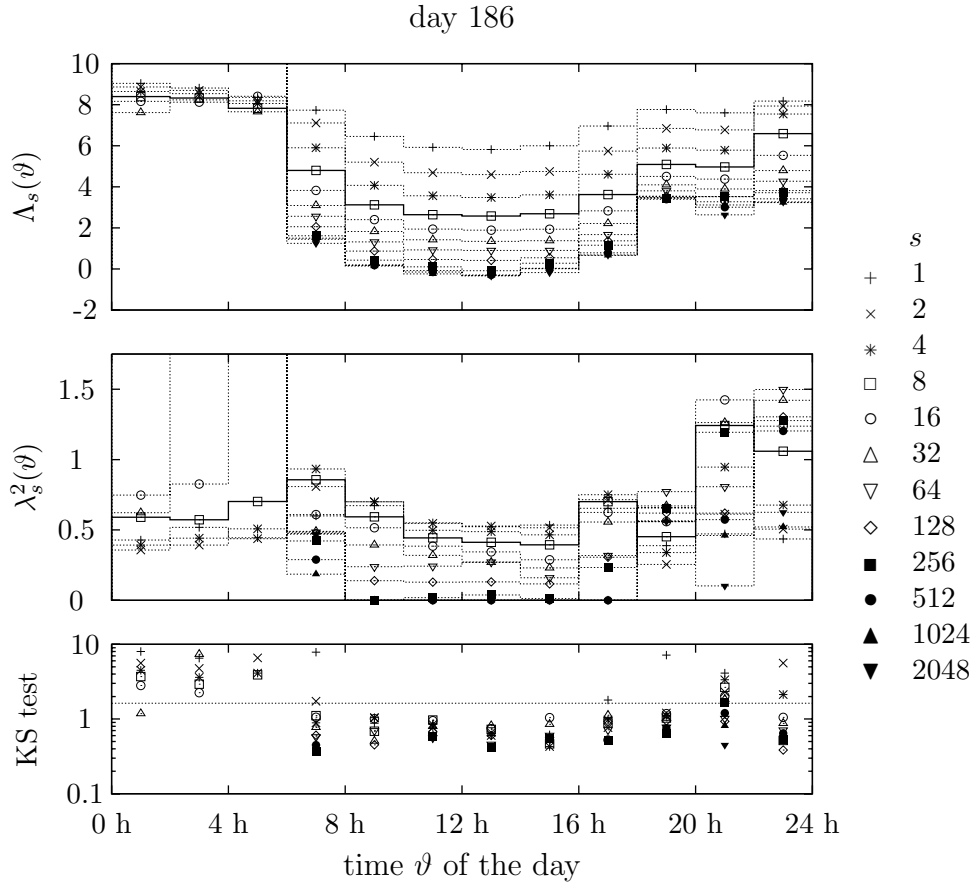
**Figure 6.19:** Superstatistical analysis of a increment series with  $s = 8$  (24 h recording) obtained at the [Lammefjord \(1987\)](#) site (day 186 with estimated  $T_s = 59$ ). **(a)**  $\Lambda$ -distribution  $h_s(\Lambda)$ . The dashed line corresponds to a Gaussian distribution with same mean and variance as  $h_s(\Lambda)$ . **(b)** Quantile-quantile plot for the  $\Lambda$ -distribution with respect to a normal distribution with same mean and variance. **(c)** Diurnal cycle of the  $\Lambda$ -series. The numbers 1 to 12 illustrate the twelve 2 h sub-samples which are individually investigated with respect to superstatistics.

displays  $\Lambda_s(\vartheta)$  and  $\lambda_s^2(\vartheta)$  as a function of time  $\vartheta$ . The time span  $8 \text{ h} \lesssim \vartheta \lesssim 18 \text{ h}$  is the only region where the  $H_0$  hypothesis (6.27) cannot be rejected on a 2 h scale with significance level 1 %: the Kolmogorov-Smirnov test value is below the critical value and the  $\lambda_s^2(\vartheta)$  goes to zero as  $s$  gets larger.<sup>9</sup> That means that Castaing’s hypothesis (6.18) is satisfied on windows of 2 h length in the mentioned time span. Outside this temporal region, the length of 2 h for the sub-samples is still too large, i.e., the resolution is too low for recovering a normally distributed  $\Lambda$ -series, or the wind activity is too calm to be called turbulent. The latter is more probable because the fluctuation analysis led to the conclusion that there is only little wind activity in the morning and evening hours of day 186, cf. Figs. 6.3 and 6.9. Additionally, due to the inverse proportionality between  $\Lambda$  and the increment volatility, a large  $\Lambda$  is equivalent to a hardly fluctuating increment. Fig. 6.19 (c) gives evidence that  $\Lambda$  is comparably large during the morning and evening hours. In any case, the increment histogram in Fig. 6.12 does not show a clear cross-over because the 24 h analysis also considers the time region in which the turbulence parameters  $\Lambda_s$  and  $\lambda_s^2$  are highly fluctuating. It should be mentioned that highly fluctuating turbulence parameters could also be an indication that there is only little wind activity in these temporal regions.

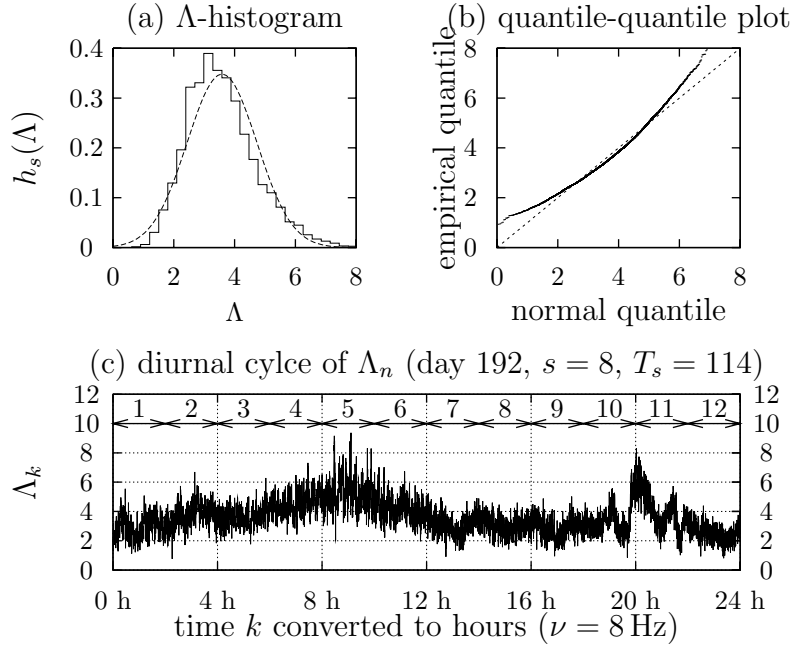
A similar behaviour is found in the data gathered at day 192. Fig. 6.21 depicts exemplarily the  $\Lambda$ -distribution for the  $s = 8$  increment series (24 h recording) and the corresponding quantile-quantile plot. It can be inferred that  $\Lambda$  is not normally distributed. Splitting the 24 h recording into twelve 2 h sub-samples, where each of which is analysed with respect to superstatistics individually, leads to the conclusion that around 8 h and 20 h the Castaing hypothesis is not satisfied. Again, these are temporal regions with low wind activity and

<sup>9</sup>Additionally, the time scale  $T_s$  is estimated to be larger than  $\tau_s$  in these sub-samples making the estimation of  $\Lambda_s$  and  $\lambda_s^2$  trustworthy.





**Figure 6.20:** Superstatistical analysis of a 2 h resolved diurnal cycle (day 186 at the [Lammefjord, 1987](#), site). The increment length is varied between  $s = 1$  and  $s = 2048$ . The solid line corresponds to  $s = 8$  and is therefore related to the graphs in Fig. 6.17. **(top)** Mean of  $\Lambda$  in sub-sample  $\vartheta$ . **(central)** Variance of  $\Lambda$  in sub-sample  $\vartheta$ . **(bottom)** Kolmogorov-Smirnov test variable  $d_{\Lambda}^{\text{KS}} \sqrt{N_{\Lambda}}$  for each sub-sample  $\vartheta$  ( $N_{\Lambda}$  denotes the number of  $\Lambda$ 's and depends on the  $T_s$  estimated in sub-sample  $\vartheta$ ). The dashed line corresponds to the critical value 1.63 for the significance level 1%.



**Figure 6.21:** Superstatistical analysis of an increment series with  $s = 8$  (24 h recording) obtained at the [Lammefjord \(1987\)](#) site (day 192 with estimated  $T_s = 114$ ). **(a)**  $\Lambda$ -distribution  $h_s(\Lambda)$ . The dashed line corresponds to a Gaussian distribution with same mean and variance as  $h_s(\Lambda)$ . **(b)** Quantile-quantile plot for the  $\Lambda$ -distribution with respect to a normal distribution with same mean and variance. **(c)** Diurnal cycle of the  $\Lambda$ -series. The numbers 1 to 12 illustrate the twelve 2 h sub-samples which are individually investigated with respect to superstatistics.

highly fluctuating turbulence parameters (such as TI,  $\Lambda_s$ , and  $\lambda_s^2$ ), cf. Figs. 6.3, 6.9, and 6.21 (c).

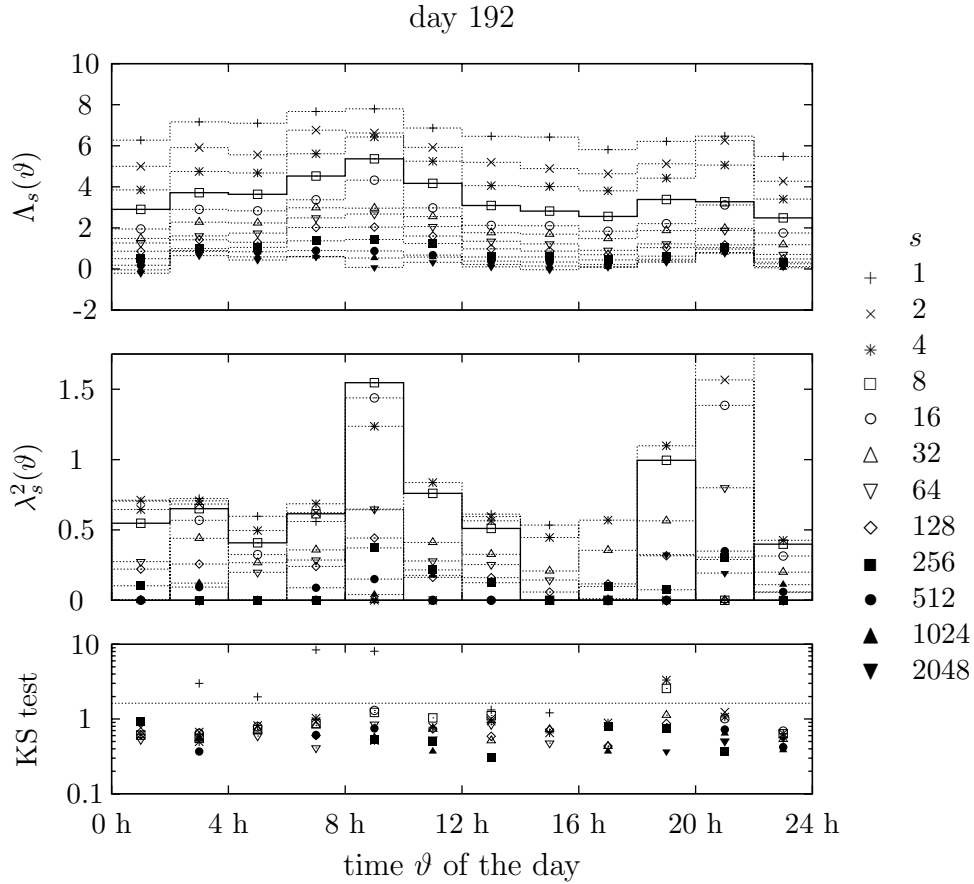
## 6.5 Summary

There is significant evidence that there are periods, in which atmospheric boundary layer wind speed has statistically similar properties to ideal turbulence. That is:

1. The fluctuation around the mean wind speed is roughly Gaussian distributed.
2. The standard deviation of the fluctuation is proportional to the mean wind speed as long as the latter is above  $5 \text{ m s}^{-1}$ .
3. The increment distribution is not (or slightly positively) skewed and in good approximation a symmetric Castaing distribution.
4. For small/large increment lengths, the increment distribution is close to a double exponential/Gaussian distribution.

The relevant parameters are

- the turbulence intensity, i.e., the proportionality factor between the standard deviation of the fluctuation and the mean wind speed, and
- the position and shape parameter of the increment distribution fitted by a symmetric Castaing distribution.



**Figure 6.22:** Superstatistical analysis of a 2 h resolved diurnal cycle (day 192 at the [Lammefjord, 1987](#), site). The increment length is varied between  $s = 1$  and  $s = 2048$ . The solid line corresponds to  $s = 8$  and is therefore related to the graphs in Fig. 6.21. **(top)** Mean of  $\Lambda$  in sub-sample  $\vartheta$ . **(central)** Variance of  $\Lambda$  in sub-sample  $\vartheta$ . **(bottom)** Kolmogorov-Smirnov test variable  $d_{\Lambda}^{\text{KS}} \sqrt{N_{\Lambda}}$  for each sub-sample  $\vartheta$  ( $N_{\Lambda}$  denotes the number of  $\Lambda$ 's and depends on the  $T_s$  estimated in sub-sample  $\vartheta$ ). The dashed line corresponds to the critical value 1.63 for the significance level 1%.

The measured turbulence intensity depends on the considered time scale  $T$  over which the moving average is computed: the value of the turbulence intensity increases with increasing  $T$ . The shape of the increment distribution depends on the considered increment length  $\tau$ : both parameters decrease with increasing  $\tau$ , the shape parameter approaches zero for sufficiently large  $\tau$ .

The turbulence parameters might be different in subsequent periods. The tools allow to quantify their time evolution. It is found that

- the periods in which the turbulence intensity can be regarded as being constant last from several hours up to 24 h and
- the Castaing parameters usually fluctuate in a diurnal cycle.

The first result suggests to transform a wind velocity recording, in which the turbulence intensity can be treated as being constant, into a first order stationary series by dividing the fluctuation by the value of the moving average. The second result can be used to explain that the increment distribution for large increment lengths is sometimes still found to be non-Gaussian (see e.g. [Boettcher \*et al.\*, 2007](#)) when analysing a recording being longer than a turbulent period. The time series can be interpreted as consisting of subsequent periods, in which the increment distribution is Gaussian for large increment length. Nevertheless, the variance, which is identical to the inverse of the position parameter, changes from one period to another so that the analysis of the full time series yields a superposition of Gaussians with different variances.

Additionally, the methods are able to indicate the existence of periods in which the above mentioned properties found in turbulent periods are most likely not satisfied. Indications are:

- The turbulence intensity is highly fluctuating. (If the proportionality between the standard deviation of the wind speed fluctuation and the mean wind velocity is not satisfied, the analysis does not yield a constant turbulence intensity.)
- There is significance that in this period the increment distribution is symmetric (or slightly positively skewed) but not of a symmetric Castaing distribution and does not tend towards a Gaussian distribution for large increment lengths.

Both indicators signal that such a period is a sequence of shorter periods which can be regarded as being turbulent with changing relevant parameters or the wind activity in this period is not turbulent.

# Chapter 7

## Discussion of the Superstatistical Algorithm

This chapter discusses the strength and shortcomings of the superstatistical algorithm based on the idea of [Beck \*et al.\* \(2005b\)](#); [Queiros \(2007\)](#) when applying it to an arbitrary time series. The extent to which this algorithm is suitable to analyse ABL wind velocity recordings as done in [Sec. 6.4](#) is discussed.

### 7.1 The algorithm in a nutshell

The time series which is to be analysed with respect to superstatistics is assumed to consist of subsequent regions in which the observable is normally distributed with zero mean. These regions are assumed to be of constant length  $T$ . Denoting the time series by  $(x_n)_{n=1}^N$ , which is one realisation of the stochastic process  $\mathbf{X} = \{X_n : n = 1, \dots, N\}$ , the first step is to estimate the scale  $T \in \mathbb{N}$  such that

$$\underbrace{(x_1, \dots, x_T)}_{\sim N_{0, \beta_1^{-1}}}, \underbrace{(x_{T+1}, \dots, x_{2T})}_{\sim N_{0, \beta_2^{-1}}}, \dots, \underbrace{(x_{(K-1)T+1}, \dots, x_N)}_{\sim N_{0, \beta_K^{-1}}} \quad (7.1)$$

with  $K = N/T$  if  $N/T \in \mathbb{N}$  or  $K = \text{int}[N/T] + 1$ . According to [Queiros \(2007\)](#), it is estimated by taking the largest  $m \in \mathbb{N}$  with

$$\underbrace{(x_1, \dots, x_m)}_{\mathbb{Kurt}\{\dots\}=\kappa_1^{(m)}}, \underbrace{(x_{m+1}, \dots, x_{2m})}_{\mathbb{Kurt}\{\dots\}=\kappa_2^{(m)}}, \dots, \underbrace{(x_{(K-1)m+1}, \dots, x_N)}_{\mathbb{Kurt}\{\dots\}=\kappa_K^{(m)}}. \quad (7.2)$$

such that  $\kappa^{(m)} = \left\langle \kappa_k^{(m)} \right\rangle_{k=1, \dots, K}$  is within numerical errors equal to three.<sup>10</sup> Then, the  $\beta$ -series  $(\beta_k)_{k=1}^K$  can be estimated by

$$\beta_k = \frac{1}{\text{Var} \{x_{(k-1)T+1}, \dots, x_{\min\{kT, N\}}\}}. \quad (7.3)$$

The algorithm is extended to check whether there is statistical significance to reject the hypothesis

$$H_0 : \quad \{\ln \beta_1, \dots, \ln \beta_K\} \sim N_{\mu_{\ln \beta}, \sigma_{\ln \beta}^2} \quad (7.4)$$

with  $\mu_{\ln \beta} = \text{Mean} \{\ln \beta_1, \dots, \ln \beta_K\}$  and  $\sigma_{\ln \beta}^2 = \text{Var} \{\ln \beta_1, \dots, \ln \beta_K\}$  by using a Kolmogorov-Smirnov test. This extension is done because this algorithm is primarily used to analyse wind speed increments which hypothetically satisfy  $H_0$ .

## 7.2 Application to an ideal series

The applicability of the algorithm is demonstrated for an ideal series, i.e., a series which is constructed to satisfy (7.1) for  $T = \Theta \in \mathbb{N}$ . The process  $\mathbf{X} = \{X_n : n = 1, \dots, N\}$  with  $N = L\Theta$ , where  $L \in \mathbb{N}$ , is defined by

$$\begin{aligned} n = 1, \dots, \Theta : & \quad X_n = \xi_n / \sqrt{B_1} \\ n = \Theta + 1, \dots, 2\Theta : & \quad X_n = \xi_n / \sqrt{B_2} \\ n = 2\Theta + 1, \dots, 3\Theta : & \quad X_n = \xi_n / \sqrt{B_3} \\ & \quad \vdots \\ n = (L-1)\Theta + 1, \dots, L\Theta : & \quad X_n = \xi_n / \sqrt{B_L}, \end{aligned} \quad (7.5)$$

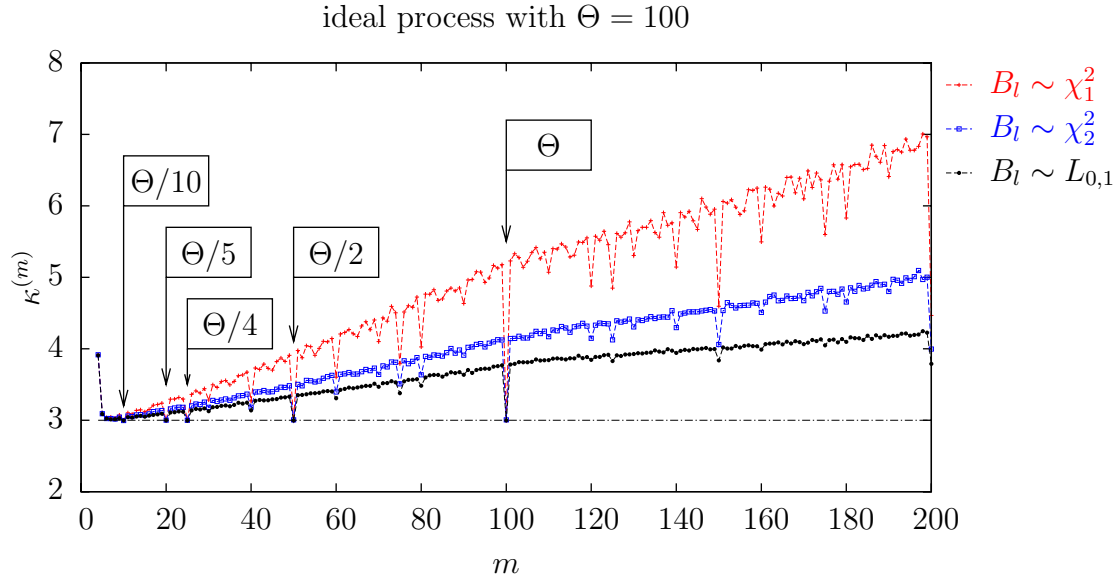
where  $\{\xi_n : n = 1, \dots, N\}$  denotes standard Gaussian white noise and  $\mathbf{B} = \{B_l : l = 1, \dots, L\}$  stands for a non-Gaussian white noise with  $B_l > 0$ .

Three different realisations of this process ( $B_l \sim L_{0,1}$ ,  $B_l \sim \chi_1^2$ , and  $B_l \sim \chi_2^2$ ) with  $\Theta = 100$  and  $N = 10^6$  are analysed with respect to superstatistics under the assumption that the true parameterisation of  $\mathbf{B}$  is unknown. Fig. 7.1 shows  $\kappa^{(m)}$  as a function of  $m$  for the process  $\mathbf{X}$ . It is evident that  $\kappa^{(m)} \approx 3$  for  $m = \Theta/j$  with  $j \in \mathbb{N}$  so that the estimation of  $T$  yields  $T = \Theta$ . The graphs also shows that  $\kappa^{(m)}$  for  $m = \Theta \pm 1$  is very different from three. Hence, it is not suitable to skip some  $m$  when estimating  $T$ .

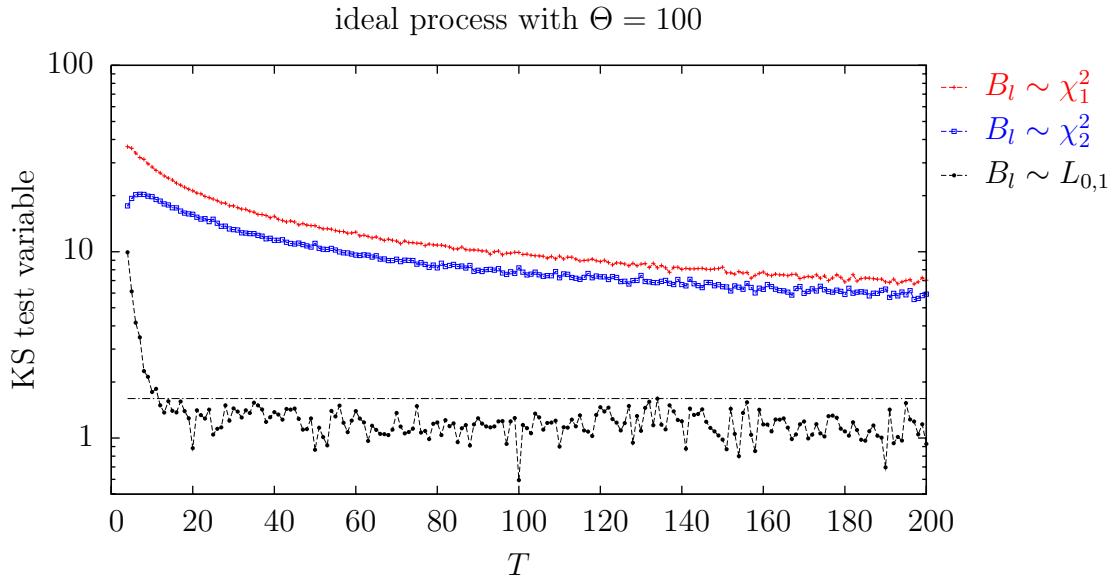
The process is designed such that for  $T = \Theta$ , the estimated  $\beta$ -series is a realisation of  $\mathbf{B}$ . If  $B_l \sim L_{0,1}$ , the series  $(\ln \beta_k)_{k=1}^K$  is normally distributed with zero mean and unit variance when  $T = \Theta$ . On the other hand, if  $B_l$  is not log-normally distributed ( $B_l \sim \chi_r^2$

---

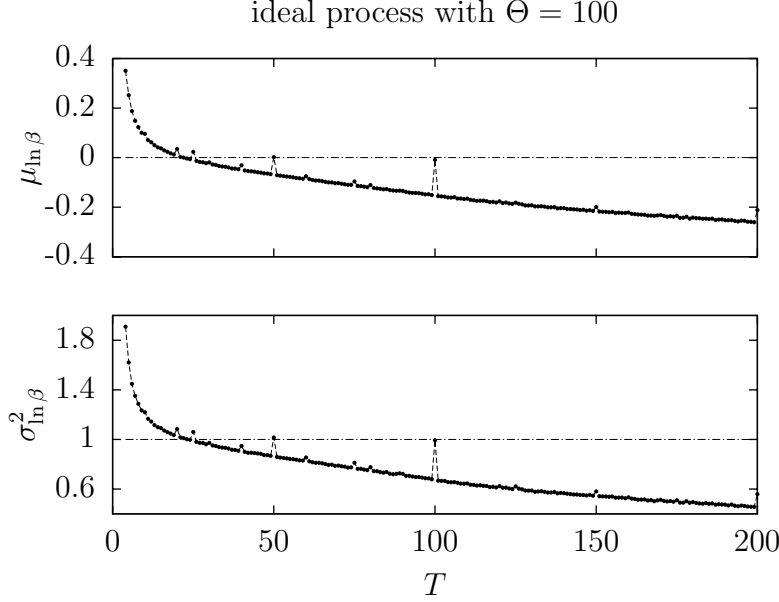
<sup>10</sup>Note that for ABL wind speed increments, there is a single  $m$  with  $\left\langle \kappa_k^{(m)} \right\rangle_{k=1, \dots, K} \approx 3$ . Hence, taking the maximum is redundant. Additionally, by physical arguments, it suffices to estimate the biased kurtosis when analysing ABL wind speed increments. For a general series, it is ‘‘safer’’ to formulate the algorithm using an unbiased kurtosis estimator.



**Figure 7.1:** Estimation of  $\kappa^{(m)}$  as a function of  $m \in \mathbb{N}$  for three ideal processes. The dashed line connects the data points for the clarity of presentation. The dash-dotted line corresponds to  $\kappa^{(m)} = 3$ .



**Figure 7.2:** Value of the Kolmogorov-Smirnov test variable as a function of  $T$ . The test checks whether there is significance for rejecting the hypothesis that  $\ln \beta$  is normally distributed. The dash-dotted line corresponds to the critical value for a significance level of 1%.



**Figure 7.3:** Estimation of  $\mu_{\ln\beta}$  and  $\sigma_{\ln\beta}^2$  as a function of  $T \in \mathbb{N}$ . The dashed line connects the data points for the clarity of presentation. The dash-dotted line reflects the true value of the corresponding parameter.

with  $r = 1, 2$  are chosen as examples), the series  $(\ln \beta_k)_{k=1}^K$  is not normally distributed at all. Regarding the robustness of the algorithm, it is interesting to analyse the extend of how much the statistics of the estimated  $\beta$ -series changes when  $T$  is not equal to  $\Theta$ . Therefore, in the following analysis the parameter  $T$  is not set to the value of the estimation using  $\kappa^{(m)}$ . Rather, it is varied between 1 and  $2\Theta$ .

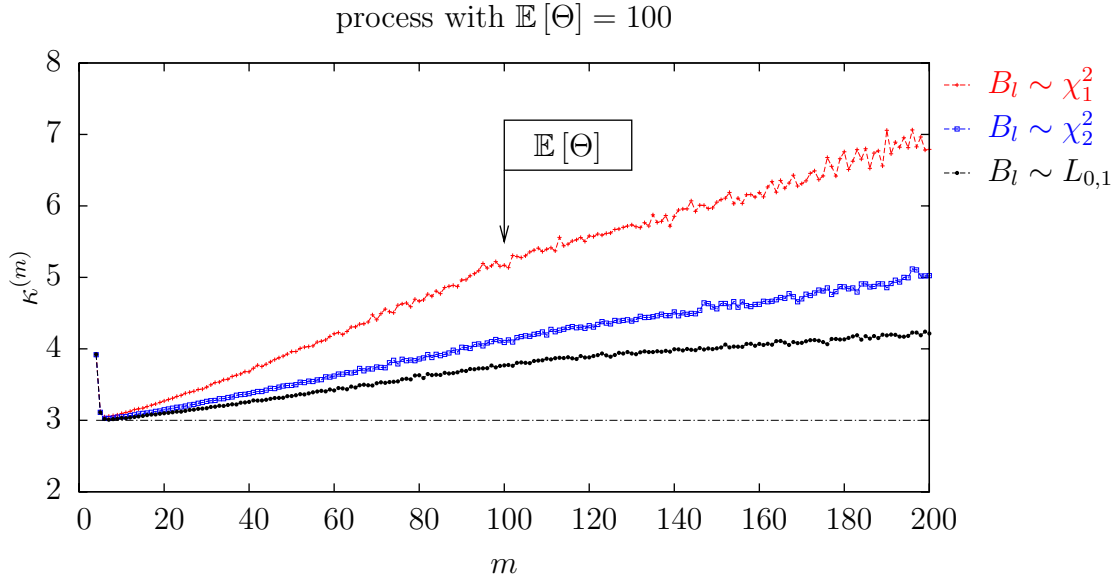
**Normality test for  $\ln \beta$ .** It is examined whether there is statistical significance to reject the hypothesis (7.4). Fig. 7.2 displays the value of the Kolmogorov-Smirnov test variable as a function of  $T$ . There is no evidence for rejecting  $H_0$  for nearly all  $T$  when  $B_l \sim L_{0,1}$ , while there is for all  $T$  when  $B_l$  is  $\chi^2$ -distributed. That is, the value of  $T$  does not have an impact on the decision whether  $H_0$  should be rejected or not: if  $H_0$  is (is not) rejected for  $T = \Theta$  is (is not) rejected for  $T \neq \Theta$ .

**Estimation of parameters.** The value of  $T$  is expected to have an impact on the estimated parameters  $\mu_{\ln\beta}$  and  $\sigma_{\ln\beta}^2$ , though. The process with log-normally distributed  $B_l$  is relevant for testing whether the Queiros (2007) algorithm is suitable to analyse wind speed increment because this process is of a symmetric Castaing distribution as well. Therefore, the estimated mean  $\mu_{\ln\beta}$  and variance  $\sigma_{\ln\beta}^2$  of the logarithm of the  $\beta$ -series obtained from the process with  $B_l \sim L_{0,1}$  are compared with 0 and 1, respectively. Fig. 7.3 shows  $\mu_{\ln\beta}$  and  $\sigma_{\ln\beta}^2$  as a function of  $T$ . It is obvious that these estimations are in agreement with the true values for  $T = \Theta$  and  $T = \Theta/2$ , but they are not for any other  $T$ .

As a conclusion, the value of the estimation of the time scale  $T$  does not have an impact on the shape of the  $\beta$ -distribution<sup>11</sup>, but on its parameterisation. In other words, in order to

<sup>11</sup>Better: on the decision if  $\ln \beta$  is of a Gaussian/non-Gaussian distribution.





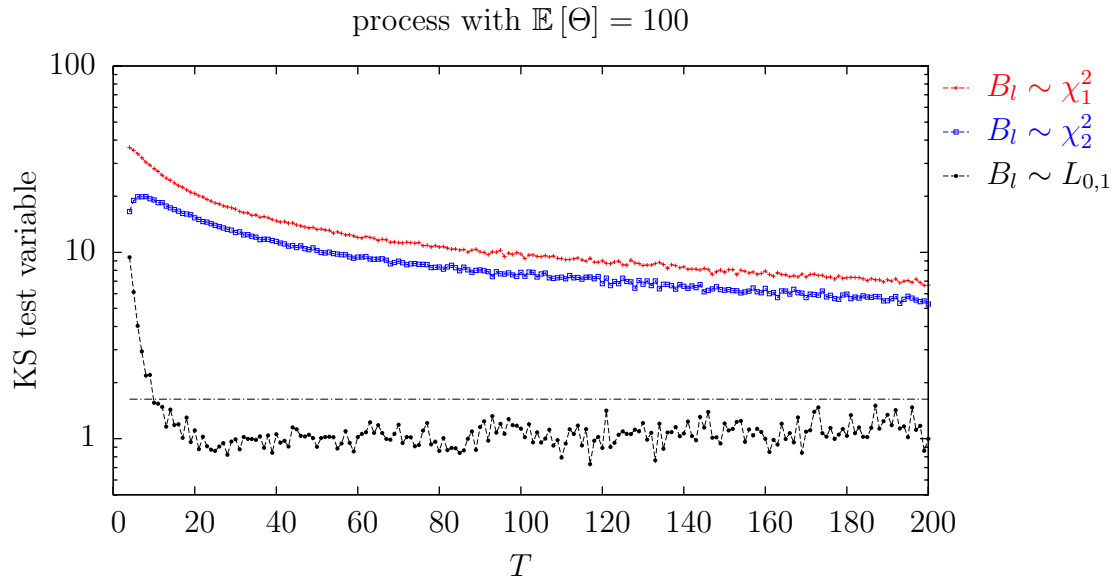
**Figure 7.4:** Estimation of  $\kappa^{(m)}$  as a function of  $m \in \mathbb{N}$  for three processes which are not ideal. The dashed line connects the data points for the clarity of presentation. The dash-dotted line corresponds to  $\kappa^{(m)} = 3$ .

find a reasonable estimation of the process parameters, the time scale  $T$  must be estimated correctly.

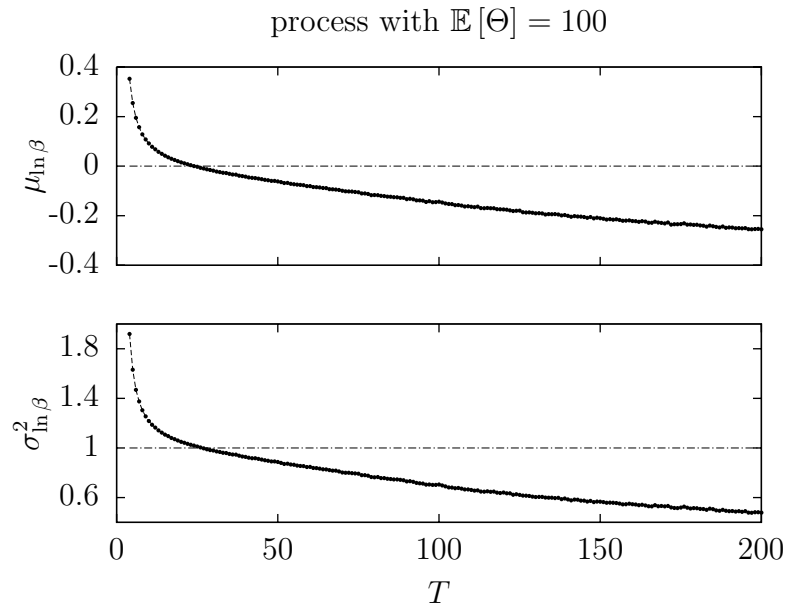
### 7.3 Application to a series which is not ideal

The suitability of the algorithm is analysed when it is applied to a process which does not satisfy the assumption that the length of the regions, in which the process is normally distributed, is constant. Consider  $\Theta = \{\Theta_l : l = 1, \dots, L\}$  to be a Poisson distributed white noise so that  $\Theta_l \in \mathbb{N}$ . The process  $\mathbf{X} = \{X_n : n = 1, \dots, N\}$  with  $N = \sum_{l=1}^L \Theta_l$  is defined by

$$\begin{aligned}
 n = 1, \dots, \Theta_1 : & \quad X_n = \xi_n / \sqrt{B_1} \\
 n = \Theta_1 + 1, \dots, \sum_{i=1}^2 \Theta_i : & \quad X_n = \xi_n / \sqrt{B_2} \\
 n = \sum_{i=1}^2 \Theta_i + 1, \dots, \sum_{i=1}^3 \Theta_i : & \quad X_n = \xi_n / \sqrt{B_3} \\
 & \quad \vdots \\
 n = \sum_{i=1}^{L-1} \Theta_i + 1, \dots, \sum_{i=1}^L \Theta_i : & \quad X_n = \xi_n / \sqrt{B_L},
 \end{aligned} \tag{7.6}$$



**Figure 7.5:** Value of the Kolmogorov-Smirnov test variable as a function of  $T$ . The test checks whether there is significance for rejecting the hypothesis that  $\ln \beta$  is normally distributed. The dash-dotted line corresponds to the critical value for a significance level of 1%.



**Figure 7.6:** Estimation of  $\mu_{\ln \beta}$  and  $\sigma_{\ln \beta}^2$  as a function of  $T \in \mathbb{N}$ . The dashed line connects the data points for the clarity of presentation. The dash-dotted line reflects the true value of the corresponding parameter.

where  $\{\xi_n : n = 1, \dots, N\}$  denotes standard Gaussian white noise and  $\mathbf{B} = \{B_l : l = 1, \dots, L\}$  stands for a non-Gaussian white noise with  $B_l > 0$ .

Fig. 7.4 indicates that the time scale estimation according to Queiros (2007) does not yield any result<sup>12</sup> because this process does not have a fixed scale  $T$ . Still, the statistics of the  $\beta$ -series can be estimated by choosing an arbitrary  $T$ . Again, the decision whether to reject  $H_0$  does not depend on  $T$ , see Fig. 7.5, but the estimation of the parameterisation of the  $\beta$ -distribution does, see Fig. 7.5.

## 7.4 Wind speed increments and conclusion

It is not expected that wind speed increments satisfy the assumption that the length of the regions, in which the increments are normally distributed, is constant, either. As concluded in the previous section, the superstatistical algorithm is reliable for testing whether the turbulence hypothesis is to be rejected or not, but the parameterisation of the  $\beta$ -statistics might be defective. However, the analysis in App. B shows that the parameterisation estimated by this algorithm is on good agreement with the parameterisation which is estimated by directly deriving the increment statistics according to Castaing *et al.* (1990).<sup>13</sup>

Concluding, the Queiros (2007) algorithm is suitable to analyse wind velocity increments, but it is not perfect. However, this imperfection does not have an influence on the decision whether a wind recording represents turbulent wind activity or not. It is merely the strength of intermittency which is sometimes slightly under- or overestimated.

---

<sup>12</sup>neglecting  $T \lesssim 10$

<sup>13</sup>Note that for wind speed increments in Sec. 6.4, the biased kurtosis is used to compute  $\kappa^{(m)}$ , which always yields a result for the time scale  $T$  due to physical reasons (intermittency).

# Chapter 8

## Fluctuation Statistics of Stochastic Processes and Wind Speed Modelling

In several applications, in particular in view of the cost efficient use of wind power, realistic models for the input wind fields are needed for numerical simulations of the flow around an obstacle. Solving the Navier-Stokes equation (4.12) by direct numerical simulation gives a realistic wind field, but for three-dimensional volumes of a realistic size the computational costs are immense. Large eddy simulation (see e.g. Frisch, 1995; Pope, 2000) is an approximated solution to the Navier-Stokes equations by not directly solving the motions of the small scales. Rather, they are modelled. Still, it requires enormous computational power. Farge *et al.* (1999) suggests to solve the Navier-Stokes equation by using wavelets (most coefficients are close to zero which reduces the computational cost). Yet, the numerics are very involved.

Another approach are stochastic models simulating turbulent wind fields, at least for the time period in which the turbulence parameters remain nearly constant. The thesis of Kleinhans (2008) discusses a broad variety of synthetic wind field generators. Notable are

- spectral models by Veers (1984); Mann (1998); Rossi *et al.* (2004) aiming to describe the process by its spectral density,
- wavelets models by Kitagawa & Nomura (2003) simulating the energy cascade by wavelets,
- multiplicative models by Cleve *et al.* (2005); Schmiegel *et al.* (2005) describing the energy dissipation rate by suitable random variables,
- Markov models by Nawroth (2007) simulating a turbulent flow by drift and diffusion processes, and
- models based on continuous time random walks by Kleinhans *et al.* (2006); Kleinhans & Friedrich (2007).

These models put their focus on a specific statistical or physical property of turbulence (cross-correlation, intermittency, etc.) It is very difficult (if at all possible) to consider all kinds of statistical properties and generate a series which perfectly reflects the atmospheric wind velocity.

Motivated by the statistical results described in Chap. 6, i.e.,

- the fluctuation of the wind velocity fluctuation is normally distributed, has zero expectation and there is a proportionality between the standard deviation of the fluctuation and the mean wind speed, and
- the wind velocity increment distribution is intermittent for small increment lengths,

this chapter discusses the usage of stochastic processes which shall have the same fluctuation statistics. That is, its fluctuation conditioned on the value of the moving average is normally distributed with zero mean and a standard deviation which is proportional to the value of the moving average. Additionally, the process shall have an intermittent increment distribution for small increment lengths.

Consequently, a method to compute the fluctuation distribution conditioned on the value of the moving average is formulated. The method is applied to several stochastic processes. It will be shown that the above mentioned fluctuation statistics not trivial. It is then discussed that under certain conditions the geometric AR(1) process<sup>14</sup> has similar fluctuation and increment statistics to atmospheric boundary layer wind speed. This similarity is investigated and the extent to which this stochastic process is a suitable model for wind speed simulation is discussed.

This chapter uses the convention that rv's are denoted by capital letters, whereas their values are denoted by small letters. Only  $N_{0,1}$ -distributed rv's are denoted by the small (Greek) letters, such as  $\xi$  and  $\eta$ . Additionally, if the rv  $X$  is  $f$ -distributed, it is written as  $X \sim f$ .

## 8.1 Conditioned fluctuation distribution

This section deals with the fluctuation statistics of a time discrete stochastic process  $\mathbf{X} = \{X_n : n \in \mathbb{N}\}$  with  $X_n \sim p_n$ . The goal is to compute the distribution of the fluctuation

$$F_n^{(m)} = X_n - \bar{X}_n^{(m)} \quad (8.1)$$

conditioned on  $\bar{X}_n^{(m)} = \bar{x}_n \in \mathbb{R}$ . The moving average over  $m = 2\tilde{m} + 1$  time steps with  $\tilde{m} \in \mathbb{N}_0$  is given by

$$\bar{X}_n^{(m)} = \frac{1}{m} \sum_{k=-\tilde{m}}^{\tilde{m}} X_{n+k}. \quad (8.2)$$

The  $F_n^{(m)}$ -distribution conditioned on  $\bar{X}_n^{(m)} = \bar{x}_n$  can be computed in three steps:

---

<sup>14</sup>A discretised geometric Ornstein-Uhlenbeck process introduced in Ex. 3.7.

- Eq. (8.2) is used to get an expression for the distribution of  $\bar{X}_n^{(m)}$  and  $\bar{X}_n^{(m)}$  conditioned on  $X_n = x_n$  by taking advantage of (8.2) reflecting a sum of rv's so that the transformation rule (2.60) can be employed. The two distributions are denoted by  $\bar{p}_n^{(m)}(\bar{x}_n)$  and  $\bar{p}_n^{(m)}(\bar{x}_n|X_n = x_n)$ , respectively.
- The relation (2.39) for conditioned distributions is used to obtain the distribution of  $X_n$  conditioned on  $\bar{X}_n^{(m)} = \bar{x}_n$ , i.e.,

$$p_n(x_n|\bar{X}_n^{(m)} = \bar{x}_n) = \frac{\bar{p}_n^{(m)}(\bar{x}_n|X_n = x_n) p_n(x_n)}{\bar{p}_n^{(m)}(\bar{x}_n)}. \quad (8.3)$$

- The fluctuation  $F_n^{(m)}$  conditioned on  $\bar{X}_n^{(m)} = \bar{x}_n$  is simply a linear transformation of  $X_n$  conditioned on  $\bar{X}_n^{(m)} = \bar{x}_n$  because

$$F_n^{(m)} \Big|_{\bar{X}_n^{(m)} = \bar{x}_n} = X_n \Big|_{\bar{X}_n^{(m)} = \bar{x}_n} - \bar{x}_n. \quad (8.4)$$

Employing the linear transformation rule (2.21) results in the  $F_n^{(m)}$ -distribution conditioned on  $\bar{X}_n^{(m)} = \bar{x}_n$ .

If the expression (8.3) is known, the mean and variance of the fluctuation conditioned on  $\bar{X}_n^{(m)} = \bar{x}_n$  can be obtained directly via

$$\mathbb{E} [F_n^{(m)}|\bar{X}_n^{(m)} = \bar{x}_n] = \mathbb{E} [X_n|\bar{X}_n^{(m)} = \bar{x}_n] - \bar{x}_n \quad (8.5a)$$

and

$$\mathbb{V}\text{ar} [F_n^{(m)}|\bar{X}_n^{(m)} = \bar{x}_n] = \mathbb{V}\text{ar} [X_n|\bar{X}_n^{(m)} = \bar{x}_n]. \quad (8.5b)$$

## 8.2 Stationary Gaussian processes

Consider a stationary Gaussian stochastic processes  $\mathbf{X} = \{X_n : n \in \mathbb{Z}\}$  with  $X_n \sim N_{\mu, \sigma^2}$  ( $n \in \mathbb{Z}$ ) and auto correlation function  $\gamma(k)$  ( $k \in \mathbb{Z}$ ).<sup>15</sup> The moving average (8.2) is again normally distributed with mean  $\mu$  and variance  $\sigma^2\theta_m^2$ , where

$$\theta_m^2 = \frac{1}{m^2} \sum_{k, k' = -\tilde{m}}^{\tilde{m}} \gamma(k - k'), \quad (8.6)$$

and  $m = 2\tilde{m} + 1$ . Regarding the distribution of  $\bar{X}_n^{(m)}$  conditioned on  $X_n = x_n$ , it is always possible to find a  $\phi_m \in \mathbb{R}$  such that the rv  $Z_n^{(m)} = \bar{X}_n^{(m)} - \phi_m X_n$  is uncorrelated to  $X_n$ . That is,

$$\mathbb{C}\text{ov} [X_n, Z_n^{(m)}] = \mathbb{C}\text{ov} [X_n, \bar{X}_n^{(m)}] - \phi_m \mathbb{C}\text{ov} [X_n, X_n] = 0 \quad (8.7)$$

---

<sup>15</sup>This discussion can be extended to non-stationary Gaussian processes by using the mean and covariance function,  $\mu(n)$  and  $R(n, m)$  ( $n, m \in \mathbb{Z}$ ), respectively, instead of mean  $\mu$ , variance  $\sigma^2$ , and auto correlation function  $\gamma(k)$ .

which leads to

$$\phi_m = \frac{1}{\sigma^2} \text{Cov} [X_n, \bar{X}_n^{(m)}] = \frac{1}{m} \sum_{k=-\tilde{m}}^{\tilde{m}} \gamma(k). \quad (8.8)$$

The rv  $Z_n^{(m)} = \bar{X}_n^{(m)} - \phi_m X_n$  is obviously normally distributed with mean  $\mu(1 - \phi_m)$  and variance  $\sigma^2(\theta_m^2 - \phi_m^2)$  because  $\text{Var} [\bar{X}_n^{(m)}] = \phi_m^2 \text{Var} [X_n] + \text{Var} [Z_n^{(m)}]$ . Consequently,

$$\bar{X}_n^{(m)} \Big|_{X_n=x_n} \sim N_{\phi_m x_n + \mu(1-\phi_m), \sigma^2(\theta_m^2 - \phi_m^2)}. \quad (8.9)$$

It is straightforward to derive the distribution of the fluctuation  $F_n^{(m)}$  conditioned on  $\bar{X}_n^{(m)} = \bar{x}_n$  by (8.3) and (8.4), see App. C.1. It is a normal distribution with

$$\mathbb{E} [F_n^{(m)} | \bar{X}_n^{(m)} = \bar{x}_n] = (\bar{x}_n - \mu) \frac{\phi_m - \theta_m^2}{\theta_m^2} \quad (8.10a)$$

and

$$\text{Var} [F_n^{(m)} | \bar{X}_n^{(m)} = \bar{x}_n] = \sigma^2 \frac{\theta_m^2 - \phi_m^2}{\theta_m^2}. \quad (8.10b)$$

That is, the mean of the fluctuation is proportional to  $\bar{x}_n - \mu$  unless  $\phi_m = \theta_m^2$ , whereas the variance of the fluctuation is always independent of  $\bar{x}_n$ .

Two examples of this process family are discussed: the Gaussian white noise and the stationary AR(1) process.

## Gaussian white noise

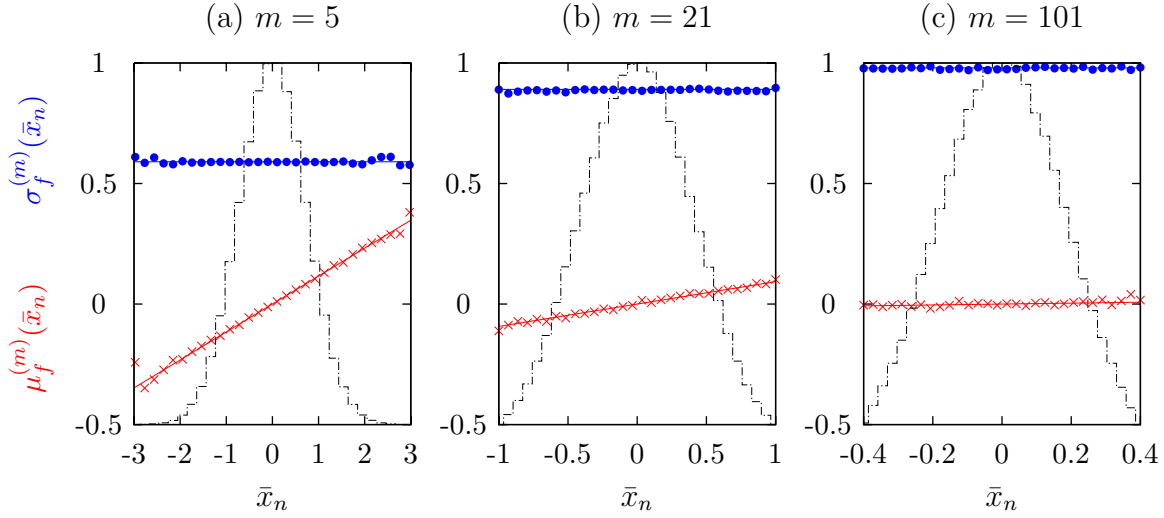
Consider  $\mathbf{X} = \{X_n : n \in \mathbb{Z}\}$  to be Gaussian white noise with mutually independent  $X_n \sim N_{\mu, \sigma^2}$ . The auto correlation function is given by  $\gamma(k) = \delta_{k,0}$  so that

$$\phi_m = \frac{1}{m} \quad \text{and} \quad \theta_m^2 = \frac{1}{m} \quad (8.11)$$

by using (8.8) and (8.6), respectively. As a consequence of (8.10), the fluctuation conditioned on  $\bar{X}_n^{(m)} = \bar{x}_n$  is normally distributed with

$$\mathbb{E} [F_n^{(m)} | \bar{X}_n^{(m)} = \bar{x}_n] = 0 \quad \text{and} \quad \text{Var} [F_n^{(m)} | \bar{X}_n^{(m)} = \bar{x}_n] = \frac{m-1}{m} \sigma^2. \quad (8.12)$$

Hence, its mean and variance are independent of  $\bar{x}_n$  and for large  $m$  the variance of the fluctuation approaches the variance  $\sigma^2$  of the whole series.



**Figure 8.1:** Estimated mean (red, crosses) and standard deviation (blue, dots) of the fluctuation conditioned on  $\bar{X}_n^{(m)} = \bar{x}_n$  with  $m = 5, 21, 101$  for a generated AR(1) series with  $a = 0.6$  and  $\sigma^2 = 1$ . The histogram in dash-dot line corresponds to an estimation of  $\bar{p}_n^{(m)}(\bar{x}_n)$  and is shown in arbitrary units. It indicates the reliability of the estimation of  $\mu_f^{(m)}(\bar{x}_n)$  and  $\sigma_f^{(m)}(\bar{x}_n)$ : the larger  $\bar{p}_n^{(m)}(\bar{x}_n)$  is, the more reliable is the estimation for  $\mu_f^{(m)}(\bar{x}_n)$  and  $\sigma_f^{(m)}(\bar{x}_n)$ .

## Stationary AR(1) process

Consider  $\mathbf{X} = \{X_n : n \in \mathbb{Z}\}$  to be a stationary AR(1) process (3.35) with

$$X_{n+1} = aX_n + \sigma\xi_n\sqrt{1-a^2}, \quad (8.13)$$

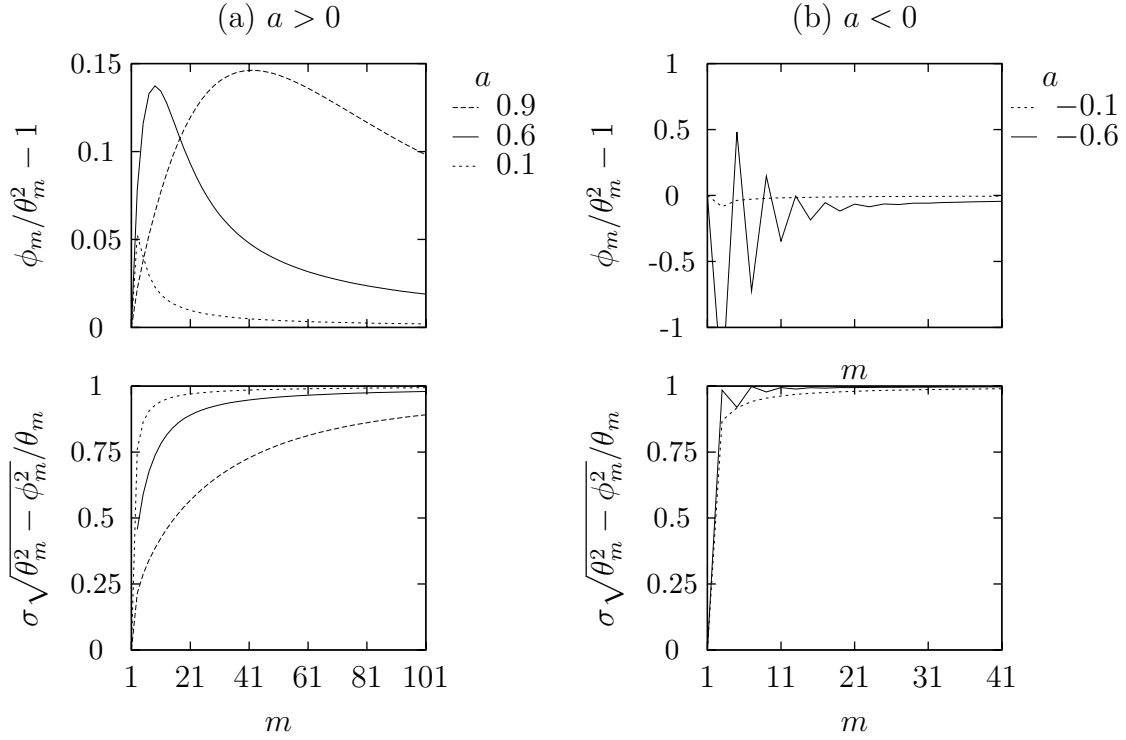
where  $a \in \mathbb{R}$  with  $|a| < 1$ ,  $\sigma \in \mathbb{R}_{>0}$ , and  $\{\xi_n : n \in \mathbb{Z}\}$  denoting standard Gaussian white noise. Thus,  $X_n$  is  $N_{0,\sigma^2}$ -distributed. Its auto correlation function is given by  $\gamma(k) = a^{|k|}$  so that

$$\phi_m = \frac{1+a}{(1-a)m} - 2a\frac{a^{(m-1)/2}}{(1-a)m} \quad \text{and} \quad \theta_m^2 = \frac{1+a}{(1-a)m} - 2a\frac{1-a^m}{(1-a)^2m^2} \quad (8.14)$$

by using (8.8) and (8.6), respectively. Inserting  $\phi_m$  and  $\theta_m^2$  into (8.10) yields the expression for the mean  $\mu_f^{(m)}(\bar{x}_n)$  and standard deviation  $\sigma_f^{(m)}(\bar{x}_n)$  of the fluctuation  $F_n^{(m)}$  conditioned on  $\bar{X}_n^{(m)} = \bar{x}_n$ .

Due to  $\mu = 0$ , the AR(1) process has  $\mu_f^{(m)}(\bar{x}_n) \propto \bar{x}$ . It is easily shown that in the limit  $m \rightarrow \infty$  the proportionality factor  $\phi_m/\theta_m^2 - 1$  becomes zero. Hence, the influence of  $\bar{x}$  vanishes with very large  $m$  causing the fluctuation statistics of an AR(1) process be similar to those of a Gaussian white noise. This can be understood by noting that the auto correlation of an AR(1) process decays exponentially, see Ex. 3.13. That is, if  $m$  is very large, the influence of auto correlation on the fluctuation statistics becomes very small making the fluctuations statistics of the process similar to the one of a  $\delta$ -correlated process.





**Figure 8.2:** The upper row shows the  $m$ -dependence of  $\phi_m/\theta_m^2 - 1$ , which corresponds to the proportionality factor between  $\mu_f^{(m)}(\bar{x}_n)$  and  $\bar{x}_n$  for generated AR(1) series with  $\sigma^2 = 1$ . The lower row depicts the  $m$ -dependence of  $\sigma\sqrt{\theta_m^2 - \phi_m^2}/\theta_m$ , which corresponds to  $\sigma_f^{(m)}(\bar{x}_n)$  as a function of  $m$ . **(a)**  $a \in \{0.1, 0.6, 0.9\}$ . **(b)**  $a \in \{-0.1, -0.6\}$ .

It is checked whether the estimation method which is used to analyse the wind speed fluctuation in Chap. 6 recovers the properties of the fluctuation when applying it to a generated AR(1) series. The fluctuation mean  $\mu_f^{(m)}(\bar{x}_n)$  and variance  $\sigma_f^{(m)}(\bar{x}_n)$  is estimated by the mean and variance of the set  $\{x_n - \bar{x}_n^{(m)}\}$  with  $n$  such that  $|\bar{x}_n^{(m)} - \bar{x}_n| \leq \Delta\bar{x}_n/2$  for reasonably small bin width  $\Delta\bar{x}_n$ . Exemplarily, an AR(1) series is generated ( $10^6$  data points) with  $a = 0.6$  and  $\sigma^2 = 1$ . The quantities  $\mu_f^{(m)}(\bar{x}_n)$  and  $\sigma_f^{(m)}(\bar{x}_n)$  are estimated as a function of  $\bar{x}_n$  for  $m \in \{5, 21, 101\}$ . Fig. 8.1 shows that this simple estimation yields agreement with  $\mu_f^{(m)}(\bar{x}_n) \propto \bar{x}_n$  and  $\sigma_f^{(m)}(\bar{x}_n) = \text{const.}$  Note that with increasing  $m$ , the value of the moving average becomes more peaked around zero in order that the range of  $\bar{x}_n$ , over which  $\mu_f^{(m)}(\bar{x}_n)$  and  $\sigma_f^{(m)}(\bar{x}_n)$  can be estimated, becomes smaller. Additionally, the proportionality factor between  $\mu_f^{(m)}(\bar{x}_n)$  and  $\bar{x}_n$  becomes smaller as  $m$  gets larger. It is given by  $\phi_m/\theta_m^2 - 1$  and depicted in Fig. 8.2 (a). For very small and reasonable large  $m$  it tends to zero and has a maximum in between. In contrast, the standard deviation  $\sigma_f^{(m)}(\bar{x}_n)$  increases with increasing  $m$  and tends to  $\sigma = 1$ , see the lower panel of Fig. 8.2 (a). The column (b) in Fig. 8.2 shows  $\phi_m/\theta_m^2 - 1$  and  $\sigma\sqrt{\theta_m^2 - \phi_m^2}/\theta_m$  for an AR(1) process with negative  $a$ . It can be seen that both quantities are oscillating and reach 0 and 1, respectively, as  $m$  becomes very large.

### 8.3 $\chi^2$ -distributed white noise

As explained in the previous section, the fluctuation of Gaussian white noise conditioned on  $\bar{X}_n^{(m)} = \bar{x}_n$  has vanishing expectation and its standard deviation is independent of  $\bar{x}_n$ . App. C.1 reveals that the latter is a consequences of the properties of the normal distribution. Being an example of a non-Gaussian process, the  $\chi^2$ -distributed white noise has a different behaviour. App. C.2 derives that the fluctuation conditioned on  $\bar{X}_n^{(m)} = \bar{x}_n$  has also vanishing mean, but its standard deviation is proportional to  $\bar{x}_n$ . The proportionality factor is given by

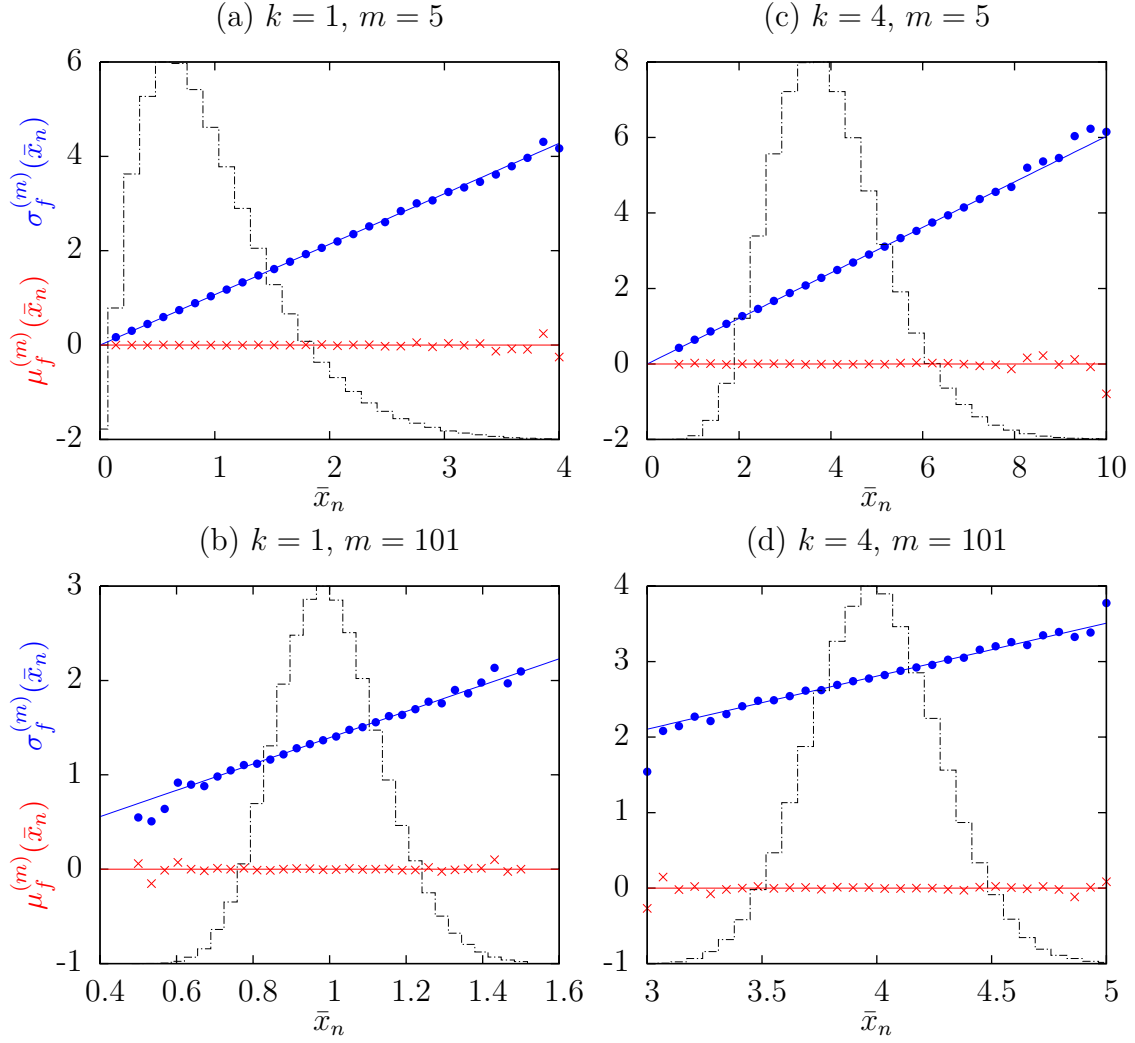
$$\alpha(m, k) = \sqrt{\frac{2(m-1)}{mk+2}} \quad (8.15)$$

which for large  $m$  tends to  $\sqrt{2/k}$ .

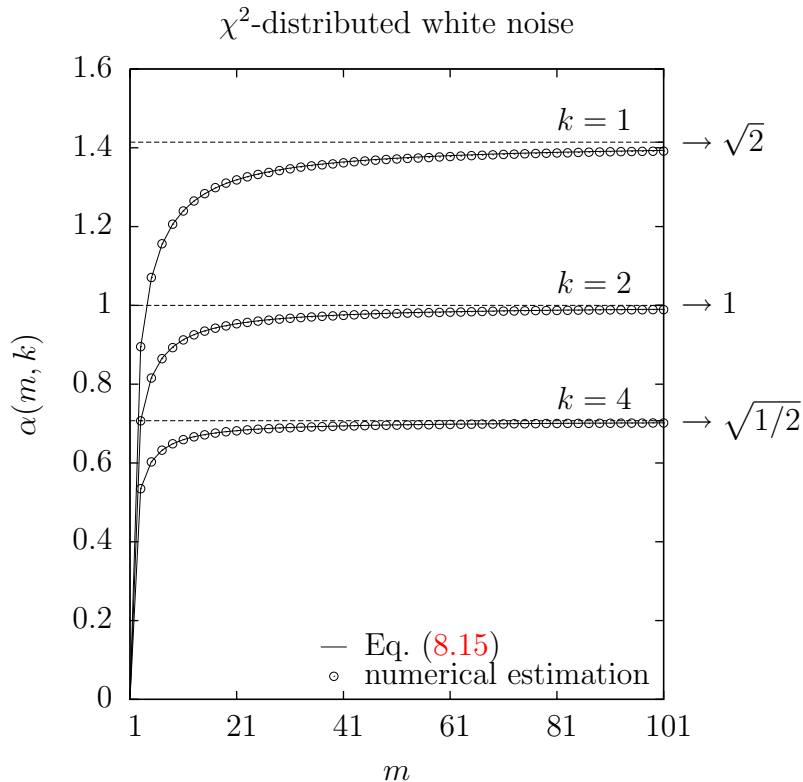
In the same spirit as for the stationary AR(1) process, the result is visualised for three generated  $\chi^2$  white noise series ( $10^6$  data points each) with  $k \in \{1, 2, 4\}$  by simple estimation. Exemplarily,  $\mu_f^{(m)}(\bar{x}_n)$  and  $\sigma_f^{(m)}(\bar{x}_n)$  are estimated as a function of  $\bar{x}_n$  for  $m \in \{5, 101\}$ . Fig. 8.3 shows agreement with  $\mu_f^{(m)}(\bar{x}_n) = 0$  (red, crosses) and  $\sigma_f^{(m)}(\bar{x}_n) \propto \bar{x}_n$  (blue, dots).

Fig. 8.4 depicts the proportionality factor between  $\sigma_f^{(m)}(\bar{x}_n)$  and  $\bar{x}_n$  as a function of  $m$ . The figure verifies agreement between the numerical estimation (the estimated standard deviation of  $F_n^{(m)}/\bar{X}_n^{(m)}$ ) and the analytical solution  $\alpha(m, k)$ . Additionally, it is evident that  $\alpha(m, k)$  approaches  $\sqrt{2/k}$  as  $m \rightarrow \infty$ .

In principle, this process could be considered as a process to simulate wind speed because it has the desired properties of vanishing  $\mu_f^{(m)}(\bar{x}_n)$  and  $\sigma_f^{(m)}(\bar{x}_n) \propto \bar{x}_n$ . If the number  $k$  of



**Figure 8.3:** Estimated mean (red, crosses) and standard deviation (blue, dots) of the fluctuation conditioned on  $\bar{X}_n^{(m)} = \bar{x}_n$  with  $m = 5, 101$  for generated  $\chi^2$ -distributed white noise series with  $k \in \{1, 4\}$ . The histogram in dash-dot line corresponds to an estimation of  $\bar{p}_n^{(m)}(\bar{x}_n)$  and is shown in arbitrary units. It indicates the reliability of the estimation of  $\mu_f^{(m)}(\bar{x}_n)$  and  $\sigma_f^{(m)}(\bar{x}_n)$ : the larger  $\bar{p}_n^{(m)}(\bar{x}_n)$  is, the more reliable is the estimation for  $\mu_f^{(m)}(\bar{x}_n)$  and  $\sigma_f^{(m)}(\bar{x}_n)$ .



**Figure 8.4:** The  $m$ -dependence of  $\alpha(m, k)$ , which corresponds to the proportionality factor between  $\sigma_f^{(m)}(\bar{x}_n)$  and  $\bar{x}_n$  for generated  $\chi^2$ -distributed white noise series with  $k \in \{1, 2, 4\}$ .

degrees of freedom is chosen to be large, the proportionality factor, which reflects the TI, becomes low. However, the fluctuation is not normally distributed as shown in App. C.2. Hence, another non-Gaussian process needs to be discussed.

## 8.4 Geometric AR(1) process

Recall that the process shall have the property that the fluctuation  $F_n^{(m)}$  conditioned on  $\bar{X}_n^{(m)} = \bar{x}_n$  is normally distributed with vanishing mean and standard deviation  $\propto \bar{x}_n$ . A first approach is to consider the process  $\mathbf{X} = \{X_n : n \in \mathbb{Z}\}$  with

$$X_{n+1} = X_n(1 + b\xi_n). \quad (8.16)$$

It has the following properties:

- $\{\xi_n : n \in \mathbb{Z}\}$  denotes standard Gaussian white noise.
- The noise strength  $b$  is so small that the fluctuation  $X_n - \bar{X}_n^{(m)}$  conditioned on  $\bar{X}_n^{(m)} = \bar{x}_n$  for given  $m = 2\tilde{m} + 1$  with  $\tilde{m} \in \mathbb{N}_0$  is roughly normal distributed due to the normality of the  $\xi_n$ 's.
- The rv  $X_n$  is not necessarily dimensionless. Rather, its unit is labeled by  $u$ .

The so-defined process is not stationary and therefore unstable for a numerical generation. A slightly modification of (8.16), namely

$$X_{n+1} = u \left( \frac{X_n}{u} \right)^a e^{b\xi_n}, \quad (8.17)$$

which for  $a = 1 - \epsilon$  with  $\epsilon \ll 1$  and  $b \ll 1$  is similar<sup>16</sup> to (8.16), is stationary for  $|a| < 1$  because its logarithm

$$\ln \frac{X_{n+1}}{u} = a \ln \frac{X_n}{u} + b\xi_n \quad (8.18)$$

corresponds to a stationary AR(1) process with variance  $\lambda^2 = b^2/(1-a^2)$ . Fixing  $m = 2\tilde{m} + 1$  and choosing the parameters  $a = 1 - \epsilon$  and  $b$  such that

$$\epsilon \ll 1/\tilde{m} \quad \text{and} \quad b^2 \ll 1/\tilde{m}, \quad (8.19)$$

it will be shown that (8.17) satisfies approximately

$$F_n^{(m)} \Big|_{\bar{X}_n^{(m)} = \bar{x}_n} \sim N_{0, \sigma_f^{(m)}(\bar{x}_n)^2} \quad (8.20)$$

with  $\sigma_f^{(m)}(\bar{x}_n) \propto \bar{x}_n$ .

Eq. (8.17) defines the *geometric AR(1) process* which corresponds to a discretised geometric Ornstein-Uhlenbeck process introduced in Ex. 3.7. Therefore,  $X_n$  is log-normally distributed with  $X_n \sim L_{u, \lambda^2}$  and  $\lambda^2 = b^2/(1-a^2) \approx b^2/(2\epsilon)$ . It should be noted that  $\lambda^2$  is not necessarily much smaller than unity because it is the quotient between two very small numbers.

Using symmetry property (C.33) of the stationary AR(1) process, the geometric AR(1) process has

$$\frac{X_{n\pm k}}{u} = \left( \frac{X_n}{u} \right)^{a^k} \exp \left[ b \sum_{l=1}^k a^{k-l} \eta_{n;\pm l} \right] \quad (8.21)$$

with mutually independent standard normally distributed  $\eta_{n;\pm 1}, \dots, \eta_{n;\pm k}$ . For  $1 \leq k \leq \tilde{m}$  the expression can be approximated<sup>17</sup> by

$$X_{n\pm k} \approx X_n \left( 1 + b \sum_{l=1}^k \eta_{n;\pm l} \right) \quad (8.22)$$

due to  $a^k \approx 1 - \epsilon k \approx 1$  and  $\text{Var} \left[ b \sum_{l=1}^k \eta_{n;\pm l} \right] = b^2 k \ll 1$ . This approximation is a generalisation of (8.16) and only holds true for  $k \leq \tilde{m}$ . Consequently, the moving average is

<sup>16</sup>Similar means in this context the following. The rv  $X_{n+1}$  conditioned on  $X_n = x_n$  is log-normally distributed with position parameter  $u(x_n/u)^a \approx x_n$  and shape parameter  $b^2$ . Due to  $\text{Var} [b\xi] = b^2 \ll 1$ , the distribution of  $e^{b\xi}$  is nearly identical to the distribution of  $1 + b\xi$  when neglecting rare tail events. Consequently,  $X_n$  conditioned on  $x_n$  is in good approximation normally distributed with mean  $x_n$  and variance  $x_n b^2$ .

<sup>17</sup>The distribution of  $e^{b\eta_{n;\pm l}}$  is approximated by the distribution of  $1 + b\eta_{n;\pm l}$  by neglecting rare tail events.

approximately

$$\bar{X}_n^{(m)} \approx X_n \left[ 1 + \frac{b}{m} \sum_{k=1}^{\tilde{m}} \sum_{l=1}^k (\eta_{m;+l} + \eta_{m;-l}) \right] = X_n [1 + Z_n^{(m)}] \quad (8.23)$$

with the dimensionless rv

$$Z_n^{(m)} = \frac{b}{m} \sum_{l=1}^{\tilde{m}} (\tilde{m} - l + 1) (\eta_{m;+l} + \eta_{m;-l}) \quad (8.24)$$

which is independent of  $X_n$ . The rv  $Z_n^{(m)}$  is normally distributed with zero mean and variance

$$\sigma_Z^2 = \mathbb{V}\text{ar} [Z_n^{(m)}] = \frac{b^2(m^2 - 1)}{12m} < \frac{b^2 m}{12} \approx \frac{b^2 \tilde{m}}{6} \ll 1. \quad (8.25)$$

Therefore, the moving average can be approximated by

$$\bar{X}_n^{(m)} \approx X_n \exp [Z_n^{(m)}] \quad (8.26)$$

and corresponds to a product of two independent log-normally distributed rv's. Due to  $X_n \sim L_{u,\lambda^2}$  and  $Z \sim N_{0,\sigma_Z^2}$ , the moving average satisfies roughly

$$\bar{X}_n^{(m)} \sim L_{u,\lambda^2+\sigma_Z^2} \quad \text{and} \quad \bar{X}_n^{(m)} \Big|_{X_n=x_n} \sim L_{x_n,\sigma_Z^2}. \quad (8.27)$$

Inserting these results into (8.3) yields the distribution of  $X_n$  conditioned on  $\bar{X}_n^{(m)} = \bar{x}_n$ :

$$p_n(x_n | \bar{X}_n^{(m)} = \bar{x}) = \frac{L_{x_n,\sigma_Z^2}(\bar{x}_n) L_{u,\lambda^2}(x_n)}{L_{u,\lambda^2+\sigma_Z^2}(\bar{x}_n)}. \quad (8.28)$$

After some manipulations, which are written in App. C.3, this becomes

$$X_n \Big|_{\bar{X}_n^{(m)}=\bar{x}_n} \sim L_{u_{\text{cond}},\lambda_{\text{cond}}^2} \quad (8.29)$$

with position parameter

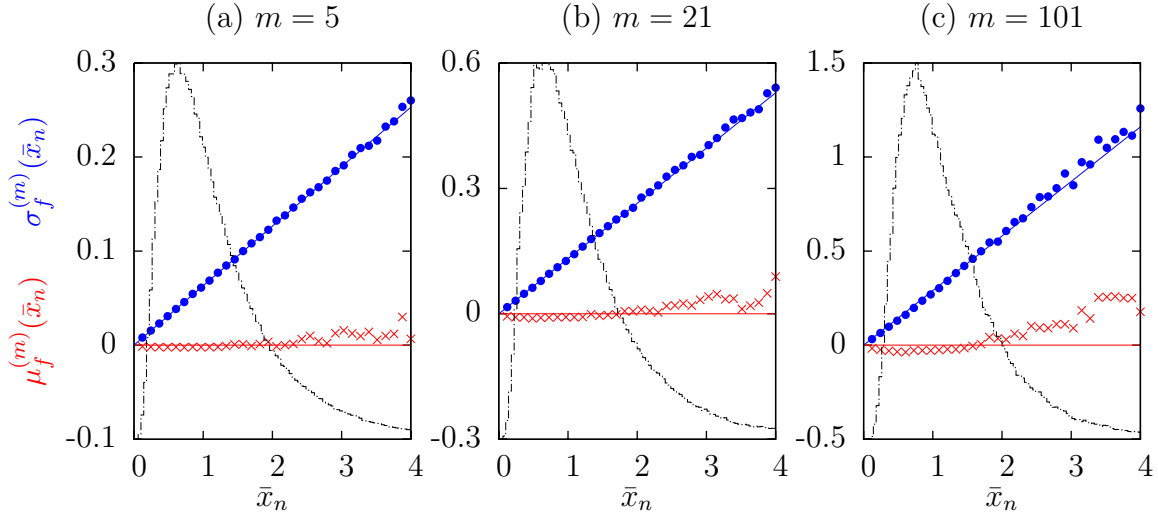
$$u_{\text{cond}} = u \left( \frac{\bar{x}_n}{u} \right)^{\lambda^2/(\lambda^2+\sigma_Z^2)} \quad (8.30)$$

and shape parameter

$$\lambda_{\text{cond}}^2 = \frac{\sigma_Z^2 \lambda^2}{\lambda^2 + \sigma_Z^2}. \quad (8.31)$$

Note that  $\sigma_Z^2 = b^2(m^2 - 1)/(12m)$  is always much smaller than  $\lambda^2 = b^2/(1 - a^2)$  because

$$\frac{\sigma_Z^2}{\lambda^2} = \frac{(m^2 - 1)(1 - a^2)}{12m} \approx \epsilon \frac{m^2 - 1}{6m} < \frac{\epsilon m}{6} \approx \frac{\epsilon \tilde{m}}{3} \ll 1. \quad (8.32)$$



**Figure 8.5:** Estimated mean (red, crosses) and standard deviation (blue, dots) of the fluctuation conditioned on  $\bar{X}_n^{(m)} = \bar{x}_n$  with  $m = 5, 21, 101$  for a generated geometric AR(1) series with  $a = 0.99$  and  $b = 0.1$ . The histogram in dash-dot line corresponds to an estimation of  $\bar{p}_n^{(m)}(\bar{x}_n)$  and is shown in arbitrary units. It indicates the reliability of the estimation of  $\mu_f^{(m)}(\bar{x}_n)$  and  $\sigma_f^{(m)}(\bar{x}_n)$ : the larger  $\bar{p}_n^{(m)}(\bar{x}_n)$  is, the more reliable is the estimation for  $\mu_f^{(m)}(\bar{x}_n)$  and  $\sigma_f^{(m)}(\bar{x}_n)$ .

Hence,  $X_n$  conditioned in  $\bar{X}_n^{(m)} = \bar{x}_n$  is in good approximation  $L_{\bar{x}_n, \sigma_Z^2}$ -distributed. As a consequence of  $\sigma_Z^2 \ll 1$ , the log-normal distribution corresponds roughly to a Gaussian distribution so that almost

$$X_n \Big|_{\bar{X}_n^{(m)} = \bar{x}_n} \sim N_{\bar{x}_n, \bar{x}_n^2 \sigma_Z^2}. \quad (8.33)$$

The fluctuation  $F_n^{(m)}$  conditioned on  $\bar{X}_n^{(m)} = \bar{x}_n$  is therefore close to normally distributed with

$$F_n^{(m)} \Big|_{\bar{X}_n^{(m)} = \bar{x}_n} \sim N_{0, \bar{x}_n^2 \sigma_Z^2}. \quad (8.34)$$

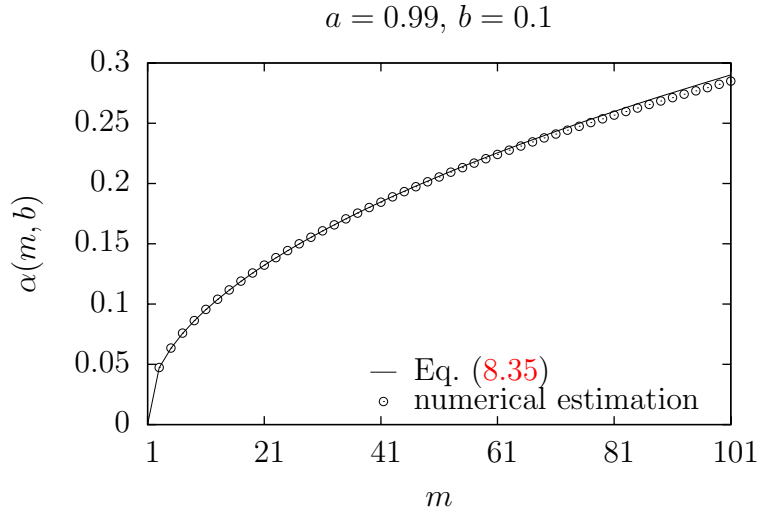
It has vanishing mean and there is a proportionality between the standard deviation and  $\bar{x}_n$ , i.e.,  $\sigma_f^{(m)}(\bar{x}_n) \propto \bar{x}_n$  with proportionality factor

$$\alpha(m, b) = \frac{b}{2} \sqrt{\frac{m^2 - 1}{3m}} \quad (8.35)$$

which grows approximately with  $\sqrt{m}$ . Of course, the proportionality is only justified for  $\tilde{m} \ll 1/b^2$  and  $\tilde{m} \ll 1/\epsilon$  (recall:  $m = 2\tilde{m} + 1$ ,  $\epsilon = 1 - a$ ).

Due to the made approximations, the result is checked numerically by generating a geometric AR(1) series ( $10^6$  data points) with  $a = 0.99$ , i.e.,  $\epsilon = 0.01$ , and  $b = 0.1$ .<sup>18</sup> Exemplarily,  $\mu_f^{(m)}(\bar{x}_n)$  and  $\sigma_f^{(m)}(\bar{x}_n)$  are estimated as a function of  $\bar{x}_n$  for  $m \in \{5, 21, 101\}$ . Due to

<sup>18</sup>As seen for the stationary AR(1) process and the  $\chi^2$ -distributed white noise, the estimation method is proven to be trustworthy.



**Figure 8.6:** The  $m$ -dependence of  $\alpha(m, b)$ , which corresponds to the proportionality factor between  $\sigma_f^{(m)}(\bar{x}_n)$  and  $\bar{x}_n$  for a generated geometric AR(1) process with  $a = 0.99$  and  $b = 0.1$ .

$1/\epsilon = 1/b^2 = 100$ , the approximation (8.34) is only justified for  $m = 2\tilde{m} + 1$  with  $\tilde{m} \ll 100$ . Fig. 8.5 (a) shows agreement with  $\mu_f^{(m)}(\bar{x}_n) \approx 0$  (red, crosses) and  $\sigma_f^{(m)}(\bar{x}_n) \propto \bar{x}_n$  (blue, dots) for  $m = 5$ . Panel (b) and (c), which depict  $\mu_f^{(m)}(\bar{x}_n)$  and  $\sigma_f^{(m)}(\bar{x}_n)$  for  $m = 21$  and  $m = 101$ , respectively, show a deviation from  $\mu_f^{(m)}(\bar{x}_n) \approx 0$  (red, crosses) but still reasonable agreement with  $\sigma_f^{(m)}(\bar{x}_n) \propto \bar{x}_n$  (blue, dots). That is, the proportionality between  $\sigma_f^{(m)}(\bar{x}_n)$  and  $\bar{x}_n$  is approximately satisfied for  $m = 2\tilde{m} + 1$  being beyond  $\tilde{m} \ll 1/\epsilon$  and  $\tilde{m} \ll 1/b^2$ .

## 8.5 Increment distribution of the geometric AR(1) process

The fact that the geometric AR(1) process, which was introduced in the previous section, has the property (8.34), does not automatically make it a suitable model to simulate wind speed. This section is intended to compare the increment distribution of a geometric AR(1) with the increment distribution derived from ABL wind speed data, see Fig. 6.12.

Sec. 6.3 gives evidence that the increment (6.14) of the horizontal component of ABL wind speed is in good approximation of a symmetric Castaing distribution (6.15). The kurtosis of the distribution depends on the increment length  $\tau$ : for small  $\tau$  the distribution is intermittent (i.e., leptokurtic) whereas for large  $\tau$  it is Gaussian (i.e., mesokurtic).

The increment series  $\mathbf{X}_s = \{X_{s;n} : n \in \mathbb{Z}\}$  of the geometric AR(1) process  $\mathbf{X} = \{X_n : n \in \mathbb{Z}\}$ , which is defined by (8.17), is given by

$$X_{s;n} = X_{n+s} - X_s, \quad (8.36)$$

where  $s \in \mathbb{N}$  denotes the increment length. If the parameters  $a = 1 - \epsilon$  and  $b$  in (8.17) are chosen such that

$$\epsilon \ll 1/s \quad \text{and} \quad b^2 \ll 1/s, \quad (8.37)$$



the increment distribution for small increment length  $s \in \mathbb{N}$  is in good approximation of a symmetric Castaing distribution. This is because the increment can be approximated by  $X_{s;n} \approx X_n b \sum_{l=1}^s \eta_{n;l}$  which is a consequence of (8.21) and (8.37). Note that  $\eta_{n;1}, \dots, \eta_{n;s}$  denote  $s$  mutually independent  $N_{0,1}$ -distributed rv's. Defining  $\zeta_{s;n} = \sum_{l=1}^s \eta_{n;l} / \sqrt{s} \sim N_{0,1}$ , the increment is written as

$$X_{s;n} \approx \underbrace{bX_n \sqrt{s}}_{\sim L_{bu\sqrt{s}, \lambda^2}} \zeta_{s;n}, \quad (8.38)$$

where it used that  $X_n \sim L_{u, \lambda^2}$  with  $\lambda^2 = b^2/(1-a^2) \approx b^2/(2\epsilon)$ . Hence, the increment corresponds to a product between a log-normally distributed rv  $bX_n \sqrt{s}$  with comparably slowly decaying auto correlation function<sup>19</sup> and an independent normally distributed rv  $\zeta_{s;n}$  with comparably quickly decaying auto correlation function<sup>20</sup>. Therefore, the increment distribution reads

$$p_{s;n}(x_{s;n}) \approx \int_0^\infty d\sigma L_{bu\sqrt{s}, \lambda^2}(\sigma) N_{0, \sigma^2}(x_{s;n}). \quad (8.39)$$

Substituting  $\sigma$  by  $\beta = 1/\sigma^2$  yields

$$p_{s;n}(x_{s;n}) \approx \int_0^\infty d\beta L_{1/(b^2 u^2 s), 4\lambda^2}(\beta) N_{0, \beta^{-1}}(x_{s;n}). \quad (8.40)$$

This corresponds to a symmetric Castaing distribution (6.15) with position and shape parameter

$$\beta_s = \frac{1}{u^2 b^2 s} \quad \text{and} \quad \lambda_s^2 = 4\lambda^2 \approx \frac{2b^2}{\epsilon}, \quad (8.41)$$

respectively.

For large increment  $s$ , the approximation  $a^s \approx 1$  and  $b^2 s \ll 1$  cannot be used, but it is possible to derive the exact form of the increment distribution for any  $a \in \mathbb{R}$  with  $|a| < 1$ ,  $b \in \mathbb{R}_{>0}$ , and  $s \in \mathbb{N}$ . Denoting the joint distribution of  $X_{n+s}$  and  $X_n$  by  $\rho_{n+s, n}(x_{n+s}, x_n)$ , the increment distribution reads

$$p_{s;n}(x_{s;n}) = \int dx_n dx_{n+s} \rho_s(x_n, x_{n+s}) \delta(x_{s;n} + x_n - x_{n+s}) \quad (8.42)$$

by applying the rules of the multivariate transformation (2.55). The joint distribution of  $X_{n+s}$  and  $X_n$  can be written as

$$\rho_{n+s, n}(x_{n+s}, x_n) = p_{n+s}(x_{n+s} | X_n = x_n) p_n(x_n) \quad (8.43)$$

by using relation (2.39) for conditioned distributions. Due to (8.21), the distribution of  $X_{n+s}$  conditioned on  $X_n = x_n$  is given by

$$p_{n+s}(x_{n+s} | X_n = x_n) = L_{u(x_n/u)^{a^s}, \lambda^2(1-a^{2s})}(x_{n+s}). \quad (8.44)$$

<sup>19</sup>It is derived in Ex. 3.7 that the auto correlation function of the geometric AR(1) process reads  $\gamma_{\mathbf{X}}(t) = [\exp(\lambda^2 a^{|t|}) - 1] / [\exp(\lambda^2) - 1]$ . Due to  $a = 1 - \epsilon$  with  $\epsilon \ll 1$ , the derivative of the auto correlation function in the limit of  $t \rightarrow \pm 0$  is approximately  $\mp \epsilon \lambda^2 \exp(\lambda^2) / [\exp(\lambda^2) - 1]$  and therefore very small in magnitude.

<sup>20</sup>The auto correlation function of  $\{\zeta_{s;n} : n \in \mathbb{Z}\}$  reads  $\mathbb{E}[\zeta_{s;n+t} \zeta_{s;n}] = \max\{1 - |t|/s, 0\}$ .

Inserting this into (8.42), integrating over  $x_{n+s}$  by taking advantage of the Dirac  $\delta$ -function, and substituting  $x_n$  by  $y = (2x_n + x_{s;n})/u$  results in

$$p_{s;n}(x_{s;n}) = \frac{1}{4\pi u \lambda^2 \sqrt{1 - a^{2s}}} \times \int_{|x_{s;n}/u|}^{\infty} \frac{dy}{y_+ y_-} \exp \left\{ -\frac{\ln^2 y_+ + \ln^2 y_- - 2a^s \ln y_+ \ln y_-}{2\lambda^2(1 - a^{2s})} \right\} \quad (8.45)$$

with  $y_{\pm} = (y \pm x_{s;n}/u)/2$ . It is a symmetric and unimodal distribution. Defining  $z(s) = 1 - a^s$ , its variance reads

$$\begin{aligned} \text{Var} [X_{s;n}] &= \text{Var} [X_{n+s}] + \text{Var} [X_n] - 2\text{Cov} [X_{n+s}, X_n] \\ &= 2u^2 e^{2\lambda^2} \left[ 1 - e^{-\lambda^2 z(s)} \right] \end{aligned} \quad (8.46)$$

and its kurtosis is given by

$$\begin{aligned} \mathbb{K}\text{urt} [X_{s;n}] &= \frac{\mathbb{E} [(X_{n+s} - X_n)^4]}{\text{Var} [X_{s;n}]^2} \\ &= 3e^{4\lambda^2} \times \frac{1 + 2e^{-\lambda^2 z(s)} + 3e^{-2\lambda^2 z(s)}}{6}, \end{aligned} \quad (8.47)$$

where a polynomial division is applied.

As a cross-check, for  $s$  being small enough such that the approximation  $\lambda^2 z(s) \approx b^2 s/2 \ll 1$  is justified, the variance and kurtosis of  $X_{s;n}$  become

$$\text{Var} [X_{s;n}] \approx u^2 b^2 s e^{2\lambda^2} \quad \text{and} \quad \mathbb{K}\text{urt} [X_{s;n}] \approx 3e^{4\lambda^2} (1 - 2b^2 s/3), \quad (8.48)$$

respectively. Approximating the distribution of  $X_{s;n}$  by a symmetric Castaing distribution with same variance and kurtosis means that the position and shape parameter need to be chosen as

$$\beta_s \approx \frac{1 - b^2 s/3}{u^2 b^2 s} \approx \frac{1}{u^2 b^2 s} \quad \text{and} \quad \lambda_s^2 \approx 4\lambda^2 - 2b^2 s/3 \approx 4\lambda^2, \quad (8.49)$$

which in zeroth order of  $b^2 s$  agrees with (8.41).

On the other hand, for very large  $s$ , the kurtosis of  $X_{s;n}$  becomes

$$\lim_{s \rightarrow \infty} \mathbb{K}\text{urt} [X_{s;n}] = 3e^{4\lambda^2} \times \frac{1 + 2e^{-\lambda^2} + 3e^{-2\lambda^2}}{6}, \quad (8.50)$$

which is always larger than three. Consequently, the increments of the geometric AR(1) process with large increment length are not Gaussian distributed as observed in ABL wind speed increments.

## 8.6 Fitting the parameters of the geometric AR(1) process to wind speed data

The previous two sections indicated that the geometric AR(1) process might be a suitable model to simulate wind speed data because:

1. The fluctuation  $F_n^{(m)}$  conditioned on  $\bar{X}_n^{(m)} = \bar{x}_n$  is in good approximation normally distributed if (8.19) is satisfied. Its mean vanishes, while there is a proportionality between its standard deviation and  $\bar{x}_n$ .
2. The distribution of the increment  $X_{s;n}$  is in good approximation of an intermittent symmetric Castaing distribution if (8.37) is satisfied.

Both are properties of ABL wind speed data  $v(t)$ . The proportionality factor between the standard deviation of the wind speed fluctuation and the mean wind speed is identical to the turbulence intensity (TI). The value of the TI depends on the chosen time scale  $T$  on which the mean wind speed is evaluated, cf. Fig. 6.11, and might be constant over a long interval such as 24 h, cf. Fig. 6.9. On the other hand, the distribution of the wind speed increment is for small increment length  $\tau$  roughly of a intermittent symmetric Castaing distribution, cf. Fig. 6.12. It is described by two parameters: the position parameter  $\beta_\tau$  and shape parameter  $\lambda_\tau^2$ . The values of both parameters also depend on the chosen increment length.

Concerning the geometric AR(1) process, it is discrete in time. The value  $x_n$  of the rv  $X_n$  is a model value for the wind speed at time  $t = n/\nu$ , where  $\nu$  denotes the model frequency. As the wind speed is measured in discrete time steps, too, it is convenient to set  $\nu$  equal to the measurement frequency. The process is characterised by three parameters:  $u$ ,  $a$ , and  $b$ . The scale  $u$  fixes the dimension of the process and has thus the dimension of velocity. In contrast, the other two parameters  $a$  and  $b$  are dimensionless.

The time scale  $T$  and increment length  $\tau$  need to be translated into dimensionless quantities  $m \in \mathbb{N}$  and  $s \in \mathbb{N}$ , respectively, reflecting  $T$  and  $\tau$  in units of time steps.

**TI time scale  $T \leftrightarrow m$ :** The mean wind speed  $\bar{v}^{(T)}(t)$  is governed by taking the average of  $v(t')$  over  $t - T/2 \leq t' \leq t + T/2$ . In discrete time steps  $n = t/\nu$  the mean wind speed  $\bar{v}^{(T)}(n/\nu)$  becomes the average of  $v(n'/\nu)$  over  $n - T\nu/2 \leq n' \leq n + T\nu/2$ . That is,  $\tilde{m} = T\nu/2$  so that  $m = 2\tilde{m} + 1 = T\nu + 1$ . Consequently,  $T$  needs to be chosen such that  $T\nu \in \mathbb{N}_0$ .

**Increment length  $\tau \leftrightarrow s$ :** The wind speed increment reads  $v(t+\tau) - v(t)$ , which in discrete time becomes  $v(n/\nu + s/\nu) - v(n/\nu)$ . Hence,  $s = \tau\nu$  so that  $\tau$  needs to be chosen such that  $\tau\nu \in \mathbb{N}$ .

The three process parameters  $u$ ,  $a$ , and  $b$  are chosen such that the generated series has same fluctuation and increment distribution as the wind speed which is to be simulated. The proportionality factor between the standard deviation of the fluctuation and the value of the

moving average reflects the TI of the generated series. Hence, Eq. (8.35) can be used to fix the value of the noise strength  $b$  by setting  $\text{TI}(T) = \alpha(T\nu + 1, b)$ , i.e.,

$$\text{TI}(T) = \frac{b}{2} \sqrt{\frac{(T\nu + 1)^2 - 1}{3(T\nu + 1)}}. \quad (8.51)$$

Strictly speaking, the  $T$  dependence of the TI is not necessarily the same as the  $T$  dependence of  $\alpha(T\nu + 1, b)$  so that the value of  $b$  depends on  $T$ .<sup>21</sup> If  $T$  is chosen such that the TI is much less than unity, squaring this equation implies that  $b^2 T\nu \ll 1$ , which ensures that the generated series has the desired proportionality between the conditioned standard deviation of the fluctuation and the value of the moving average. The remaining two process parameters,  $a$  and  $u$ , can be fixed by fitting the increment distribution parameters (8.41), i.e.,

$$u^2 = \frac{1}{\beta_\tau b^2 \tau \nu} \quad \text{and} \quad 1 - a = \epsilon = \frac{2b^2}{\lambda_\tau^2}. \quad (8.52)$$

The  $\tau$ -dependence does not vanish ( $\beta_\tau \tau \neq \text{const}$  and  $\lambda_\tau^2 \neq \text{const}$ ) so that the value of  $u$  and  $a$  depend on  $\tau$  and  $T$  (through  $b$ ). If the increment length  $\tau$  is chosen to be  $\tau \ll 1/b^2$ , i.e., the wind speed increment distribution is intermittent, the equation implies that  $\epsilon\tau \ll 1$  and satisfies condition (8.37). Thus, it is ensured that the increments of the generated series are of an intermittent symmetric Castaing distribution.

The previous chapter concluded that the day 191 data gathered at the [Lammefjord \(1987\)](#) site provide a suitable 24 h turbulent ABL wind speed data set, i.e., the turbulence parameters can be treated as being constant over 24 h. It has measurement frequency of  $\nu = 8\text{ Hz}$ . Setting  $T = 5\text{ s}$ , the TI is estimated to be 6.94%, see Fig. 6.11. The corresponding geometric AR(1) process has noise strength  $b = 0.0376$ . Regarding the increments, fixing  $\tau = 1\text{ s}$  yields an intermittent increment distribution with  $\beta_\tau = 1.383\text{ s}^2\text{ m}^{-2}$  and  $\lambda_\tau^2 = 0.385$ , see Fig. 6.13. That is, the corresponding geometric AR(1) process has  $u = 8.00\text{ m s}^{-1}$  and  $a = 0.99265$ . The statistical properties of this process are compared with the properties of the day 191 wind speed data. Therefore,  $5 \times 10^6$  data points of a geometric AR(1) series with  $u = 8.00\text{ m s}^{-1}$ ,  $a = 0.99265$ , and  $b = 0.0376$  are generated. The value of  $X_n$ ,  $\bar{X}_n^{(m)}$ , and  $F_n^{(m)}$  are denoted by  $x_n$ ,  $\bar{x}_n^{(m)}$ , and  $f_n^{(m)}$ , respectively.

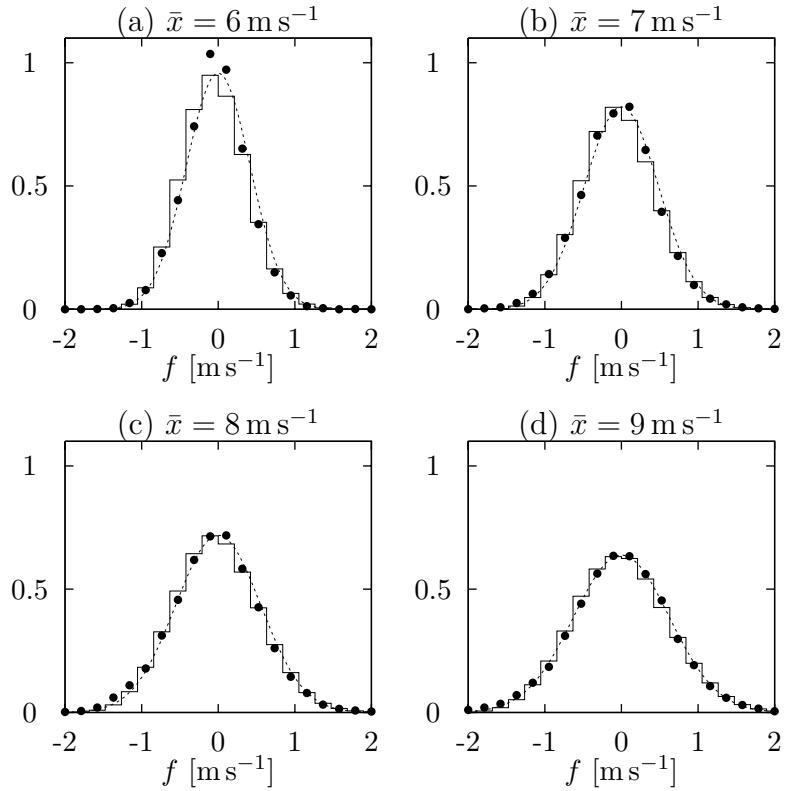
**Fluctuation Statistics.** The histograms of the fluctuation  $F_n^{(m)}$  with  $m = T\nu + 1 = 41$  conditioned on  $\bar{X}_n^{(m)} = \bar{x}$  are estimated by estimating the histogram of the set  $\{f_n^{(m)} : n \in \mathcal{N}_{\Delta\bar{x}}^{(m)}(\bar{x})\}$  with  $\mathcal{N}_{\Delta\bar{x}}^{(m)}(\bar{x}) = \{n : \bar{x}_n^{(m)} \in [\bar{x} - \Delta\bar{x}/2, \bar{x} + \Delta\bar{x}/2]\}$ . The bin width of  $\bar{x}$  is set to  $\Delta\bar{x} = 0.25\text{ m s}^{-1}$ . Fig. 8.7 displays the histograms for  $6\text{ m s}^{-1} \leq \bar{x} \leq 9\text{ m s}^{-1}$  together with the day 191 data histograms from Fig. 6.4 and symmetric normal distributions. It is evident that the histograms are in good agreement.

**Increment Statistics.** The histogram of the increment  $x_{n+s} - x_n$  with  $s = \tau\nu = 8$  is estimated and shown in Fig. 8.8. Additionally, the day 191 increment histogram from

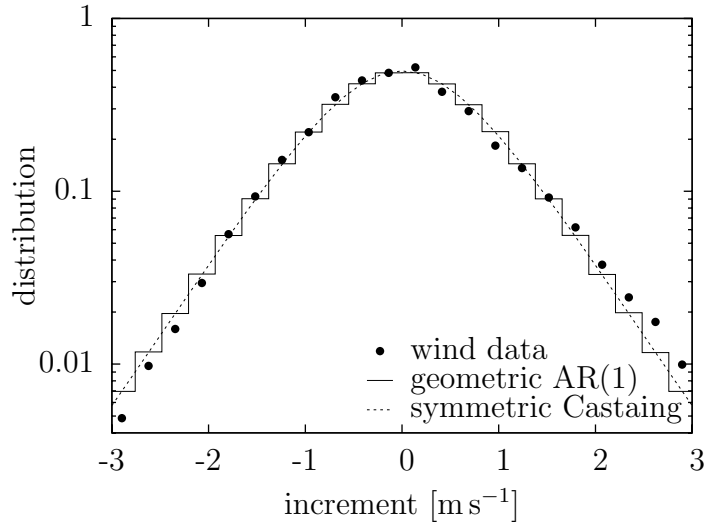
---

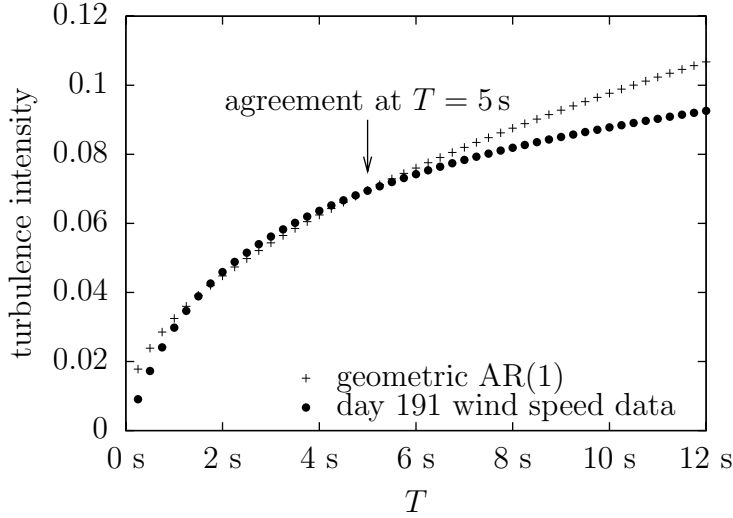
<sup>21</sup>For instance, Fig. 6.11 shows that the ABL TI does not grow with the square root of  $T$ .

**Figure 8.7:** Estimation of the fluctuation distributions conditioned on mean  $\bar{x}$ . For numerical evaluation,  $\Delta\bar{x}$  is chosen to be  $0.25 \text{ m s}^{-1}$ . The scale  $T$  is fixed to be  $5 \text{ s}$  with estimated turbulence intensity  $6.94\%$ . The solid histograms represent the generated geometric AR(1) series with  $a = 0.99265$ ,  $b = 0.0376$ , and  $u = 8.00 \text{ m s}^{-1}$ . The dots represent the day 191 data obtained at the [Lammefjord \(1987\)](#) site, cf. Fig. 6.4. The dashed lines correspond to symmetric normal distributions with standard deviation being proportional to  $\bar{x}$ .

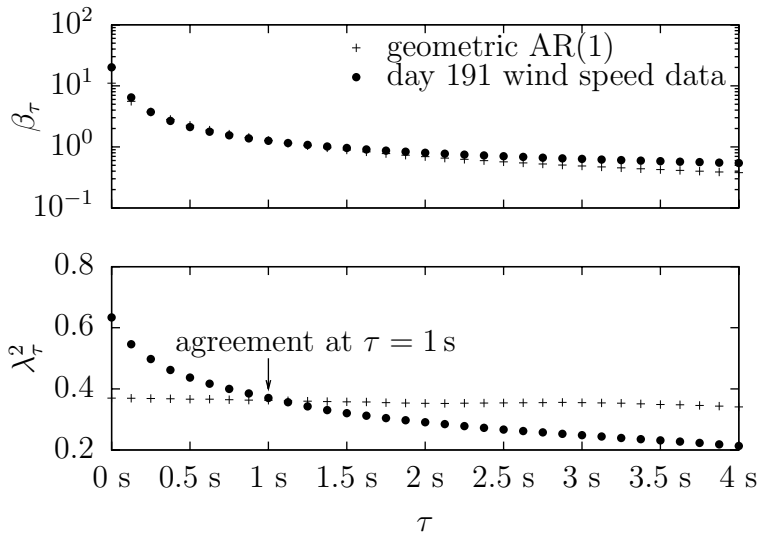


**Figure 8.8:** Estimation of the increment distributions with increment length  $\tau = 1 \text{ s}$ . The histograms are plotted semi-logarithmically. The solid histogram represents the generated geometric AR(1) series with  $a = 0.99265$ ,  $b = 0.0376$ , and  $U = 8.00 \text{ m s}^{-1}$ . The dots represent the day 191 data obtained at the [Lammefjord \(1987\)](#) site, cf. Fig. 6.12. The dashed line corresponds to the symmetric Castaing distribution.





**Figure 8.9:** Estimation of the turbulence intensity as a function of time scale  $T$ .



**Figure 8.10:** Estimation of  $\beta_\tau = \sqrt{\mathbb{Kurt}[X_{\tau\nu;n}]/3} \times \text{Var}[X_{\tau\nu;n}]^{-1}$  and  $\lambda_\tau^2 = \ln(\mathbb{Kurt}[X_{\tau\nu;n}]/3)$  as a function of increment length  $\tau$ .

Fig. 6.12 and the corresponding symmetric Castaing distribution are plotted. The figure shows good agreement between the histograms.

As a cross-check, the turbulence intensity of the generated series is estimated numerically as a function of scale  $T$  and depicted in Fig. 8.9. Additionally, the figure shows the  $T$ -dependence of the reference wind speed data. As expected, they are in agreement for  $T = 5$  s, but also in good agreement for  $T < 5$  s. The same cross-check is done for the increment distribution parameters  $\beta_\tau$  and  $\lambda_\tau^2$ . Fig. 8.10 shows that  $\beta_\tau$  is in good agreement with its corresponding value obtained from wind speed data even for  $\tau \neq 1$  s. In contrast, the figure also shows that  $\lambda_\tau^2$  estimated from the generated series is only in agreement with the corresponding value obtained from wind speed data for  $\tau = 1$  s.

# Chapter 9

## Summary and Outlook

### 9.1 Summary

The presented thesis analyses the statistical properties of atmospheric boundary layer wind speed recordings as well as the extent to which it can be simulated by a stochastic process. This section summarises the results grouped by the corresponding topic: wind velocity statistics, the suitability of the superstatistical algorithm, and the wind velocity simulation with stochastic processes.

#### Wind velocity statistics

The air motion in the atmospheric boundary layer is strongly influenced by thermal effects. Unlike the flow in an laboratory experiment, the wind driving parameters change with time. Motivated by this non-stationarity which appears in wind velocity recordings, Chap. 6 develops a variety of statistical tools which allow to verify that a wind velocity recording has periods in which the wind speed has statistically similar properties to ideal turbulence. The methods are based on

- fluctuation statistics in order to analyse the wind velocity and
- superstatistics in order to analyse the wind velocity increments.

The statistical properties, which were found, are:

1. The fluctuation around the mean wind speed is roughly Gaussian distributed and has vanishing mean.
2. The standard deviation of the fluctuation is proportional to the mean wind speed, as long as the latter is above  $5 \text{ m s}^{-1}$ .
3. The increment distribution is not (or slightly positively) skewed and in good approximation a symmetric [Castaing \*et al.\* \(1990\)](#) distribution.

4. For small/large increment lengths, the increment distribution is close to a double exponential/Gaussian distribution.

Additionally, the methods allow to quantify the time evolution of the relevant parameters. These are

- the turbulence intensity, i.e., the proportionality factor between the standard deviation of the fluctuation and the mean wind speed, and
- the position parameter and shape parameter of the increment distribution fitted by a symmetric Castaing distribution.

The measured turbulence intensity depends on the considered time scale  $T$  over which the moving average is computed: the value of the turbulence intensity increases with increasing  $T$ . Even for large  $T$ , such as several minutes, the proportionality between the standard deviation of the fluctuation and the value of the moving average is statistically satisfied. The shape of the increment distribution depends on the considered increment length  $\tau$ : both parameters decrease with increasing  $\tau$ , the shape parameter approaches zero for sufficiently large  $\tau$ . It is found that

- the periods in which the turbulence intensity can be regarded as being constant last from several hours up to 24 h and
- the Castaing parameters usually fluctuate in a diurnal cycle.

The first result suggests to transform a wind velocity recording, in which the turbulence intensity can be treated as being constant, into a first order stationary series by dividing the fluctuation by the value of the moving average. The second result can be used to explain that the increment distribution for large increment lengths is sometimes still found to be non-Gaussian when analysing a recording being longer than a turbulent period (see e.g. [Boettcher et al., 2007](#)). The time series can be interpreted as consisting of subsequent periods in which the increment distribution is Gaussian for large increment length. Nevertheless, the variance, which is identical to the inverse of the position parameter, changes from one period to another so that the analysis of the full time series yields a superposition of Gaussians with different variances.

Additionally, the methods are able to indicate the existence of periods in which the above mentioned properties found in turbulent periods are most likely not satisfied. Indications are:

- The turbulence intensity is highly fluctuating. (If the proportionality between the standard deviation of the wind speed fluctuation and the mean wind velocity is not satisfied, the analysis does not yield a constant turbulence intensity.)
- There is significance that in this period the increment distribution is symmetric (or slightly positively skewed) but not of a symmetric Castaing distribution and does not tend towards a Gaussian distribution for large increment lengths.



Both indicators signal that such a period is a sequence of shorter periods which can be regarded as being turbulent with changing relevant parameters or the wind activity in this period is not turbulent.

## Superstatistical algorithm

Chap. 7 discusses the superstatistical algorithm in more detail. The algorithm is proposed by Beck *et al.* (2005b) and Queiros (2007) and is extended in this thesis in order to analyse the turbulent nature of wind velocity increments. The algorithm approximates the series which is to be analysed by equally long periods in which the series is symmetrically Gaussian distributed. The length of these periods is estimated by computing the mean period kurtosis. The inverse variance estimated in each period leads to the  $\beta$ -series, which for wind velocity data is hypothetically log-normally distributed.

The algorithm is applied to artificial time series which per construction consist of equally long periods in which the series is symmetrically Gaussian distributed. It is shown that the parameter estimation of the  $\beta$ -statistics strongly depends on the estimated value of the period length, but the decision whether  $\beta$  is log-normally distributed does not.

Furthermore, the algorithm is applied to artificial time series which per construction consist of periods in which the series is symmetrically Gaussian distributed, but the periods are not of equal length. Rather, the length of the periods is Poisson distributed. Indeed, the algorithm finds that there are no equally long periods and cannot estimate their length. Still, it is demonstrated that the estimation of the  $\beta$ -distribution parameters strongly depends on the period length by setting the latter manually as internal parameter. However, the decision whether  $\beta$  is log-normally distributed does not depend on the value of this particular parameter.

Concerning atmospheric boundary layer wind velocity increment recordings, the approximation by equally long Gaussian periods might be too rough. The effect of not finding a period length is avoided by using a biased kurtosis estimator instead of an unbiased one. Whether this is a suitable estimation procedure for the period length is discussed in App. B. It compares the parameters of the  $\beta$ -distribution with the parameters gained directly from the wind velocity increment statistics according to Castaing *et al.* (1990). It is found that only the estimation of the shape parameter is slightly biased.

Concluding, the Queiros (2007) algorithm is suitable to analyse wind velocity increments, but it is not perfect. However, this imperfection does not have an influence on the decision whether a wind recording represents turbulent wind activity or not. It is merely the strength of intermittency which is sometimes slightly under- or overestimated.

## Simulation with stochastic processes

Chap. 8 deals with stochastic processes that might be suitable candidates to simulate atmospheric wind velocity during a period with constant relevant parameters. The process is required to have the same fluctuation statistics conditioned on the value of the moving average as wind velocity recordings. That is, the fluctuation

- is normally distributed,
- has vanishing mean, and
- has a proportionality between its standard deviation and the value of the moving average.

A method which allows to evaluate the fluctuation distribution of an arbitrary process is introduced. This tool is successfully applied to a variety of discrete stochastic processes. It is found that a process with the above mentioned fluctuation statistics is non-trivial. So for instance, Gaussian white noise has normally distributed fluctuation, but lacks the proportionality, whereas  $\chi^2$ -distributed white noise shows the proportionality, but lacks the normal distribution.

The geometric AR(1) process however has the desired property when choosing its parameters accordingly and when the time scale  $T$  over which the moving average is computed is small. Moreover, the increment distribution of this process is in good approximation a symmetric Castaing distribution when the increment length  $\tau$  is small. Hence, the geometric AR(1) process is a suitable candidate to generate series which have the corresponding fluctuation statistics for small scale  $T$  and intermittent increment distributions for small increment length  $\tau$ . The process is described by three parameters and they can be chosen such that the process has the same turbulence intensity and increment distribution as the wind speed series which is to be simulated. Unfortunately, the  $T$ -dependence of the turbulence intensity and the  $\tau$ -dependence of the increment distribution parameters are different between the data generated by a geometric AR(1) process and boundary layer wind velocity data. Hence, in order to fit the parameters, the time scale  $T$  and increment length  $\tau$  need to be fixed. Estimating the turbulence intensity at scale  $T$  and the two increment distribution parameters for increments with increment length  $\tau$ , the geometric AR(1) process parameter can be fitted by (8.51) and (8.52). The turbulence intensity of the generated series well agrees with the wind turbulence intensity for smaller time scales than  $T$ . The main drawback of this process class is that it does not provide the transition from an intermittent increment distribution for small increment lengths to a Gaussian increment distribution for large increment length.

## 9.2 Outlook

As mentioned in the beginning of this thesis, the understanding of turbulence is a gradual process. Therefore, based on the found results and the methods used in this thesis, the following issues are suggestions for further studies. They are again grouped by the corresponding topic.

### Wind velocity statistics

- It is found that the turbulence parameters change with time. It would be interesting to see whether there is a connection between the change of the turbulence parameters and the change of other meteorological parameters.

- The analysis suggested that some parts of a wind velocity recording do not reflect turbulent wind activity. The question which remains to be answered is whether the flow is laminar-like or the parameters of the turbulence generating mechanism fluctuate on a smaller time scale.
- The proposed method to obtain a first order stationary wind recording allows to analyse the linear cross-correlation between several measurements because the estimation of the linear correlation requires a constant mean and variance of the series. The results of the correlation analysis are essential for constructing realistic wind field models, which simulate the wind velocity at several spatial points simultaneously.
- The search for significant wind gust patterns can be improved by suitable re-scaling. That is, the amplitude of a gust is statistically proportional to the mean wind speed.
- In order to improve the comparison between atmospheric and ideal turbulence, the longitudinal wind velocity increments need to be measured. This however requires a measurement at several spatial points at the same time but has the advantage that the assumption of the Taylor hypothesis is not needed.

### Superstatistical algorithm

- The mentioned imperfections of the superstatistics algorithm can be avoided by a different approach and by working directly with the wind velocity increment distribution. Keeping in mind that the increment distribution represents a superposition of Gaussian distributions with different variance ( $= \beta^{-1}$ ), the  $\beta$ 's do not have to be estimated from the recording as done by Beck *et al.* (2005b); Queiros (2007). Rather, their distribution might be optimised such that the superposition of the normal distributions gets close to the empirically found wind velocity increment distribution. This approach still allows to test the turbulence hypothesis by testing whether  $\beta$  is log-normally distributed or not. An optimisation method based on the Metropolis *et al.* (1953) algorithm might be appropriate.

### Simulation with stochastic processes

- The introduced recipe to calculate the distribution of the fluctuation conditioned on the value of the moving average for any process might be used to formulate the corresponding criteria for a process to have any desired fluctuation property. This would help in the development of wind velocity models based on stochastic processes exhibiting more realistic characteristics.
- It might be advantageous to develop stochastic processes with changing turbulence parameters in order to simulate atmospheric turbulence over a long period.

More generally, understanding the nature of boundary layer turbulence several tens of metres above ground might be relevant for man-made constructions such as houses, bridges,

wind turbines, but also for objects from the biosphere such as trees or the flight of insects. Moreover, wind gust prediction algorithms can be improved by taking into account the non-stationarity of the turbulence parameters. Last, but not least, Richard Feynman once said that turbulence is the most important unsolved problem of classical physics. Further studies based on the results of this thesis may be some of many puzzle pieces to solve this problem.

**Part III**  
**Appendix**

# Appendix A

## Frequently used Distributions and their Properties

This appendix is a collection of distribution functions to fix the notation which is used throughout this thesis. Additionally, it illustrates the meaning of skewness and kurtosis.

The works of [Johnson \*et al.\* \(1994\)](#) and [Johnson \*et al.\* \(2005\)](#) serve as reference for pdf's and pmf's, respectively, and their properties.

### A.1 Binomial distribution

The *binomial distribution* is an example of a pmf of a discrete rv with range  $\mathcal{X} = \{0, 1, 2, 3, \dots, n\}$  and defined by

$$B_{n,p}(k) = \binom{n}{k} p^k (1-p)^{n-k} \quad (\text{A.1})$$

for  $k = 0, 1, 2, \dots, n$  and  $0 \leq p \leq 1$ . For  $p = 1/2$  this equation turns into (2.4). The expected value, variance, skewness, and kurtosis read

$$\mathbb{E}[X] = np, \quad (\text{A.2a})$$

$$\text{Var}[X] = np(1-p), \quad (\text{A.2b})$$

$$\text{Skew}[X] = \frac{1-2p}{\sqrt{np(1-p)}}, \quad \text{and} \quad (\text{A.2c})$$

$$\mathbb{K}\text{urt}[X] = 3 + \frac{1-6p(1-p)}{np(1-p)}, \quad (\text{A.2d})$$

respectively.

If  $n$  is very large, the skewness vanishes and the kurtosis tends to three. The *De Moivre-Laplace theorem* states that for the *standardised binomial variable*  $\hat{X} = (X - np) / \sqrt{np(1-p)}$  the probability

$$\lim_{n \rightarrow \infty} \mathbb{P}[\alpha < \hat{X} < \beta] = \frac{1}{\sqrt{2\pi}} \int_{\alpha}^{\beta} du e^{-u^2/2} \quad (\text{A.3})$$

which give a connection to the normal distribution explained in the next section.

## A.2 Gaussian distribution

The *Gaussian distribution* or *normal distribution* is defined by

$$N_{\mu,\sigma^2}(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left[-\frac{(x-\mu)^2}{2\sigma^2}\right], \quad (\text{A.4})$$

where  $\mu$  and  $\sigma^2$  are real numbers. Note that  $\sigma^2$  must be positive. It is very tractable analytically, symmetric about  $\mu$  (hence, its median is  $\mu$ ), and can be used to approximate a large variety of distributions in large samples (see e.g. the binomial distribution in the previous section), which is known as the *central limit theorem*.

A normally distributed rv  $X$  has

$$\mathbb{E}[X] = \mu, \quad (\text{A.5a})$$

$$\text{Var}[X] = \sigma^2, \quad (\text{A.5b})$$

$$\text{Skew}[X] = 0, \quad \text{and} \quad (\text{A.5c})$$

$$\mathbb{Kurt}[X] = 3. \quad (\text{A.5d})$$

The  $n^{\text{th}}$  central moment reads

$$\mathbb{E}[(x-\mu)^n] = \begin{cases} (n-1)!! \sigma^n & \text{for even } n, \\ 0 & \text{for odd } n, \end{cases} \quad (\text{A.6})$$

where  $(\cdot)!!$  denotes the double factorial (2.34). The cdf of the normal distribution is given by

$$\Phi_{\mu,\sigma^2}(x) = \frac{1}{2} \text{erfc}\left(-\frac{x-\mu}{\sqrt{2\sigma^2}}\right), \quad (\text{A.7})$$

where  $\text{erfc}(\cdot)$  denotes the *complementary error function*. It is defined by  $\text{erfc}(x) = 1 - \text{erf}(x)$ , where

$$\text{erf}(x) = \frac{2}{\sqrt{\pi}} \int_0^x dt e^{-t^2} \quad (\text{A.8})$$

denotes the *error function*, and takes the value one if and only if its argument is zero. This proves that the median of the normal distribution is  $\mu$ .

It is shown in Ex. 2.4 that if  $X$  is normally distributed, its linear transformation  $aX + b$  with  $a \neq 0$  is also of a normal distribution, i.e.,

$$X \sim N_{\mu,\sigma^2} \Rightarrow aX + b \sim N_{\mu+b, a^2\sigma^2}, \quad (\text{A.9})$$

which follows from Eq. (2.21).

The characteristic function of the normal distribution reads

$$\mathbb{E}[e^{itX}] = \exp\left[i\mu t - \frac{\sigma^2}{2}t^2\right], \quad (\text{A.10})$$

where  $X \sim N_{\mu,\sigma^2}$ .

### A.3 $\chi^2$ -distribution

If  $Y \sim N_{0,1}$ , the rv  $X = Y^2$  is  $\chi^2$ -distributed. More generally, if  $Y_i \sim N_{0,1}$  are independent identically distributed rv's for  $i = 1, 2, \dots, k$ , the rv  $X = \sum_{i=1}^k Y_i^2$  is of a  $\chi^2$ -distribution with  $k$  number of degrees of freedom. It is defined by

$$\chi_k^2(x) = \begin{cases} \frac{1}{2\Gamma(k/2)} \left(\frac{x}{2}\right)^{k/2-1} e^{-x/2} & \text{for } x > 0, \\ 0 & \text{for } x \leq 0, \end{cases} \quad (\text{A.11})$$

where  $\Gamma(\cdot)$  denotes the *Gamma function*

$$\Gamma(z) = \int_0^\infty dt t^{z-1} e^{-t}. \quad (\text{A.12})$$

A  $\chi_k^2$ -distributed rv  $X$  is characterised by

$$\mathbb{E}[X] = k, \quad (\text{A.13a})$$

$$\text{Var}[X] = 2k, \quad (\text{A.13b})$$

$$\text{Skew}[X] = \sqrt{8/k}, \quad \text{and} \quad (\text{A.13c})$$

$$\mathbb{Kurt}[X] = 3 + 12/k. \quad (\text{A.13d})$$

It should be noted that the *standardised*  $\chi^2$ -variable rv  $\hat{X} = (X - k)/\sqrt{2k}$ , which has zero expectation and unit variance, has also skewness  $\sqrt{8/k}$  and kurtosis  $3 + 12/k$ . If  $k \rightarrow \infty$ , its skewness vanishes and its kurtosis becomes three, which corresponds to the skewness and kurtosis of the normal distribution, respectively. In fact, it can be shown that for  $k \rightarrow \infty$  the rv  $\hat{X}$  is  $N_{0,1}$ -distributed.

### A.4 Log-normal distribution

The log-normal distribution plays a central role in various branches of sciences. [Limpert et al. \(2001\)](#) discusses its meaning in e.g. geology and mining, human medicine, environment and atmospheric sciences, ecology, and food technology, to name a few.

If  $Y \sim N_{\mu, \sigma^2}$ , its exponentially transformed rv  $X = e^Y$  is said to be *log-normally distributed*. Applying Eq. (2.19), the pdf of  $X$  is given by

$$L_{x_0, \lambda^2}(x) = \begin{cases} \frac{1}{x\sqrt{2\pi\lambda^2}} \exp\left[-\frac{1}{2\lambda^2} \ln^2 \frac{x}{x_0}\right] & \text{for } x > 0, \\ 0 & \text{for } x \leq 0, \end{cases} \quad (\text{A.14})$$

where  $x_0 = e^\mu > 0$  and  $\lambda^2 = \sigma^2$  are called *position* and *shape parameter*, respectively, because they are directly related to the median and kurtosis, respectively. A  $L_{x_0, \lambda^2}$ -distributed rv



has median  $x_0$ , mode  $x_0 e^{-\lambda^2}$  as well as

$$\mathbb{E}[X] = x_0 e^{\lambda^2/2}, \quad (\text{A.15a})$$

$$\text{Var}[X] = x_0^2 e^{\lambda^2} (e^{\lambda^2} - 1), \quad (\text{A.15b})$$

$$\text{Skew}[X] = (e^{\lambda^2} + 2) \sqrt{e^{\lambda^2} - 1}, \quad \text{and} \quad (\text{A.15c})$$

$$\mathbb{K}\text{urt}[X] = e^{4\lambda^2} + 2e^{3\lambda^2} + 3e^{2\lambda^2} - 3. \quad (\text{A.15d})$$

The higher moments are

$$\mathbb{E}[X^n] = x_0^n e^{n^2 \lambda^2/2}. \quad (\text{A.16})$$

It is worth mentioning that

$$\mathbb{E}[\ln X] = \ln x_0, \quad (\text{A.17a})$$

$$\text{Var}[\ln X] = \lambda^2, \quad (\text{A.17b})$$

$$\text{Skew}[\ln X] = 0, \quad \text{and} \quad (\text{A.17c})$$

$$\mathbb{K}\text{urt}[\ln X] = 3 \quad (\text{A.17d})$$

because  $\ln X \sim N_{\ln x_0, \lambda^2}$ .

An important special case is when the shape parameter is much less than unity. Consider the rv  $x_0 e^{\lambda \xi}$  with  $\xi \sim N_{0,1}$ . It is  $L_{x_0, \lambda^2}$ -distributed whilst its expansion  $x_0 e^{\lambda \xi} \approx x_0 + x_0 \lambda \xi \sim N_{x_0, x_0^2 \lambda^2}$ . That is, a log-normally distributed rv with very small shape parameter can be treated like a Gaussian distributed rv. Additionally, in the limit for  $\lambda^2 \rightarrow 0$  the log-normal distribution becomes

$$\lim_{\lambda^2 \rightarrow 0} L_{x_0, \lambda^2}(x) = \delta(x - x_0), \quad (\text{A.18})$$

where  $\delta(\cdot)$  denotes the Dirac  $\delta$ -function.

It should be mentioned, that if  $X \sim L_{x'_0, \lambda'^2}$ , its linear transformed rv  $aX + b$  with  $a > 0$  and  $b \neq 0$  is not log-normally-distributed. It is said to be *shifted log-normally distributed*. Applying Eq. (2.21) leads to the three-parameter pdf

$$L_{x_0, \lambda^2, \delta}(x) = \begin{cases} \frac{1}{(x-\delta)\sqrt{2\pi\lambda^2}} \exp\left[-\frac{1}{2\lambda^2} \ln^2 \frac{x-\delta}{x_0}\right] & \text{for } x > \delta, \\ 0 & \text{for } x \leq \delta \end{cases} \quad (\text{A.19})$$

with  $x_0 = ax'_0$ ,  $\lambda^2 = \lambda'^2$ , and  $\delta = b$ .

In contrast to the linear transformation, the nonlinear transformation  $aX^b$  with  $a > 0$  and  $b \neq 0$  is log-normally distributed:

$$X \sim L_{x_0, \lambda^2} \Rightarrow aX^b \sim L_{ax_0^b, b^2 \lambda^2}. \quad (\text{A.20})$$

This is shown as follows. The logarithm of  $aX^b$  corresponds to  $\ln a + b \ln X$ , which according to (A.9) is normally distributed with mean  $\ln(x_0 a)$  and variance  $b^2 \lambda^2$ . Hence,  $aX^b \sim L_{ax_0^b, b^2 \lambda^2}$ .

## A.5 Cauchy distribution

The *Cauchy distribution*, which is also known as *Lorentz* or *Breit-Wigner distribution*, is defined by

$$\text{Cy}_{x_0, \gamma} = \frac{\gamma}{\pi[(x - x_0)^2 + \gamma^2]}. \quad (\text{A.21})$$

It has no well-defined mean and variance. In fact, the improper integrals  $\int_{-\infty}^{\infty} dx x^n \text{Cy}_{x_0, \gamma}(x)$  do not converge for  $n = 1, 2, \dots$ . Nevertheless, the Cauchy distribution is symmetric about  $x_0$  and has mode and median  $x_0$ . Its characteristic function reads

$$\mathbb{E} [e^{itX}] = \exp [ix_0t - \gamma|t|] \quad (\text{A.22})$$

where  $X \sim \text{Cy}_{x_0, \gamma}$ .

The Cauchy distribution is related to the normal distribution as follows. If the two iid rv's  $X$  and  $Y$  are normally distributed with zero mean and unit variance, their ration  $X/Y$  is Cauchy distributed, i.e.,

$$X, Y \sim N_{0,1} \text{ (iid)} \quad \Rightarrow \quad \frac{X}{Y} \sim \text{Cy}_{0,1}. \quad (\text{A.23})$$

## A.6 Lévy distribution

The *Lévy distribution* is defined by

$$\text{Lv}_{x_0, c} = \frac{\exp \left[ -\frac{c}{2(x-x_0)} \right]}{(x - x_0)^{3/2}} \times \sqrt{\frac{c}{2\pi}}. \quad (\text{A.24})$$

The moments are not defined as the improper integrals  $\int_{-\infty}^{\infty} dx x^n \text{Lv}_{x_0, c}(x)$  do not converge for  $n = 1, 2, \dots$ . Its characteristic function reads

$$\begin{aligned} \mathbb{E} [e^{itX}] &= \exp \left[ ix_0t - \sqrt{-2ict} \right] \\ &= \exp \left[ ix_0t - (1 - i \operatorname{sgn}(t)) \sqrt{c|t|} \right]. \end{aligned} \quad (\text{A.25})$$

## A.7 Uniform distribution

The *uniform distribution* is defined by

$$U_{x_1, x_2} = \begin{cases} \frac{1}{x_2 - x_1} & \text{for } x_1 \leq x \leq x_2 \\ 0 & \text{otherwise.} \end{cases} \quad (\text{A.26})$$

A  $U_{x_1, x_2}$ -distributed rv  $X$  has

$$\mathbb{E}[X] = \frac{x_1 + x_2}{2}, \quad (\text{A.27a})$$

$$\text{Var}[X] = \frac{(x_1 - x_2)^2}{12}, \quad (\text{A.27b})$$

$$\text{Skew}[X] = 0, \quad \text{and} \quad (\text{A.27c})$$

$$\mathbb{Kurt}[X] = 9/5 = 1.8. \quad (\text{A.27d})$$

A uniformly distributed rv remains uniformly distributed under linear transformation, i.e.,

$$X \sim U_{x_1, x_2} \Rightarrow aX + b \sim U_{ax_1+b, ax_2+b} \quad (\text{A.28})$$

with  $a \neq 0$ , which follows from Eq. (2.21).

## A.8 Exponential distribution

If  $Y \sim U_{0,1}$ , the rv  $X = -x_0 \ln Y$  is of an *exponential distribution* with pdf

$$E_{x_0}(x) = \begin{cases} \frac{1}{x_0} \exp\left[-\frac{x}{x_0}\right] & \text{for } x \geq 0, \\ 0 & \text{for } x < 0, \end{cases} \quad (\text{A.29})$$

which follows from Eq. (2.19). The parameter  $x_0$  is called the scale parameter. An  $E_{x_0}$ -distributed rv  $X$  has

$$\mathbb{E}[X] = x_0, \quad (\text{A.30a})$$

$$\text{Var}[X] = x_0^2, \quad (\text{A.30b})$$

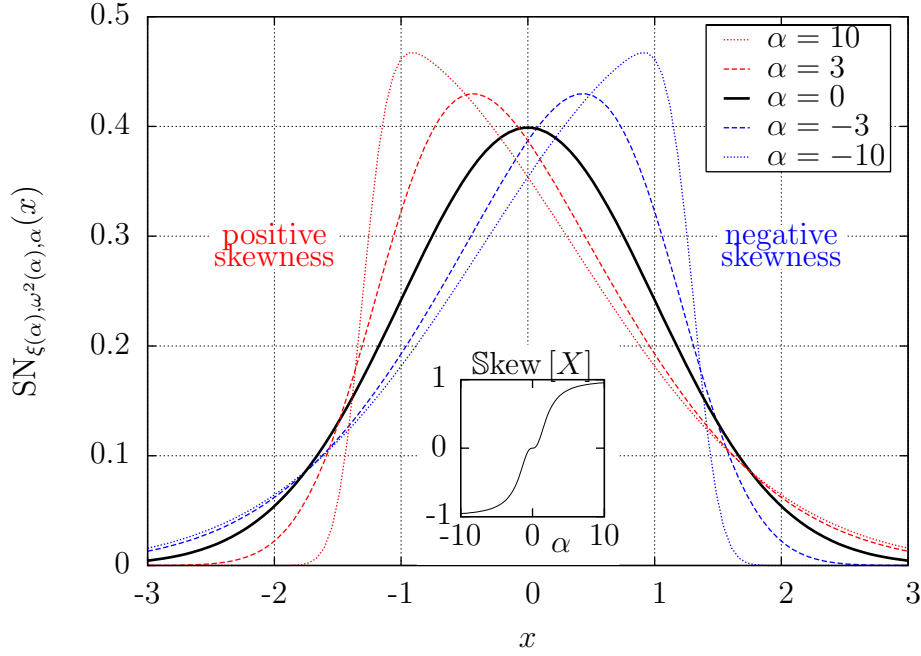
$$\text{Skew}[X] = 2, \quad \text{and} \quad (\text{A.30c})$$

$$\mathbb{Kurt}[X] = 9. \quad (\text{A.30d})$$

## A.9 Interpretation of skewness and kurtosis

In the previous sections, a variety of distributions were introduced. This section illustrates the meaning of skewness and kurtosis exemplified by two “generalised normal distributions”:

- the skew normal distribution  $\text{SN}_{\xi, \omega^2, \alpha}$  recovering the normal distribution for  $\alpha = 0$  and
- a general distribution  $\text{KN}_{\alpha, \beta, \gamma}$ , which recovers the normal distribution for  $\gamma = 2$ .



**Figure A.1:** The skew normal distribution, where the parameters  $\xi(\alpha)$  and  $\omega^2(\alpha)$  are chosen such that it has zero mean and unit variance, for a variety of  $\alpha$ . The red (blue) curves correspond to  $\alpha > 0$  ( $\alpha < 0$ ) and hence positive (negative) skewness. The inner panel depicts the skewness of a  $\text{SN}_{\xi, \omega^2, \alpha}$ -distributed rv  $X$  as a function of  $\alpha$ .

### A.9.1 Skewness

The meaning skewness is best exemplified by the *skew normal distribution* where the skewness can be changed by a third parameter. [Azzalini \(1985\)](#) introduced

$$\text{SN}_{\xi, \omega^2, \alpha}(x) = N_{\xi, \omega^2}(x) \times \text{erfc} \left( -\alpha \frac{x - \xi}{\sqrt{2\omega^2}} \right) \quad (\text{A.31})$$

as a distribution, which obviously corresponds to the normal distribution with mean  $\xi$  and variance  $\omega^2$  if  $\alpha = 0$ . Defining

$$\delta = \frac{\alpha}{\sqrt{1 + \alpha^2}} \times \sqrt{\frac{2}{\pi}}, \quad (\text{A.32})$$

the mean and variance of a  $\text{SN}_{\xi, \omega^2, \alpha}$ -distributed rv  $X$  are given by

$$\mathbb{E}[X] = \xi + \omega\delta \quad \text{and} \quad (\text{A.33a})$$

$$\text{Var}[X] = \omega^2(1 - \delta^2), \quad (\text{A.33b})$$

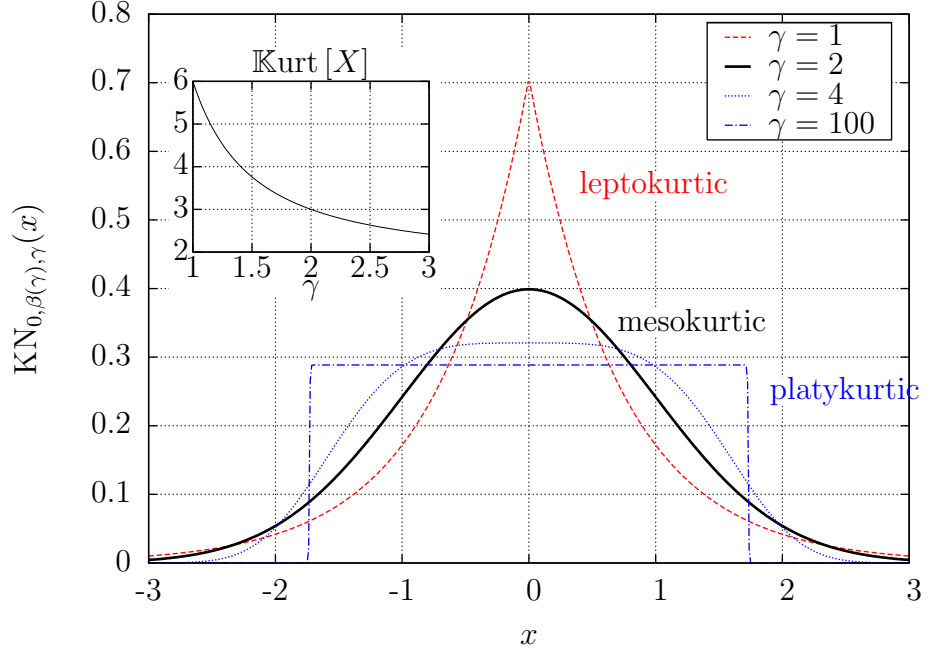
respectively. Its skewness and kurtosis read

$$\text{Skew}[X] = \frac{4 - \pi}{2} \frac{\delta^3}{(1 - \delta^2)^{3/2}} \quad \text{and} \quad (\text{A.33c})$$

$$\mathbb{K}\text{urt}[X] = 3 + 2(\pi - 3) \left[ \frac{\delta^2}{1 - \delta^2} \right]^2, \quad (\text{A.33d})$$

**Figure A.2:**

The  $\text{KN}_{0,\beta(\gamma),\gamma}$ -distribution with zero mean and unit variance for a variety of  $\gamma$ . The red curve ( $\gamma = 1$ ) has kurtosis smaller than three whereas the blue curves ( $\gamma = 4, 100$ ) have kurtosis larger than three. The inner panel shows the kurtosis of a  $\text{KN}_{0,\beta(\gamma),\gamma}$ -distributed rv  $X$  as a function of  $\gamma$ .



respectively, and only depend on  $\delta$ , which is a function of  $\alpha$ . It can be seen that for vanishing  $\delta$ , i.e., vanishing  $\alpha$ , the skewness becomes zero and the kurtosis becomes three.

Setting  $\xi(\alpha) = -\delta(\alpha)/\sqrt{1-\delta^2(\alpha)}$  and  $\omega^2(\alpha) = 1/(1-\delta^2(\alpha))$ , the skew normal distribution  $\text{SN}_{\xi(\alpha),\omega^2(\alpha),\alpha}$  has zero mean and unit variance. The only parameter left is  $\alpha$ , which controls the skewness and kurtosis. For  $\alpha < 0$  ( $\alpha > 0$ ), and therefore  $\delta < 0$  ( $\delta > 0$ ), the skewness is negative (positive). Fig. A.1 displays a variety of skew normal distributions with zero mean and unit variance for a variety of  $\alpha$ . The red (blue) curves with  $\alpha > 0$  ( $\alpha < 0$ ) have positive (negative) skewness. The black curve, which has  $\alpha = 0$ , corresponds to a Gaussian distribution with vanishing skewness. It can be seen that for a positive (negative) skewness the right (left) tail is longer, i.e., the mass of the distribution is concentrated on the left (right) of the distribution and therefore the median is smaller (larger) than the expected value.

## A.9.2 Kurtosis

The meaning kurtosis is best exemplified by a distribution which has zero skewness and an additional parameter to change the kurtosis. [Box \(1953\)](#) and [Turner \(1960\)](#) considered the three-parameter pdf

$$\text{KN}_{\alpha,\beta,\gamma}(x) = \frac{\gamma}{2\beta\Gamma(1/\gamma)} \exp\left[-\frac{|x-\alpha|^\gamma}{\beta^\gamma}\right] \quad (\text{A.34})$$

with  $\beta, \gamma > 0$ . There are four special cases:

1.  $\gamma = 1$  leads to the *double exponential distribution*  $\text{KN}_{\alpha,\beta,1}(x) = \frac{1}{2\beta} e^{-|x-\alpha|/\beta}$ .
2.  $\gamma = 2$  leads to the normal distribution  $\text{KN}_{\alpha,\beta,2}(x) = N_{\alpha,\beta^2/2}(x)$ .

3.  $\gamma = q$ , where  $q$  is an even integer, leads to the  $q^{\text{th}}$  *power distribution*.
4.  $\gamma \rightarrow \infty$  leads to the uniform distribution  $\lim_{\gamma \rightarrow \infty} \text{KN}_{\alpha, \beta, \gamma}(x) = U_{\alpha - \beta, \alpha + \beta}(x)$ .

This distribution is symmetric about its expectation value  $\alpha$  so that its skewness vanishes. A  $\text{KN}_{\alpha, \beta, \gamma}$ -distributed rv  $X$  has variance

$$\text{Var}[X] = \beta^2 \frac{\Gamma(3/\gamma)}{\Gamma(1/\gamma)} \quad (\text{A.35})$$

and kurtosis

$$\mathbb{K}\text{urt}[X] = \frac{\Gamma(5/\gamma) \Gamma(1/\gamma)}{\Gamma(3/\gamma)^2}. \quad (\text{A.36})$$

That is, setting  $\beta(\gamma) = \sqrt{\Gamma(1/\gamma)/\Gamma(3/\gamma)}$ , the distribution  $\text{KN}_{0, \beta(\gamma), \gamma}$  has zero mean, unit variance and a kurtosis determined by  $\gamma$ . Fig. A.2 depicts four such distributions:  $\gamma = 1$  (double exponential distribution),  $\gamma = 2$  (normal distribution),  $\gamma = 4$  (power distribution), and  $\gamma = 100$  (nearly uniform distribution). The distributions with  $\gamma < 2$  have larger kurtosis than three and are more peaked than the normal distribution, i.e., leptokurtic. The distributions with  $\gamma > 2$  have smaller kurtosis than three and are less peaked than the normal distribution, i.e., platykurtic.

# Appendix B

## Influence of the $T_s$ -estimation

In Sec. 6.4, the ABL wind speed increment series  $(u_{s;n})_{n=1}^N$  with increment length  $s$  is analysed with respect to superstatistics in order to check whether the intermittency hypothesis (6.18) is satisfied. This appendix discusses the influence of the estimated value  $T_s$  on the outcome of the superstatistical analysis. Hence, the time scale  $T_s$  is treated as a free parameter and not estimated by the Queiros (2007) method. The rest of the analysis remains the same, i.e.,

1. estimate the  $\beta$ -series by

$$\underbrace{(u_{s;1}, \dots, u_{s;T_s})}_{1/\text{Var}\{\dots\}=\beta_1}, \underbrace{(u_{s;T_s+1}, \dots, u_{s;2T_s})}_{1/\text{Var}\{\dots\}=\beta_2}, \dots, \underbrace{(u_{s;(K-1)T_s+1}, \dots, u_{s;N})}_{1/\text{Var}\{\dots\}=\beta_K}, \quad (\text{B.1})$$

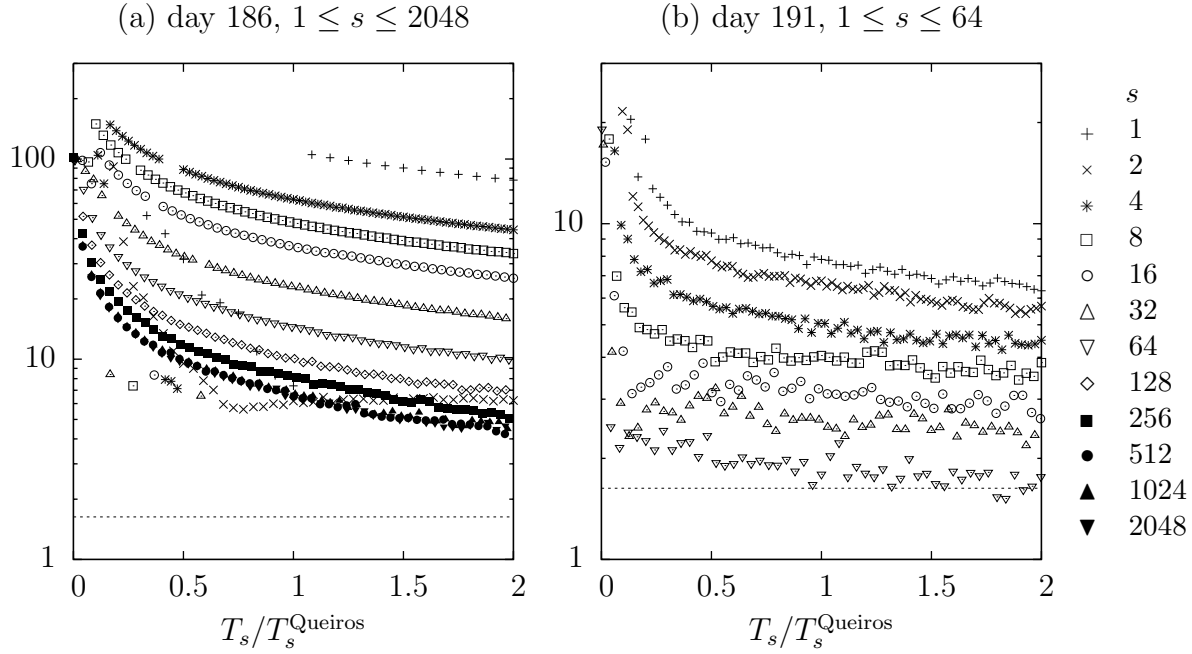
2. estimate the cdf of  $\Lambda = \ln \beta$  by treating the wind speed data as being dimensionless and
3. perform a Kolmogorov-Smirnov normality test on the distribution of  $\Lambda$  to check hypothesis

$$H_0 : \begin{cases} \Lambda \sim N_{\Lambda_s, \lambda_s^2} & \text{for all } s \text{ and} \\ \lambda_s^2 \rightarrow 0 & \text{for large } s \end{cases} \quad (\text{B.2})$$

with  $\Lambda_s$  and  $\lambda_s^2$  denoting the estimated mean and variance, respectively, of the  $\Lambda$ -series.

In Sec. 6.4 the time scale  $T_s$  is estimated by Queiros (2007). It is found that for the three analysed 24 h recordings the hypothesis is not satisfied, whereas it is on some of the 2 h recordings. The questions, which this appendix tries to answer, are

- Does the outcome of the Kolmogorov-Smirnov test change for a  $T_s \neq T_s^{\text{Queiros}}$ .<sup>22</sup>
- If not: Does the estimation of the parameters  $\Lambda_s$  and  $\lambda_s^2$  depend on the value of  $T_s$ ?
- Are the estimated parameters  $\Lambda_s$  and  $\lambda_s^2$  (nearly) identical to the estimated position and shape parameter, respectively, of the corresponding symmetric Castaing distribution when setting  $T_s = T_s^{\text{Queiros}}$ .



$s$	$T_s^{\text{Queiros}}$	$s$	$T_s^{\text{Queiros}}$
1	12	64	583
2	22	128	1,099
4	36	256	1,584
8	59	512	2,176
16	98	1024	2,063
32	240	2048	2,048

$s$	$T_s^{\text{Queiros}}$	$s$	$T_s^{\text{Queiros}}$
1	30	64	1,679
2	42	128*	97,445
4	67	256*	486,587
8	116	512*	614,283
16	223	1024*	656,959
32	459	2048*	665,541

**Figure B.1:** The test value of the Kolmogorov-Smirnov normality test for the  $\Lambda$ -series as a function of  $T_s$ . The dashed line reflects the critical value for a significance level of 1%. The size of the corresponding  $\Lambda$ -series is roughly  $6.9 \times 10^5 / T_s$ . The table list the value of  $T_s^{\text{Queiros}}$  as a function of  $s$ . The star marks the value of  $s$  for which  $T_s^{\text{Queiros}}$  is so large that the corresponding increment series is not regarded as being intermittent. The results for those  $s$  are not plotted. (a) Day 186 data for  $1 \leq s \leq 2048$ . (b) Day 191 data for  $1 \leq s \leq 64$ .



Analysing the increments of two 24 h recordings gathered at the [Lammefjord \(1987\)](#) site (day 186 and 191), [Fig. B.1](#) shows that the outcome of the Kolmogorov-Smirnov test, i.e., the decision whether to reject the hypothesis  $\Lambda \sim N_{\Lambda_s, \lambda_s^2}$ , does not depend on  $T_s$ . Panel (a) depicts the test variable as a function of  $T_s/T_s^{\text{Queiros}}$  for a variety of  $s$  for the day 186 data. Panel (b) depicts the same for a smaller variety of  $s$  for the day 191 data.<sup>23</sup> In both cases the test suggests to reject the hypothesis on a significance level of 1% for any value of  $T_s$  between 0 and  $2T_s^{\text{Queiros}}$ . The results of [Sec. 6.4](#) indicates that the hypothesis  $H_0$  might be satisfied on a 2 h recording when setting  $T_s = T_s^{\text{Queiros}}$ . For instance, the day 186 data show in [Fig. 6.20](#) that the hypothesis cannot be rejected between 12–14 h, whereas it is rejected between 20–22 h for small  $s$ . The decision of these two 2 h recordings is analysed again as a function of  $T_s$  and shown in [Fig. B.2](#). It yields nearly the same outcome for  $0 \leq T_s \leq 2T_s^{\text{Queiros}}$ . That is, if the hypothesis is rejected/not rejected for  $T_s = T_s^{\text{Queiros}}$ , it is (mostly<sup>24</sup>) rejected/not rejected for  $T_s \neq T_s^{\text{Queiros}}$ . Even though this is not a comprehensive analysis, it is assumed that the value of  $T_s$  does not have a strong influence on the decision whether to reject/accept the hypothesis.

The second question is then: Does the value of  $T_s$  have an influence of the estimated parameters  $\Lambda_s$  and  $\lambda_s^2$ ? [Fig. B.3](#), which depicts the mean  $\Lambda_s$  and variance  $\lambda_s^2$  of the estimated  $\Lambda$ -series as a function of  $T_s$  exemplarily for the day 186 increments (between 12–14 h), suggests that with increasing  $T_s$  the estimation of  $\Lambda_s$  and  $\lambda_s^2$  decreases. However, the  $T_s$ -dependence of  $\lambda_s^2$  is much stronger than the  $T_s$ -dependence of  $\Lambda_s$ . This behaviour is observed in other 2 h recordings as well. In other words, the value of  $T_s$  has a small impact on the estimation of  $\Lambda_s$  and a large impact on the estimation of  $\lambda_s^2$ .

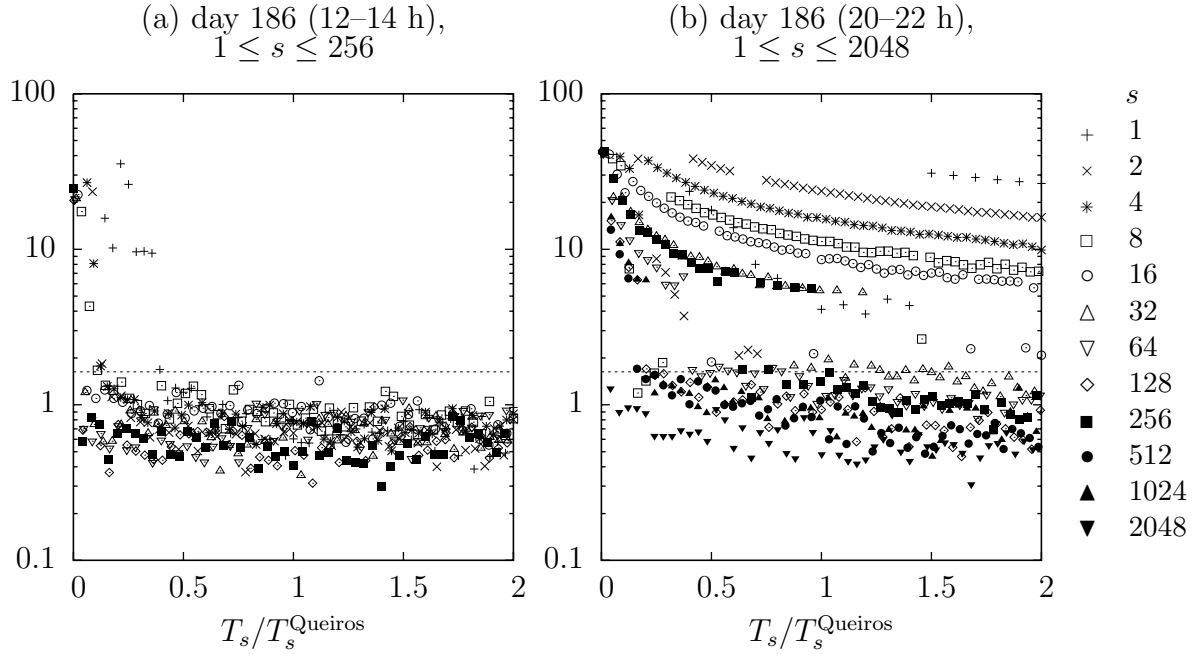
So, the third question is whether the time scale  $T_s^{\text{Queiros}}$  estimated by the [Queiros \(2007\)](#) method yields an estimation of  $\Lambda_s$  and  $\lambda_s^2$  which is (nearly) identical to the position and shape parameter, respectively, of the corresponding symmetric Castaing distribution. That means that the estimation of  $\Lambda_s$  and  $\lambda_s^2$  for the twelve 2 h recordings of a 24 h series by using the [Queiros \(2007\)](#) and [Castaing et al. \(1990\)](#) method are compared. In this analysis the increment length is varied between 1 and 2048.

**Day 186 recordings in [Fig. B.4](#).** The 2 h recordings for which the superstatistical analysis in [Fig. 6.20](#) suggests to accept/reject the hypothesis  $H_0$  are marked with  $\bullet/\times$ . It is evident that  $\Lambda_s$  estimated by the [Queiros \(2007\)](#) method is nearly identical to  $\Lambda_s$  estimated by the [Castaing et al. \(1990\)](#) method. On the other hand, there are some deviations between the two  $\lambda_s^2$ s estimated by the two methods. The deviation for the 2 h recordings which apparently satisfy the hypothesis  $H_0$  is generally smaller than the deviation for the recordings in which  $H_0$  is rejected.

<sup>22</sup>  $T_s^{\text{Queiros}}$  denotes the time scale  $T_s$  which is obtained by the [Queiros \(2007\)](#) method.

<sup>23</sup> The smaller variety is chosen because the test only makes sense if the size, which is approximately  $6.9 \times 10^5/T_s$ , of the  $\Lambda$  series is sufficiently large. As shown in [Fig. 6.16](#), the time scale  $T_s^{\text{Queiros}}$ , around which  $T_s$  is varied, increases with  $s$  so that the size of the  $\Lambda$ -series becomes too small when changing  $T_s$  up to  $2T_s^{\text{Queiros}}$ . In fact, an estimated  $T_s$ , which is very large, means that the series is not very intermittent, i.e.,  $\lambda_s^2$  is very small. This in turn makes the test  $\Lambda \sim N_{\Lambda_s, \lambda_s^2}$  redundant.

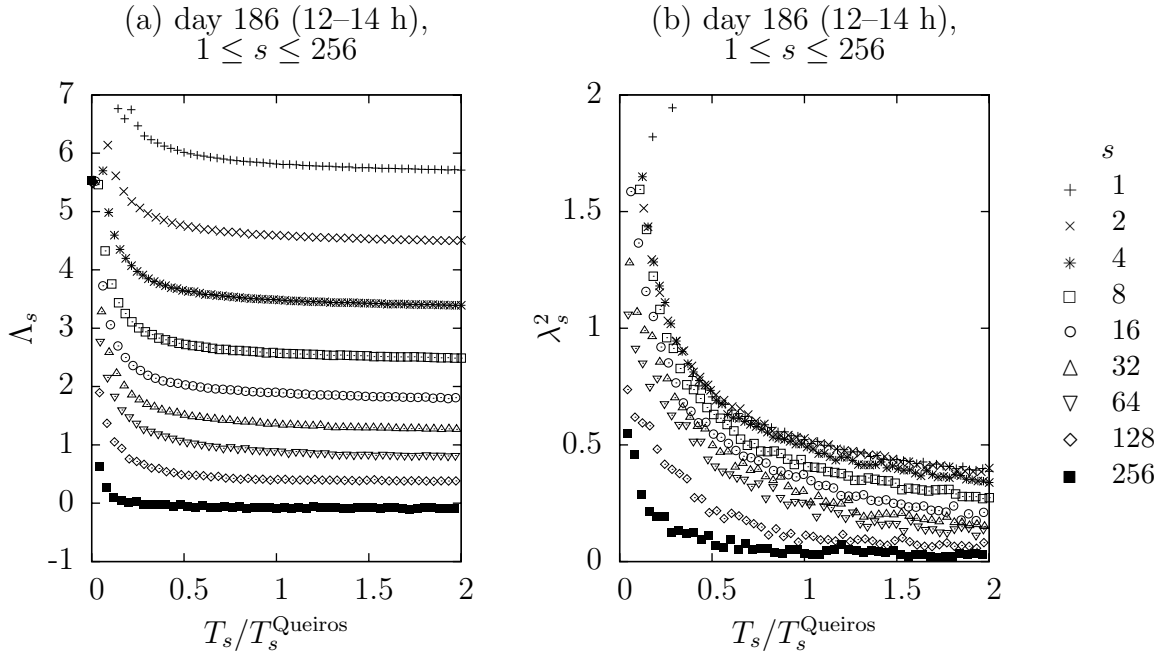
<sup>24</sup> [Fig. B.2](#) (b) shows some cases where the hypothesis is rejected for  $T_s \lesssim T_s^{\text{Queiros}}$  and accepted for  $T_s \gtrsim T_s^{\text{Queiros}}$ .



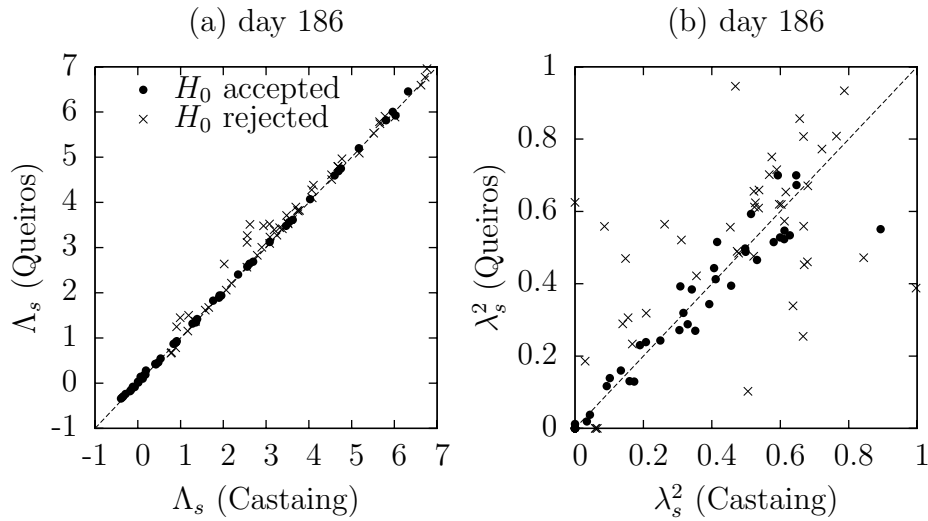
$s$	$T_s^{\text{Queiros}}$	$s$	$T_s^{\text{Queiros}}$
1	28	64	561
2	46	128	1,543
4	65	256	7,823
8	110	512*	57,089
16	197	1024*	56,577
32	318	2048*	55,553

$s$	$T_s^{\text{Queiros}}$	$s$	$T_s^{\text{Queiros}}$
1	10	64	518
2	24	128	879
4	47	256	255
8	79	512	1160
16	112	1024	786
32	256	2048	3750

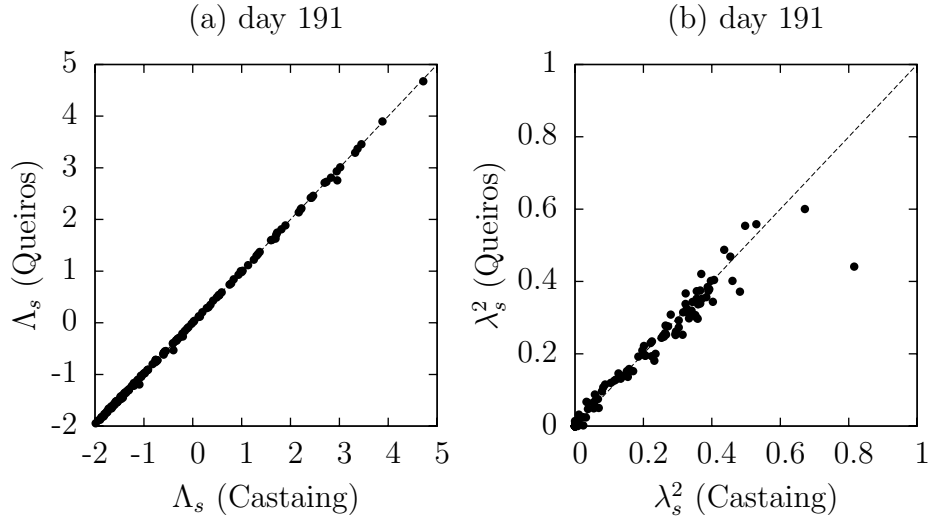
**Figure B.2:** The test value of the Kolmogorov-Smirnov normality test for the  $\Lambda$ -series as a function of  $T_s$ . The dashed line reflects the critical value for a significance level of 1%. The size of the corresponding  $\Lambda$ -series is roughly  $57600/T_s$ . The table list the value of  $T_s^{\text{Queiros}}$  as a function of  $s$ . The star marks the value of  $s$  for which  $T_s^{\text{Queiros}}$  is so large that the corresponding increment series is not regarded as being intermittent. The results for those  $s$  are not plotted. **(a)** Day 186 data (12–14 h,  $H_0$  is satisfied) for  $1 \leq s \leq 256$ . **(b)** Day 186 data (20–22 h,  $H_0$  is mostly rejected) for  $1 \leq s \leq 2048$ .



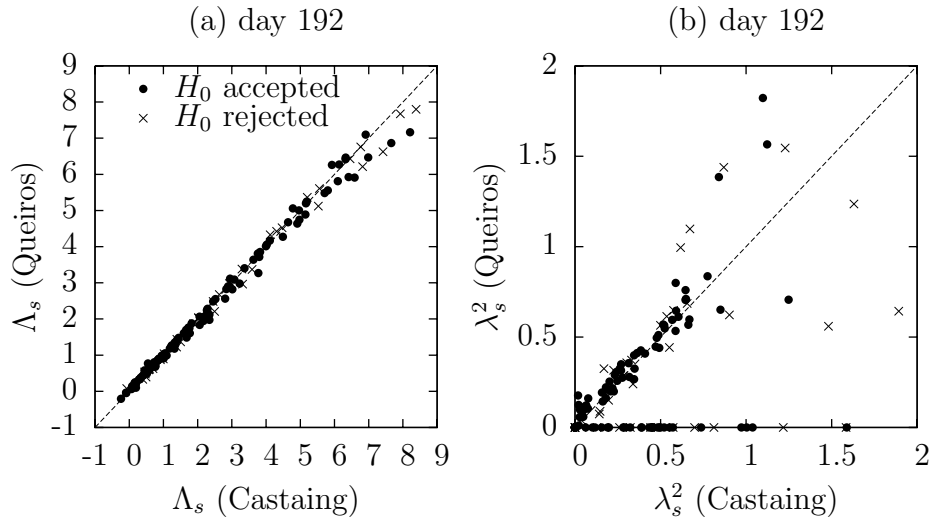
**Figure B.3:** Estimation of the parameters  $\Lambda_s$  and  $\lambda_s^2$  as a function of the value of  $T_s$  for the day 186 (12–14 h) data.



**Figure B.4:** Comparison between the parameters estimated by the [Queiros \(2007\)](#) method and those estimated by the [Castaing \*et al.\* \(1990\)](#) method. The methods are applied to 2 h recordings of the day 186 data. The values obtained from the recording, for which  $H_0$  is accepted/rejected, are marked with  $\bullet/\times$ . The increment length is varied between 1 and 2048.



**Figure B.5:** Comparison between the parameters estimated by the [Queiros \(2007\)](#) method and those estimated by the [Castaing \*et al.\* \(1990\)](#) method. The methods are applied to 2 h recordings of the day 191 data. The increment length is varied between 1 and 2048.



**Figure B.6:** Comparison between the parameters estimated by the [Queiros \(2007\)](#) method and those estimated by the [Castaing \*et al.\* \(1990\)](#) method. The methods are applied to 2 h recordings of the day 192 data. The values obtained from the recording, for which  $H_0$  is accepted/rejected, are marked with  $\bullet/\times$ . The increment length is varied between 1 and 2048.

**Day 191 recordings in Fig. B.5.** It is evident that  $\Lambda_s$  estimated by the [Queiros \(2007\)](#) method is nearly identical to  $\Lambda_s$  estimated by the [Castaing \*et al.\* \(1990\)](#) method. On the other hand, there are some deviations between the two  $\lambda_s^2$ 's estimated by the two methods.

**Day 192 recordings in Fig. B.6.** The 2 h recordings for which the superstatistical analysis in Fig. 6.22 suggests to accept/reject the hypothesis  $H_0$  are marked with  $\bullet/\times$ . It is evident that  $\Lambda_s$  estimated by the [Queiros \(2007\)](#) method is nearly identical to  $\Lambda_s$  estimated by the [Castaing \*et al.\* \(1990\)](#) method. On the other hand, there are some deviations between the two  $\lambda_s^2$ 's estimated by the the two methods. The deviation for the 2 h recordings which apparently satisfy the hypothesis  $H_0$  are not necessarily smaller than the deviation for the recordings in which  $H_0$  is rejected. Additionally, there are some  $\lambda_s^2$  which by the [Queiros \(2007\)](#) method are estimated to be zero and which by the [Castaing \*et al.\* \(1990\)](#) method are estimated to be positive.

Exemplified by these three time series, it is shown that the [Queiros \(2007\)](#) estimation of  $\Lambda_s$  is acceptable. This is because the estimation of  $\Lambda_s$  does not strongly depend on the value of  $T_s$ . On the other hand, the estimation of  $\lambda_s^2$  can (but does not have to) deviate from the shape parameter of the corresponding symmetric Castaing distribution.

As a conclusion, the [Queiros \(2007\)](#) estimation of the time scale  $T_s$  is reasonable but not perfect. Its value does not have major influence on the decision whether to reject the turbulence hypothesis in form of  $H_0$  or not. It does however have an impact on the estimation of  $\lambda_s^2$ , i.e., on the strength of intermittency.

# Appendix C

## Derivation of Fluctuation Statistics

### C.1 Stationary Gaussian processes

Recall that the Gaussian stochastic process  $\mathbf{X} = \{X_n : n \in \mathbb{Z}\}$  has the following properties:

$$X_n \sim N_{\mu, \sigma^2}, \quad (\text{C.1a})$$

$$\bar{X}_n^m \sim N_{\mu, \sigma^2 \theta_m^2}, \quad \text{and} \quad (\text{C.1b})$$

$$\bar{X}_n^m |_{X_n=x_n} \sim N_{\phi_m x_n + \mu(1-\phi_m), \sigma^2(\theta_m^2 - \phi_m^2)} \quad (\text{C.1c})$$

with

$$\theta_m^2 = \frac{1}{m^2} \sum_{k, k'=-\tilde{m}}^{\tilde{m}} \gamma(k - k') \quad \text{and} \quad \phi_m = \frac{1}{m} \sum_{k=-\tilde{m}}^{\tilde{m}} \gamma(k), \quad (\text{C.2})$$

where  $\gamma(k)$  denotes the auto correlation function. Consequently, Eq. (8.3) yields

$$p_x(x_n | \bar{X}_n^m = \bar{x}_n) = \frac{N_{\phi_m x_n + \mu(1-\phi_m), \sigma^2(\theta_m^2 - \phi_m^2)}(\bar{x}_n) N_{\mu, \sigma^2}(x_n)}{N_{\mu, \sigma^2 \theta_m^2}(\bar{x}_n)} \quad (\text{C.3})$$

Defining  $z_n = (x_n - \mu)/\sigma$  and  $\bar{z}_n = (\bar{x}_n - \mu)/\sigma$ , the numerator reads

$$\begin{aligned} & N_{\phi_m x_n + \mu(1-\phi_m), \sigma^2(\theta_m^2 - \phi_m^2)}(\bar{x}_n) N_{\mu, \sigma^2}(x_n) \\ &= \frac{1}{2\pi\sigma^2 \sqrt{\theta_m^2 - \phi_m^2}} \exp \left[ -\frac{(\bar{z}_n - \phi_m z_n)^2}{2(\theta_m^2 - \phi_m^2)} - \frac{z_n^2}{2} \right] \\ &= \frac{1}{2\pi\sigma^2 \sqrt{\theta_m^2 - \phi_m^2}} \exp \left[ -\frac{\theta_m^2}{2(\theta_m^2 - \phi_m^2)} \left( z_n - \frac{\bar{z}_n \phi_m}{\theta_m^2} \right)^2 \right] \exp \left[ -\frac{\bar{z}_n^2}{2\theta_m^2} \right] \\ &= \frac{1}{\sigma^2} N_{0, \frac{\theta_m^2 - \phi_m^2}{\theta_m^2}} \left( z_n - \frac{\bar{z}_n \phi_m}{\theta_m^2} \right) N_{0, \theta_m^2}(\bar{z}_n) \\ &= N_{\mu + (\bar{x}_n - \mu) \frac{\phi_m}{\theta_m^2}, \sigma^2 \frac{\theta_m^2 - \phi_m^2}{\theta_m^2}}(x_n) N_{\mu, \sigma^2 \theta_m^2}(\bar{x}_n) \end{aligned} \quad (\text{C.4})$$

concluding that  $X_n$  conditioned on  $\bar{X}_n^{(m)} = \bar{x}_n$  is normally distributed with

$$\mathbb{E} [X_n | \bar{X}_n^{(m)} = \bar{x}_n] = \mu + (\bar{x}_n - \mu) \frac{\phi_m}{\theta_m^2} \quad (\text{C.5a})$$

and

$$\text{Var} [X_n | \bar{X}_n^{(m)} = \bar{x}_n] = \sigma^2 \frac{\theta_m^2 - \phi_m^2}{\theta_m^2}. \quad (\text{C.5b})$$

Regarding the fluctuation  $F_n^{(m)}$  conditioned on  $\bar{X}_n^{(m)} = \bar{x}_n$ , it is normally distributed with

$$\mathbb{E} [F_n^{(m)} | \bar{X}_n^{(m)} = \bar{x}_n] = (\bar{x}_n - \mu) \frac{\phi_m - \theta_m^2}{\theta_m^2} \quad (\text{C.6a})$$

and

$$\text{Var} [F_n^{(m)} | \bar{X}_n^{(m)} = \bar{x}_n] = \sigma^2 \frac{\theta_m^2 - \phi_m^2}{\theta_m^2} \quad (\text{C.6b})$$

due to the linear transformation rule (A.9) of the normal distribution.

## C.2 $\chi^2$ -distributed white noise

The  $\chi^2$ -distributed white noise is defined by  $\mathbf{X} = \{X_n : n \in \mathbb{Z}\}$  with  $X_n \sim \chi_k^2$ , where  $k$  denotes the degrees of freedom. The moving average over  $m = 2\tilde{m} + 1$  ( $\tilde{m} \in \mathbb{N}_0$ ) time steps reads

$$\bar{X}_n^{(m)} = \frac{1}{m} \underbrace{\sum_{k=-\tilde{m}}^{\tilde{m}} X_{n+k}}_{\sim \chi_{m k}^2}. \quad (\text{C.7})$$

The sum of  $\chi^2$ -distributed rv's is again  $\chi^2$ -distributed by adding the degrees of freedom. The distribution of  $\bar{X}_n^{(m)}$  is given by

$$\bar{p}_n^{(m)}(\bar{x}_n) = m \chi_{m k}^2(m \bar{x}_n), \quad (\text{C.8})$$

where the linear transformation rule (2.21) is used. Furthermore, the conditioned distribution of the moving average

$$\bar{X}_n^{(m)} \Big|_{X_n=x_n} = \frac{x_n}{m} + \frac{1}{m} \underbrace{\sum_{k=1}^{\tilde{m}} [X_{n-k} + X_{n+k}]}_{\sim \chi_{(m-1)k}^2} \quad (\text{C.9})$$

is obtained by again using (2.21) and reads

$$\bar{p}_n^{(m)}(\bar{x}_n | X_n = x_n) = m \chi_{(m-1)k}^2(m\bar{x}_n - x_n) \quad (\text{C.10})$$

with  $0 \leq x_n \leq m\bar{x}_n$ . Inserting these distributions into (8.3) yields after some algebra

$$p_n(x_n | \bar{X}_n^{(m)} = \bar{x}_n) = \frac{A_{m,k}}{m\bar{x}_n} \left[ \left( \frac{x_n}{m\bar{x}_n} \right)^{k-2} \left( 1 - \frac{x_n}{m\bar{x}_n} \right)^{(m-1)k-2} \right]^{1/2} \quad (\text{C.11})$$

with normalisation constant

$$A_{m,k} = \frac{\Gamma\left(\frac{mk}{2}\right)}{\Gamma\left(\frac{k}{2}\right) \Gamma\left(\frac{mk}{2} - \frac{k}{2}\right)}. \quad (\text{C.12})$$

According to (8.4), the distribution of the fluctuation  $F_n^{(m)}$  conditioned on  $\bar{X}_n^{(m)} = \bar{x}_n$  is given by applying the linear transformation rule (2.21). Its mean and variance are given by (8.5) and are related to the mean and variance of  $p_n(x_n | \bar{X}_n^{(m)} = \bar{x}_n)$ . Therefore, the moments  $\mathbb{E} \left[ X_n^r | \bar{X}_n^{(m)} = \bar{x}_n \right]$  for  $r \in \{0, 1, 2\}$  need to be computed. They are given by

$$\int_0^{m\bar{x}_n} dx_n x_n^r p_n(x_n | \bar{X}_n^{(m)} = \bar{x}_n) = A_{m,k} C_{r,m,k} \times m^r \bar{x}_n^r \quad (\text{C.13})$$

with

$$C_{r,m,k} = \int_0^1 dt t^r [t^{k-2} (1-t)^{(m-1)k-2}]^{1/2}. \quad (\text{C.14})$$

The integral  $C_{r,m,k}$  can be evaluated as follows. Setting  $r = 0$  in (C.13) leads to  $1 = A_{m,k} C_{0,m,k}$  and thus  $C_{0,m,k} = 1/A_{m,k}$ . Noting that  $C_{r,m,k} = C_{0,m'_r,k'_r}$  with  $m'_r = 1 + (m-1)k/k'_r$  and  $k'_r = k + 2r$ , the integral  $C_{r,m,k}$  reads

$$C_{r,m,k} = \frac{\Gamma\left(\frac{k}{2} + r\right) \Gamma\left(\frac{mk}{2} - \frac{k}{2}\right)}{\Gamma\left(\frac{mk}{2} + r\right)}. \quad (\text{C.15})$$

Using the properties of the  $\Gamma$ -function, a convenient formulation for  $r \geq 1$  can be found by

$$C_{r,m,k} = C_{0,m,k} \prod_{l=0}^{r-1} \frac{\frac{k}{2} + l}{\frac{mk}{2} + l} \quad (\text{C.16})$$

so that

$$\mathbb{E} \left[ X_n^r | \bar{X}_n^{(m)} = \bar{x}_n \right] = m^r \bar{x}_n \frac{C_{r,m,k}}{C_{0,m,k}} = m^r \bar{x}_n \prod_{l=0}^{r-1} \frac{k + 2l}{mk + 2l} \quad (\text{C.17})$$

As a consequence of (8.5), the mean and variance of the conditioned fluctuation read

$$\mathbb{E} \left[ F_n^{(m)} | \bar{X}_n^{(m)} = \bar{x}_n \right] = 0 \quad \text{and} \quad \text{Var} \left[ F_n^{(m)} | \bar{X}_n^{(m)} = \bar{x}_n \right] = 2\bar{x}_n^2 \frac{m-1}{mk+2}. \quad (\text{C.18})$$



That is, it has vanishing expectation value, whereas its standard deviation is proportional to  $\bar{x}_n$ . The proportional factor is given by

$$\alpha(m) = \sqrt{\frac{2(m-1)}{mk+2}} \quad (\text{C.19})$$

which for large  $m$  tends to  $\sqrt{2/k}$ .

### C.3 Geometric AR(1) process

Inserting the definition of the log-normal distribution, the numerator of

$$p_n(x_n | \bar{X}_n^{(m)} = \bar{x}) = \frac{L_{x_n, \sigma_Z^2}(\bar{x}_n) L_{u, \lambda^2}(x_n)}{L_{u, \lambda^2 + \sigma_Z^2}(\bar{x}_n)} \quad (\text{C.20})$$

can be written as

$$L_{x_n, \sigma_Z^2}(\bar{x}_n) L_{u, \lambda^2}(x_n) = \frac{1}{2\pi\sigma_Z\lambda x\bar{x}_n} \exp\left[-\frac{1}{2\sigma_Z^2} \ln^2 \frac{\bar{x}_n}{x_n} - \frac{1}{2\lambda^2} \ln^2 \frac{x_n}{u}\right]. \quad (\text{C.21})$$

The exponent can be written as

$$\begin{aligned} & -\frac{1}{2\sigma_Z^2} \ln^2 \frac{x_n}{\bar{x}_n} - \frac{1}{2\lambda^2} \ln^2 \frac{x_n}{u} \\ &= -\frac{1}{2\sigma_Z^2} \underbrace{\ln^2\left(\frac{x_n}{u} \frac{u}{\bar{x}_n}\right)}_{[\ln(x_n/u) - \ln(\bar{x}_n/u)]^2} - \frac{1}{2\lambda^2} \ln^2 \frac{x_n}{u} \\ &= -\frac{1}{2\sigma_Z^2\lambda^2} \left[ (\lambda^2 + \sigma_Z^2) \ln^2 \frac{x_n}{u} - 2\lambda^2 \ln\left(\frac{x_n}{u}\right) \ln\left(\frac{\bar{x}_n}{u}\right) + \lambda^2 \ln^2 \frac{\bar{x}_n}{u} \right] \\ &= -\frac{\lambda^2 + \sigma_Z^2}{2\sigma_Z^2\lambda^2} \left[ \ln^2 \frac{x_n}{u} - 2\frac{\lambda^2}{\lambda^2 + \sigma_Z^2} \ln\left(\frac{x_n}{u}\right) \ln\left(\frac{\bar{x}_n}{u}\right) + \frac{\lambda^2}{\lambda^2 + \sigma_Z^2} \ln^2 \frac{\bar{x}_n}{u} \right] \\ &= -\frac{\lambda^2 + \sigma_Z^2}{2\sigma_Z^2\lambda^2} \left[ \left( \ln \frac{x_n}{u} - \frac{\lambda^2}{\lambda^2 + \sigma_Z^2} \ln \frac{\bar{x}_n}{u} \right)^2 + \frac{\lambda^2\sigma_Z^2}{(\lambda^2 + \sigma_Z^2)^2} \ln^2 \frac{\bar{x}_n}{u} \right] \\ &= -\frac{\lambda^2 + \sigma_Z^2}{2\sigma_Z^2\lambda^2} \ln^2 \left[ \frac{x_n}{u} \left( \frac{u}{\bar{x}_n} \right)^{\frac{\lambda^2}{\lambda^2 + \sigma_Z^2}} \right] - \frac{1}{2(\lambda^2 + \sigma_Z^2)} \ln^2 \frac{\bar{x}_n}{u}. \end{aligned} \quad (\text{C.22})$$

Defining

$$u_{\text{cond}} = u \left( \frac{\bar{x}_n}{u} \right)^{\frac{\lambda^2}{\lambda^2 + \sigma_Z^2}} \quad \text{and} \quad \lambda_{\text{cond}}^2 = \lambda^2 \frac{\sigma_Z^2}{\lambda^2 + \sigma_Z^2}, \quad (\text{C.23})$$

the exponent becomes

$$-\frac{1}{2\sigma_Z^2} \ln^2 \frac{x_n}{\bar{x}_n} - \frac{1}{2\lambda^2} \ln^2 \frac{x_n}{u} = \frac{1}{2\lambda_{\text{cond}}^2} \ln^2 \frac{x_n}{u_{\text{cond}}} - \frac{1}{2(\lambda^2 + \sigma_Z^2)} \ln^2 \frac{\bar{x}_n}{u} \quad (\text{C.24})$$

so that the numerator reads

$$\begin{aligned} & \frac{1}{2\pi\sigma_Z\lambda x\bar{x}_n} \exp\left[-\frac{1}{2\sigma_Z^2} \ln^2 \frac{\bar{x}_n}{x_n} - \frac{1}{2\lambda^2} \ln^2 \frac{x_n}{u}\right] \\ & = L_{u_{\text{cond}}, \lambda_{\text{cond}}^2}(x_n) L_{u, \lambda^2 + \sigma_Z^2}(\bar{x}_n). \end{aligned} \quad (\text{C.25})$$

As a consequence,

$$p_n(x_n | \bar{X}_n^{(m)} = \bar{x}) = L_{u_{\text{cond}}, \lambda_{\text{cond}}^2}(x_n). \quad (\text{C.26})$$

## C.4 Symmetry of the stationary AR(1) process

The stationary AR(1) process  $\mathbf{X} = \{X_n : n \in \mathbb{Z}\}$  has an important symmetry: the distribution of  $X_{n-k}$  conditioned on  $X_n = x_n$  is the same as the distribution of  $X_{n-k}$  conditioned on  $X_n = x_n$  for any  $k \in \mathbb{N}$ .

This is shown as follows. The stationary AR(1) process is defined by

$$X_{n+1} = aX_n + \sigma\xi_n\sqrt{1-a^2} \quad (\text{C.27})$$

with  $|a| < 1$ ,  $\sigma^2 > 0$ , and  $\{\xi_n : n \in \mathbb{Z}\}$  denoting standard Gaussian white noise. The distribution of  $X_{n+1}$  conditioned on  $X_n = x_n$  can directly be read off from Eq. (C.27). It is given by

$$p_{n+1}(x_{n+1} | X_n = x_n) = N_{ax_n, \sigma^2(1-a^2)}(x_{n+1}). \quad (\text{C.28})$$

Due to stationary, this distribution is equivalent to the distribution of  $X_n$  conditioned on  $X_{n-1} = x_{n-1}$ . Hence, the distribution of  $X_{n-1}$  conditioned on  $X_n = x_n$  reads

$$\begin{aligned} p_{n-1}(x_{n-1} | X_n = x_n) & \stackrel{(2.39)}{=} \frac{p_n(x_n | X_{n-1} = x_{n-1}) p_{n-1}(x_{n-1})}{p_n(x_n)} \\ & \stackrel{\text{stationarity}}{=} \frac{N_{ax_{n-1}, \sigma^2(1-a^2)}(x_n) N_{0, \sigma^2}(x_{n-1})}{N_{0, \sigma^2}(x_n)}. \end{aligned} \quad (\text{C.29})$$

After a bit of algebra, this equation becomes

$$p_{n-1}(x_{n-1} | X_n = x_n) = N_{ax_n, \sigma^2(1-a^2)}(x_{n-1}) \quad (\text{C.30})$$

so that

$$X_{n-1} \Big|_{X_n=x_n} \sim X_{n+1} \Big|_{X_n=x_n}. \quad (\text{C.31})$$

Therefore, it is convenient to write

$$X_{n\pm 1} = aX_n + \sigma\eta_{n;\pm 1}\sqrt{1-a^2} \quad (\text{C.32})$$

with  $\eta_{n;\pm 1}$  being two independent  $N_{0,1}$ -distributed rv's. Iterating this equation  $k \in \mathbb{N}$  times results in

$$X_{n\pm k} = a^k X_n + \sigma\sqrt{1-a^2} \sum_{l=1}^k a^{k-l} \eta_{n;\pm l}, \quad (\text{C.33})$$

where the rv's  $\eta_{n;\pm 1}, \dots, \eta_{n;\pm k}$ , each of which is  $N_{0,1}$ -distributed, are mutually independent and independent from  $X_n$ .

# Bibliography

- ABE, S. & THURNER, S. 2005 Complex networks emerging from fluctuating random graphs: Analytic formula for the hidden variable distribution. *Phys. Rev. E* **72**, 036102.
- ABRAMOWITZ, M. & STEGUN, I. A. 1964 *Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables*. Dover.
- ABUL-MAGD, A.Y. 2006 Superstatistics in random matrix theory. *Physica A* **361**, 41–54.
- AL-HUSSAINI, E.K. & TURNER, L.F. 1979 Asymptotic performance of 2-sample nonparametric detectors when detecting non-fluctuating signals in non-gaussian noise performance of 2-sample nonparametric detectors when detecting non-fluctuating signals in non-gaussian noise. *IEEE Trans. Inf. Theory* **25**, 124–127.
- ANSELMET, F., GAGNE, Y., HOPFINGER, E.J. & ANTONIA, R.A. 1984 High-order velocity structure functions in turbulent shear flows. *J. Fluid Mech.* **140**, 63–89.
- AUSLOOS, M. & IVANOVA, K. 2003 Dynamical model and nonextensive statistical mechanics of a market index on large time windows. *Phys. Rev. E* **68**, 046122.
- AZZALINI, A. 1985 A class of distributions which includes the normal ones. *Scand. J. Statist.* **12**, 171–178.
- BAAQUIE, B.E. 2004 *Quantum Finance*. Cambridge University Press.
- BAIESI, M., PACZUSKI, M. & STELLA, A.L. 2006 Intensity thresholds and the statistics of the temporal occurrence of solar flares. *Phys. Rev. Lett.* **96**, 051103.
- BATCHELOR, G.K. & TOWNSEND, A.A. 1949 The nature of turbulent motion at large wave-numbers. *Proc. R. Soc. London A* .
- BEAULIEU, N. C., ABUDAYYA, A. A. & MCLANE, P. J. 1993 On approximating the distribution of a sum of independent lognormal random-variables. In *IEEE WESCANEX 93: Communications, Computers and Power in the Modern Environment*, pp. 72–79. Saskatoon, Canada.
- BECK, C. 2004a Generalized statistical mechanics of cosmic rays. *Physica A* **331**, 173–181.

- BECK, C. 2004*b* Superstatistics in hydrodynamic turbulence. *Physica D* **193**, 195–207.
- BECK, C. 2004*c* Superstatistics: theory and applications. *Continuum mechanics and thermodynamics* **16** (3), 293 – 304.
- BECK, C. 2007 Statistics of three-dimensional lagrangian turbulence. *Phys. Rev. Lett.* **98**, 064502.
- BECK, C. & COHEN, E.G.D. 2003 Superstatistics. *Physica A* **322**, 267–275.
- BECK, C., COHEN, E.G.D. & RIZZO, S. 2005*a* Atmospheric turbulence and superstatistics. *Europhys. News* **36**, 181–183.
- BECK, C., COHEN, E.G.D. & SWINNEY, H.L. 2005*b* From time series to superstatistics. *Phys. Rev. E* **72**, 056133.
- BOETTCHER, F. 2005 Statistische analyse der atmosphärischen turbulenz und allgemeiner stochastischer prozesse. PhD thesis, Carl von Ossietzky Universität Oldenburg.
- BOETTCHER, F., BARTH, S. & PEINKE, J. 2007 Small and large scale fluctuations in atmospheric wind speeds. *Stoch. Environ. Res. Risk Assess.* **21**, 299–308.
- BOETTCHER, F., RENNER, C, WALDL, H. P. & PEINKE, J. 2003 On the statistics of wind gusts. *Bound.-Layer Meteorol.* **108**, 163–173.
- BOX, G.E.P. 1953 A note on regions for tests of kurtosis. *Biometrika* **40**, 465–468.
- BRIGGS, K. & BECK, C. 2007 Modelling train delays with q-exponential functions. *Physica A* **378**, 498–504.
- BRONSTEIN, I.N., SEMENDJAJEW, K.A., MUSIOL, G. & MÜHLIG, H. 1999 *Taschenbuch der Mathematik*. Harri Deutsch.
- BURTON, T., SHARPE, D., JENKINS, N. & BOSSANYI, E. 2004 *Wind Energy Handbook*. John Wiley.
- CARDIERI, P. & RAPPAPORT, T.S. 2000 Statistics of the sum of lognormal variables in wireless communications. In *IEEE Vehicular Technology Conference Proceedings 2000*, pp. 1823–1827. Tokyo, Japan.
- CASELLA, G. & BERGER, R.L. 2002 *Statistical Inference*. Duxbury.
- CASTAING, B., GAGNE, Y. & HOPFINGER, E. J. 1990 Velocity probability density-functions of high reynolds-number turbulence. *Physica D* **46**, 177–200.
- CHAVANIS, P.H. 2006 Coarse-grained distributions and superstatistics. *Physica A* **359**, 177–212.

- CLEVE, J., DZIEKAN, T., SCHMIEGEL, J., BARNDORFF-NIELSEN, O.E., PEARSON, B.R., SREENIVASAN, K.R. & GREINER, M. 2005 Finite-size scaling of two-point statistics and the turbulent energy cascade generators. *Phys. Rev. E* **71**, 026309.
- DANIEL, W.W. 1990 *Applied Nonparametric Statistics*. PWS-KENT.
- DANIELS, K.E., BECK, C. & BODENSCHATZ, E. 2004 Defect turbulence and generalized statistical mechanics. *Physica D* **193**, 208–217.
- DONNELLY, R.J. 1999 Cryogenic fluid dynamics. *J. Phys.-Condens. Matter* **11**, 7783–7834.
- ECKHARDT, B., FAISST, H., SCHMIEGEL, A. & SCHNEIDER, T.M. 2008 Dynamical systems and the transition to turbulence in linearly stable shear flows. *Phil. Trans. R. Soc. A* **366**, 1297–1315.
- ECKHARDT, B. & SCHNEIDER, T.M. 2008 How does flow in a pipe become turbulent? *Eur. Phys. J. B* **64**, 457–462.
- FARGE, M., KEVLAHAN, N.K.R., PERRIER, V. & SCHNEIDER, K. 1999 Turbulence analysis, modeling and computing using wavelets. In *Wavelets in Physics* (ed. J.C. van den Berg). Cambridge University Press.
- FEFFERMAN, C.F. 2000 Existence and smoothness of the navier-stokes equation, [http://www.claymath.org/millennium/navier-stokes\\_equations](http://www.claymath.org/millennium/navier-stokes_equations).
- FRISCH, URIEL 1995 *Turbulence*. Cambridge University Press.
- GRANGER, C. W. J. & NEWBOLD, P. 1976 Forecasting transformed series. *J. R. Stat. Soc. B* **38**, 189–203.
- HARTUNG 2005 *Statistik*. Oldenbourg Wissenschaftsverlag GmbH.
- VAN DER HOVEN, I. 1957 Power spectrum of horizontal wind speed in the frequency range from 0.0007 to 900 cycles per hour. *J. Meteorol.* **14**, 160–164.
- JOHNSON, N.L., KEMP, A.W. & KOTZ, S. 2005 *Univariate Discrete Distributions*. John Wiley.
- JOHNSON, N.L., KOTZ, S. & BALAKRISHAN, N. 1994 *Continuous Univariate Distributions*. John Wiley.
- JUNG, S. & SWINNEY, H.L. 2005 Velocity difference statistics in turbulence. *Phys. Rev. E* **72**, 026304.
- KANTZ, H. & SCHREIBER, T. 2004 *Nonlinear Time Series Analysis*. Cambridge University Press.

- KASSAM, S.A. & THOMAS, J.B. 1976 Asymptotically robust detection of a known signal in contaminated non-gaussian noise. *IEEE Trans. Inf. Theory* **22**, 22–26.
- KHINTCHINE, A.Y. & LÉVY, P. 1936 Sur les lois stable. *C. R. Acad. Sci. Paris* **202**, 374–376.
- KHOLMYANSKY, M., MORICONI, L. & TSINOBER, A. 2007 Large-scale intermittency in the atmospheric boundary layer. *Phys. Rev. E* **76**, 026307.
- KITAGAWA, T. & NOMURA, T. 2003 A wavelet based method to generate artificial wind fluctuation data. *J. Wind Eng. Ind. Aerodyn.* **9**, 943–964.
- KLEINHANS, D. 2008 Stochastische modellierung komplexer systeme. PhD thesis, Westfälische Wilhelms-Universität Münster.
- KLEINHANS, D. & FRIEDRICH, R. 2007 Continuous time random walks: Simulation of continuous trajectories. *Phys. Rev. E* **76**, 061102.
- KLEINHANS, D., FRIEDRICH, R., GONTIER, H. & SCHAFFARCZYK, A.P. 2006 Simulation of intermittent wind fields: A new approach. In *Proceedings of DEWEK 2006*.
- KOLMOGOROV, A. N. 1941a Dissipation of energy in locally isotropic turbulence. *Dokl. Akad. Nauk SSSR* **32**, 16–18.
- KOLMOGOROV, A. N. 1941b The local structure of turbulence in incompressible viscous fluid for very large reynolds number. *Dokl. Akad. Nauk SSSR* **30**, 299–303.
- KOLMOGOROV, A. N. 1941c On degeneration (decay) of isotropic turbulence in an incompressible viscous liquid. *Dokl. Akad. Nauk SSSR* **31**, 538–540.
- KOLMOGOROV, A. N. 1962 A refinement of previous hypotheses concerning the local structure of turbulence in a viscous incompressible fluid at high reynolds number. *J. Fluid Mech.* **13** (1), 82–85.
- LAMB, H. 1932 *Hydrodynamics*, 6th edn. Dover.
- Lammeffjord 1987 data are obtained from the Risø National Laboratory in Denmark, <http://www.risoe.dk/vea>, through <http://www.winddata.com>.
- LÉVY, P. 1925 *Calcul des Probabilités*. Gauthier-Villars.
- LIMPERT, E., STAHEL, W.A. & ABBT, M. 2001 Log-normal distributions across the sciences: Keys and clues. *BioScience* **51**, 341–352.
- LOVEJOY, S., SCHERTZER, D. & STANWAY, J.D. 2001 Direct evidence of multifractal atmospheric cascades from planetary scales down to 1 km. *Phys. Rev. Lett.* **86**, 5200–5203.

- LUMLEY, J.L. 1965 Interpretation of time spectra measured in high-intensity shear flows. *Phys. Fluids* **8**, 1056–1062.
- MANDELBROT, B. 1972 Possible refinement of the lognormal hypothesis concerning the distribution of energy dissipation in intermittent turbulence. In *Statistical Models and Turbulence* (ed. M. Rosenblatt & C. Van Atta), pp. 333–351. Springer.
- MANN, J. 1998 Wind field simulation. *Prob. Eng. Mech* **13**, 269–282.
- MENEVEAU, C. & SREENIVASAN, K.R. 1991 The multifractal nature of turbulent energy dissipation. *J. Fluid Mech.* **224**, 429–484.
- METROPOLIS, N., ROSENBLUTH, A.W., ROSENBLUTH, M.N., TELLER, A.H. & TELLER, E. 1953 Equation of state calculations by fast computing machines. *J. Chem. Phys* **21**, 1087–1092.
- MILLER, L.H. 1956 Table of percentage points of kolmogorov statistics. *J. Amer. Statist. Assoc.* **51**, 111–121.
- MONTGOMERY, D.C., C.L., JENNINGS & KULAHCI, M. 2008 *Introduction to Time Series Analysis and Forecasting*. John Wiley.
- NAVIER, C.L.M.H. 1823 Mémoire sur les lois du mouvement des fluides. *Mem. Acad. Roy. Sci.* **6**, 389–440.
- NAWROTH, A. 2007 Stochastische Analyse und Modellierung von Finanz- und Turbulenzzeitreihen. PhD thesis, Carl von Ossietzky Universität Oldenburg.
- NEUMANN, T., ENDER, C., MOLLY, J.P., NEDDERMANN, B., WINKLER, W. & STRACK, M. 2002 Weiterer Ausbau der Windenergienutzung im Hinblick auf den Klimaschutz – Teil 2. *Tech. Rep.*. Deutsches Windenergie-Institut GmbH, [www.dewi.de](http://www.dewi.de).
- NOBACH, H., XU, H. & BODENSCHATZ, E. 2008 Göttingen high pressure turbulence facility. In *Max Planck Institute for Dynamics and Self-Organization Research Report 2008* (ed. K. Mölter). Goltze Druck.
- NOLAN, J. P. 2009 *Stable Distributions – Models for Heavy Tailed Data*. Boston: Birkhäuser, in progress, Chapter 1 online at [academic2.american.edu/~jpnolan](http://academic2.american.edu/~jpnolan).
- NOVIKOV, E.A. 1969 Scale similarity for random fields. *Sov. Phys. Dokl.* pp. 104–107.
- NOVIKOV, E.A. 1971 Intermittency and scale similarity in the structure of a turbulent flow. *Prikl. Mat. Mech.* **35**, 266–277.
- NOVIKOV, E.A. 1990 The effects of intermittency on statistical characteristics of turbulence and scale similarity of breakdown coefficients. *Phys. Fluids A* **2**, 814–820.

- NOVIKOV, E.A. & STEWART, R.W. 1964 Intermittency of turbulence and the spectrum of fluctuations of energy dissipation. *Izv. Akad. Nauk SSSR Ser. Geofiz.* **3**, 408–413.
- OBUKHOV, A.M. 1941*a* On the distribution of energy in the spectrum of turbulent flow. *Dokl. Akad. Nauk SSSR* **32**, 22–24.
- OBUKHOV, A.M. 1941*b* Spectral energy distribution in a turbulent flow. *Izv. Akad. Nauk SSSR Ser. Geogr. Geofiz.* **5**, 453–466.
- OBUKHOV, A.M. 1962 Some specific features of atmospheric turbulence. *J. Fluid Mech.* **13**, 77–81.
- PINTON, J.F. & LABBÉ, R. 1994 Correction to the Taylor hypothesis in swirling flows. *J. Phys. II France* **4**, 1461–1468.
- POPE, STEPHEN B. 2000 *Turbulent Flows*. Cambridge University Press.
- PORPORATO, A., VICO, G. & FAY, P.A. 2006 Superstatistics of hydro-climatic fluctuations and interannual ecosystem productivity. *Geophys. Res. Lett.* **33**, L15402.
- PULLEN, A., QIAO, L.M. & SAWYER, S. 2008*a* Global wind 2008 report. *Tech. Rep.*. Global Wind Energy Council, [www.gwec.net](http://www.gwec.net).
- PULLEN, A., SAWYER, S., TESKE, S. & AUBREY, C. 2008*b* Global wind energy outlook 2008. *Tech. Rep.*. Greenpeace International, Global Wind Energy Council, [www.gwec.net](http://www.gwec.net).
- QUEIROS, S. M. D. 2007 On new conditions for evaluate long-time scales in superstatistical time series. *Physica A* **385**, 191–198.
- RASCH, D. 1989 *Einführung in die Mathematische Statistik*. VEB Deutscher Verlag der Wissenschaften.
- RENNER, C., PEINKE, J. & FRIEDRICH, R. 2001 Experimental indications for Markov properties of small-scale turbulence. *J. Fluid Mech.* **433**, 383–409.
- RENZO, M.D. & GRAZIOSI, F. 2009 Approximating the linear combination of log-normal rvs via Pearson type IV distribution for UWB performance analysis. *IEEE Trans. Commun.* **57**, 388–403.
- REYNOLDS, A.M. 2003 Superstatistical mechanics of tracer-particle motions in turbulence. *Phys. Rev. Lett.* **91**, 084503.
- REYNOLDS, O. 1883 An experimental investigation of the circumstances which determine whether the motion of water shall be direct or sinuous, and of the law of resistance in parallel channels. *Phil. Trans. R. Soc. Lond.* **174**, 935–982.
- RICHARDSON, L.F. 1920 The supply of energy from and to atmospheric eddies. *Proc. R. Soc. London A* **97**, 354–373.



- RICHARDSON, L.F. 1922 *Weather Prediction by Numerical Process*. Cambridge University Press.
- RICHARDSON, L.F. 1926 Atmospheric diffusion shown on a distance-neighbour graph. *Proc. R. Soc. London A* **110**, 709–737.
- RIZZO, S. & RAPISARDIA, A. 2005 Application of superstatistics to atmospheric turbulence. In *Complexity, Metastability and Nonextensivity* (ed. C. Beck, G. Benedek, A. Rapisarda & C. Tsallis). World Scientific.
- ROSS, S.M. 1996 *Stochastic Processes*. John Wiley.
- ROSSI, R., LAZZARI, M. & VITALIANI, R. 2004 Wind field simulation for structural engineering purposes. *Int. J. Numer. Methods Eng.* **61**, 738–763.
- SAGAUT, P. & CAMBON, C. 2008 *Homogeneous Turbulence Dynamics*. Cambridge University Press.
- SATTIN, F. & SALASNICH, L. 2002 Multiparameter generalization of nonextensive statistical mechanics. *Phys. Rev. E* **65**, 035106.
- SCHMIEGEL, J., CLEVE, J., EGGERS, H.C., PEARSON, B.R. & GREINER, M. 2005 Stochastic energy-cascade model for  $(1 + 1)$ -dimensional fully developed turbulence. *Phys. Lett. A* **320**, 247–253.
- SCHNEIDER, T.M. & ECKHARDT, B. 2008 Lifetime statistics in transitional pipe flow. *Phys. Rev. E* **78**, 046310.
- SCHNEIDER, T.M., GIBSON, J.F., LAGHA, M., DE LILLO, F. & ECKHARDT, B. 2008 Laminar-turbulent boundary in plane couette flow. *Phys. Rev. E* **78**, 037301.
- SHE, Z.S. & LÉVÊQUE, E. 1994 Universal scaling laws in fully developed turbulence. *Phys. Rev. Lett.* **72**, 336–339.
- SHE, Z.S. & WAYMIRE, E.C. 1995 Quantized energy cascade and log-poisson statistics in fully developed turbulence. *Phys. Rev. Lett.* **74**, 262–265.
- SHESKIN, D.J. 1997 *Handbook of Parametric and Nonparametric Statistical Procedures*. CRC Press.
- SOBCZYK, K. 1991 *Stochastic Differential Equations*. Kluwer.
- SREENIVASAN, K.R. & ANTONIA, R.A. 1997 The phenomenology of small-scale turbulence. *Ann. Rev. Fluid Mech.* **29**, 435–472.
- STEELE, J.M. 2001 *Stochastic calculus and financial applications*. Springer.
- STUDENT 1927 Errors of routine analysis. *Biometrika* **19**, 151–164.

- STULL, R.B. 1988 *An Introduction to Boundary Layer Meteorology*. Springer.
- TAYLOR, G.I. 1938 The spectrum of turbulence. *Proc. R. Soc. London A* **164**, 476–490.
- TOUCHETTE, H. & BECK, C. 2005 Asymptotics of superstatistics. *Phys. Rev. E* **71**, 016131.
- TURNER, M.E. 1960 On heuristic estimation methods. *Biometrika* **16**, 299–301.
- VAN DYKE, M. 1982 *An Album of Fluid Motion*. The Parabolic Press.
- VEERS, P.S. 1984 Modeling stochastic wind loads on vertical axis wind turbines. *Tech. Rep.*. Sandia National Laboratories.
- VEERS, P.S. & BUTTERFIELD, S. 2001 Extreme loads estimation for wind turbines: Issues and opportunities for improved practice. *Tech. Rep.*. AIAA.
- VINCENT, A. & MENEGUZZI, M. 1991 The spatial structure and statistical properties of homogeneous turbulence. *J. Fluid Mech.* **225**, 1–20.
- WALLACE, J.M. & HOBBS, P.V. 2006 *Atmospheric Science*, 2nd edn. Academic Press Elsevier.
- WEISSTEIN, E. 2003 *CRC Concise Encyclopedia of Mathematics*. CRC Press.
- WILLIS, G.E. & DEARDORFF, J.W. 1976 On the use of Taylor's translation hypothesis for diffusion in the mixed layer. *Q. J. R. Meteorol. Soc.* **102**, 817–822.
- WITTE, R. S. & WITTE, J. W. 2001 *Statistics*. Harcourt.
- YOSIDA, K. 1980 *Functional Analysis*. Springer.

# Acknowledgements

At this place I want to thank those people who have helped me in various ways and whose support has been invaluable during the period of my postgraduate studies.

Firstly, I would like to thank my supervisor Holger Kantz for giving me the opportunity to study atmospheric wind, a topic which I am very interested in. I am very thankful for the innumerable discussions, his suggestions and support.

This study was part of the project about atmospheric turbulence and wind energy headed by Joachim Peinke. I want to thank the project members for the productive collaboration and additionally Germany's Federal Ministry of Education for funding.

I benefited from various constructive discussions and collaborations with Fatemeh Ghasemi, Julia Gottschall, Sarah Hallerberg, Max Little, David Kleinhans, Hervier Gontier, Jochen Bröcker, Rafael Vilela, Markus Niemman, and Friedrich Lenz. Therefore, many thanks go to them.

I am very grateful to the colleagues and staff of the Max Planck Institute for the Physics of Complex Systems for providing the pleasant working atmosphere and helping with practical issues.

I thank my friends and family for their encouragement and many thanks go to my two cats for their relaxing purring after stressful working days. Last, but not least, I want to say thank you to my wife Nan for her love and support, especially during the time writing the thesis. Without her I would never have got so far at all.

# Versicherung

Hiermit versichere ich, dass ich die vorliegende Arbeit ohne unzulässige Hilfe Dritter und ohne Benutzung anderer als der angegebenen Hilfsmittel angefertigt habe; die aus fremden Quellen direkt oder indirekt übernommenen Gedanken sind als solche kenntlich gemacht. Die Arbeit wurde bisher weder im Inland, noch im Ausland in gleicher oder ähnlicher Form einer anderen Prüfungsbehörde vorgelegt.

Die vorliegende Dissertation wurde am Max-Planck-Institut für Physik komplexer Systeme unter der wissenschaftlichen Betreuung von Prof. Dr. Holger Kantz angefertigt.

Es haben keine früheren Promotionsverfahren von mir stattgefunden.

Hiermit erkenne ich die Promotionsordnung des Fachbereiches Mathematik und Naturwissenschaften der Bergischen Universität Wuppertal an.

Dresden, 2. Juni 2009, Thomas Laubrich

# Index

- advection term, 39
- AR, *see* auto regressive stochastic process
- ARMA, *see* auto regressive moving average stochastic process
- auto correlation function, 29
- axioms of probability, 6
  
- binomial coefficient, 8
- body force, 38
- Boltzmann factor, 50
- Borel field, 6
- Brownian motion, *see* Wiener process
  
- cdf, *see* cumulative distribution function
- central limit theorem, 114
- characteristic function, 20
- coefficient of variation, 10
- continuity equation, 37
- continuum hypothesis, 37
- convolution, 19
- convolution theorem, 19
- correlation, 16
  - average, 17
- correlation function, 29
- correlation matrix, 16
- covariance, 16
- covariance function, 30
- cumulative distributions function, 8
  
- De Moivre–Laplace Theorem, 113
- discretisation time, 33
- distribution
  - asymmetric Castaing, 49
  - binomial, 113
  - Breit-Wigner, 117
  - Cauchy, 117
  - chi-square, 115
  - double exponential, 120
  - exponential, 118
  - Gaussian, 114
  - Lévy, 117
  - Lévy skew  $\alpha$ -stable, 22
  - leptokurtic, 11
  - log-normal, 115
  - Lorentz, 117
  - mesokurtic, 11
  - multimodal, 10
  - normal, 114
  - platykurtic, 11
  - shifted log-normal, 116
  - skew normal, 119
  - stable, 21
  - symmetric, 11
  - symmetric Castaing, 48
  - uniform, 117
  - unimodal, 10
- double factorial, 14
  
- energy dissipation rate, 41
- enstrophy, 41
- error function, 114
  - complementary, 114
- expectation value, 10
  - estimated, 23
- expected value, *see* expectation value
  
- four-fifths law, 45
  
- Gamma function, 115
- Gaussian white noise
  - geometric, 33
  - standard, 30

helicity, 41  
 histogram, 23  
 hypothesis, 24  
  
 iid, *see* independent identically distributed random variable  
 increment length, 63  
 index set, 28  
 integral length, 43  
 intermittency correction, 47  
  
 Jensen's inequality, 10  
  
 K41 theory, *see* Kolmogorov 1941 theory  
 K62 theory, *see* Kolmogorov 1962 theory  
 kinematic viscosity, 39  
 Kolmogorov 1941 theory, 44  
 Kolmogorov 1962 theory, 45  
 Kolmogorov dissipation length, 43  
 Kolmogorov's axioms, 6  
 Kolmogorov-Smirnov test, 25  
 kurtosis, 11  
  
 laminar flow, 43  
 level of significance, 24  
 linear correlation, 16  
 linear filter, 34  
 log-normal model, 47  
 longitudinal structure function, 44  
  
 MA, *see* moving average stochastic process  
 mass conservation, 37  
 median, 10  
 mgf, *see* moment generating function  
 mode, 10  
 modified pressure, 39  
 moment, 10  
     central, 10  
     raw, 10  
 moment generating function, 13  
  
 Navier-Stokes equation, 39  
     conservation laws, 41  
 NSE, *see* Navier-Stokes equation  
  
 Ornstein-Uhlenbeck process, 31  
     discrete, 34  
     geometric, 33  
  
 parameter  
     position (log-normal), 115  
     scale (exponential), 118  
     shape (log-normal), 115  
 pdf, *see* probability density function  
 periodic boundary condition, 39  
 periodicity box, 39  
 pmf, *see* probability mass function  
 power distribution, 121  
 probability, 7  
 probability density function, 8  
     conditioned, 15  
     joint, 14  
     marginal, 15  
 probability function, 6  
 probability mass function, 8  
     conditioned, 15  
     joint, 14  
     marginal, 15  
  
 random event, 6  
 random experiment, 6  
 random variable, 8  
     anti-correlated  $\sim$ s, 16  
     continuous, 8  
     correlated  $\sim$ s, 16  
     identically distributed  $\sim$ s, 9  
     independent  $\sim$ s, 15  
     independent identically distributed  $\sim$ s, 15  
     standardised binomial, 113  
     standardised chi-square, 115  
     uncorrelated  $\sim$ s, 16  
 Reynolds number, 43  
     similarity, 43  
 rv, *see* random variable  
  
 sample path, 28  
 sample space, 6  
 sigma algebra, 6  
 skewness, 11

- solenoidal vector field, 38
- standard deviation, 10
- stochastic process, 28
  - $\delta$ -correlated, 29
  - auto regressive, 36
  - auto regressive moving average, 36
  - discretised continuous, 33
  - Gaussian, 30
  - geometric AR(1), 96
  - geometric Gaussian, 32
  - moving average, 35
  - normal, 30
  - stationary of order  $p$ , 28
  - strictly stationary, 28
- stress tensor, 38
- superstatistics, 50
- surface force, 38
  
- Taylor hypothesis, 55
- TI, *see* turbulence intensity
- transformation
  - exponential, 13
  - linear, 12
- turbulence intensity, 55
- turbulent flow, 43
- two-thirds law, 45
- type I error, 24
- type II error, 25
  
- variance, 10
  - estimated, 24
- velocity increment
  - longitudinal, 44
- vorticity, 40
  
- Wiener process
  - discrete, 33
  - general, 31
  - geometric, 33
  - standard, 31