

---

# Characterization of Jump-Diffusion Stochastic Dynamics: Analysis and Applications on Real World Data

---

Von der Fakultät für Mathematik und Naturwissenschaften  
der Carl von Ossietzky Universität Oldenburg  
zur Erlangung des Grades und Titels eines

DOKTORS DER NATURWISSENSCHAFTEN

DR. RER. NAT.

angenommene Dissertation von Herrn

PYEI PHYO LIN

geboren am 10. August 1989 in Yangon, Myanmar



1. Gutachter: Prof. Dr. Joachim Peinke
  2. Gutachter: Prof. Dr. Pedro G. Lind
- Tag der Abgabe: 11. Januar 2023
- Tag der Disputation: 06. April 2023

# Contents

<b>Abstract</b>	<b>i</b>
<b>Zusammenfassung</b>	<b>iii</b>
<b>1 Introduction</b>	<b>1</b>
<b>2 Fundamentals of Stochastic Analysis</b>	<b>7</b>
2.1 Stochastic Processes and Markov Property . . . . .	8
2.2 Kramers-Moyal Expansion . . . . .	10
2.3 Fokker-Planck and Langevin Equation . . . . .	11
2.4 Wiener Process . . . . .	12
2.5 Stochastic Processes with Discontinuities . . . . .	14
2.5.1 Jump-Diffusion Process . . . . .	14
2.6 Poisson Process . . . . .	16
2.6.1 Compound Poisson Process . . . . .	18
2.7 Ornstein-Uhlenbeck Process in the Presence of Jumps . .	19
2.7.1 Numerical Simulations . . . . .	21
<b>3 Stochastic Analysis of Micro-cone Penetration Tests in Snow</b>	<b>27</b>
3.1 Introduction . . . . .	28

3.2	Stochastic Method . . . . .	31
3.2.1	Langevin Equation . . . . .	31
3.2.2	Jump-diffusion Dynamics . . . . .	32
3.2.3	Synthetic Examples . . . . .	34
3.3	Application to Snow Measurements . . . . .	37
3.3.1	Laboratory Samples . . . . .	37
3.3.2	Application to Field Snow Data . . . . .	44
3.4	Discussion . . . . .	47
3.5	Conclusions . . . . .	50
<b>4</b>	<b>Discontinuous Jump Behavior of the Energy Conversion in Wind Energy Systems</b>	<b>51</b>
4.1	Introduction . . . . .	52
4.2	Stochastic data analysis of wind energy system . . . . .	55
4.2.1	Data description . . . . .	55
4.2.2	Power Conversion Process Described by Stochastic Dynamics . . . . .	55
4.2.3	Results . . . . .	57
4.3	Conclusion and Outlook . . . . .	64
4.4	Appendix . . . . .	65
4.4.1	Median as a Robust Estimator . . . . .	65
4.4.2	Drift Potential of Rotational Speed . . . . .	66
<b>5</b>	<b>Characterization of Continuous and Discontinuous Stochastic Processes</b>	<b>69</b>
5.1	Expansion of Conditional Moments at Finite Time Step . . . . .	71
5.1.1	Conditional Moments of Diffusion and Jump-Diffusion Processes . . . . .	72

---

5.1.2	Conditional Moments of Pure Noises . . . . .	76
5.2	Methods of Distinguishing Diffusive and Jumpy Stochastic Behaviors . . . . .	77
5.2.1	$Q$ -criterion to Distinguish Diffusive and Jumpy Behaviors . . . . .	77
5.2.2	$\Theta$ - and $\Lambda$ -criterion to Distinguish Diffusive and Jumpy Behaviors . . . . .	80
5.2.3	$Q$ -, $\Theta$ - and $\Lambda$ -criterion for General Diffusion and Jump-Diffusion Processes . . . . .	84
5.3	Consequences of Downsampling of the Data . . . . .	86
5.4	Transition from Jumpy to Diffusive Behavior with Changing Jump Amplitude . . . . .	95
5.5	Application on Real World Data . . . . .	97
<b>6</b>	<b>Conclusion and Outlooks</b>	<b>103</b>
	<b>References</b>	<b>105</b>
	<b>Curriculum</b>	<b>I</b>
	<b>List of Publications</b>	<b>III</b>
	<b>Acknowledgement</b>	<b>V</b>
	<b>Erklärung</b>	<b>VII</b>



## Abstract

Randomly fluctuating or stochastic behavior can be observed in nature in the context of the interactions of nonlinear complex dynamical systems. These can be described by mathematical models and their stochastic nature can generally be separated into continuous and discontinuous contributions. In this thesis, the continuous contribution is considered to be a classical Brownian motion or Wiener process and the discontinuous part as a compound Poisson process. With these assumptions, a jump-diffusion stochastic differential equation is applied. The basic analytic solutions of simple stochastic processes are derived and validated with numerical simulations. These methods are then applied in the field of snow physics and the energy conversion processes in wind turbines. With these advanced stochastic analysis applied to micro-cone penetration tests in snow hardness data, one can characterize the different types of snow by the contributions of continuous diffusion and discontinuous jump noise. Similarly, the same concept of noise contributions can be used on the data of wind energy systems to specify different operational states of wind turbines associated with the control system. The results provide additional insights to the understanding of the complex nature of real world systems. In addition to these applications, more thorough mathematical characterization is performed on the simulated diffusion and jump-diffusion processes in order to introduce robust criteria to distinguish the continuous or discontinuous nature of the given data. These criteria associated with the nature of different time scales present in the stochastic processes are also studied. Finally, these methods are applied on selected real world data. Our studies approach the complex dynamics from a new perspective which allows us not only to have a better understanding of their complex nature but also to pave the way for future research on new phenomena.





## Zusammenfassung

Zufällig fluktuierendes oder stochastisches Verhalten kann in der Natur oft im Zusammenhang mit Wechselwirkungen von nichtlinearen komplexen dynamischen Systemen beobachtet werden. Diese können durch mathematische Modelle beschrieben werden, wobei ihre stochastischen Eigenschaften allgemein in stetige und unstetige Beiträge getrennt werden. In dieser Arbeit wird der stetige Beitrag als klassische Brownsche Bewegung oder Wiener-Prozess und der unstetige Anteil als zusammengesetzter Poisson-Prozess betrachtet. Mit diesen Annahmen wird eine stochastische Sprung-Diffusions-Differentialgleichung angesetzt. Die grundlegenden analytischen Lösungen einfacher stochastischer Prozesse werden hergeleitet und mit numerischen Simulationen validiert. Diese Methoden werden dann im Bereich der Schneephysik und der Energieumwandlungsprozesse in Windenergieanlagen angewendet. Mit dieser fortgeschrittenen stochastischen Analyse, die zunächst auf die Mikro-Drucksondierung in Schneefestigkeitsdaten angewendet wird, werden die verschiedenen Schneearten durch Beiträge von stetigem Diffusionsrauschen und unzeitigem Sprungrauschen charakterisiert. In ähnlicher Weise wird das gleiche Konzept von Rauschbeiträgen auf die Daten von Windenergiesystemen angewandt, um unterschiedliche Betriebszustände von Windenergieanlagen zu identifizieren, welche der Regelung zugeordnet sind. Die Ergebnisse liefern neue Erkenntnisse zum Verständnis der komplexen Natur dieser Systeme. Zusätzlich zu diesen Anwendungen wird eine grundlegende mathematische Charakterisierung von simulierten Diffusions- und Sprung-Diffusions-Prozessen durchgeführt, um robuste Kriterien einzuführen, welche eine Unterscheidung der stetigen oder unzeitigigen Natur der gegebenen Daten ermöglichen. Diese Kriterien selbst werden ebenfalls im Zusammenhang mit den unterschiedlichen Zeitskalen untersucht, welche die stochastischen Prozessen charakterisieren. Schließlich werden diese Methoden auf einige ausgewählte Daten aus der realen Welt angewendet. Die Studien dieser Arbeit nähern sich der komplexen Dynamik aus einer neuen Perspektive, die es ermöglicht, nicht nur deren komplexe Natur besser zu verstehen, sondern auch den Weg für die zukünftige Erforschung neuer Phänomene zu ebnet.



# Chapter 1

## Introduction

Our world exists as the interactions of the nonlinear complex dynamical systems which can be described by mathematical models. In the context of the complex systems in nature, the randomly fluctuating or stochastic behaviors can usually be observed. The study of random nature traces back to 19<sup>th</sup> century when the botanist Robert Brown observed the pollen grains of the plant suspended in water move in irregular zigzag motion under microscope [13]. This phenomenon is later known as Brownian motion, named after him.

In 1900, the mathematician Louis Bachelier modelled the Brownian motion while studying the dynamical behavior of the Paris stock market [7, 8, 18, 41]. In 1905, Albert Einstein published his famous paper on Brownian motion where he argued that the microscopically visible particle suspended in the fluid is moving randomly due to the sum of several collisions with the molecules in the fluid [19]. In 1908, Paul Langevin formulated it into simpler version [46] that the force acting on a suspended particle consists of the viscous resistance or friction force and the random forces caused by the collision with the surrounding micro-particles as in Eq. (1.1),

$$m \frac{dv}{dt} = -\lambda v + \Gamma, \quad (1.1)$$

where  $v$  is the velocity of the particle,  $m$  its mass,  $\lambda$  the friction con-

stant and  $\Gamma$  the random force. In 1923, Norbert Wiener introduced the mathematical description of the Brownian motions [41] which is also known as Wiener process. In 1944, Kiyoshi Itô introduced the stochastic calculus with the Wiener process [40]. Besides Itô's integral, another alternative form is introduced by Ruslan Stratonovich which is known as Stratonovich's integral [80]. However, throughout this thesis, we will only apply the Itô's convention.

Itô introduced the stochastic differential equation (SDE) of the stochastic process  $x$  in the form which resembles the Langevin description of Brownian motion. We call this type of SDE as the Langevin equation which is

$$dx(t) = D^{(1)}(x, t) dt + \sqrt{D^{(2)}(x, t)} dW_t. \quad (1.2)$$

This equation has a deterministic or drift term. A SDE consists of at least one term which is a stochastic process, typically Wiener process  $W_t$ . This type of process with the Wiener process is generally known as the diffusion process which has continuous sample path. The stochastic term in this case is also called diffusion term. We will discuss about it in detail in upcoming chapters.

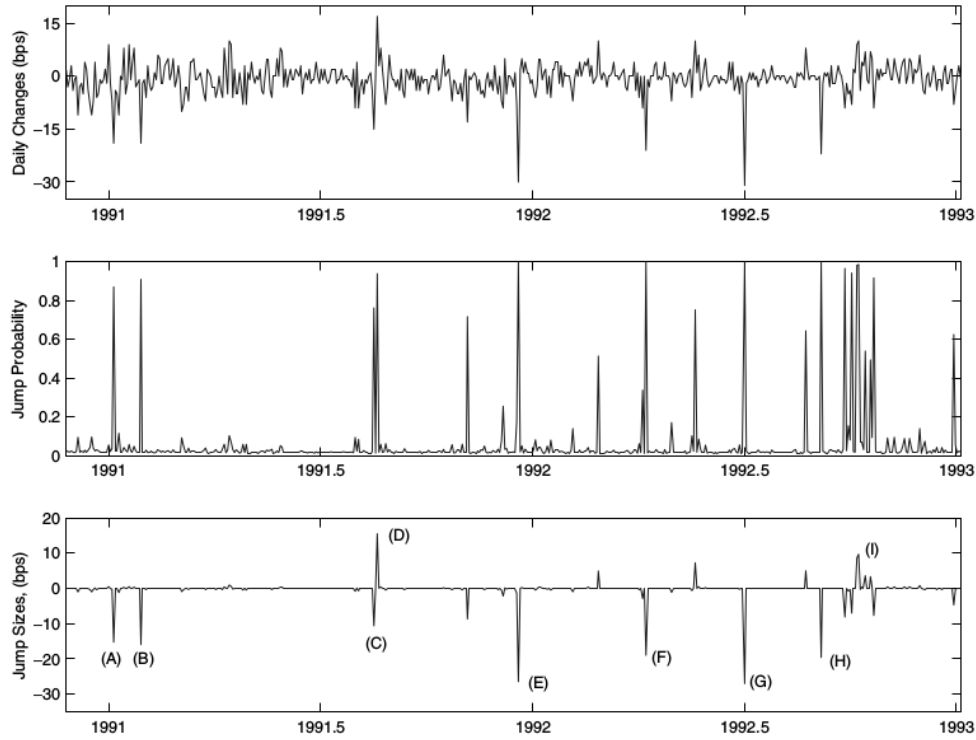
With the advancement of stochastic calculus, the stochastic differential equations (SDEs) are used to model many real world complex systems [23], ranging from natural sciences to social sciences. The example applications of the advanced methods are in physical systems such as renewable wind energy conversion [2, 29, 56, 57], geoscience like seismology [51], biomedical systems including cardiology [26] and epileptic brain dynamics [10, 68, 69], epidemiology [1], finance and economics [22, 25, 36]. The famous Black-Scholes option pricing formula is derived based on the geometric Brownian motion which is the Langevin equation with linear drift and diffusion [12].

Brownian motion or diffusion process alone may not be sufficient to describe the complex systems. It is the continuous process and in reality there are processes with sudden changes which resemble the discontinuous sample paths. These discontinuities can be modeled as jump process

which can, for example, be the Poisson or Lévy jumps. Applications can be found in many different fields, for example, in physics such as neutron scattering from the liquid [16], in meteorology the jumpy behaviour of solar clear sky index [6], in biomedical systems the iEEG brain dynamics [5] and heart beat dynamics [35] and in epidemiology [20]. Besides the natural science fields, it is commonly used in the economic and finance [11, 42, 79]. The Nobel prize winning Black-Scholes model is extended into Black-Scholes-Merton model in the presence of Poisson jumps [54].

In 2004, Michael Johannes showed that the jump events resemble the major macroeconomic events of our history [42]. In the period of 1991 to 1993, the jump events can be determined from the changes in short rates. In his study, he evaluated the jump probability and the jump sizes which occur locally within this time period. Fig. 1.1 shows that the major events marked with the markers A-I generated the significant jump events both in probability and size. The detailed dates of jump events and major events occurred on these dates are as follow:

- (A) 09.01.1991, the outbreak of the Gulf War,
- (B) 01.02.1991, a U.S. unemployment announcement and comments by the Federal Reserve,
- (C) 19.08.1991, the Kremlin coup and the collapse of the Soviet Union,
- (D) 21.08.1991, the emergence of Boris Yeltsin as leader of the remnants of the Soviet Union,
- (E) 20.12.1991, the Federal Reserve lowers the discount rate,
- (F) 09.04.1992, large Japanese equity market decline,
- (G) 02.07.1992, the Federal Reserve lowers the discount rate;
- (H) 04.09.1992, a U.S. unemployment announcement and
- (I) 10.1992, the Bush-Clinton presidential debate.



**Figure 1.1:** The time series of the changes of short rates, locally estimated jump probability and the jump sizes in the period of 1991-1993. The markers A-I show the occurrence of the major events: (A) 09.01.1991, the outbreak of the Gulf War, (B) 01.02.1991, a U.S. unemployment announcement and comments by the Federal Reserve, (C) 19.08.1991, the Kremlin coup and the collapse of the Soviet Union, (D) 21.08.1991, the emergence of Boris Yeltsin as leader of the remnants of the Soviet Union, (E) 20.12.1991, the Federal Reserve lowers the discount rate, (F) 09.04.1992, large Japanese equity market decline, (G) 02.07.1992, the Federal Reserve lowers the discount rate; (H) 04.09.1992, a U.S. unemployment announcement and (I) 10.1992, the Bush-Clinton presidential debate. This figure is reprinted from [42] with permission from Wiley.

From this example in [42], we can therefore observe the importance of the jumps or discontinuities in the stochastic processes. In this thesis, we analyze the diffusion and jump noise present in other real world systems such as the geoscience and the renewable energy systems, and characterize them accordingly which gives more additional insights to the better understanding of the complex systems from a new perspective. In addi-

tion to the applications, more rigorous characterizations of the diffusion and jump-diffusion processes are performed based on simulations and analytical calculations of the properties of the stochastic processes such as simple Ornstein-Uhlenbeck process, also in the presence of jump noise.

This thesis is organized as follow. In upcoming Chapter 2, the fundamental principles of the stochastic analysis applied in this thesis are discussed in details. Starting from the probabilistic description of stochastic processes and Markov properties to the stochastic calculus, the mathematical properties of the basic stochastic processes which are used in the corresponding SDEs. Theoretical calculations and numerical simulations are also compared.

Applying these mathematical descriptions of stochastic processes, the highly-resolved snow hardness measurement data with the spatial resolution of a few micrometers from the micro-cone penetration tests are analyzed in Chapter 3. The failure of ice bonds and pushing aside of snow grains results in the stochastic signals at this high resolution. The contribution of diffusion or jump noises is determined which can be used to differentiate different snow types.

Next, the method is applied again to the energy conversion process in wind energy system in Chapter 4. The analysis is performed to study how different operational regions of a wind turbine can be characterized by the stochastic signals. Diffusion or jump noise contributions are determined across the whole operational range of the wind turbine which could be explained by its control strategies.

After these two applications, more rigorous mathematical characterization is performed on simulated diffusion and jump-diffusion processes in Chapter 5. In addition to the study of [45], more robust criteria to distinguish whether the stochastic process is diffusive or non-diffusive from the data are introduced. The nature of different time scales present in the stochastic processes associated with these criteria are also described. They are tested with the simulated stochastic processes as well as selected real world data.

Finally, the conclusion of this thesis with the outlooks are given in Chapter 6.





## Chapter 2

# Fundamentals of Stochastic Analysis

In this chapter, the fundamental principles of stochastic processes are discussed. The basic properties and definitions are summarized based on the following literature [24, 33, 66, 75, 81] with the additional examples, such as Ornstein-Uhlenbeck (OU) process in the presence of jump noise. Throughout this thesis, the following notations are used for the key parameters:

- $M^{(j)}$  :  $j^{\text{th}}$  order conditional moments,
- $K^{(j)}$  :  $j^{\text{th}}$  order Kramers-Moyal coefficients,
- $D^{(1)}$  : Drift coefficient
- $D^{(2)}$  : Diffusion coefficient.

The definitions or conventions of these parameters will be discussed throughout this chapter.

## 2.1 Stochastic Processes and Markov Property

A stochastic process can be viewed as a time or space dependent variable which fluctuates randomly. In this section, we are going to discuss about the time-dependent stochastic process  $x(t)$ . Since it fluctuates randomly, the exact time evolution cannot be repeatable. Therefore, we need a probabilistic description of the stochastic process.

A complete statistical description of the stochastic process  $x(t)$  can be uniquely defined by the joint probability density function (joint PDF) as follow.

$$p(x_N, t_N; x_{N-1}, t_{N-1}; \dots; x_1, t_1), \text{ with } N \text{ discrete time } t_1 < t_2 < \dots < t_N, \quad (2.1)$$

In terms of the joint PDF, we can define the conditional probability density function (conditional PDF) as

$$p(x_N, t_N | x_{N-1}, t_{N-1}; \dots; x_1, t_1) = \frac{p(x_N, t_N; x_{N-1}, t_{N-1}; \dots; x_1, t_1)}{p(x_{N-1}, t_{N-1}; \dots; x_1, t_1)}, \quad (2.2)$$

where  $p(x_N, t_N | x_{N-1}, t_{N-1}, \dots, x_1, t_1)$  is the probability of  $x_N$  at time  $t_N$  conditioned on all previous states.

For a purely random process, the samples  $x_i$  are independent such that

$$p(x_N, t_N | x_{N-1}, t_{N-1}; \dots; x_1, t_1) = p(x_N, t_N), \quad (2.3)$$

as a consequence, the joint PDF becomes

$$p(x_N, t_N; x_{N-1}, t_{N-1}; \dots; x_1, t_1) = \prod_{i=1}^N p(x_i, t_i), \quad (2.4)$$

Besides the purely random process, another process in which the future state can be described by the information of present state is known as Markov process. In other word, the variable  $x_i$  at time  $t_i$  depends only

on  $x_{i-1}$  at time  $t_{i-1}$  and it can be described as

$$p(x_N, t_N | x_{N-1}, t_{N-1}; \dots; x_1, t_1) = p(x_N, t_N | x_{N-1}, t_{N-1}) . \quad (2.5)$$

The joint PDF for Markov process can be written as,

$$p(x_N, t_N; x_{N-1}, t_{N-1}; \dots; x_1, t_1) = p(x_N, t_N | x_{N-1}, t_{N-1}) p(x_{N-1}, t_{N-1}; \dots; x_1, t_1) . \quad (2.6)$$

Repeating the same concept for  $p(x_{N-1}, t_{N-1}; \dots; x_1, t_1)$ , we can rewrite the joint PDF as,

$$p(x_N, t_N; x_{N-1}, t_{N-1}; \dots; x_1, t_1) = p(x_1, t_1) \prod_{i=2}^N p(x_i, t_i | x_{i-1}, t_{i-1}) . \quad (2.7)$$

Markov process is often said to be “memoryless” since it does not depend on the past but on the present state. This does not mean that there is no long-time correlation.

For three-point probability distribution, we have the relation

$$p(x_3, t_3) = \int dx_2 p(x_3, t_3 | x_2, t_2) p(x_2, t_2) . \quad (2.8)$$

Similarly, for the conditional PDF, we can write

$$p(x_3, t_3 | x_1, t_1) = \int dx_2 p(x_3, t_3 | x_2, t_2; x_1, t_1) p(x_2, t_2 | x_1, t_1) . \quad (2.9)$$

For a Markov process, Eq. (2.9) becomes

$$p(x_3, t_3 | x_1, t_1) = \int dx_2 p(x_3, t_3 | x_2, t_2) p(x_2, t_2 | x_1, t_1) , \quad (2.10)$$

which is known as the Chapman-Kolmogorov equation.

## 2.2 Kramers-Moyal Expansion

For a Markov process, the PDF  $p(x, t + \tau)$  at time  $t + \tau$  and  $p(x', t)$  at time  $t$  can be related by

$$p(x, t + \tau) = \int p(x, t + \tau | x', t) p(x', t) dx' , \quad (2.11)$$

where  $p(x, t + \tau | x', t)$  is the conditional PDF. The  $j^{\text{th}}$ -order conditional moments  $M^{(j)}(x', t)$  can be defined as follow,

$$\begin{aligned} M^{(j)}(x', t, \tau) &= \left\langle (x(t + \tau) - x(t))^j \Big|_{x(t)=x'} \right\rangle \\ &= \int (x - x')^j p(x, t + \tau | x', t) dx \end{aligned} \quad (2.12)$$

From these definitions, we can derive the Kramers-Moyal (KM) expansion which describe the probabilistic time evolution of the stochastic process  $x(t)$ . It reads [75]

$$\frac{\partial}{\partial t} p(x, t) = \sum_{j=1}^{\infty} \frac{1}{j!} \left( -\frac{\partial}{\partial x} \right)^j [K^{(j)}(x, t) p(x, t)] \quad (2.13)$$

where  $K^{(j)}(x, t)$  is called Kramer-Moyal (KM) coefficient which is

$$K^{(j)}(x, t) = \lim_{\tau \rightarrow 0} \frac{M^{(j)}(x, t)}{\tau}. \quad (2.14)$$

KM expansion can be written in operator form,

$$\frac{\partial}{\partial t} p(x, t) = \mathcal{L}_{\text{KM}} p(x, t) \quad (2.15)$$

where  $\mathcal{L}_{\text{KM}}$  is the Kramers-Moyal (KM) operator.

$$\mathcal{L}_{\text{KM}} = \sum_{j=1}^{\infty} \frac{1}{j!} \left( -\frac{\partial}{\partial x} \right)^j K^{(j)}(x, t) \quad (2.16)$$

The detailed derivation can be seen in [75].

## 2.3 Fokker-Planck and Langevin Equation

Pawula theorem [62] states if  $K^{(j)}$ , as defined by Eq. (2.14), exists for all  $j$ , and if  $K^{(j)} = 0$  for some even  $j$ , then  $K^{(j)} = 0$  for all  $j \geq 3$ . In simple terms, if we can prove that  $K^{(4)} = 0$ , all the coefficients of order  $j \geq 3$  are also vanishing.

Then, the KM expansion becomes the Fokker-Planck (FP) equation which is

$$\frac{\partial}{\partial t} p(x, t) = -\frac{\partial}{\partial x} (D^{(1)}(x, t) p(x, t)) + \frac{1}{2} \frac{\partial^2}{\partial x^2} (D^{(2)}(x, t) p(x, t)) \quad , \quad (2.17)$$

where  $D^{(1)}(x, t) = K^{(1)}(x, t)$  is the drift or deterministic term and  $D^{(2)}(x, t) = K^{(2)}(x, t)$  the diffusion term. FP equation is the special case of the KM expansion.

FP equation can be written in operator form,

$$\frac{\partial}{\partial t} p(x, t) = \mathcal{L}_{\text{FP}} p(x, t) \quad (2.18)$$

where  $\mathcal{L}_{\text{FP}}$  is the Fokker-Planck (FP) operator.

$$\mathcal{L}_{\text{FP}} = -\frac{\partial}{\partial x} D^{(1)}(x, t) + \frac{1}{2} \frac{\partial^2}{\partial x^2} D^{(2)}(x, t) \quad (2.19)$$

The FP equation is also valid for the conditional PDF  $p(x, t|x', t')$  of the stochastic process  $x(t)$ , whose PDF  $p(x, t)$  fulfils the FP equation, such that

$$\frac{\partial}{\partial t} p(x, t|x', t') = -\frac{\partial}{\partial x} D^{(1)}(x, t) p(x, t|x', t') + \frac{1}{2} \frac{\partial^2}{\partial x^2} D^{(2)}(x, t) p(x, t|x', t'), \quad (2.20)$$

with the initial condition  $p(x, t|x', t) = \delta(x - x')$ . For small time step  $\tau$ , the short-time propagator  $p(x, t + \tau|x', t)$  has the formal solution,

$$p(x, t + \tau | x', t) = \exp(\tau \mathcal{L}_{\text{FP}}) \delta(x - x'), \quad (2.21)$$

For sufficiently small time step  $\tau$ ,  $D^{(1)}$  and  $D^{(2)}$  can be considered as constant. Eq. (2.21) can further be derived which gives

$$p(x, t + \tau | x', t) = \frac{1}{\sqrt{2\pi\tau D^{(2)}(x', t)}} \exp\left(-\frac{(x - x' - \tau D^{(1)}(x', t))^2}{2\tau D^{(2)}(x', t)}\right), \quad (2.22)$$

which yields the Gaussian distribution with mean  $(x' + \tau D^{(1)}(x', t))$  and variance  $\tau D^{(2)}(x', t)$ . The detailed derivation can be seen in [75, 81].

Equivalent to the probabilistic description of the stochastic process  $x(t)$  with FP-equation, it can also be written by stochastic dynamical equation or stochastic differential equation (SDE). It is the continuous diffusion process which is also called Langevin equation,

$$dx(t) = D^{(1)}(x, t) dt + \sqrt{D^{(2)}(x, t)} dW_t, \quad (2.23)$$

where  $W_t$  is the Wiener process. The Wiener increment is  $dW_t = \Gamma_t \cdot dt$  where  $\Gamma_t$  is Gaussian white noise, i.e.  $\langle \Gamma_t \rangle = 0$  and  $\langle \Gamma_t \Gamma_{t'} \rangle = \delta(t - t')$ . More detailed definitions and properties of Wiener process will be discussed next.

## 2.4 Wiener Process

The Wiener process  $W_t$  is a mathematical realization of the Brownian motion. It is a Markov process and has the following basic properties:

- $W_0 = 0$  at  $t = 0$ .
- $W_t$  is a continuous process such that  $W_{t-} = W_t = W_{t+}$ .  
 $W_{t^\pm} = \lim_{t \rightarrow t^\pm} W_t$ .
- $W_t$  has stationary and independent increments. The Wiener increments,

$$\Delta W_t = W_{t+\Delta t} - W_t, \quad (2.24)$$

are mutually independent for all non-overlapping  $t$ .

- The increment  $\Delta W_t$  has a Gaussian distribution with mean zero and variance  $\Delta t$ ,  $\Delta W_t \sim N(0, \Delta t)$  which means the PDF is

$$p_{\Delta W_t}(x) = \frac{1}{\sqrt{2\pi\Delta t}} \exp\left(-\frac{x^2}{2\Delta t}\right), \quad (2.25)$$

In general, the Wiener process is a Gaussian distributed random variable with zero mean and variance  $t > 0$ , i.e.  $W_t \sim N(0, t)$ . From these properties, the covariance of the Wiener process can be deduced that

$$\text{Cov}(W_s, W_t) = \langle W_s W_t \rangle = \min(s, t), \quad (2.26)$$

and the covariance of the Wiener (differential) increment consequently becomes

$$\text{Cov}(dW_s, dW_t) = \langle dW_s dW_t \rangle = \delta(s - t) ds dt, \quad (2.27)$$

which means that  $\text{Cov}(dW_s, dW_t) = 0$  for  $s \neq t$  due to the independent increments for all non-overlapping time. For  $s = t$ , the covariance becomes  $\text{Cov}(dW_s, dW_t) = \text{Var}(dW_t) = dt$ .

The further properties are that higher than second order of Wiener increment and the product of time increment and Wiener increment are negligible in  $dt$  precision, such that

$$(dW_t)^n = 0, \text{ for } n \geq 3 \quad (2.28)$$

$$(dt)^i (dW_t)^j = 0, \text{ for } i, j \geq 1 \quad (2.29)$$

The final remark is that the Wiener process is nowhere differentiable even though it is a continuous process.

## 2.5 Stochastic Processes with Discontinuities

The Wiener process and the Langevin equation are the continuous or diffusive processes which satisfy the Lindeberg's continuity condition [24, 81],

$$C(t) = \lim_{\Delta t \rightarrow 0} \frac{\text{Prob}[|\Delta x(t)| > \delta |_{x(t)=x}]}{\Delta t} = 0, \quad (2.30)$$

where  $\Delta x(t) = x(t + \Delta t) - x(t)$ .

KM coefficients from third-order and higher are vanishing in continuous process. If the higher-order KM coefficients are not negligible, it will be an indication of discontinuity in the stochastic dynamics. The detailed proof of continuity condition can be seen in [81].

### 2.5.1 Jump-Diffusion Process

When the signal of a stochastic process typically presents sharp changes at some instant (jump or discontinuous events), higher order KM coefficients are non-negligible. An extension to the Langevin-type modeling with the additional jump noise is needed which is known as a jump-diffusion dynamics [5, 11, 42, 79, 81, 82] given by the following stochastic differential equation:

$$dx(t) = D^{(1)}(x, t) dt + \sqrt{D^{(2)}(x, t)} dW_t + \xi dJ_t, \quad (2.31)$$

where, again,  $D^{(1)}(x, t)$  and  $D^{(2)}(x, t)$  are the drift and the diffusion coefficients, respectively, and  $W_t$  is the Wiener process. The quantity  $\xi$  is the size of the jump noise and we take it as a normally distributed random variable with zero mean and variance  $\sigma_\xi^2(x, t)$ , i.e.,  $\xi \sim N(0, \sigma_\xi^2)$ . The variance  $\sigma_\xi^2(x, t)$  is also called jump amplitude. The term  $J_t$  is the Poisson jump process, which is the zero-one jump process with a jump rate (or intensity)  $\lambda(x, t)$ . The term  $dZ_t = \xi dJ_t$  is also known as compound Poisson process. More detailed definitions and properties of Poisson and



compound Poisson process will be discussed in next section.

For jump-diffusion processes, the drift and diffusion coefficients ( $D^{(1)}$ ,  $D^{(2)}$ ), the jump rate  $\lambda$  and amplitude  $\sigma_\xi^2$  are related to the KM coefficients  $K^{(j)}(x, t)$  such that

$$D^{(1)}(x, t) = K^{(1)}(x, t), \quad (2.32)$$

$$D^{(2)}(x, t) + \lambda(x, t)\langle \xi^2 \rangle = K^{(2)}(x, t), \quad (2.33)$$

$$\lambda(x, t)\langle \xi^j \rangle = K^{(j)}(x, t), \quad \text{for } j > 2. \quad (2.34)$$

The estimate of the drift coefficient is the same for the diffusion process (Eq. 2.23) and the jump-diffusion process (Eq. 2.31). Jump amplitude  $\sigma_\xi^2$  and rate  $\lambda$  can be estimated by using Eq. (2.34) with  $j = 4$  and  $j = 6$ , and the Wick's theorem for Gaussian random variables which states that  $\langle \xi^{2n} \rangle = \frac{(2n)!}{2^n n!} \langle \xi^2 \rangle^n$ :

$$\sigma_\xi^2(x, t) = \frac{K^{(6)}(x, t)}{5K^{(4)}(x, t)}, \quad (2.35)$$

$$\lambda(x, t) = \frac{K^{(4)}(x, t)}{3\sigma_\xi^4(x, t)}. \quad (2.36)$$

To improve the estimation of KM coefficients  $K^{(j)}(x, t)$  in particular of high-order coefficients, the Nadaraya-Watson estimator, which is a kernel estimator, can be used [61, 85]:

$$K^{(j)}(x, t) = \lim_{\Delta t \rightarrow 0} \frac{\sum_i k\left(\frac{x_{i\Delta t} - x}{h}\right) (x_{(i+1)\Delta t} - x_{i\Delta t})^j}{\sum_i k\left(\frac{x_{i\Delta t} - x}{h}\right) \Delta t}, \quad (2.37)$$

where  $k(u)$  is the kernel function. With the kernel-based method the conditional moments can be calculated more smoothly by controlling the kernel bandwidth  $h$  [44]. These results allow us to estimate the model parameters directly from the given data, such as snow hardness data in Chapter 3 and wind power data in Chapter 4.

Next, the definitions and properties of Poisson process and compound Poisson process will be discussed in detail.

## 2.6 Poisson Process

Along with the continuous Wiener process, the Poisson process  $J_t$  is also an important mathematical tool to describe the discontinuous jumps in the stochastic processes. It is also a Markov process with the following basic properties:

- $J_0 = 0$  at  $t = 0$ .
- $J_t$  is right continuous and has a unit jump at time  $T_k > 0$  such that  $J_{T_k^+} = J_{T_k^-} + 1$ . Again,  $J_{T_k^\pm} = \lim_{t \rightarrow T_k^\pm} J_t$ .
- $J_t$  is right-continuous such that  $J_{t^+} = J_t \geq J_{t^-}$ ,  $t > 0$ .
- $J_t$  has stationary and independent increments. The Poisson increments,

$$\Delta J_t = J_{t+\Delta t} - J_t, \quad (2.38)$$

are mutually independent for all non-overlapping  $t$ .

- $J_t$  has a Poisson distribution with mean and variance  $\lambda t$ ,  $J_t \sim \mathcal{P}(\lambda t)$  such that

$$\text{Prob}(J_t = k) = p_{J_t}(\lambda t, k) = \frac{(\lambda t)^k}{k!} e^{-\lambda t}, \quad (2.39)$$

for the integer  $k = 0, 1, 2, \dots$ , jump rate  $\lambda > 0$  and  $t > 0$ .

- The increment  $\Delta J_t$  has a Poisson distribution with mean and variance  $\lambda \Delta t$ ,  $\Delta J_t \sim \mathcal{P}(\lambda \Delta t)$ .

$$\text{Prob}(\Delta J_t = k) = p_{\Delta J_t}(\lambda \Delta t, k) = \frac{(\lambda \Delta t)^k}{k!} e^{-\lambda \Delta t}, \quad (2.40)$$

for the integer  $k = 0, 1, 2, \dots$ , jump rate  $\lambda > 0$  and  $\Delta t > 0$ .

- The time between each jump or inter-jump time  $\Delta T_j = T_{j+1} - T_j$  where  $T_j$  is the  $j^{\text{th}}$  jump time for  $j = 0, 1, 2, \dots$  of a simple Poisson

process  $J_t$  with fixed jump rate  $\lambda$  is exponentially distributed such that

$$\text{Prob}(\Delta T_j \leq \Delta t | T_j) = 1 - e^{-\lambda \Delta t} \approx \lambda \Delta t. \quad (2.41)$$

From these properties, the covariance of the Poisson process can be deduced that

$$\text{Cov}(J_s, J_t) = \langle \hat{J}_s \hat{J}_t \rangle = \lambda \min(s, t), \quad (2.42)$$

where  $\hat{J}_t = J_t - \langle J_t \rangle$  is the compensated Poisson process.

The covariance of the Poisson (differential) increment consequently becomes

$$\text{Cov}(dJ_s, dJ_t) = \langle d\hat{J}_s d\hat{J}_t \rangle = \lambda \delta(s - t) ds dt, \quad (2.43)$$

which means that  $\text{Cov}(dJ_s, dJ_t) = 0$  for  $s \neq t$  due to the independent increments for all non-overlapping time. For  $s = t$ , the covariance becomes  $\text{Cov}(dJ_s, dJ_t) = \text{Var}(dJ_t) = \lambda dt$ .

The further properties are that higher order of Poisson increment in  $dt$  precision is the Poisson increment such that

$$(dJ_t)^n = dJ_t, n \geq 1 \quad (2.44)$$

The product of time increment and Poisson increment as well as that of Wiener increment and Poisson increment are negligible in  $dt$  precision, i.e.

$$(dt)^i (dJ_t)^j = 0, \text{ for } i, j \geq 1 \quad (2.45)$$

$$(dW_t)^i (dJ_t)^j = 0, \text{ for } i, j \geq 1 \quad (2.46)$$

The final remark on the Poisson process is that for short time asymptotic ( $\Delta t \rightarrow 0^+$ ), the Poisson process obeys zero-one jump law, such that

$$\begin{aligned}\text{Prob}(\Delta J_t = 0) &= 1 - \lambda\Delta t + \mathcal{O}(\lambda\Delta t)^2, \\ \text{Prob}(\Delta J_t = 1) &= \lambda\Delta t + \mathcal{O}(\lambda\Delta t)^2, \\ \text{Prob}(\Delta J_t > 1) &= \mathcal{O}(\lambda\Delta t)^2,\end{aligned}\tag{2.47}$$

which means that there is either zero- or one-jump event within a short-time step  $\Delta t$ .

### 2.6.1 Compound Poisson Process

A pure Poisson process has constant unit jump. In order to have a more realistic model, the random jump sizes can be considered which is known as compound Poisson process.

Let  $\xi_i$  for  $i = 1, 2, \dots$  be the independent and identically distributed sequence of random variables and also independent of the Poisson process  $J_t$ . The compound Poisson process  $Z_t$  can then be defined as

$$Z_t = \sum_{i=1}^{J_t} \xi_i, t \geq 0.\tag{2.48}$$

Compound Poisson process has the following basic properties:

- $Z_0 = 0$  at  $t = 0$ .
- $Z_t$  has stationary and independent increments. The Poisson increments,

$$\Delta Z_t = \xi \Delta J_t,\tag{2.49}$$

are mutually independent for all non-overlapping  $t$ .

- The mean and variance of  $Z_t$  are

$$\langle Z_t \rangle = \langle \xi \rangle \lambda t, \quad (2.50)$$

$$\text{Var}(Z_t) = \langle \xi^2 \rangle \lambda t. \quad (2.51)$$

Let us consider the Gaussian distributed jump size with zero mean and the variance  $\sigma_\xi^2$ , such that  $\xi \sim N(0, \sigma_\xi^2)$  as an example. Then, the corresponding compound Poisson process has zero mean and the variance  $\lambda \sigma_\xi^2 t$ .

From these properties, the covariance of the compound Poisson process with zero mean Gaussian jump size can be deduced that

$$\text{Cov}(Z_s, Z_t) = \langle Z_s Z_t \rangle = \lambda \sigma_\xi^2 \min(s, t). \quad (2.52)$$

The covariance of its (differential) increment consequently becomes

$$\text{Cov}(dZ_s, dZ_t) = \langle dZ_s dZ_t \rangle = \lambda \sigma_\xi^2 \delta(s - t) ds dt, \quad (2.53)$$

which means that  $\text{Cov}(dZ_s, dZ_t) = 0$  for  $s \neq t$  due to the independent increments for all non-overlapping time. For  $s = t$ , the covariance becomes  $\text{Cov}(dZ_s, dZ_t) = \text{Var}(dZ_t) = \lambda \sigma_\xi^2 dt$ .

## 2.7 Ornstein-Uhlenbeck Process in the Presence of Jumps

In the presence of compound Poisson jump noise  $Z_t = \xi J_t$  with constant jump amplitude  $\sigma_\xi^2$  and jump rate  $\lambda$ , Ornstein-Uhlenbeck (OU) Process  $x = x(t)$  can be written as

$$dx = -\gamma x dt + \sqrt{D} dW_t + \xi dJ_t. \quad (2.54)$$

By introducing the substituting variable  $y = xe^{\gamma t}$ , using Eq (2.29) and (2.45), and applying the Itô product rule,  $d(f \cdot g) = f \cdot dg + g \cdot df + df \cdot dg$ , we can solve Eq. (2.54). The solution in  $dt$  precision is

$$x = x_0 e^{-\gamma t} + \sqrt{D} \int_0^t e^{-\gamma(t-t')} dW_{t'} + \int_0^t e^{-\gamma(t-t')} dZ_{t'}, \quad (2.55)$$

where  $x_0 = x(0)$ .

The mean of the OU process can be calculated as

$$\langle x \rangle = x_0 e^{-\gamma t}. \quad (2.56)$$

and  $\langle x \rangle = 0$  for sufficiently long simulation time  $t \rightarrow \infty$ .

The covariance of OU process with jump can be calculated using Eq. (2.27), (2.46) and (2.53) such that

$$C_{xx}(t, \tau) = \text{Cov}(x(t)x(t + \tau)) = \frac{D + \lambda\sigma_\xi^2}{2\gamma} [e^{-\gamma\tau} - e^{-\gamma(\tau+2t)}]. \quad (2.57)$$

We can calculate the variance of the OU process with jump from  $C_{xx}(t, \tau = 0)$  which is

$$\text{Var}(x) = \langle x^2 - \langle x \rangle^2 \rangle = \frac{D + \lambda\sigma_\xi^2}{2\gamma} e^{-2t}. \quad (2.58)$$

For sufficiently long simulation time  $t \rightarrow \infty$ , the covariance becomes

$$C_{xx}(\tau) = \frac{D + \lambda\sigma_\xi^2}{2\gamma} e^{-\gamma\tau}, \quad (2.59)$$

and the variance

$$\text{Var}(x) = \frac{D + \lambda\sigma_\xi^2}{2\gamma}. \quad (2.60)$$

Next, the numerical simulations of these stochastic processes will be performed. With the help of the discussed definitions and properties, the choice of model parameters for the simulations becomes much simpler to obtain the desired statistics of the processes.

### 2.7.1 Numerical Simulations

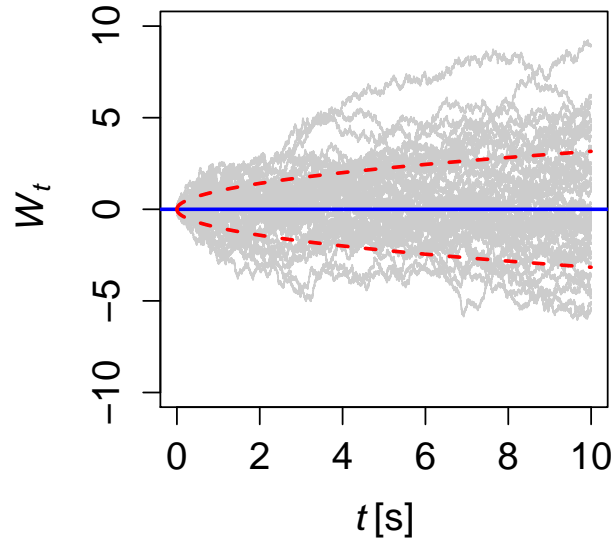
The OU process in the presence of jumps can be simulated numerically using Euler's method.

$$x(t + \tau) = x(t) - \gamma x(t) \tau + \sqrt{D} \Delta W_t + \xi \Delta J_t . \quad (2.61)$$

According to Eq. (2.25), the Wiener increment  $\Delta W_t$  has the same distribution as  $\eta\sqrt{\tau}$  where  $\eta \sim N(0, 1)$  is the standard normal random variable so that

$$\Delta W_t = \eta\sqrt{\tau} \quad (2.62)$$

can be used for simulation of Wiener increment. With this, Wiener process can be reconstructed and 50 trajectories are plotted with ensemble mean (0) and standard deviation ( $\sqrt{t}$ ) in Fig. 2.1.



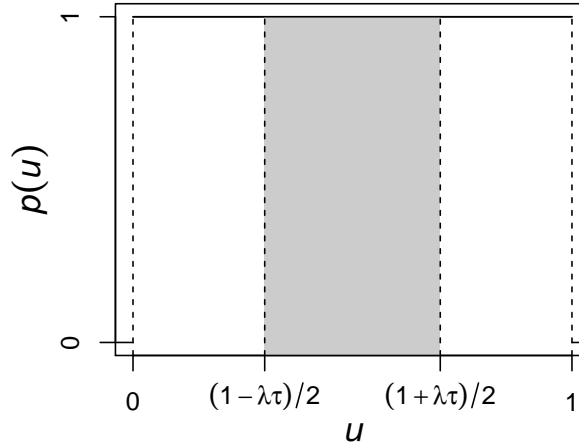
**Figure 2.1:** 50 trajectories of Wiener process are plotted in gray lines with ensemble mean (0) in blue solid line and lower and upper bound of standard deviation ( $\sqrt{t}$ ) in red dotted lines.

Poisson increment obeying zero-one jump law described in Eq. (2.47)

can numerically be simulated by a standard uniform number generator and the method of acceptance-rejection [28, 33, 65]. The open interval  $(0, 1)$  is partitioned into the central interval  $\left[\frac{1 - \lambda\tau}{2}, \frac{1 + \lambda\tau}{2}\right]$  and the complement. When the generated standard uniform random number  $u$  is within the central interval there is a jump ( $\Delta J_t = 1$ ). When it is outside of the central interval (complement) there is no jump ( $\Delta J_t = 0$ ). It can be summarized such that

$$\Delta J_t = \begin{cases} 1, & \frac{1 - \lambda\tau}{2} \leq u \leq \frac{1 + \lambda\tau}{2}, \\ 0, & \text{else.} \end{cases} \quad (2.63)$$

The graphical visualization of the method is shown in Fig. 2.2 where the PDF of uniform distribution is plotted and the central interval is represented in gray shade.

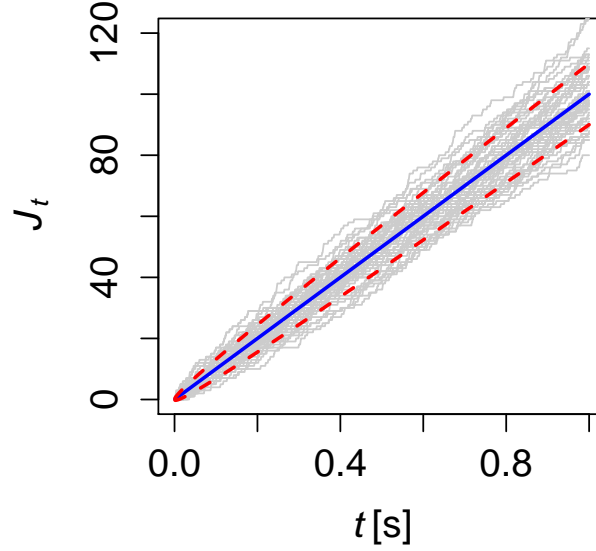


**Figure 2.2:** The PDF of uniform distribution is shown. The open interval  $(0, 1)$  is partitioned into the central interval  $\left[\frac{1 - \lambda\tau}{2}, \frac{1 + \lambda\tau}{2}\right]$  (gray shade) and the complement. If the generated standard uniform random number  $u$  is landed in the gray area, there is a jump ( $\Delta J_t = 1$ ). If it is outside of the gray area, there is no jump ( $\Delta J_t = 0$ ).

With this simulation of Poisson increment, Poisson process with  $\lambda = 100 \text{ s}^{-1}$  ( $\lambda\Delta t = 0.1$ ) can be reconstructed and 50 trajectories are plotted

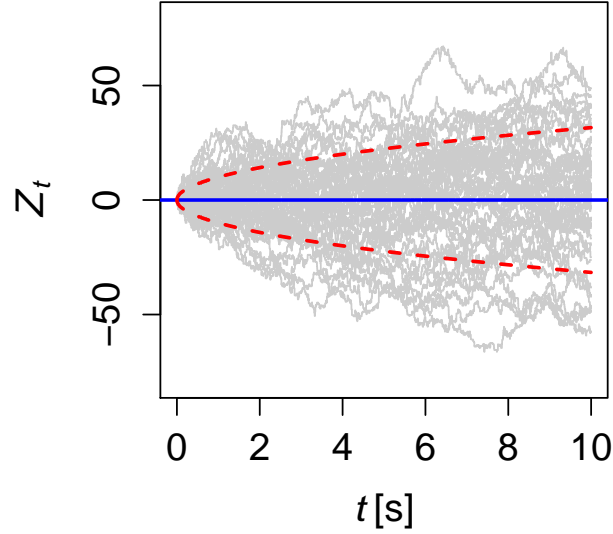


with ensemble mean  $(\lambda t)$  and standard deviation  $(\sqrt{\lambda t})$  in Fig. 2.3.



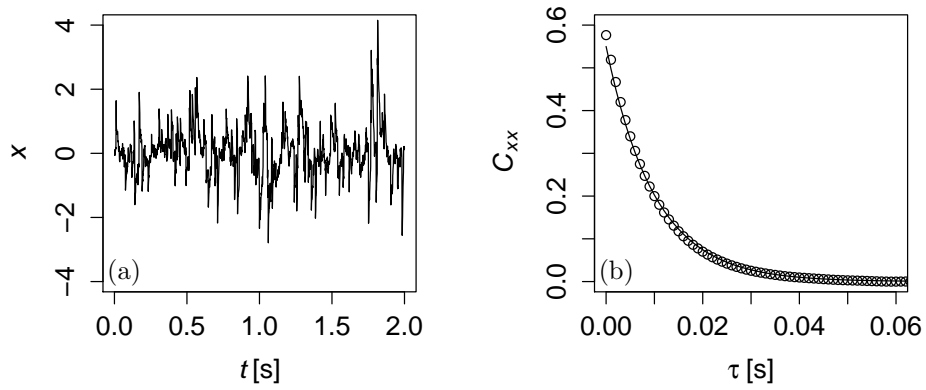
**Figure 2.3:** 50 trajectories of Poisson process with  $\lambda = 100 \text{ s}^{-1}$  ( $\lambda\Delta t = 0.1$ ) are plotted in gray lines with ensemble mean  $(\lambda t)$  in blue solid line and lower and upper bound of standard deviation  $(\sqrt{\lambda t})$  in red dotted lines.

The simulation of compound Poisson process can be performed by integrating  $dZ_t = \xi dJ_t$ . The parameters  $\lambda = 100 \text{ s}^{-1}$  ( $\lambda\Delta t = 0.1$ ) and Gaussian distributed jump size with variance or jump amplitude  $\sigma_\xi^2 = 1$  are used. 50 trajectories are again generated and plotted together with ensemble mean  $(0)$  and standard deviation  $(\sqrt{\lambda\sigma_\xi^2 t})$  in Fig. 2.4.



**Figure 2.4:** 50 trajectories of compound Poisson process with  $\lambda = 100 \text{ s}^{-1}$  ( $\lambda\Delta t = 0.1$ ) and Gaussian distributed jump size with variance or jump amplitude  $\sigma_\xi^2 = 1$  are plotted in gray lines with ensemble mean (0) in blue solid line and lower and upper bound of standard deviation ( $\sqrt{\lambda\sigma_\xi^2 t}$ ) in red dotted lines.

Using Euler's scheme Eq. (2.61), a synthetic time series of the OU jump-diffusion process is generated for  $\Delta t = 10^{-3} \text{ s}$  with  $\gamma = 100 \text{ s}^{-1}$ ,  $D = 10 \text{ s}^{-1}$  and for the additional jump terms with  $\lambda = 100 \text{ s}^{-1}$  and  $\sigma_\xi^2 = 1$ . In Fig. 2.5 the timeseries of the OU jump-diffusion processes is shown. Its covariance function  $C_{xx}(\tau)$  is also evaluated and plotted together with the theoretical result from Eq. (2.59).



**Figure 2.5:** Timeseries of OU jump-diffusion processes, (a), and its covariance function  $C_{xx}(\tau)$ , (b). The evaluated  $C_{xx}(\tau)$  is plotted in open circles and its theoretical result from Eq. (2.59) are plotted with solid line.

In Chapter 3 and 5, these numerical simulations are applied to characterize the contribution of the diffusion and jump noise as well as the nature of different time scales present in the (jump-)diffusion stochastic dynamics.



## Chapter 3

# Stochastic Analysis of Micro-cone Penetration Tests in Snow<sup>1</sup>

### Abstract

Cone penetration tests have long been used to characterize snowpack stratigraphy. With the development of sophisticated digital penetrometers such as the SnowMicroPen, vertical profiles of snow hardness can now be measured at a spatial resolution of a few micrometers. By using small penetrometer tips at this high vertical resolution, further details of the penetration process are resolved, leading to many more stochastic signals. An accurate interpretation of these signals regarding snow characteristics requires advanced data analysis. Here, the failure of ice connections and the pushing aside of separated snow grains during cone penetration lead to a combination of (a) diffusive noise, as in Brownian motion, and (b) jumpy noise, as proposed by previous dedicated inversion methods. The determination of the Kramers–Moyal coefficients enables differentiating between diffusive and jumpy behaviors and determining the functional resistance dependencies of these stochastic contributions. We show how different snow types can be characterized by this com-

---

<sup>1</sup>published in “The Cryosphere” [47].

bination of highly resolved measurements and data analysis methods. In particular, we show that denser snow structures exhibited a more collective diffusive behavior supposedly related to the pushing aside of separated snow grains. On less dense structures with larger pore space, the measured hardness profile appeared to be characterized by stronger jump noise, which we interpret as related to breaking of a single cohesive bond. The proposed methodology provides new insights into the characterization of the snowpack stratigraphy with micro-cone penetration tests.

### 3.1 Introduction

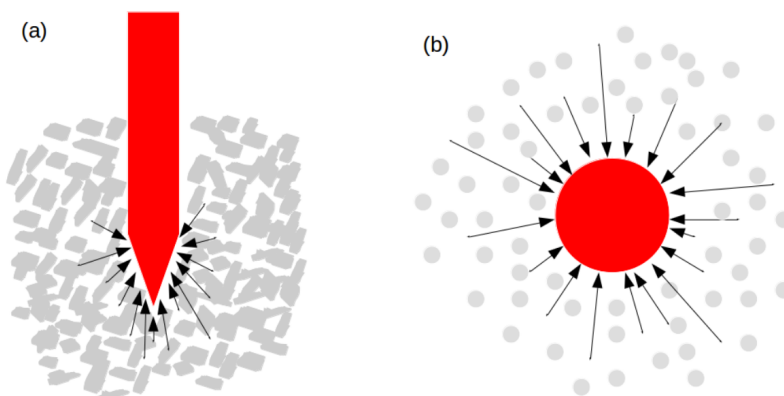
Snow is an essential component of our environment and can significantly impact our lives: from the wishful dream of a white Christmas to the misfortune of avalanche accidents. Having a closer look at snow, one discovers many microstructural patterns and realizes that snow on the ground undergoes constant evolution [17]. The snow microstructure fully controls its physical and mechanical properties, which are essential for diverse applications, such as avalanche forecasting [78]. A snowpack is typically structured in numerous layers composed of different snow types, where such stratigraphy will determine the snowpack stability. Cone penetration tests have long been used to characterize the snowpack stratigraphy [9]. The SnowMicroPen (SMP) can perform cone penetration tests of snow in the field [77]. It measures the force needed to drive a cylinder with a millimetric conic tip into the snowpack. With its high resolution (250 measurements  $\text{mm}^{-1}$ ), the measured force or hardness is supposedly linked to the snow microstructure [43]. A typical consequence of such a high-precision measurement is that more and more details of the penetration process can be resolved, leading to many more stochastic signals. In this context, it is of particular interest to employ advanced data analysis to find out how different kinds of stochastic signals are related to different snow types. There are also other penetrometers used for cone penetration tests in the field of snow; for example [53] used a large-scale penetrometer with a tip diameter of 36.7 mm in polar snow. For our

analysis, we specifically used the measurement data of micro-cone penetration tests from [77].

[43] developed the first model to estimate micromechanical properties of snow from measured penetration profiles. They assumed that the material compaction is negligible and that the penetration resistance is only composed of friction between the penetrometer and snow grains and a superposition of spatially uncorrelated and identical brittle failures of individual snow microstructural elements (e.g., the bonds between the snow grains). [52] extended the theory of [43] to account for simultaneous ruptures by Monte Carlo simulations. They precisely resolved the snow micromechanical parameters, such as the deflection length, rupture force and rupture intensity. [49] re-stated the pioneering idea of [43] and described the fluctuating penetration hardness as a Poisson shot noise process. In their model, the micromechanical parameters can be simply derived from the cumulants and the co-variance of the penetration signal. [64] further extended the homogeneous Poisson process of [49] so that the scale of variation in the rupture intensity could be decoupled from the scale of variations in the deflection length and the rupture force. These models are now commonly used to characterize the snowpack stratigraphy from SMP measurements. Indeed, [67] related the micromechanical parameters derived from SMP measurements to some of the most critical snow characteristics, namely density, specific surface area and structural correlation length. These relations are now routinely used to quantify the snowpack stratigraphy [15]. Besides, [71] estimated the elastic modulus and fracture energy from the micromechanical parameters, which can then be used to compute point snow stability for avalanche hazard assessment [72].

Here, we consider the measured fluctuating hardness as a consequence of summing up the interactions between the penetrometer tip and individual snow particles. We describe this penetration process in analogy to the well-known Brownian motion [19], where a microscopically visible particle suspended in fluid is moving randomly. Due to the sum of several collisions with the molecules in the fluid as illustrated in Fig. 3.1, the large red particle undergoes a motion described by a stochastic process. Such a stochastic process is driven by white noise and is known

as a Langevin process (see [24], and Sect. 3.2). While very similar elementary collision events are assumed for classical Brownian motion, we also need to consider brittle failures of individual snow microstructural elements (bonds between snow grains, crushing of grain clusters), which cause sharp declines in penetration hardness. These brittle failures were modeled as a Poisson shot noise process by [49] and [64]. Jump noise acts as discontinuous paths inside the diffusion process, and low-jump events can be considered Poisson distributed noise (which corresponds to the abovementioned shot noise). The idea of this work is to employ a method that allows estimating the underlying stochastic differential equation from empirical data and differentiating between a Langevin (pure diffusive) and a jump-diffusion process [5]. Via this advanced analysis, we seek more detailed snow characterization from micro-cone penetration test resistance data.<sup>2</sup>



**Figure 3.1:** (a) Penetration resistance caused by the interactions of the snow particles (gray) with the penetrometer tip (red); (b) Brownian motion where a microscopically visible particle (red) suspended in the fluid is moving randomly due to the collisions with the molecules (gray) in the fluid.

The article is organized as follows. In Sect. 2, we summarize the stochastic analysis method and show how it is possible to distinguish between diffusive and jump noise. In Sect. 3, the method is applied to centimetric snow samples whose microstructure is also captured by tomography. In

<sup>2</sup>A direct comparison of our stochastic approach with the works based on shot noise is out of the scope of this paper.



Sect. 4, as a proof of concept, the SMP profile of a natural snowpack is analyzed with this technique.

## 3.2 Stochastic Method

A stochastic process  $x(t)$  can be described through stochastic differential equations. This section explains the equations used to model cone penetration tests in snow. Since the SMP is driven by a motor with constant speed  $u = \frac{dz}{dt}$  ( $z$  is depth;  $t$  is time) and samples the measurement every  $4 \mu\text{m}$ , the measured penetration force or snow hardness  $R$  is considered the depth dynamics  $R(z(t))$  and handled like a stochastic variable  $x(t)$ .

### 3.2.1 Langevin Equation

A diffusive process  $x(t)$ , which is a continuous stochastic process, follows the Langevin equation, where for a small step size  $dt$  has the following expression [23, 74, 81]:

$$dx(t) = D^{(1)}(x, t) dt + \sqrt{D^{(2)}(x, t)} dW_t, \quad (3.1)$$

where  $D^{(1)}(x, t)$  and  $D^{(2)}(x, t)$  are the drift and the diffusion coefficients, respectively, and  $W_t$  is the Wiener process. The drift term  $D^{(1)}$  describes how fluctuations relax to the local mean values of  $x$ , defined by  $D^{(1)}(x, t) = 0$ . The diffusion term  $D^{(2)}$  represents the amplitude of the noise. The coefficients  $D^{(1)}$  and  $D^{(2)}$  are also known as the first- and second-order Kramers–Moyal (KM) coefficients, respectively. In general, KM coefficients can be directly determined from the given data  $x(t)$  using their definitions of the conditional incremental average [23, 81]; i.e.,

$$K^{(j)}(x, t) = \lim_{\Delta t \rightarrow 0} \frac{\langle (x(t + \Delta t) - x(t))^j |_{x(t)=x} \rangle}{\Delta t}. \quad (3.2)$$

Further details on methods of this estimation can be found in [23] and

[73]<sup>3</sup>. The Langevin equation describes a continuous diffusion process where  $K^{(j)}(x, t) = 0$  for  $j \geq 3$  and  $D^{(j)}(x, t) = K^{(j)}(x, t)$ . According to Pawula's theorem, all KM coefficients  $K^{(j)}$  are vanishing for  $j \geq 3$  if  $K^{(4)}(x, t) = 0$  [62, 74]. In our case, however, the higher-order coefficients are not always vanishing; hence we extend the discussion to the jump-diffusion process (see Sect. 3.2.2).

For the SMP data considered, the drift term  $D^{(1)}$  in Eq. (3.1) describes how hardness fluctuations relax to the local mean values of hardness  $R$ , defined by  $D^{(1)}(R, z) = 0$ . The diffusion  $D^{(2)}$  term represents the amplitude of the hardness fluctuations. The coefficients  $D^{(1)}$  and  $D^{(2)}$  are  $z$ -dependent for non-stationary (inhomogeneous) processes. Here, we assume that for a chosen small depth interval ( $z$ ),  $D^{(1)}(R, z)$  and  $D^{(2)}(R, z)$  only depend on  $R$  (similarly to [49], for their shot noise model).

### 3.2.2 Jump-diffusion Dynamics

Typically, when the signal of a stochastic process presents sharp changes at some instant (discontinuity events), higher-order Kramers–Moyal coefficients (especially  $K^{(4)}(x, t)$ ) become non-negligible. In this case, an extension to Langevin-type modeling with additional jump noise is needed [5, 11, 42, 79, 81, 82]. Such a jump-diffusion dynamic is given by the following stochastic differential equation:

$$dx(t) = D^{(1)}(x, t) dt + \sqrt{D^{(2)}(x, t)} dW_t + \xi dJ_t, \quad (3.3)$$

where, again,  $D^{(1)}$  and  $D^{(2)}$  are the drift and the diffusion coefficients, respectively, and  $W_t$  is the Wiener process. The quantity  $\xi$  is the size of the jump noise and is assumed to be normally distributed; i.e.,  $\xi \sim N(0, \sigma_\xi^2)$ , where  $\sigma_\xi^2(x, t)$  is the so-called jump amplitude. The term  $J_t$

---

<sup>3</sup>KM coefficients for the Langevin equation are defined as  $K^{(j)}(x, t) = D^{(j)}(x, t) = \frac{1}{j!} \lim_{\Delta t \rightarrow 0} \frac{\langle (x(t + \Delta t) - x(t))^j |_{x(t)=x} \rangle}{\Delta t}$  in [23] and [73]. In order to make it consistent with the jump-diffusion process, our definition differs by a factor of  $\frac{1}{j!}$ , in which  $dW_t = \Gamma(t) \cdot dt$  where  $\langle \Gamma(t) \rangle = 0$  and  $\langle \Gamma(t) \Gamma(t') \rangle = \delta(t - t')$ . The corresponding Fokker–Planck equation will be  $\frac{\partial}{\partial t} p(x, t) = -\frac{\partial}{\partial x} [D^{(1)}(x, t) p(x, t)] + \frac{1}{2} \frac{\partial^2}{\partial x^2} [D^{(2)}(x, t) p(x, t)]$ .

is the Poisson jump process, which is the zero–one jump process with a jump rate (or intensity)  $\lambda(x, t)$  [33, 81].

For jump-diffusion processes, the drift and diffusion coefficients ( $D^{(1)}$ ,  $D^{(2)}$ ), the jump rate  $\lambda$  and amplitude  $\sigma_\xi^2$  are related to the KM coefficients as [5]:

$$D^{(1)}(x, t) = K^{(1)}(x, t), \quad (3.4)$$

$$D^{(2)}(x, t) + \lambda(x, t)\langle \xi^2 \rangle = K^{(2)}(x, t), \quad (3.5)$$

$$\lambda(x, t)\langle \xi^j \rangle = K^{(j)}(x, t), \quad \text{for } j > 2. \quad (3.6)$$

The estimate of the drift coefficient is the same for the diffusion process (Eq. 3.1) and the jump-diffusion process (Eq. 3.3). Jump amplitude  $\sigma_\xi^2$  and rate  $\lambda$  can be estimated by using Eq. (3.6) with  $j = 4$  and  $j = 6$  and Wick's theorem [39, 86] for Gaussian random variables, which states that  $\langle \xi^{2n} \rangle = \frac{(2n)!}{2^n n!} \langle \xi^2 \rangle^n$ :

$$\sigma_\xi^2(x, t) = \frac{K^{(6)}(x, t)}{5K^{(4)}(x, t)}, \quad (3.7)$$

$$\lambda(x, t) = \frac{K^{(4)}(x, t)}{3\sigma_\xi^4(x, t)}. \quad (3.8)$$

To improve the estimation of KM coefficients  $K^{(j)}(x, t)$  and in particular of high-order coefficients, the Nadaraya–Watson estimator, which is a kernel estimator, can be used [61, 85]:

$$K^{(j)}(x, t) = \lim_{\Delta t \rightarrow 0} \frac{\sum_i k\left(\frac{x_{i\Delta t} - x}{h}\right) (x_{(i+1)\Delta t} - x_{i\Delta t})^j}{\sum_i k\left(\frac{x_{i\Delta t} - x}{h}\right) \Delta t}, \quad (3.9)$$

where here we use a Gaussian kernel  $k(u)$  [81]. With the kernel-based method the conditional moments can be calculated more smoothly by controlling the kernel bandwidth  $h$  [44], where here we use the kernel bandwidth  $h = 0.3$ .

For the SMP measurements considered, we also assume that for a chosen small depth interval ( $z$ ),  $D^{(1)}(R, z)$ ,  $D^{(2)}(R, z)$ ,  $\sigma_\xi^2(R, z)$  and  $\lambda(R, z)$  only

depend on  $R$ . The stochastic differential equation for the jump-diffusion process thus reads, at small depth intervals, as

$$dR(z) = D^{(1)}(R) dz + \sqrt{D^{(2)}(R)} dW_z + \xi dJ_z, \quad (3.10)$$

with an interpretation of the drift and diffusion terms ( $D^{(1)}$ ,  $D^{(2)}$ ) analogous to the purely diffusive process (Eq. 3.1) but now extended by a jump noise term. This is the same as Eq. (3.3) in terms of depth  $z$  instead of time  $t$ . The jump rate  $\lambda$  has the dimension of  $\frac{1}{[Z]}$  and can be related to the shot noise intensity described by [49] and [64]. Typically,  $\lambda dz$  corresponds to the stochastic average of the number of jumps for a penetration increment of depth  $dz$  [5, 81]. The jump amplitude  $\sigma_\xi^2$  represents the square of the typical size of a jump. Note that the jump can be negative (failure of a microstructural element) or positive (loading of a microstructural element). Here, we do not consider any progressive loading of a microstructural element as described by [49] and [64] with the microstructural deflection length  $\delta$ . Here, the loading of a microstructural element and its contribution to penetration hardness are considered instantaneous.

### 3.2.3 Synthetic Examples

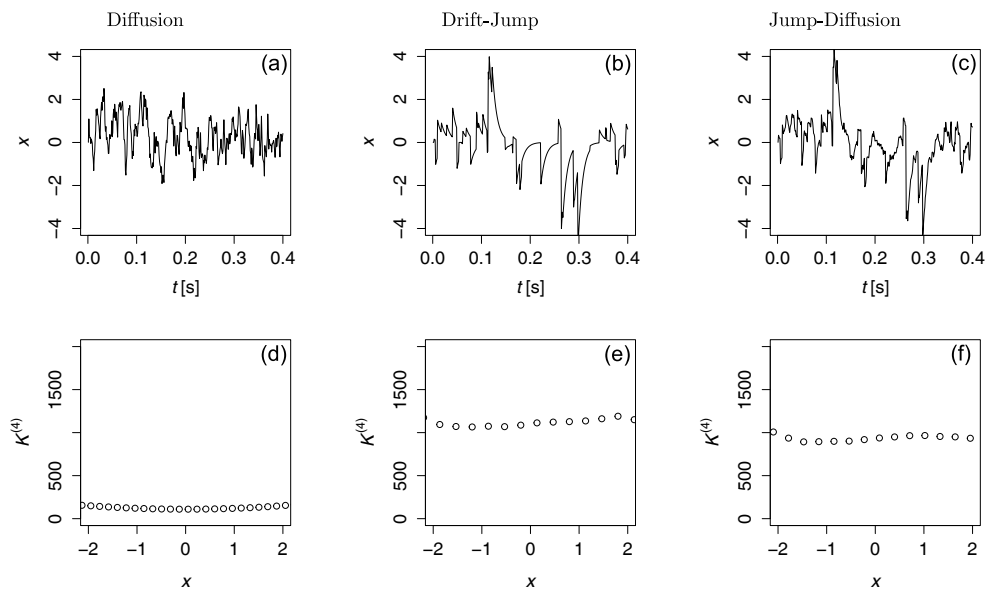
In this section, we illustrate how diffusive and jump noises affect the stochastic fluctuations on a synthetic example. An Ornstein–Uhlenbeck (OU) process  $x$  is described by a stochastic differential equation (SDE) with linear relaxation dynamics and an additive uncorrelated noise. With an additional jump term, it is defined as

$$dx = -\gamma x dt + \sqrt{D} dW_t + \xi dJ_t. \quad (3.11)$$

where  $\gamma$  is the relaxation rate,  $D$  the constant diffusion coefficient and  $W_t$  a scalar Wiener process. The noise  $\xi \sim N(0, \sigma_\xi^2)$  is assumed to be normally distributed with the constant variance or jump amplitude  $\sigma_\xi^2$ .  $J_t \sim P(\lambda t)$  is the Poisson jump process, which is a zero–one jump process

with constant jump rate  $\lambda$ . Here, we have triply stochastic processes  $W_t$ ,  $J_t$  and  $\xi$ , which are assumed to be independent of each other.

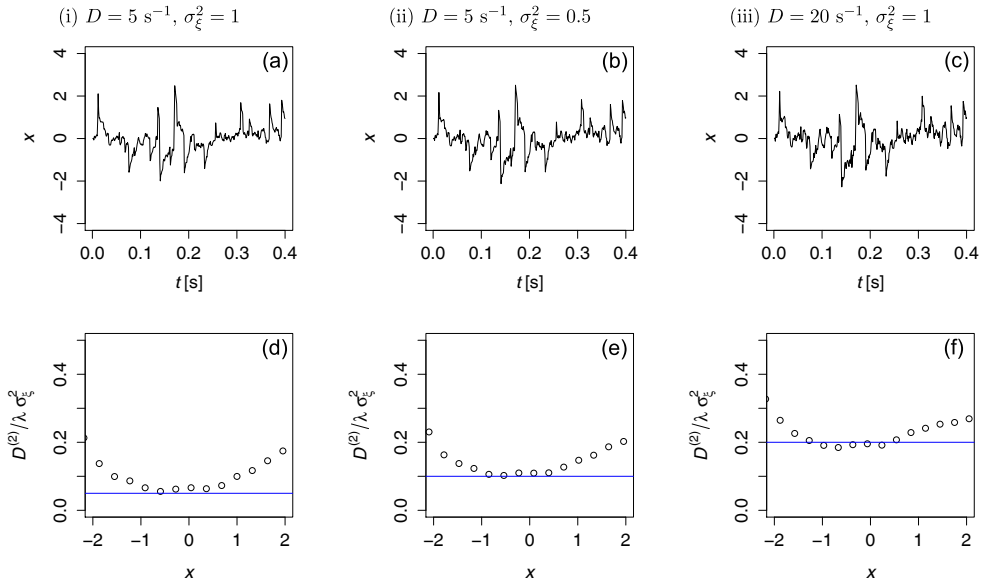
Three synthetic time series of the OU jump-diffusion process were generated for  $\Delta t = 10^{-3}$  s with  $\gamma = 100$  s $^{-1}$ ,  $D = 10$  s $^{-1}$  and for the additional jump terms with  $\lambda = 100$  s $^{-1}$  and  $\sigma_\xi^2 = 1$ . The generated data were normalized with their respective standard deviation. In Fig. 3.2 the normalized time series of the OU jump-diffusion processes are shown. A pure diffusion process (left), a drift-jump process (middle) and the combined jump-diffusion process (right) are shown. The fourth-order KM coefficients  $K^{(4)}(x)$  of each time series are also plotted in Fig. 3.2, bottom row.  $K^{(4)}$  of the diffusion process is negligibly small compared to drift-jump and jump-diffusion processes.



**Figure 3.2:** Normalized time series of Ornstein–Uhlenbeck processes with only diffusion, only jump (drift-jump) and jump-diffusion terms (a, b, c) and their fourth-order KM coefficients  $K^{(4)}$  (d, e, f).  $K^{(4)}$  of the diffusion process is negligible compared to drift-jump and jump-diffusion processes. For all three examples, we used the same noises in the stochastic part of the stochastic differential equation.

For a jump-diffusion process, another parameter that we considered was the ratio of diffusion and jump noise  $\frac{D^{(2)}}{\lambda\sigma_\xi^2}$ , which becomes here for our

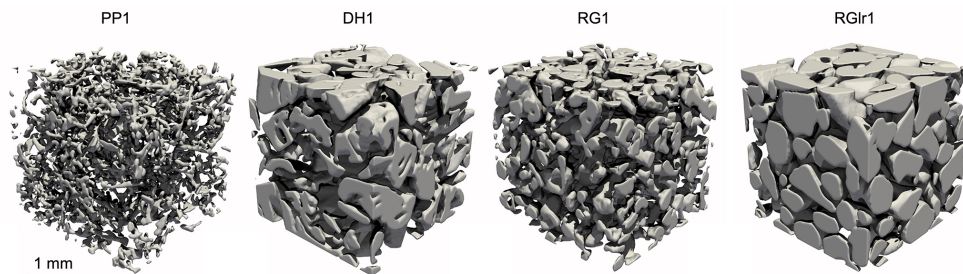
OU jump-diffusion process  $\frac{D}{\lambda\sigma_\xi^2}$ . To validate our method, based on the KM coefficients of Eq. (3.5) to Eq. (3.8), three pairs of parameters were chosen: (i)  $D = 5 \text{ s}^{-1}$ ,  $\sigma_\xi^2 = 1$ ; (ii)  $D = 5 \text{ s}^{-1}$ ,  $\sigma_\xi^2 = 0.5$ ; and (iii)  $D = 20 \text{ s}^{-1}$ ,  $\sigma_\xi^2 = 1$ . The other parameters are the same as in the previous example. The normalized time series of these examples are plotted in Fig. 3.3, top row. In Fig. 3.3, bottom row, the ratios of diffusion and jump noise  $\frac{D}{\lambda\sigma_\xi^2}$  estimated from each time series were compared to the expected values (blue lines). As we used normalization and the same noises in simulation, all time series are very similar; however, one can observe clearly the much noisier fine structure in case (iii) where the diffusion coefficient is larger.



**Figure 3.3:** Normalized time series of OU jump-diffusion processes with  $\Delta t = 10^{-3} \text{ s}$ ,  $\gamma = 100 \text{ s}^{-1}$  and  $\lambda = 100 \text{ s}^{-1}$  for (i)  $D = 5 \text{ s}^{-1}$ ,  $\sigma_\xi^2 = 1$ ; (ii)  $D = 5 \text{ s}^{-1}$ ,  $\sigma_\xi^2 = 0.5$ ; and (iii)  $D = 20 \text{ s}^{-1}$ ,  $\sigma_\xi^2 = 1$  (**a**, **b**, **c**) and the corresponding ratio of diffusion and jump noise  $\frac{D^{(2)}}{\lambda\sigma_\xi^2} = \frac{D}{\lambda\sigma_\xi^2}$  (**d**, **e**, **f**). Dots are results from the KM coefficients, and the blue line denotes the theoretical values given by the constants. For all three examples, we used the same noises in the stochastic part of the SDE.

### 3.3 Application to Snow Measurements

In this section, our main aim is to show how the jump-diffusion model can be used to distinguish snow types from hardness profiles measured with the SMP. Firstly, small snow samples whose microstructure was also fully characterized by tomography before being measured by the SMP were used to test the developed methodology. Secondly, as a proof of concept, we analyzed one penetration profile of a snowpack measured in the field and we provided the subsequent profile of microstructural parameters. Last, the results were discussed.



**Figure 3.4:** Three-dimensional view of the microstructure of some representative samples: precipitation particles (PP1), depth hoar (DH1), rounded grains (RG1) and large rounded grains (RGl1). The ice matrix is shown in gray; the pore space is transparent. Sub-samples shown are cubic with a side length of 3 mm. Details on the data acquisition can be found in [63].

#### 3.3.1 Laboratory Samples

##### Measurement Data

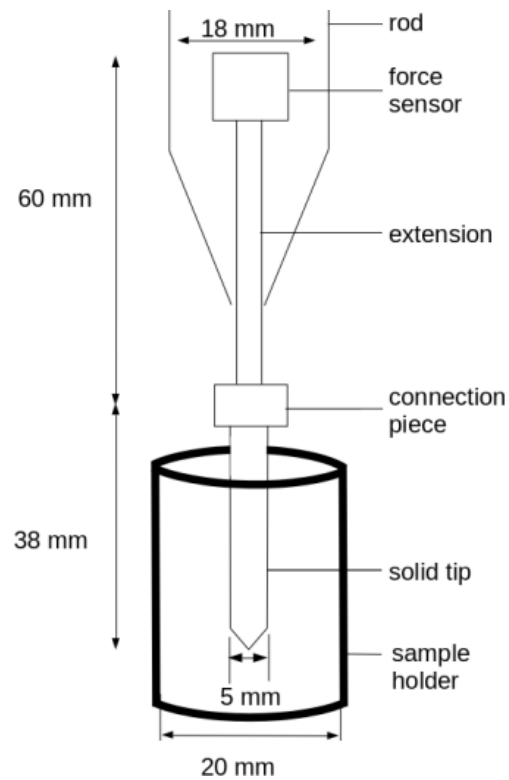
We tested several snow samples composed of four different natural snow types, namely precipitation particles (PP), depth hoar (DH), rounded grains (RG) and large rounded grains (RGl) as classified in [21]. The samples were prepared by sieving snow into small sample holders (diameter and height of 20 mm) and letting them sinter for a couple of days at  $-10$  °C. Their microstructure was captured with X-ray tomography at a nominal resolution of  $15$   $\mu\text{m}$  (Fig. 3.4). The cone penetration test was conducted with a modified version of the SMP, as shown in Fig. 3.5.

More information on sample preparation and the SMP measurement can be found in the study of [63]. The main sample properties are summarized in Table 3.1, and the measured hardness profiles (one example for each snow type) are plotted in Fig. 3.6. The first 4 mm is affected by the progressive penetration of the conic tip and was not considered in the stochastic analysis [64]. The remaining profiles were divided into smaller segments of a depth of 10 mm.

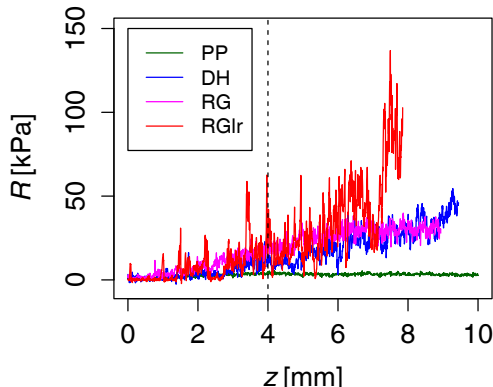
**Table 3.1:** Overview of the detailed properties of the snow samples used. Snow types are classified according to the international classification of snow on the ground [21]. The density and specific surface area (SSA) are derived from the tomographic images from [63]. Additionally, the standard deviations  $\sigma_R$  of the detrended profiles are calculated.

Sample name	Snow type	Sieve size (mm)	Density ( $\text{kg m}^{-3}$ )	SSA ( $\text{m}^2 \text{kg}^{-1}$ )	$\sigma_R$ (kPa)
PP1	Precipitation particles	1.6	92	53.5	0.55
PP2	Precipitation particles	1.6	137	41	0.81
DH1	Depth hoar	1.6	345	16.9	4.78
DH2	Depth hoar	1.6	364	15.9	3.67
DH3	Depth hoar	1.6	364	16.5	3.89
RG1	Rounded grains	1.6	289	23.0	3.29
RG2	Rounded grains	1.6	304	23.7	3.81
RG3	Rounded grains	1.6	325	20.6	3.63
RGr1	Large rounded grains	1	530	10.1	13.20
RGr2	Large rounded grains	1	544	10.3	11.78
RGr3	Large rounded grains	1.6	557	9.9	8.99
RGr4	Large rounded grains	1	542	9.3	14.01
RGr5	Large rounded grains	1	541	9.7	20.49
RGr6	Large rounded grains	1	526	10.1	17.22





**Figure 3.5:** Setup of micro-cone penetration test [63]. The cone penetration tests (CPTs) were conducted by inserting a cylinder of diameter 5 mm, with a conical tip of an apex angle of  $60^\circ$ , into the snow samples. The samples were placed in the cylinder sample holder with a diameter of 20 mm. The cone was inserted vertically at a constant speed of  $20 \text{ mm s}^{-1}$ . The SMP force sensor (Kistler 9207) measures forces in the range of  $[0, 40] \text{ N}$  with a resolution of 0.01 N.



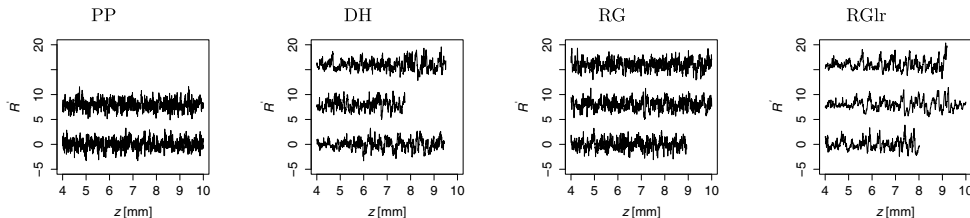
**Figure 3.6:** Segments of snow hardness profiles of PP1, DH1, RG1 and RGl1. These four different types occur as natural snow types. Precipitation particles (PP) have the smallest trend and fluctuation force; large rounded grains (RGl1) are the largest, while depth hoar (DH) and rounded grains (RG) have similar trends and fluctuation forces between those of PP and RGl1. The first 4 mm is affected by the progressive penetration of the conic tip and was not considered in the stochastic analysis.

To work out the significance of advanced stochastic features for snow, we focused on the fluctuations in the hardness profiles. Each profile was first detrended. The trend  $\bar{R}$  was computed as the convolution of the original signal with a Gaussian kernel with a standard deviation of 0.6 mm.<sup>4</sup> The fluctuation amplitude  $\sigma_R$  was computed as the standard deviation of  $R - \bar{R}$ . The detrended profiles are defined as  $R' = \frac{R - \bar{R}}{\sigma_R}$ . The detrended profiles  $R'$  are characterized by a zero mean and a standard deviation of 1. The average value of  $\bar{R}$  on the segment and the value of the standard deviation  $\sigma_R$  are shown for each segment in Table 3.1. The detrended profiles  $R'$  are shown in Fig. 3.7, for all four snow types.

To estimate errors, we divided the detrended and normalized data into different sub-samples. Given are two PP, three DH, three RG and six RGl1 measurement profiles. These profiles were separated into smaller segments, which finally gives four PP, five DH, five RG and six RGl1 samples. We estimated the KM coefficients of each sample using Eq. (3.2) and averaged them over the sub-samples of each snow type. Thus, drift,

<sup>4</sup>The results do not change significantly if the kernel widths are changed between 0.14 and 0.66 mm, where this range corresponds to the range of average grain sizes of the snow types.

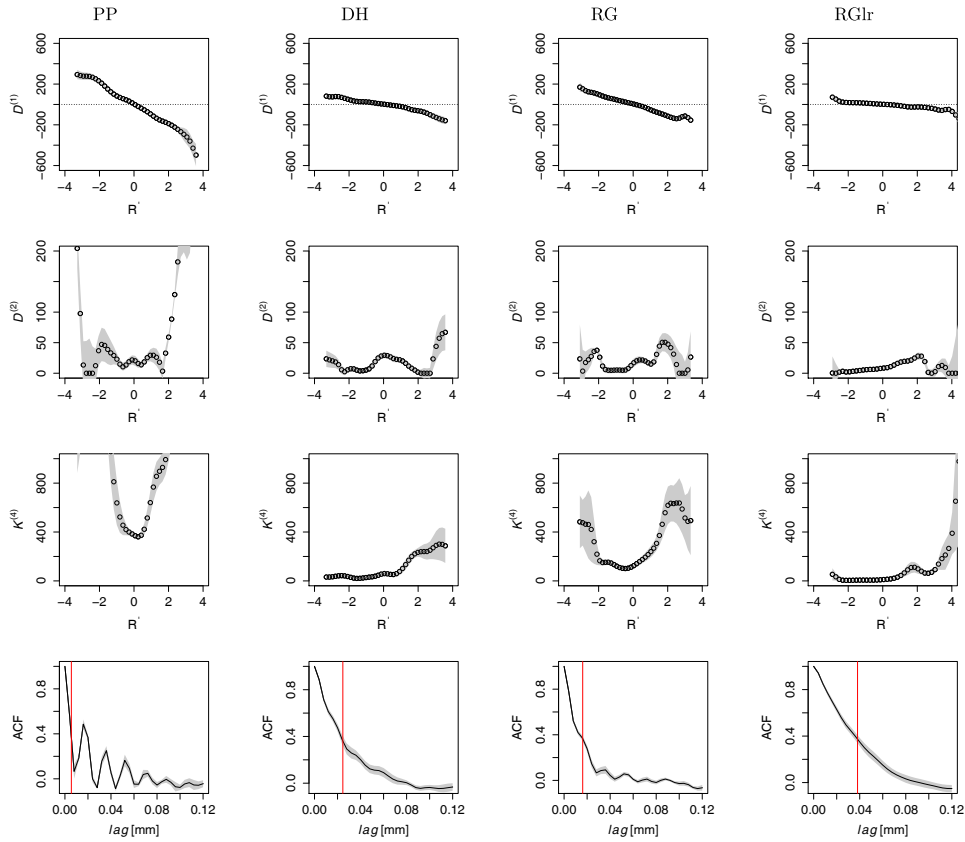
diffusion functions and jump parameters and their errors were estimated. The errors were reported as the standard error of the means.



**Figure 3.7:** Detrended snow hardness profiles of four different snow types, PP, DH, RG and RGlr. For the snow type RGlr, we only show the samples RGlr4, RGlr5 and RGlr6, as described in Table 3.1. The detrended profiles  $R'(z)$  of each snow type are shifted vertically for better visualization.

## Results

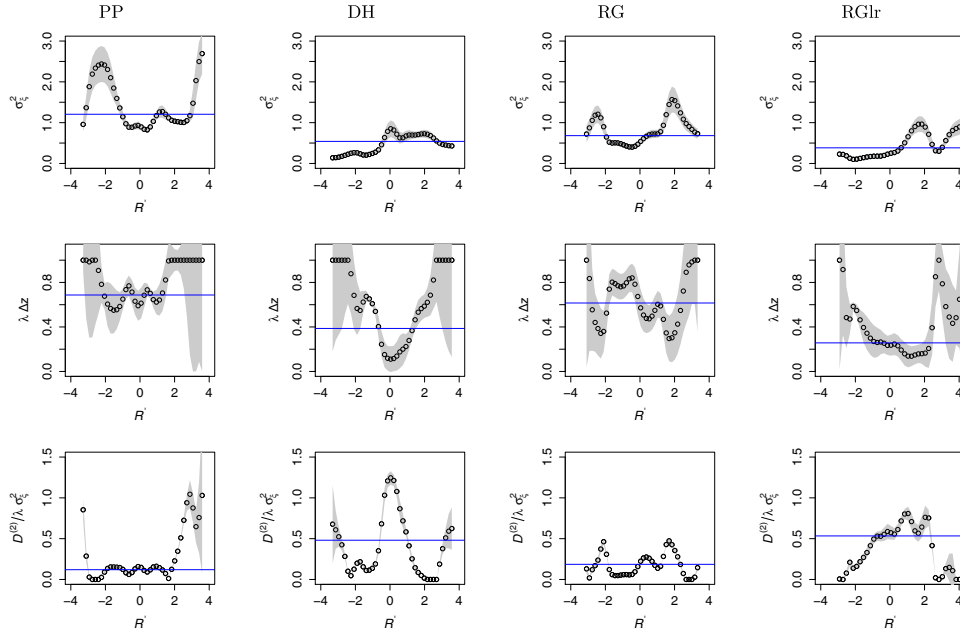
According to the description provided in Sect. 3.2, the drift  $D^{(1)}(R')$  and diffusion  $D^{(2)}(R')$ , as well as the fourth-order KM coefficients  $K^{(4)}(R')$ , for normalized data were determined for the four different snow types, PP, DH, RG and RGlr as shown in Fig. 3.8. In addition, the autocorrelation function (ACF) was determined from the signals. If the fluctuations in  $R'(z)$  belong to the diffusion processes, one would expect that  $K^{(4)}(R') = 0$ . However, we find that in general,  $K^{(4)}(R') \neq 0$  and the higher-order KM coefficients are not negligible. This indicates the presence of discontinuities in the snow hardness profile, so the jump-diffusion model is considered. Therefore, we estimated jump parameters, i.e., jump amplitudes  $\sigma_{\xi}^2(R')$  and jump rates  $\lambda(R')$  from data of  $R'(z)$  in Fig. 3.9.



**Figure 3.8:** State-dependent drift  $D^{(1)}(R')$ , diffusion  $D^{(2)}(R')$ , fourth-order KM coefficients  $K^{(4)}(R')$  and their respective autocorrelation functions (ACFs) of four different snow types, PP, DH, RG and RGlR (left to right). The errors are shown as gray-shaded background. The red lines in the ACF plots indicate the correlation length scales determined from ACFs. Comparing the correlation length scales  $L_C = \frac{1}{\gamma}$  where  $D^{(1)} = -\gamma R'$  with those of the autocorrelation functions (ACFs), we find that both length scales have the same ordering of their values for all snow types.

As shown in Fig. 3.8, the drift coefficients,  $D^{(1)}(R', z)$ , are mostly linear functions with negative slopes, which describe how fast the system tends back toward the stable fixed point. Due to our normalization, the fixed point of dynamics is located at the origin, i.e.,  $R' = 0$ . Taking  $D^{(1)} = -\gamma R'$ , the correlation length scale is given by  $L_C = \frac{1}{\gamma}$ . For each snow type,  $L_C$  was determined for  $-2 < R' < 2$ :  $\{PP, DH, RG, RGlR\} = \{0.01, 0.04, 0.02, 0.08\}$  mm. Comparing  $L_C$

with the correlation lengths of the autocorrelation functions,  $-L_{ACF}$ ,  $\{\text{PP, DH, RG, RGlR}\} = \{0.006, 0.025, 0.016, 0.038\}$  mm – we find that both length scales have the same ordering of their values for all snow types. The snow types PP and RG have the shorter correlation length scale, in comparison to the snow types DH and RGlR.



**Figure 3.9:** Jump amplitudes  $\sigma_{\xi}^2(R')$ , jump probabilities  $\lambda(R') \Delta z$ , and diffusion and jump ratios  $\frac{D^{(2)}}{\lambda \sigma_{\xi}^2}$  of four different snow types, PP, DH, RG and RGlR (left to right). More data are present in the range of  $-2 < R' < 2$ ; we focus our statistical analysis in this range with fewer uncertainties. The errors are shown as gray-shaded background. The ratios of diffusion and jump noise  $\frac{D^{(2)}}{\lambda \sigma_{\xi}^2}$  for PP and RG are minimum near zero and maximum for DH and RGlR, which means that the larger the ice structure, the stronger the diffusion noise, and vice versa. The horizontal blue lines show the mean values of the respective parameters in the range of  $-2 < R' < 2$ .

The jump amplitudes,  $\sigma_{\xi}^2(R')$ , and the jump probabilities,  $\lambda(R') \Delta z$ , are shown in Fig. 3.9. The jump amplitude,  $\sigma_{\xi}^2(R')$ , indicates how large the jump noise for different  $R'$  values is. The jump probability describes how probable jumps or discontinuities in forces can occur. To analyze whether

diffusion or jump noise is dominating, the dimensionless ratio of diffusion and jump noise  $\frac{D^{(2)}}{\lambda\sigma_\xi^2}$  (Fig. 3.9, bottom row) was calculated. For a rough estimation, the mean values were determined in the range of  $-2 < R' < 2$  and are plotted as horizontal blue lines. The mean of  $\bar{\lambda}$  can be used to define the second characteristic length scale  $L_J = \frac{1}{\lambda}$  (apart from  $\frac{1}{\gamma}$ ). For the abovementioned range of  $R'$  we obtained  $\{\text{PP, DH, RG, RGlR}\} = \{0.006, 0.01, 0.007, 0.02\}$  mm. Results are summarized in Table 3.2, and we discuss these in Sect. 3.4.

**Table 3.2:** Summary of the results for the correlation length scales  $L_C = \frac{1}{\gamma}$ ,  $L_J = \frac{1}{\lambda}$ , jump amplitude  $\sigma_\xi^2$ , and diffusion and jump ratio  $\frac{D^{(2)}}{\lambda\sigma_\xi^2}$  for all snow types analyzed. The results were evaluated in the range of  $-2 < R' < 2$ .

Snow type	$L_{\text{ACF}}$ (mm)	$L_C = \frac{1}{\gamma}$ (mm)	$L_J = \frac{1}{\lambda}$ (mm)	$\overline{\sigma_\xi^2}$	$\frac{D^{(2)}}{\lambda\sigma_\xi^2}$
PP	$(0.006 \pm 0.004)$	$(0.0101 \pm 0.0009)$	$(0.0058 \pm 0.0002)$	$(1.21 \pm 0.09)$	$(0.120 \pm 0.008)$
DH	$(0.025 \pm 0.004)$	$(0.040 \pm 0.009)$	$(0.010 \pm 0.001)$	$(0.54 \pm 0.05)$	$(0.48 \pm 0.09)$
RG	$(0.016 \pm 0.004)$	$(0.020 \pm 0.002)$	$(0.0065 \pm 0.0004)$	$(0.68 \pm 0.06)$	$(0.19 \pm 0.03)$
RGlR	$(0.038 \pm 0.004)$	$(0.08 \pm 0.02)$	$(0.016 \pm 0.001)$	$(0.38 \pm 0.07)$	$(0.53 \pm 0.04)$

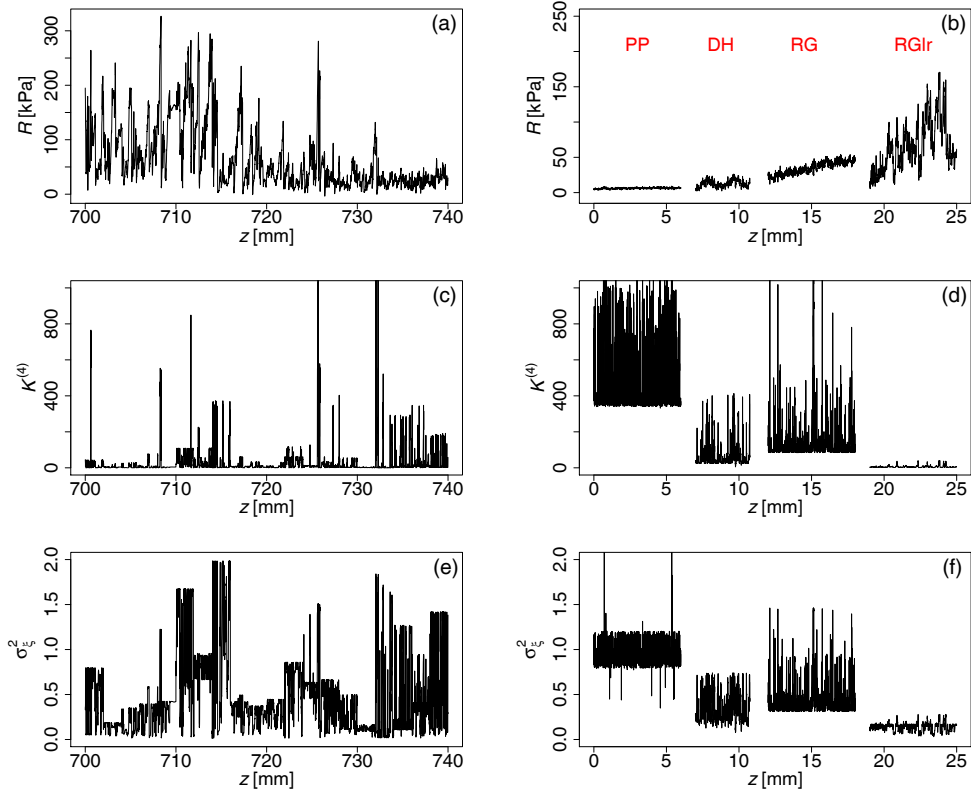
### 3.3.2 Application to Field Snow Data

#### Measurement Data

Next, measurements from a field campaign are presented [32]. The measurements were also performed with an SMP, but the tip had a slightly different shape corresponding to the standard version of the SMP [43]. The spatial sampling is again  $4 \mu\text{m}$ . This difference in the measurement methods was irrelevant, as we subsequently show with these preliminary results that in principle, the stochastic methodology can also be applied to real snow data and that qualitatively comparable results are obtained.

## Results

The snow hardness profile of a field measurement is shown in the top left of Fig. 3.10. The measurement profile is strongly inhomogeneous; therefore, we used the Nadaraya–Watson estimator to determine the local characteristics of the profile. Using the moving-window technique, the profile was separated into non-overlapping windows of 500 data points (2 mm), and the detrending was performed on each window by the Gaussian kernel with a kernel size of 0.6 mm, normalized with its standard deviation as in our previous analysis of laboratory data. For each depth value  $z$  and the corresponding value  $R'(z)$ , the local values of the fourth-order KM coefficient  $K^{(4)}(z)$  and the jump amplitude  $\sigma_{\xi}^2(z)$  can be determined, as shown in Fig. 3.10. The local characteristic of each snow type from the previous section is plotted in the right column of Fig. 3.10 for better comparison with the field measurement data. Interpretation of these results will be discussed next.



**Figure 3.10:** The snow hardness profile of a field measurement together with its fourth-order KM coefficients  $K^{(4)}$  and the jump amplitude  $\sigma_\xi^2(z)$  is shown in (a), (c) and (e). These parameters were determined using non-overlapping moving window with 500 data points (2 mm) by means of the Nadaraya–Watson estimator. Snow hardness profiles of the laboratory-prepared snow types and their local parameters are also plotted to enable better comparison (b, d, f); they are shifted horizontally for better visualization. The results of field measurements are shown at the depth of  $700 \text{ mm} < z < 740 \text{ mm}$ . With reference to the local characteristic snow types from laboratory measurements, we can see in the dynamics that mixtures of different snow types are present in this section of measurement. For  $732 \text{ mm} < z < 740 \text{ mm}$ , the high  $K^{(4)}$  and  $\sigma_\xi^2$  indicate the presence of small and less dense structures of snow which resemble the RG-like snow types.



### 3.4 Discussion

Our work is based on the proposed analogy of Brownian motion and the SMP penetration process, as illustrated in Fig. 3.1. The events of bond breaking or of collision with molecules are summed up in a mean force and noise. For continuous Brownian noise, we need an integration over sufficient micro-scale events, as discussed by [19] in his original paper. In our interpretation, it is found that sufficiently large particles lead to this integration; see Fig. 3.1b. To our interpretation, this integration over discrete single events of bond breaking in the immediate surroundings of the SMP, in addition to the pushing aside of loose snow grains during the penetration process, forms continuous Brownian noise. The jump noise may represent the bond-breaking events occurring directly at the tip of the SMP, and the amplitude of the jump noise should depend on the strength of the ice bonds and void sizes. From this interpretation, it is clear that snow type morphology, shown in Fig. 3.4, is essential for effective stochastic analysis as outlined herein.

We start the discussion with the mean values of  $R$  and the standard deviation  $\sigma_R$  (see Fig. 3.6 and Table 3.1). The less dense PP and the dense RGlR snow can be well separated, whereas the differences are less clear for DH and RG. In the following, we discuss the measured SMP penetration profiles based on our stochastic results. Since we now focus on a stochastic investigation of the fluctuations in the penetration profiles, the detrended and normalized data are  $R'$  used. Furthermore, we can note that the normalization of the snow profiles affects neither the correlation length scales  $L_C = \frac{1}{\gamma}$  from the drift coefficients nor the jump characteristic length scales  $L_J = \frac{1}{\lambda}$ . Because this analysis depends on the number of available data, our discussion of the estimated KM coefficients is limited to the range  $-2 < R' < 2$ .

The drift terms  $D^{(1)}$ , Fig. 3.8, are all monotonously decaying with increasing  $R'$  and can be approximated by a linear decay,  $D^{(1)} = -\gamma R'$ . The slope indicates how fast the signals relax to a fixed point located at  $R' = 0$ . The magnitudes of the slopes are  $PP > RG > DH > RGlR$ ; thus the PP snow has the fastest relaxation or shortest correlation length scale

$L_C$ . We find that the larger the ice structures, the slower (or longer) the relaxation. If we compare this result with the snow structures shown in Fig. 3.4, we conclude that  $L_C$  or  $\gamma$  is clearly related to the size of the snow structures.

The results for  $D^{(2)}$  show that about the same diffusive noise amplitude is found for all snow types. In contrast, we see clear differences for the fourth-order KM coefficients  $K^{(4)}$ . Although  $K^{(4)} \neq 0$  is always the case, clear differences in the magnitude of this KM coefficient are found.  $K^{(4)}$  is the largest in PP, followed by RG, DH and RGlR.

The amplitudes of jump noise,  $\sigma_\xi^2$ , are the highest for PP, followed by RG, DH and RGlR. For the jump probabilities  $\lambda \Delta z$ , we distinguish a group composed of PP and RG, with higher values, and one composed of DH and RGlR, with lower values. One can interpret this finding such that for the precipitation particles (PP, recent snow) with very small ice structures and high porosity, the breaking occurs easily and frequently, which explains that the jump probability has the largest contribution here. Similarly, RG is also less dense with smaller ice structures than DH and RGlR. Thus, we can also interpret this finding such that the smaller the ice structures and the less densely the snow is packed, the stronger the jump noise. For the densely packed snow with larger ice structures, the breaking of the ice structures is less frequent, which explains a lower jump probability. The mean of the jump rate  $\bar{\lambda}$  in the range of  $-2 < R' < 2$  can be used to define the second characteristic length scale  $L_J = \frac{1}{\bar{\lambda}}$ . Similarly to the correlation length scale  $L_C = \frac{1}{\gamma}$ , PP has the smallest length, followed by RG, DH and RGlR.

Besides the features of the different terms in the stochastic processes, the contributions of the diffusive and the jump noise can be compared by the dimensionless quotient of  $\frac{D^{(2)}}{\lambda \sigma_\xi^2}$ , i.e., the relation between the two noise contributions. Consistent with our earlier discussion, the smallest values for  $\frac{D^{(2)}}{\lambda \sigma_\xi^2}$  are obtained for PP; i.e., the jump noise dominates due to the frequent fracture of small (soft) ice structures. For the other snow types, we see that within the range of  $-2 < R' < 2$ , the values of  $\frac{D^{(2)}}{\lambda \sigma_\xi^2}$

increase with larger ice structures, in accordance with Fig. 3.4. The quotient  $\frac{D^{(2)}}{\lambda\sigma_\xi^2}$  is smaller for PP and RG and larger for DH and RGlr. For RGlr, the diffusive noise dominates in a broad range of  $R'$  values. For the densely packed snow with larger grain sizes, it also takes much force to push the ice grains on the side but not necessarily to break the cohesive bonds close to the tip. Therefore the penetration signal is dominated by Brownian noise. This result is also consistent with the value of  $K^{(4)}$ , which is relatively small for RGlr. It is interesting to see that  $\frac{D^{(2)}}{\lambda\sigma_\xi^2}$  and  $K^{(4)}$  enable differentiation between RG and DH. In contrast, according to the classical statistical features of the snow signals shown in Table 3.1, the differences for DH and RG are less clear.

In Sect. 3.3.2, we analyze field measurement data which are highly inhomogeneous. With reference to the laboratory measurements, we see dynamics that suggest mixtures of different snow types within this depth segment. For  $732 \text{ mm} < z < 740 \text{ mm}$ , the high  $K^{(4)}$  and  $\sigma_\xi^2$  indicate the presence of small and less dense structures of snow which resemble the RG-like snow types. Based on these preliminary results of real field data, the developed methodology appears promising for interpreting cone penetration tests in the field, but further quantitative evaluation is required.

## 3.5 Conclusions

In conclusion, we observe that the advanced stochastic analysis of SMP measurements of snow layers allows differentiation of snow types. The diffusive and jump-noise contribution can be quantified and gives new insights into the stochastic behaviors of the cone penetration test in snow. For different snow types, we find an interesting mixture of diffusive- and jump-like noise. We propose the interpretation that the dominant diffusive process is due to the pushing aside of many snow grains, whereas the breaking of ice structures leads to dominant jump noise. Our results show that the denser structures typical of DH and RGl<sub>r</sub> lead to a more collective diffusive behavior, whereas for the highly porous snow structures of PP and RG, the single breaking events lead to a relatively strong jump noise. For this interpretation, all  $R$  values were detrended and normalized; thus the absolute values of the snow hardness  $R$  are not essential but more the resulting collective behavior of the snow types.

Finally, we would like to point out that our characterization of a complex material, snow, by a penetration process should have the potential to be generalized to, for example, biological tissue or ground layers. Last but not least, we would like to point out that our work provides additional insight into analyzing and modeling the complex nature of snow types, complementing existing methods.

## Chapter 4

# Discontinuous Jump Behavior of the Energy Conversion in Wind Energy Systems<sup>1</sup>

### Abstract

The power conversion process of a wind turbine can be characterized by a stochastic differential equation (SDE) of the power output conditioned to certain fixed wind speeds. An analogous approach can also be applied to the mechanical loads on a wind turbine, such as generator torque. The constructed SDE consists of the deterministic and stochastic terms, the latter corresponding to the highly fluctuating behavior of the wind turbine. Here we show how advanced stochastic analysis of the noise contribution can be used to show different operating modes of the conversion process of a wind turbine. The parameters of the SDE, known as Kramers-Moyal (KM) coefficients, are estimated directly from the measurement data. Clear evidence is found that both, continuous diffusion noise and discontinuous jump noise are present. The difference in the noise contributions indicates different operational regions. In particular, we observe that the jump character or discontinuity in power production

---

<sup>1</sup>submitted as P. P. LIN, M. WÄCHTER, M. R. R. TABAR, J. PEINKE: Discontinuous Jump Behavior of the Energy Conversion in Wind Energy Systems, to “PRX Energy”.

has a significant contribution in the regions where the control system switches strategies. We find that there is a high increase in jump amplitude near the transition to the rated region, and the switching strategies cannot result in a smooth transition. The proposed analysis provides new insights to the control strategies of the wind turbine.

## 4.1 Introduction

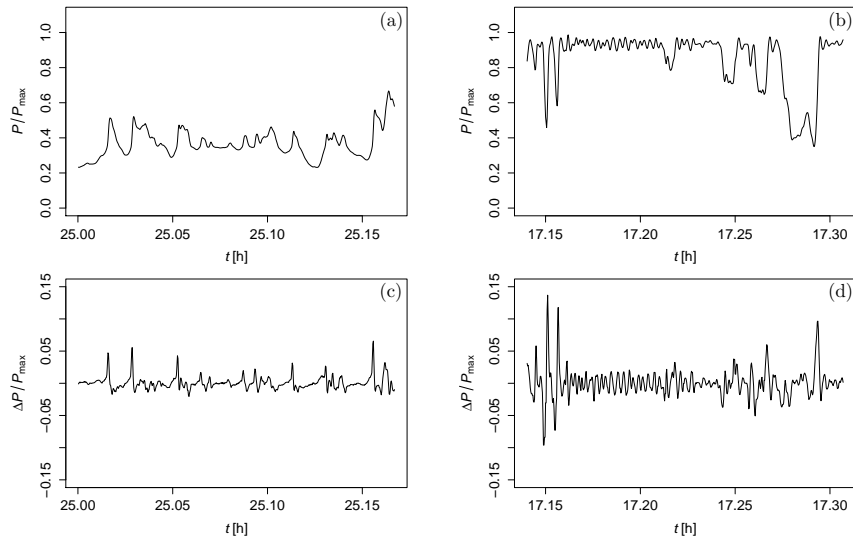
Wind energy is one of the most promising contributions to the global energy transition from fossil fuels to clean and sustainable energy. Europe could install around 105 GW of new wind energy capacity in the period of 2021–2025 as reported in [87]. However, the complex and intermittent nature of wind makes wind energy production difficult to predict, which is important for a stable energy supply [4, 14, 59, 76]. Furthermore this nature of wind may cause premature mechanical fatigue failure [60, 83]. It is known that the commonly used industry standard by the International Electrotechnical Commission [38] is not describing properly the variability of wind and wind power [31, 83]. In particular, if one investigates time series of the power output of a wind turbine, one can find very rapid power fluctuations, which can become larger than 50% of the rated power [56]. Such short time power fluctuations in the range of MW will represent special loads for the drive train and also for the power grid, as these fluctuations seem to some add up in a wind farm instead of being averaged out [31, 56].

In this contribution, we focus on a statistically advanced description of the power fluctuations of a wind turbine. In recent years, it has been shown that the power conversion process of a wind turbine can be modeled by a stochastic Langevin differential equation of the power output  $P$  conditioned to certain fixed wind speeds  $u$  [2, 30, 56]. An analogous approach has also been used to model the mechanical loads on a wind turbine such as generator torque  $T$  [48]. The advantage of this approach is that the model equations (in form of the Langevin equation) can be extracted directly from given data. This model can reproduce the stochastic, turbulent and intermittent nature of wind power [56]. In the

Langevin modeling of conversion dynamics of a wind turbine, the focus was up-to now on the deterministic part of the power time series, while the question remained open how to correctly capture the abrupt large power fluctuations mentioned above.

The Langevin equation describes a diffusion process with continuous trajectory. It consists of the deterministic term and the continuous stochastic term which is modeled by a Wiener process or a Brownian motion. Its two model parameters which are the drift and diffusion coefficients can be estimated directly from the measurement data [23, 73, 81]. These parameters are also known as Kramers-Moyal (KM) coefficients which are considered up to second order in the Langevin equation. For the continuous process, the coefficients higher than third order are negligible.

Looking at the temporally high-resolved wind power data, one can see portions of time series, which look like a diffusive process (see Fig. 4.1 (a)), but there are also periods where sudden big jumps of the delivered power become obvious, see Fig. 4.1 (b).



**Figure 4.1:** Wind power time series, spanning the period of ten minutes. (a) shows the period where the power changes are not very large. (b) shows the period where the power changes are very large, up to about 40% of the rated power. Increments  $\Delta P := P(t + \tau) - P(t)$  emphasize the fluctuations and are shown for sampling period  $\tau = 1$  s in (c) and (d).

In this contribution, we aim to investigate how far these jumps make it necessary to extend the stochastic description. If the higher order KM coefficients ( $> 3$ ) are not negligible, they would be an indicator of non-continuity in the process [5, 81]. One possibility to model this behavior is the extension of the Langevin diffusion process to a jump-diffusion process. An additional discontinuous stochastic term for the jump process is introduced for which we assume that it can be modeled by a Poisson process. In this more general stochastic approach two more parameters arise which are the jump rate and the jump amplitude. We show how they can be estimated from the higher order KM coefficients. This analysis aims to give a more realistic stochastic description of the power output of a wind turbine. Relation to control strategies, or the use for improved modeling of the wind energy resource in a power grid, will be discussed in this paper.

Our aim is to show in detail the procedure to estimate the general stochastic jump-diffusion process with Wiener and Poisson noise to achieve an advanced stochastic characterization and modelling of the wind power conversion dynamics of a wind turbine. Here, we analyze the SCADA data from a wind turbine with resolution of 1 Hz. The article is organized as follows. At first, we describe the analysed data. Next, the stochastic analysis method is summarized, and it is shown how it is possible to quantify and separate the contributions of diffusion and jump fluctuations. Finally, the results of the data analysis for power output conditioned to wind speed are presented. In addition, we investigate the stochastic relation between generator torque and generator rotational speed.



## 4.2 Stochastic data analysis of wind energy system

### 4.2.1 Data description

The measurement data are extracted from a wind turbine of a wind farm. The wind farm is installed onshore over an area covering roughly  $4 \text{ km}^2$  and is surrounded by flat rural terrain with 12 identical variable-speed, pitch-regulated wind turbines. The rated power of each turbine is in the order of 2 MW. The values were made anonymous to keep the confidentiality of the data. Thus, all the data are normalized with their corresponding maximum for our analysis.

The measured quantities are the net electrical power output,  $P$ , generated by the wind turbine, the wind speed,  $u$ , measured on the nacelle by a cup anemometer and the rotational speed or rpm,  $\Omega$ , of the generator. The torque,  $T$ , on the generator is calculated from the power and rpm of the generator using the relation

$$T = \frac{60 \text{ s}}{2\pi} \frac{P}{\Omega} . \quad (4.1)$$

All measurements were performed at a sampling frequency  $f_s = 1 \text{ Hz}$ . The measurement campaign was conducted over a period of eight months, from June 2009 till February 2010. The same data were used also in the study of [57].

### 4.2.2 Power Conversion Process Described by Stochastic Dynamics

Assuming the validity of a diffusive process, the power conversion process of a wind turbine can be modelled as a stochastic Langevin equation of the power output  $P$  conditioned to certain fixed wind speed  $u$  [2, 30, 56],

$$dP(t, u) = D^{(1)}(P|u) dt + \sqrt{D^{(2)}(P|u)} dW_t , \quad (4.2)$$

where  $W_t$  is a Wiener process, a scalar Brownian motion. The general non-linear functions  $D^{(1)}(P|u)$  and  $D^{(2)}(P|u)$  are the drift and the diffusion functions, which in case of the Langevin equation (4.2) are identical to the first and second order Kramers-Moyal (KM) coefficients. In general, the  $j$ -th order KM coefficients,  $K^{(j)}(P|u)$ , can be directly determined from given data  $P$  for each wind speed  $u$ , using their definitions in terms of conditional incremental averaging, cf. [23, 81], as

$$K^{(j)}(P|u) = \lim_{\Delta t \rightarrow 0} \frac{1}{\Delta t} \langle (P(t + \Delta t) - P(t))^j |_{P(t)=P, u(t)=u} \rangle. \quad (4.3)$$

The Langevin equation describes a continuous diffusion process where  $K^{(j)}(P|u) = 0$  for  $j \geq 3$  and  $D^{(j)}(P|u) = K^{(j)}(P|u)$  for  $j = 1, 2$ . Further details on methods of this estimation can be found in [23, 73].<sup>2</sup> All the higher order KM coefficients vanish when the fourth order KM coefficient  $K^{(4)}(P|u)$  is negligible according to the Pawula theorem [75]. When the signal of a stochastic process has sharp changes, or discontinuities, at some instants, typically higher order Kramers-Moyal coefficients and especially  $K^{(4)}(P|u)$  are not negligible anymore. In this case, an extension of the Langevin-type modeling with an additional jump noise is needed, see [5, 11, 42, 79, 81, 82]. Such a jump-diffusion dynamics for a power conversion process is given by

$$dP(t, u) = D^{(1)}(P|u) dt + \sqrt{D^{(2)}(P|u)} dW_t + \xi dJ_t, \quad (4.4)$$

where again  $W_t$  is a Wiener process,  $D^{(1)}(P|u)$  and  $D^{(2)}(P|u)$  are the drift and the diffusion functions. In the following we assume that  $\xi dJ_t$  is a Poisson jump process. The coefficient  $\xi$  is the jump size, which is assumed to be normally distributed,  $\xi \sim N(0, \sigma_\xi^2)$ , with zero mean and

---

<sup>2</sup>KM coefficients of a Langevin process in  $x(t)$  are defined for  $j = 1, 2$  as  $K^{(j)}(x, t) = D^{(j)}(x, t) = \frac{1}{j!} \lim_{\Delta t \rightarrow 0} \frac{1}{\Delta t} \langle (x(t + \Delta t) - x(t))^j |_{x(t)=x} \rangle$  in [23, 73]. In order to stay consistent with the jump-diffusion process, our definition differs by a factor of  $\frac{1}{j!}$ , and  $dW_t = \int_t^{t+dt} \Gamma(\tau) \cdot d\tau$  where  $\langle \Gamma(t) \rangle = 0$  and  $\langle \Gamma(t)\Gamma(t') \rangle = \delta(t - t')$ . The corresponding Fokker-Planck equation will be  $\frac{\partial}{\partial t} p(x, t) = -\frac{\partial}{\partial x} [D^{(1)}(x, t) p(x, t)] + \frac{1}{2} \frac{\partial^2}{\partial x^2} [D^{(2)}(x, t) p(x, t)]$ .

variance  $\sigma_\xi^2$ .  $\xi$  is also known as jump amplitude.  $J_t$  is a Poisson jump process which is a zero-one jump process with jump rate  $\lambda(P|u)$  [33, 81]. The drift and diffusion coefficients and the jump rate are now related to the KM coefficients  $K^{(j)}(P|u)$  in the following way [5]:

$$D_j^{(1)}(P|u) = K^{(1)}(P|u), \quad (4.5)$$

$$D_j^{(2)}(P|u) + \lambda(P|u)\langle\xi^2\rangle = K^{(2)}(P|u), \quad (4.6)$$

$$\lambda(P|u)\langle\xi^j\rangle = K^{(j)}(P|u) \quad \text{for } j > 2. \quad (4.7)$$

Since  $\xi$  has zero mean, its second order moment is  $\langle\xi^2\rangle = \sigma_\xi^2$ . From Eq. (4.5), it can be seen that the estimation of the drift coefficient is the same for the diffusion process, which obeys the Langevin equation, and the jump-diffusion process. Here we go into more details of the noise part and do not assume anymore a vanishing  $K^{(4)} = 0$ .

Jump amplitude  $\sigma_\xi^2$  and jump rate  $\lambda$  can be estimated by using Eq. (4.7) with  $j = 4$  and  $6$  and Wick's theorem [39, 86] for Gaussian random variables, i.e.,  $\langle\xi^{2n}\rangle = \frac{(2n)!}{2^n n!} \langle\xi^2\rangle^n$ ,

$$\sigma_\xi^2(P|u) = \frac{K^{(6)}(P|u)}{5K^{(4)}(P|u)}, \quad (4.8)$$

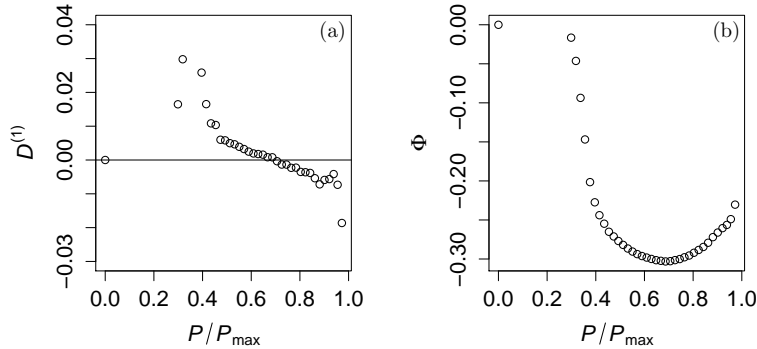
$$\lambda(P|u) = \frac{K^{(4)}(P|u)}{3\sigma_\xi^4(P|u)}. \quad (4.9)$$

### 4.2.3 Results

#### Results for Electrical Power Output

First, we analyze the relation between wind speed and power. For chosen fixed wind speed values with bin sizes of  $0.5 \text{ ms}^{-1}$ , the KM coefficients  $K^{(j)}(P|u)$  are determined with the assumption of stationarity within the corresponding wind speed bin. Firstly, the drift coefficients are determined. The zero-crossings of the drift coefficient,  $D^{(1)}(P|u) = 0$ , correspond to the stable fixed points or equilibria of each wind speed bin if the slope of  $D^{(1)}$  is negative [2, 29]. Zero-crossings with positive slope are

unstable fixed points. Alternatively, this can be expressed by a drift potential, which is defined as  $\Phi = - \int_P D^{(1)}(P|u) dP$ . The zero-crossings with negative slope of the drift correspond correspond to minima of the drift potential. An example of a drift coefficient and corresponding potential for the wind speed of  $u = 0.41 u_{\max}$  is shown in Fig. 4.2.

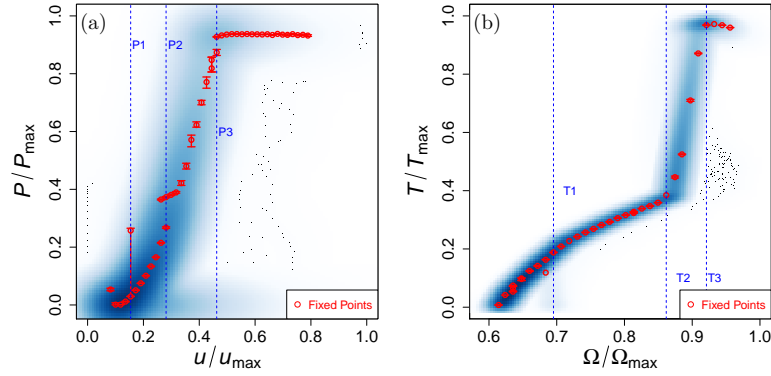


**Figure 4.2:** Drift coefficient  $D^1(P|u)$  (a) and corresponding potential  $\Phi$  (b) for the wind speed of  $u = 0.41 u_{\max}$ . Zero-crossings of the drift coefficient  $D^1(P|u) = 0$  or local minima of drift potential  $\Phi$  are stable fixed points which describe the equilibrium dynamics.

For each wind speed bin, there can be single or multiple fixed points. With these stable fixed points, we can reconstruct the characteristic power curve, which we call Langevin Power Curve (LPC) [55, 84], as shown in Fig. 4.3 (a). These stable fixed points can already be used for a definition of different operational states of the wind turbine. In our case, we mark three distinct states (P1, P2 and P3) which separate the operational regions by blue dotted lines. Multiple fixed points are found to be at these states. Near operation point P2 in Fig. 4.3 (a), we observe the shifting of fixed points in a discontinuous way. Such details cannot be detected by the standard averaging procedure of power curve defined by [38].

The fixed point analysis and characterization of power output of a wind turbine by Langevin equation (4.2) or diffusion process have been studied

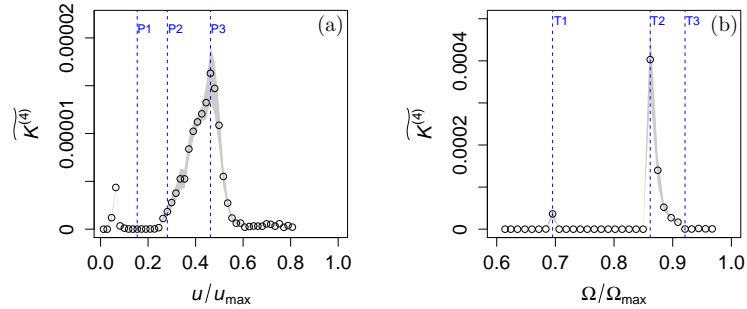
by [2, 30, 57]. In their works, they extensively focused on the drift coefficient. Higher order KM coefficients were not considered. In our work here, we focus on the noisy part and evaluate the higher order KM coefficients.



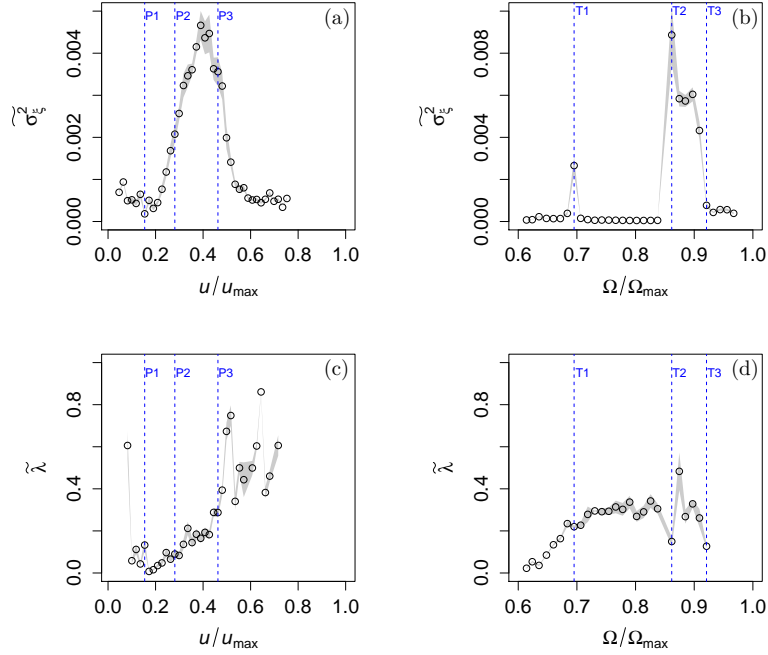
**Figure 4.3:** Characteristic power curve, (a), determined from the zero crossings of the drift coefficients at each wind speed bin and characteristic torque curve, (b), at each rotational speed rpm bin, presented in red open circles. They are also called Langevin Power Curve (LPC) and Langevin Torque Curve (LTC), respectively. The blue background shows the density scatter plot of the measurement data and darker regions indicate more data points are available. The black dots are the outliers of the density scatter plot. Three distinct states (P1, P2 and P3) for LPC and (T1, T2 and T3) for LTC which separate the operational regions are marked by blue dotted lines.

As explained in Sec. 4.2.2, first the fourth-order KM coefficient  $K^{(4)}(P|u)$  is estimated to see if jump noise matters. An example of  $K^{(4)}(P|u)$  and the jump amplitude  $\sigma_{\xi}^2(P|u)$  for the wind speed of  $u = 0.41 u_{\max}$  with their medians (solid black lines) is shown in Fig. 4.7. The fixed point for this wind speed bin is located at  $P = 0.7 P_{\max}$ , see Fig. 4.2. Statistically, we can obtain more accurate results near the fixed point due to the better coverage of data, whereas for regions with less data (farther away from the fixed point) the results become more noisy and outliers are seen. A robust method to estimate is to use the medians instead of the means [37]. Some examples are shown in Appendix 4.4.1.

In the following, we investigate details of the median values. Thus, we simplify the process to those with constant parameters  $\sigma_\xi^2$  and  $\lambda$  for the jump process. The  $P$ -dependence of these parameters can also be studied, which we do not do here to keep the discussion simpler, see Fig. 4.7. Fig. 4.4 (a) shows that there is an increase of  $\widetilde{K}^{(4)}(P|u)$  near the state P3, the transition point to the rated power. This behavior of  $\widetilde{K}^{(4)}$  shows that not only diffusive noise is present, thus we proceed to analyze also the higher order KM-coefficients from which we can determine the jump amplitude  $\widetilde{\sigma}_\xi^2$  and jump rate  $\widetilde{\lambda}$  for each wind speed as shown in Fig. 4.5 (a) and (c). The jump amplitude  $\sigma_\xi^2$  is highest between the states P2 and P3 which is just below the transition to the rated power region where the switching of the control strategy play a major role. The jump rate  $\widetilde{\lambda}$  is highest in the region of rated power.



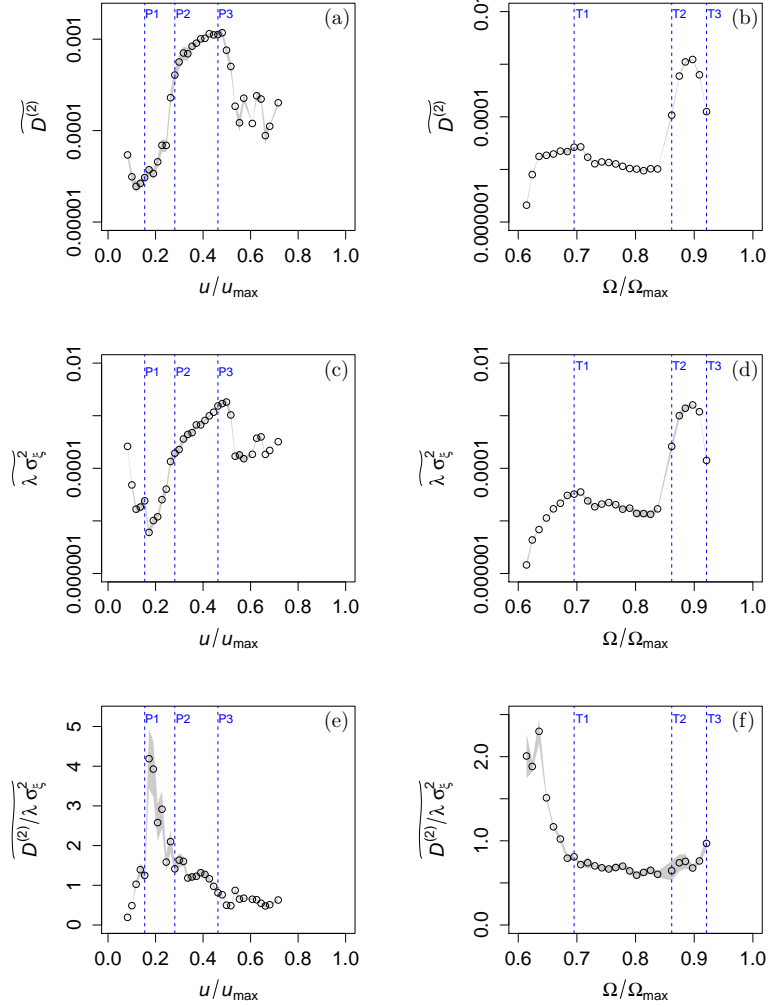
**Figure 4.4:** The median of the fourth-order KM coefficient,  $\widetilde{K}^{(4)}(u)$  over the power bins for each wind speeds bin (a) and over the torque bins for each rpm bin (b). There is an increase of  $\widetilde{K}^{(4)}(u)$  near the transition to the rated region. Statistical uncertainties are shown as gray-shaded background. Blue dotted lines are the distinct states observed from the fixed point analysis, see Fig. 4.3.



**Figure 4.5:** The median of jump amplitude,  $\widetilde{\sigma}_\xi^2$ , over the power bins for each wind speed bin (a) and over the torque bins for each rpm bin (b). There is a significant increase in jump amplitude near the transition to the rated region. The median of jump rate,  $\widetilde{\lambda}$ , over the power bins for each wind speed bin (c) and over the torque bins for each rpm bin (d). There is a significant increase in jump rate after the transition to the rated region and a small increase near the cut-in region for power analysis. The jump rate for torque analysis is similar in between all three distinct states shown by blue dotted lines as in Fig. 4.3. Statistical uncertainties are shown as gray-shaded background.

In order to quantify the overall jump contribution, we determine the product  $\widetilde{\lambda\sigma}_\xi^2$  as shown in Fig. 4.6 (c). Again we see that the jump contribution  $\widetilde{\lambda\sigma}_\xi^2$  is highest around the state P3. Moreover, we also determine the median of the diffusion to jump ratio  $\frac{\widetilde{D}^{(2)}}{\lambda\sigma_\xi^2}$  over the power bins for each wind speed bin obtained from the resolved  $\frac{D^{(2)}}{\lambda\sigma_\xi^2}$  values as shown in Fig. 4.6 (e) to find out whether diffusive or jump noise is dominating. If the ratio is large, there is more diffusive noise and vice versa. The jump noise is dominant in the rated power region after the state P3. The

diffusive noise is dominant between the states P1 and P2, which also coincides with the lowest jump contribution  $\lambda\sigma_\xi^2$ , Fig. 4.6 (c).



**Figure 4.6:** The median of diffusion coefficient,  $\widetilde{D}^{(2)}$ , (a), overall jump contribution,  $\widetilde{\lambda\sigma_\xi^2}$ , (c), and the diffusion to jump ratio,  $\widetilde{D^{(2)}/\lambda\sigma_\xi^2}$ , (e), over the power bins for each wind speed bin, and respective quantities over the torque bins for each rpm bin in (b), (d), and (f). Sub-figures (a-d) are plotted in semi-logarithmic scale for better visualization. Statistical uncertainties are shown as gray-shaded background. Blue dotted lines are the distinct states observed from fixed point analysis, see Fig. 4.3.



The analysis in this subsection shows two important points. First, a jump process is present and should be included in an advanced stochastic description or, respectively, model. Second, below rated power, first a diffusive stochastic behavior is dominating, while jumpy noise becomes important for the considered wind turbine close to the transition to rated power.

### Results for Generator Torque

Apart from the dynamical dependence of power on the wind speed, another central characteristic of the wind turbine is the generator torque  $T$  vs. rotational speed or rpm  $\Omega$  [48]. The generator torque was calculated from power and rpm data according to Eq. (4.1), whereas the rpm was measured independently. A similar stochastic approach like for the power and wind speed is applied to the torque and rpm dynamics  $T(t, \Omega)$ , next. Based on the drift coefficient, a characteristic curve like the Langevin Power Curve is calculated as shown in Fig. 4.3 (b). We refer to it as Langevin Torque Curve (LTC). We again mark three distinct states (T1, T2 and T3) with blue dotted lines which can be deduced from the fixed point analysis.

Next, we also determine the jump contribution  $\lambda\sigma_\xi^2$  and the diffusion to jump ratio  $\frac{D^{(2)}}{\lambda\sigma_\xi^2}$  for each rpm bin, see Fig. 4.6 (d) and (f). In Fig. 4.6 (d), we see that the median of the jump contribution  $\widetilde{\lambda\sigma_\xi^2}$  becomes significant in the region between T2 and T3, cf. Fig. 4.3 (b), at relatively high rpm values. Outside this regime, jump noise does not dominantly contribute to the dynamics. On the other hand, the median of the diffusion to jump ratio  $\frac{\widetilde{D^{(2)}}}{\lambda\sigma_\xi^2}$  takes its largest values at low values of rotational speed rpm. At the rotational speeds rpm below the state T1, diffusive behavior is much more dominant.

The conclusions from the previous subsection 4.2.3 apply also to the torque analysis in an analogous way. Similar to the power characteristic, a jump process is also present for the torque characteristic, which could be expected as we compute the torque from power and rotational speed.

Diffusive stochastic behavior is dominating for low rpm, but in the region from T1 on ( $\Omega/\Omega_{\max} \gtrsim 0.7$ ) diffusive and jumpy behavior seem to be more balanced for the torque case.

### 4.3 Conclusion and Outlook

In our work, we investigate the contribution of the higher-order KM coefficients to the stochastic conversion dynamics of a wind turbine. As described in Sec. 4.2.2, these higher-order coefficients allow to quantify the contributions of diffusive behavior and jump noise, and indicate that discontinuities in the trajectory of the measurement data are due to the stochastic jump noise. The main results are that we can quantify with our proposed method how the amplitudes and the ratio of the two noise contributions change in different operating ranges of a wind turbine. The region below rated power seems to provide the highest values of the amplitudes ( $D^{(2)}$  and  $\sigma_{\xi}^2$ ). Sometimes the maximal values are found for the transition states defined by the fixed point characteristics. The ratio between the contributions of the diffusive and jumpy noise shows that at low wind speed and low power a diffusive noise is dominating whereas for higher power more jump noise is present, with some detailed differences for power and torque. All this indicates that it is the interplay between the stochastic driving wind speed and the reacting control system that determines the noise contribution. In particular, the jump contribution is closely linked to the control system as one can see, to our interpretation, in the rapid changes of  $\sigma_{\xi}^2$  in Fig. 4.5 (b). Interestingly, this is more prominent in the torque signal than in the power signal. It is well known that the control system is not operating directly with the wind signal but with torque  $T$  and the rotational speed  $\Omega$ .

Near the states T1, T2 and T3, there are three distinct operational rotational speeds  $\Omega$  which the control system prefers to approach. It is a common control strategy to avoid certain resonance frequencies of the structure in order to mitigate excessive loads. We show this by evaluating the drift potential  $\Phi(\Omega)$  of the rotational speed. The minima of this potential correspond to the preferred rotational speeds, see Appendix 4.4.2.

Moreover, looking at Fig. 4.3 (b), between the states T2 and T3, there is a steep gradient which enforces a large change in generator torque  $T$  at only a small regime in rotational speed  $\Omega$ . In this range, we also observe that there is a huge increase in both diffusive and jump noise by more than two orders of magnitude, see Fig. 4.6 (b) and (d).

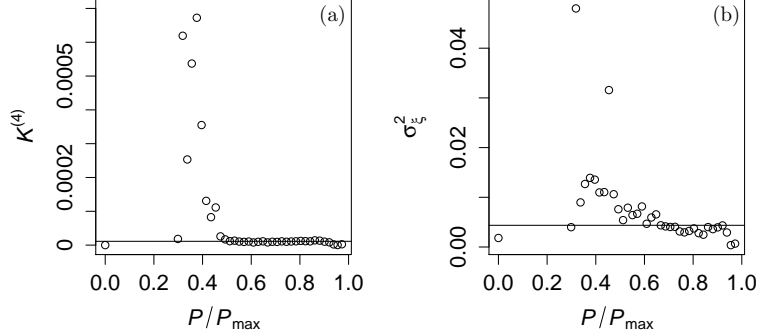
So far we used the stochastic methods to characterize the dynamics of the wind energy conversion process. It goes without saying that the characterization can be used to compare quantitatively different turbines. Potential failures in the control system should be detectable by comparison of the different stochastic terms. One may see how with time some noise contribution changes as the system gets old, or one may show how different wind turbines or different control strategies perform differently in a dynamic sense. Together with detailed knowledge of a specific turbine this should also be useful for monitoring, e.g., performance or structural health.

At last we would like to point out that besides this characterization the stochastic methods presented here also deliver the explicit form of the stochastic differential equations. Thus, it is also possible to use our results as very efficient dynamics models for power and torque. Long time simulations can be done easily. Such models are of interest for the simulation of the contribution of wind energy to the power grid and for the simulation of loads.

## 4.4 Appendix

### 4.4.1 Median as a Robust Estimator

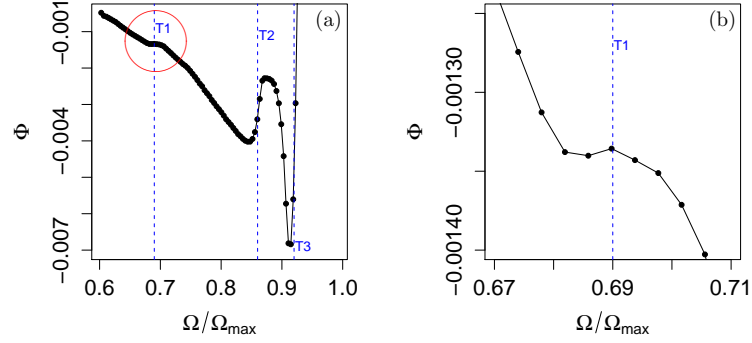
Statistically, we can obtain more accurate results near the fixed point due to the better coverage of data, whereas for regions with less data (farther away from the fixed point) the results become more noisy and outliers are seen. A robust method to estimate the typical value of  $K^{(4)}(P|u)$  and  $\sigma_\xi^2(P|u)$ , as examples, is to use the medians  $\widetilde{K}^{(4)}(u)$  and  $\widetilde{\sigma}_\xi^2(u)$  as shown by solid lines in Fig. 4.7.



**Figure 4.7:** Fourth-order KM coefficient  $K^{(4)}(P|u)$  (a) and the jump amplitude  $\sigma_{\xi}^2(P|u)$  (b) for the wind speed of  $u = 0.41 u_{\max}$ . (For the corresponding drift term see Fig. 4.2). The solid black lines are their respective medians  $\widetilde{K}^{(4)}(u)$  and  $\widetilde{\sigma}_{\xi}^2(u)$ . The fixed point for this wind speed bin is  $P = 0.7 P_{\max}$ . Statistically, more accurate results can be obtained near the fixed point due to the better coverage of data. By using the median, our results are more robust to outliers far away from the fixed points.

#### 4.4.2 Drift Potential of Rotational Speed

We analyzed all the data of rotational speed  $\Omega$  in the range of  $0.6 \Omega_{\max}$  and  $\Omega_{\max}$  without any conditioning or binning on other variables. We evaluated the drift coefficient  $D^{(1)}(\Omega)$  and then determined the drift potential which is  $\Phi(\Omega) = - \int_{\Omega} D^{(1)}(\Omega) d\Omega$  which is plotted in Fig. 4.8. Minima of the drift potential corresponds to the stable fixed points or equilibria. Here we can observe three minima around the three states T1, T2 and T3 which the control system prefers to approach. It is a common control strategy to avoid certain resonance frequencies of the structure in order to mitigate excessive loads as shown in Fig. 4.3.



**Figure 4.8:** Drift potential  $\Phi(\Omega)$  determined in the range of  $0.6 \Omega_{\max}$  and  $\Omega_{\max}$ , (a). Minima of the drift potential corresponds to the stable fixed points or equilibria. Here we can observe three minima around the three states T1, T2 and T3 as shown in Fig. 4.3. The minimum around the state T1 which is presented with the red circle is elaborated in (b).

In Fig. 4.8 (b), we can clearly observe a minimum around the state T1. From our results in Fig. 4.4, 4.5 and 4.6, there is also a slight increase in noises around this state. This indicates that the control system of the wind turbine starts switching the strategies at this state around T1. As a remark, we calculated the deterministic potential only which reflects the mechanical and control mechanism of the wind turbine.



## Chapter 5

# Characterization of Continuous and Discontinuous Stochastic Processes

Recent empirical and theoretical studies have shown that jumps (discontinuous / non-Brownian events) exist in many complex systems and may have important impact on the risk management of these systems. These analyses have received considerable attention in the last few years [33, 81]. In general, jumps are discontinuous alterations in time series in which the jump activity will cause higher uncertainties in the stochastic features of the underlying system.

If there exists jump discontinuities in the trajectory of the stochastic process, a straightforward approximation is to add an additional noise term, which is called jump noise, to the Langevin equation. The corresponding dynamical equations are known as jump-diffusion processes [5, 33, 81]. In this picture, jump noise acts as the discontinuous paths inside the diffusion process. In the following, the jump events is considered as Poisson processes, which follows the zero-one jump law. This means that there is either zero- or one-jump event within a short-time step  $\tau$  which was discussed in Chapter 2.

The measured data sampled in discrete time appears as a sequence of discontinuous hops (jumps), even if the underlying trajectory is continuous. In [45], the so-called  $Q$ -criterion has been derived from the expansion of the Kramers–Moyal conditional moments, which allows us to differentiate whether a stochastic time series obtained from some dynamical equations has been generated by a diffusive (continuous) or by a jumpy (discontinuous) process.

This  $Q$ -criterion is an approximation, which is valid for the time-scales, which are less than the correlation time scale  $T_C$  for the diffusion process and the average time scale between the jumps  $T_J$  for jump-diffusion process. Based on a given dynamical equation, it is possible to formulate a set of similar stochastic processes with different time scales  $T_C$  and  $T_J$  in order to validate the  $Q$ -criterion for the corresponding time series. For real world data, however, we have in general no previous knowledge about the time scales, specially of  $T_J$ .

Here, it is the challenge to derive stochastic measures to distinguish between continuous and discontinuous trajectories. In this chapter, the behavior of  $Q$ -criterion at different time scales is discussed. More robust additional criteria, which are again based on the expansion of the Kramers–Moyal conditional moments, are introduced. They also allow us by downsampling of the given data to see whether the underlying process has a diffusive or jumpy behavior for a given time series. The substantial potential of these methods are shown by their application to synthetic time series from different dynamical systems as well as to various real world experimental data.



## 5.1 Expansion of Conditional Moments at Finite Time Step

The conditional probability distribution of the process  $x = x(t)$  which satisfies the Kramers-Moyal differential equation can be written as [45, 81],

$$\frac{\partial p(x, t|x', t')}{\partial t} = \mathcal{L}_{\text{KM}} p(x, t|x', t') \quad (5.1)$$

with initial condition  $p(x, t|x', t) = \delta(x - x')$  and the Kramers-Moyal (KM) operator  $\mathcal{L}_{\text{KM}}$  is given by,

$$\mathcal{L}_{\text{KM}} = \sum_{n=1}^{\infty} \frac{1}{n!} \left( -\frac{\partial}{\partial x} \right)^n K^{(n)}(x, t) \quad . \quad (5.2)$$

The formal solution of (5.1) reads

$$p(x, t + \tau|x', t) = \exp\{\tau \mathcal{L}_{\text{KM}}\} \delta(x - x') \quad . \quad (5.3)$$

The  $n^{\text{th}}$ -order conditional moments  $M^{(n)}(x, t, \tau)$  with finite  $\tau$  can be written as,

$$\begin{aligned} M^{(n)}(x_i, t, \tau) &= \int_{-\infty}^{\infty} (x - x_i)^n \exp\{\tau \mathcal{L}_{\text{KM}}\} \delta(x - x_i) dx \\ &= \exp\{\tau \mathcal{L}_{\text{KM}}^\dagger\} (x - x_i)^n |_{x=x_i} \end{aligned} \quad (5.4)$$

where  $\mathcal{L}_{\text{KM}}^\dagger$  is the adjoint operator of  $\mathcal{L}_{\text{KM}}$  and is given by,

$$\mathcal{L}_{\text{KM}}^\dagger = \sum_{n=1}^{\infty} \frac{1}{n!} K^{(n)}(x, t) \left( \frac{\partial}{\partial x} \right)^n \quad . \quad (5.5)$$

where  $K^{(n)}(x, t) = \lim_{\tau \rightarrow 0} \frac{1}{\tau} M^{(n)}(x, t, \tau)$  are the  $n^{\text{th}}$ -order Kramers-Moyal coefficients and  $M^{(n)}(x, t, \tau)$  are given by Eq. (5.4). Next, the explicit cases of the diffusion and jump-diffusion processes are discussed.

### 5.1.1 Conditional Moments of Diffusion and Jump-Diffusion Processes

- Diffusion process, in general, can be described by Langevin equation such that

$$dx = D^{(1)}(x, t) dt + \sqrt{D^{(2)}(x, t)} dW_t, \quad (5.6)$$

where  $D^{(1)}(x, t)$  and  $D^{(2)}(x, t)$  are the drift and the diffusion coefficients, respectively, and  $W_t$  is the Wiener process. The Wiener increment is  $dW_t = \eta_t \cdot dt$  where  $\eta_t$  is the Gaussian white noise, i.e.  $\langle \eta_t \rangle = 0$  and  $\langle \eta_t \eta_{t'} \rangle = \delta(t - t')$ .

We can evaluate its conditional moments  $M_d^{(n)}(x, \tau)$  using Eq. (5.4). For diffusion process or Langevin equation, the adjoint KM operator  $\mathcal{L}_{\text{KM}}^\dagger$  becomes the adjoint Fokker-Planck (FP) operator  $\mathcal{L}_{\text{FP}}^\dagger$ . For better readability, let the drift be  $D^{(1)}(x, t) = K^{(1)}(x, t) = a$  and the diffusion  $D^{(2)}(x, t) = K^{(2)}(x, t) = b^2$  and the adjoint FP-operator becomes

$$\mathcal{L}_{\text{FP}}^\dagger = a \frac{\partial}{\partial x} + \frac{1}{2} b^2 \frac{\partial^2}{\partial x^2}. \quad (5.7)$$

In this formulation,  $a$  and  $b$  can still be the function of  $x$  and  $t$ . The second-, fourth- and sixth-order conditional moments up to first non-vanishing term can be derived from Eqs. (5.4) and (5.7) as follow [45]:

$$\begin{aligned} M_d^{(2)}(x, \tau) &= b^2 \tau + \mathcal{O}(\tau^2), \\ M_d^{(4)}(x, \tau) &= 3b^4 \tau^2 + \mathcal{O}(\tau^3), \\ M_d^{(6)}(x, \tau) &= 15b^6 \tau^3 + \mathcal{O}(\tau^4). \end{aligned} \quad (5.8)$$

In order to distinguish from the jump-diffusion process which will be discussed next, the subscript “d” is used here for the diffusion process.

- Ornstein-Uhlenbeck (OU) process is considered here to be more concrete such that

$$dx = -\gamma x dt + \sqrt{D} dW_t . \quad (5.9)$$

where  $\gamma$  and  $D$  are positive real constants. Expansion of  $M_d^{(n)}(x, t, \tau)$  of OU process for small  $\tau$  up to  $\mathcal{O}(\tau^3)$  (up to  $\mathcal{O}(\tau^4)$  for sixth-order conditional moment) reads

$$\begin{aligned} M_d^{(2)}(x, \tau) &= D\tau + (\gamma^2 x^2 - \gamma D) \tau^2 + \mathcal{O}(\tau^3), \\ M_d^{(4)}(x, \tau) &= 3D^2\tau^2 + \mathcal{O}(\tau^3), \\ M_d^{(6)}(x, \tau) &= 15D^3\tau^3 + \mathcal{O}(\tau^4). \end{aligned} \quad (5.10)$$

- Jump-diffusion process, in general, can be written as

$$dx = D^{(1)}(x, t) dt + \sqrt{D^{(2)}(x, t)} dW_t + \xi dJ_t, \quad (5.11)$$

where, again,  $D^{(1)}(x, t)$  and  $D^{(2)}(x, t)$  are the drift and the diffusion coefficients, respectively, and  $W_t$  is the Wiener process. The quantity  $\xi$  is the size of the jump noise which is assumed to be a normally distributed random variable with zero mean and variance  $\sigma_\xi^2(x)$ , i.e.,  $\xi \sim N(0, \sigma_\xi^2)$ . The variance  $\sigma_\xi^2(x)$  is also called jump amplitude. The term  $J_t$  is the Poisson jump process, which is the zero-one jump process with a jump rate (or intensity)  $\lambda(x)$ . The KM-coefficients of such jump-diffusion are

$$\begin{aligned} K^{(1)}(x, t) &= D^{(1)}(x, t) \\ K^{(2)}(x, t) &= [D^{(2)}(x, t) + \langle \xi^2 \rangle \lambda(x)] \\ K^{(2n)}(x, t) &= \langle \xi^{2n} \rangle \lambda(x), \quad \text{for } 2n > 2 \end{aligned} \quad (5.12)$$

Therefore, the adjoint KM-operator  $\mathcal{L}_{\text{KM}}^\dagger$  for jump-diffusion process becomes

$$\begin{aligned}
 \mathcal{L}_{\text{KM}}^\dagger &= \underbrace{D^{(1)}(x, t)}_{\mathbf{A}} \frac{\partial}{\partial x} + \underbrace{\frac{[D^{(2)}(x, t) + \langle \xi^2 \rangle \lambda(x)]}{2!}}_{\mathbf{B}} \frac{\partial^2}{\partial x^2} \\
 &+ \underbrace{\frac{\langle \xi^4 \rangle \lambda(x)}{4!}}_{\mathbf{C}} \frac{\partial^4}{\partial x^4} + \underbrace{\frac{\langle \xi^6 \rangle \lambda(x)}{6!}}_{\mathbf{D}} \frac{\partial^6}{\partial x^6} \\
 &+ \underbrace{\frac{\langle \xi^8 \rangle \lambda(x)}{8!}}_{\mathbf{E}} \frac{\partial^8}{\partial x^8} + \dots \quad . \quad (5.13)
 \end{aligned}$$

Again for better readability, we use the substitution of the coefficients with  $\mathbf{A}, \mathbf{B}, \mathbf{C}, \mathbf{D}, \mathbf{E}, \dots$ , which can still be the function of  $x$  and  $t$  in this formulation. With it, we can derive the conditional moments of the jump-diffusion equation Eq. (5.11) for second-, fourth and sixth-orders of the time interval  $\tau$  using Eq. (5.4) and (5.13) such that

$$\begin{aligned}
 M_j^{(2)}(x, \tau) &= 2\mathbf{B}\tau + \mathcal{O}(\tau^2), \\
 M_j^{(4)}(x, \tau) &= 4!\mathbf{C}\tau + \mathcal{O}(\tau^2), \\
 M_j^{(6)}(x, \tau) &= 6!\mathbf{D}\tau + \mathcal{O}(\tau^2), \quad (5.14)
 \end{aligned}$$

where the subscript “j” denotes the jump-diffusion process [45]. Here, they are derived up to first non-vanishing term.

- OU process with an additional jump term with constant jump rate and jump amplitude is now considered. It is a linear jump-diffusion process and one finds,

$$dx = -\gamma x dt + \sqrt{D} dW_t + \xi dJ_t . \quad (5.15)$$

Adding the jump term in OU process creates the discontinuities in the trajectory.

For finite time interval  $\tau$ , the conditional moments  $M_j^{(n)}(x, \tau)$  of an OU jump-diffusion process can be determined up to  $\mathcal{O}(\tau^4)$  using Eqs. (5.4) and (5.13) as follows:

$$\begin{aligned}
 M_j^{(2)}(x, \tau) &= (D + \lambda\sigma_\xi^2)\tau + (\gamma^2 x^2 - \gamma(D + \lambda\sigma_\xi^2))\tau^2 \\
 &\quad + \left( \frac{1}{3}\gamma^2(D + \lambda\sigma_\xi^2) - \frac{1}{2}\gamma^3 x^2 \right)\tau^3 + \mathcal{O}(\tau^4), \\
 M_j^{(4)}(x, \tau) &= 3\lambda\sigma_\xi^4\tau + 3((D + \lambda\sigma_\xi^2)^2 - 2\gamma\lambda\sigma_\xi^4)\tau^2 \\
 &\quad + (6x^2\gamma^2(D + \lambda\sigma_\xi^2) + 8\gamma^2\lambda\sigma_\xi^4 \\
 &\quad - 6\gamma(D + \lambda\sigma_\xi^2)^2)\tau^3 + \mathcal{O}(\tau^4), \\
 M_j^{(6)}(x, \tau) &= 15\lambda\sigma_\xi^6\tau + 45(\lambda\sigma_\xi^4(D + \lambda\sigma_\xi^2) - \gamma\lambda\sigma_\xi^6)\tau^2 \\
 &\quad + (45\gamma^2\lambda\sigma_\xi^4 x^2 + 15(D + \lambda\sigma_\xi^2)^3 \\
 &\quad + 540\gamma^2\lambda\sigma_\xi^6 - 810\gamma\lambda\sigma_\xi^4(D + \lambda\sigma_\xi^2))\tau^3 + \mathcal{O}(\tau^4),
 \end{aligned} \tag{5.16}$$

Next, the explicit expansions of the conditional moments of pure noises are discussed.

### 5.1.2 Conditional Moments of Pure Noises

- For a Gaussian distributed white noise with zero mean and variance  $\sigma_\eta^2$ ,  $x = \eta \sim N(0, \sigma_\eta^2)$ , the  $n^{\text{th}}$  order conditional moments can be calculated by Eq. (2.12) such that

$$\begin{aligned}
 M_g^{(n)}(x, \tau) &= \langle (x(t+\tau) - x(t))^n |_{x(t)=x} \rangle, \\
 &= \int_{-\infty}^{\infty} (x' - x)^n \cdot p(x', t' | x, t) \cdot dx', \\
 &= \int_{-\infty}^{\infty} (x' - x)^n \cdot p(x', t') \cdot dx'. \tag{5.17}
 \end{aligned}$$

For  $x = 0$  and  $p(x', t') = p(x') = \frac{1}{\sqrt{2\pi\sigma_\eta^2}} \exp\left(-\frac{x'^2}{2\sigma_\eta^2}\right)$ , we can use Wick's theorem [39, 86] to solve integral in Eq. (5.17). The  $(2n+1)^{\text{th}}$  and the  $2n^{\text{th}}$  conditional moments of Gaussian white noise become

$$\begin{aligned}
 M_g^{(2n+1)}(x=0, \tau) &= 0, \\
 M_g^{(2n)}(x=0, \tau) &= \frac{(2n)!}{2^n n!} \sigma_\eta^{2n}, \tag{5.18}
 \end{aligned}$$

where  $\sigma_\eta^2$  is its variance.

- For a Poisson distributed jump noise  $x = \Delta J \sim \mathcal{P}(\lambda\tau)$ , the  $n^{\text{th}}$  order conditional moments can be calculated by using the discrete definition of conditional moment such that

$$\begin{aligned}
 M_p^{(n)}(x, \tau) &= \langle (\Delta x(t))^n |_{x(t)=x} \rangle, \\
 &= \sum_{k=1}^{\infty} k^n \cdot \text{Prob}[\Delta x(t) = k], \\
 &= \sum_{k=1}^{\infty} k^n \frac{(\lambda\tau)^k}{k!} \exp(-\lambda\tau). \tag{5.19}
 \end{aligned}$$

For Poisson jump noise, following zero-one jump law, the probability

of having a jump event becomes  $\text{Prob}[\Delta x(t) = 1]$ . Applying Taylor expansion of  $\exp(\lambda\tau)$ , Eq. (5.19) becomes

$$M_p^{(n)}(x, \tau) = \lambda\tau + \mathcal{O}(\tau^2) . \quad (5.20)$$

With these expansions, the methods to distinguish diffusive and jumpy stochastic behaviors can be determined which will be discussed next.

## 5.2 Methods of Distinguishing Diffusive and Jumpy Stochastic Behaviors

To show how different stochastic behaviors can be distinguished the following four cases are defined such that

- (i)  $x$ : OU process according to Eq. (5.9) which is diffusive,
- (ii)  $x$ : OU process with jumps according to Eq. (5.15) which is jumpy,
- (iii)  $x = \eta \sim N(0, \sigma_\eta^2)$ : pure Gaussian white noise and
- (iv)  $x = \Delta J \sim \mathcal{P}(\lambda\tau)$ : pure Poisson noise for given time step  $\tau$ ,

as the examples. Note the last two cases are pure independent uncorrelated noise data.

### 5.2.1 $Q$ -criterion to Distinguish Diffusive and Jumpy Behaviors

As a first criterion, the  $Q$ -criterion is introduced in [45] to distinguish whether given synthetic data simulated with different integration time step  $\tau_s$  are diffusive or jumpy. Here, this criterion is also applied to distinguish between uncorrelated Gaussian white noise and Poisson noise.

For a small  $\tau$ , we obtain the function  $Q(x, \tau)$  for the four defined cases such that

$$Q(x, \tau) = \frac{M^{(6)}(x, \tau)}{5M^{(4)}(x, \tau)} \approx \begin{cases} D^{(2)}(x) \tau, & \text{diffusive,} \\ \sigma_{\xi}^2(x), & \text{jumpy,} \\ \sigma_{\eta}^2, & \text{Gaussian white noise, } x = 0, \\ \frac{1}{5}, & \text{pure Poisson jumps,} \end{cases} \quad (5.21)$$

considering up to the first non-vanishing order term [45] which is derived from the fourth- and sixth-order conditional moments. Note that  $x = 0$  is required only for Gaussian white noise, otherwise the relations hold for all  $x$ .

For diffusive process, case (i), the function  $Q$  has a linear relationship with  $\tau$  while it is constant in first order approximation in  $\tau$  for jump-diffusion process, case (ii). Additionally, the function  $Q$  is also constant for the Gaussian white noise, case (iii), and Poisson jump noise, case (iv), which can be derived by direct computations of conditional moments.

To validate the  $Q$ -criterion for synthetic data, we numerically integrate the diffusion process, case (i), with  $\gamma = 1 \text{ s}^{-1}$ ,  $D = 1 \text{ s}^{-1}$  and the jump-diffusion process, case (ii), with the additional jump parameters  $\lambda = 100 \text{ s}^{-1}$  and  $\sigma_{\xi}^2 = 1$  for different integration time step  $\tau_s$ . For the cases (iii) and (iv) of pure noisy data, we use standard normal Gaussian white noise (zero mean and variance  $\sigma^2 = 1$ ) and the pure Poisson jump noise with  $\lambda = 100 \text{ s}^{-1}$ .

From the simulated data, we determine  $Q = Q(x = 0, \tau = \tau_s)$  for different  $\tau_s$  as shown in Fig. 5.1. Theoretical predictions as described in Eq. (5.21) are also plotted with the black solid lines. For jump-diffusion process Eq. (5.21) is valid for  $\tau_s < \frac{1}{\lambda} = T_J$ , which is the average time scale between the jumps. If the integration time step  $\tau_s > T_J$ , the jump-diffusion process behaves like diffusion process with the apparent diffusion coefficient  $\widetilde{D}^{(2)}$  which has the relation,  $\widetilde{D}^{(2)}\tau_s = D^{(2)}\tau_s + \sigma_{\xi}^2$ . It is shown with the blue dashed curve in Fig. 5.1 (ii).



To prove this relation, the jump-diffusion equation can be written using Euler's method as

$$x(t + \tau_s) - x(t) = D^{(1)}(x) \tau_s + \sqrt{D^{(2)}(x)} \Delta W_t + \xi \Delta J_t. \quad (5.22)$$

For  $\tau_s > \frac{1}{\lambda}$ ,  $\Delta J_t = 1$  due to zero-one jump law discussed in Chapter 2, the Eq. (5.22) becomes

$$x(t + \tau_s) - x(t) = D^{(1)}(x) \tau_s + \sqrt{D^{(2)}(x) \tau_s} \eta + \xi. \quad (5.23)$$

$\sqrt{D^{(2)}(x) \tau_s} \eta + \xi$  is the sum of two Gaussian random variables which is equivalent to  $\sqrt{\widetilde{D}^{(2)}(x) \tau_s} \widetilde{\eta} = \sqrt{D^{(2)}(x) \tau_s + \sigma_\xi^2} \widetilde{\eta}$ .  $\widetilde{\eta} \sim N(0, 1)$  is the standard normal Gaussian white noise.

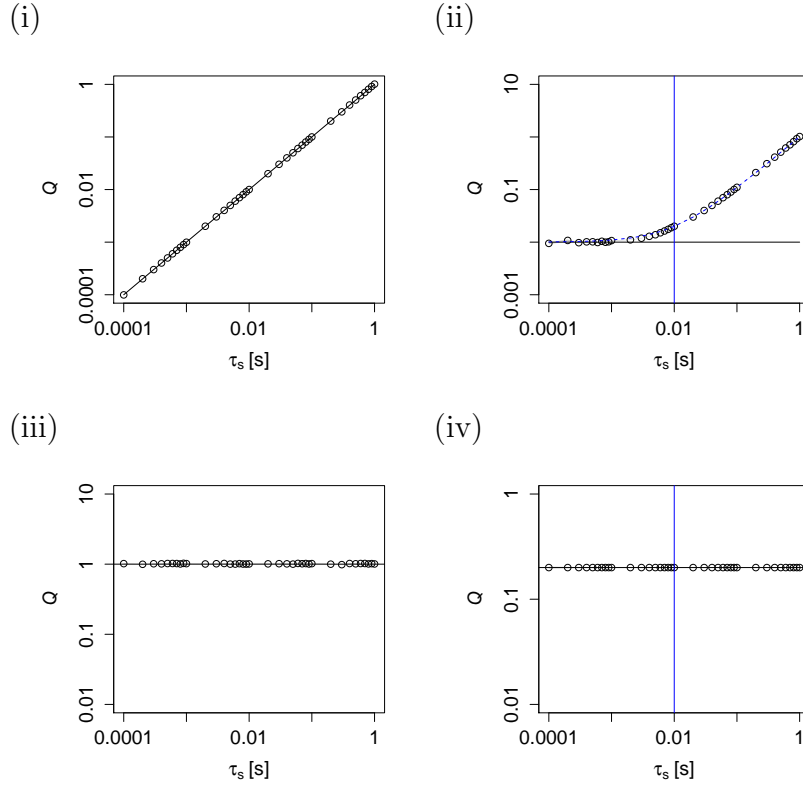
Therefore, the jump-diffusion equation for  $\lambda > \frac{1}{\tau_s}$  becomes

$$x(t + \tau_s) - x(t) = D^{(1)}(x) \tau_s + \sqrt{\widetilde{D}^{(2)}(x) \tau_s} \widetilde{\eta}, \quad (5.24)$$

respectively,

$$dx = D^{(1)}(x) dt + \sqrt{\widetilde{D}^{(2)}(x)} d\widetilde{W}_t, \quad (5.25)$$

where  $\widetilde{D}^{(2)}$  is the apparent diffusion coefficient and  $\widetilde{W}_t$  is the Wiener process. From Fig. 5.1, all those estimates are rather close to the preset values and emphasize the accuracy of the predictions of Q-criterion for four types of the synthetic data.



**Figure 5.1:** The  $Q$ -criterion for OU process (i), OU jump-diffusion process (ii), standard normal Gaussian white noise (iii) and pure Poisson jump noise (iv) are evaluated and plotted in double logarithmic scale. The OU processes are generated for different integration time steps  $\tau_s$  with  $\gamma = 1 \text{ s}^{-1}$ ,  $D = 1 \text{ s}^{-1}$  and also with the additional jump terms with  $\lambda = 100 \text{ s}^{-1}$  and  $\sigma_\xi^2 = 0.01$ . The vertical blue lines show the jump time scale  $T_J = \frac{1}{\lambda}$ . For each  $\tau = \tau_s$ ,  $Q = Q(x = 0, \tau)$  is calculated. The theoretical predictions according to Eq. (5.21) are presented with the black solid lines. The blue dashed curve in case (ii) shows  $Q = D\tau_s + \sigma_\xi^2$ . Standard errors are smaller than the symbol size.

### 5.2.2 $\Theta$ - and $\Lambda$ -criterion to Distinguish Diffusive and Jumpy Behaviors

If the jump-diffusion process has a large jump rate  $\lambda$ , the time scale  $T_J$  will be small so that constant  $Q$  behavior cannot be clearly observed. Instead, it shows the linear relation for  $\tau > T_J$  which makes it difficult

to distinguish from diffusive behavior since  $Q$  also has linear relation for diffusion process. Therefore, new robust criteria complement to  $Q$ -criterion are needed. Similar to the derivation of  $Q$ -criterion, we define the functions  $\Theta(x, \tau)$  and  $\Lambda(x, \tau)$  in Eqs. (5.26) and (5.27) such that

$$\Theta(x, \tau) = 3 \frac{(M^{(2)}(x, \tau))^2}{M^{(4)}(x, \tau)} \approx \begin{cases} 1, & \text{diffusive} \\ \frac{(D^{(2)}(x) + \lambda(x)\sigma_\xi^2(x))^2}{\lambda(x)\sigma_\xi^4(x)}\tau, & \text{jumpy} \\ 1, & \text{white noise at } x = 0 \\ 3\lambda\tau, & \text{pure Poisson jumps,} \end{cases} \quad (5.26)$$

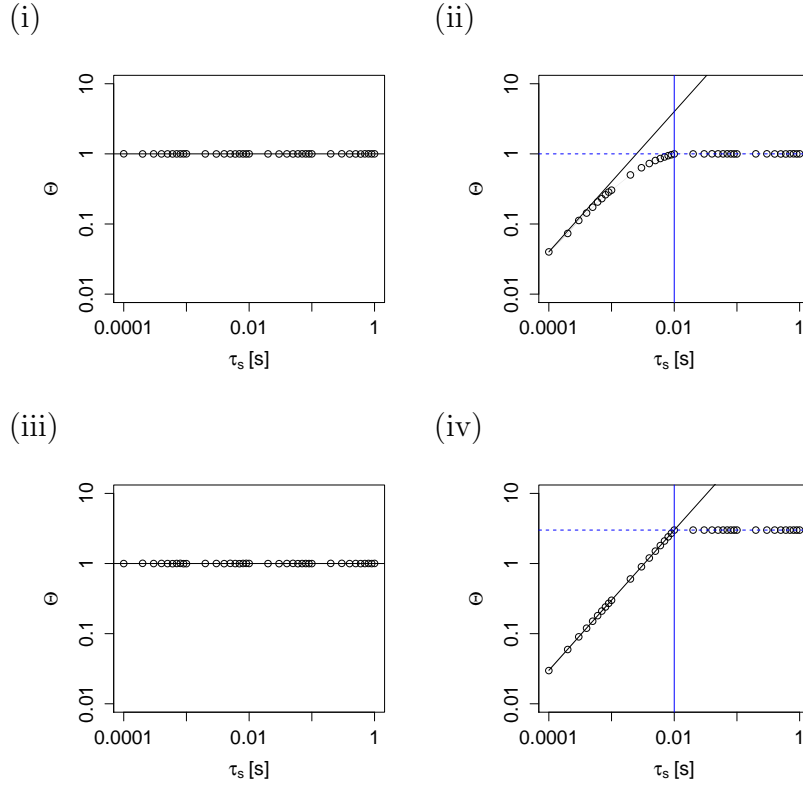
and

$$\Lambda(x, \tau) = \frac{25 (M^{(4)}(x, \tau))^3}{3 (M^{(6)}(x, \tau))^2} \approx \begin{cases} 1, & \text{diffusive} \\ \lambda(x)\tau, & \text{jumpy} \\ 1, & \text{white noise at } x = 0 \\ \frac{25}{3}\lambda\tau, & \text{pure Poisson jumps,} \end{cases} \quad (5.27)$$

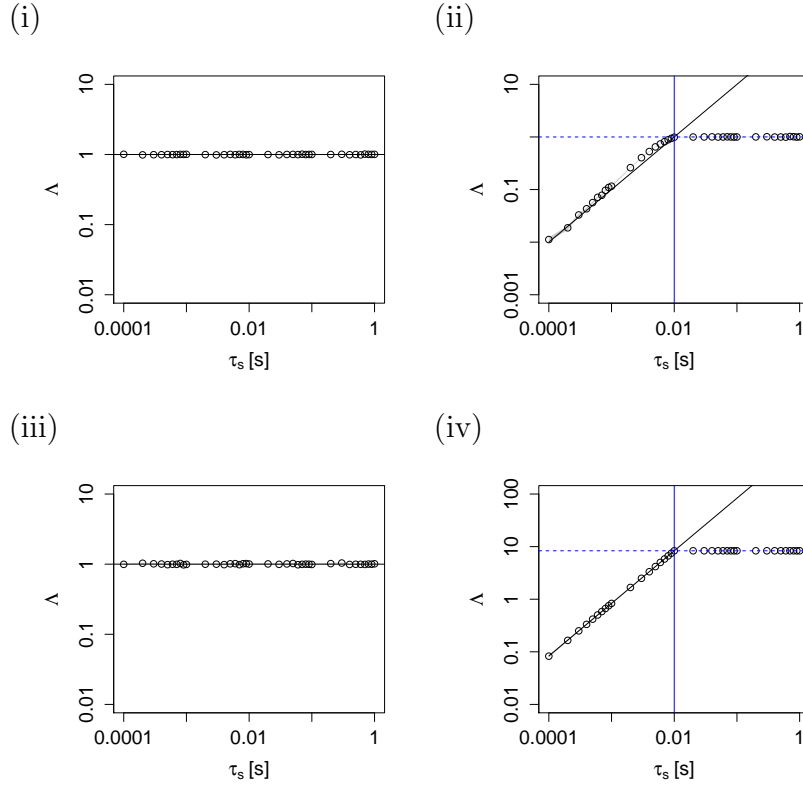
which can be derived from the second-, fourth- and sixth-order conditional moments Eqs. (5.8), (5.14), (5.18) and (5.20). The relation  $M^{(4)}(x, \tau) = 3 (M^{(2)}(x, \tau))^2$  for diffusion process corresponding to  $\Theta$ -criterion has also been shown in [45]. The analytic results for  $\Theta$ - and  $\Lambda$ -criterion of the same four cases are given. The explicit  $\tau$ -dependence for the jumpy case can be obtained from Eq. (5.14) and considered up to the first non-vanishing order term. Again only for the case of Gaussian white noise,  $\Theta$  and  $\Lambda$  has to be determined for  $x = 0$ . From Eq. (5.26) and (5.27), we can see the similar functional behavior of  $\Theta$  and  $\Lambda$ .  $\Lambda$  is also associated with the jump probability  $\lambda\tau$  in jump-diffusion process for  $\tau < \frac{1}{\lambda} = T_J$ .

To validate these  $\Theta$ - and  $\Lambda$ -criterion, the same simulations are performed as in the validation of the  $Q$ -criterion. The results are shown in Fig. 5.2 and 5.3. Theoretical predictions described in Eq. (5.26) and (5.27) are also plotted. From these plots, we can observe that approximations in

Eq. (5.26) and (5.27) for jump processes are valid for  $\tau_s < \frac{1}{\lambda} = T_J$ . For  $\tau_s > T_J$ , the jump-diffusion process apparently becomes diffusion process which gives  $\Theta = 1$  and  $\Lambda = 1$ . The zero-one Poisson jump noise for  $\tau_s > T_J$  apparently has the jump probability of one,  $\text{Prob}[\Delta x(t) = 1] = \lambda\tau = 1$  as discussed in Chapter 2, which gives  $\Theta = 3$  and  $\Lambda = \frac{25}{3}$ .



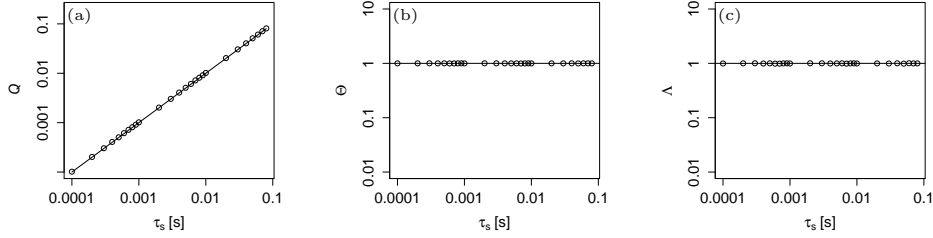
**Figure 5.2:** The  $\Theta$ -criterion for OU process (i), OU jump-diffusion process (ii), standard normal Gaussian white noise (iii) and pure Poisson jump noise (iv) are evaluated and plotted in double logarithmic scale. The OU processes are generated for different integration time steps  $\tau_s$  with  $\gamma = 1 \text{ s}^{-1}$ ,  $D = 1 \text{ s}^{-1}$  and also with the additional jump terms with  $\lambda = 100 \text{ s}^{-1}$  and  $\sigma_\xi^2 = 0.01$ . The vertical blue lines show the jump time scale  $T_J = \frac{1}{\lambda}$ . For each  $\tau = \tau_s$ ,  $\Theta = \Theta(x = 0, \tau)$  is calculated. The theoretical predictions according to Eq. (5.26) are presented with the black solid lines. The blue dashed lines  $\Lambda = 1$  in case (ii) and  $\Lambda = 3$  in case (iv) show the behavior after  $\tau_s > T_J = \frac{1}{\lambda}$ . Standard errors are smaller than the symbol size.



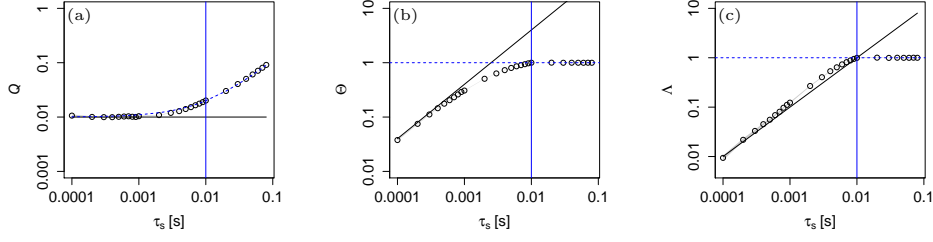
**Figure 5.3:** The  $\Lambda$ -criterion for OU process (i), OU jump-diffusion process (ii), standard normal Gaussian white noise (iii) and pure Poisson jump noise (iv) are evaluated and plotted in double logarithmic scale. The OU processes are generated for different integration time steps  $\tau_s$  with  $\gamma = 1 \text{ s}^{-1}$ ,  $D = 1 \text{ s}^{-1}$  and also with the additional jump terms with  $\lambda = 100 \text{ s}^{-1}$  and  $\sigma_\xi^2 = 0.01$ . The vertical blue lines show the jump time scale  $T_J = \frac{1}{\lambda}$ . For each  $\tau = \tau_s$ ,  $\Lambda = \Lambda(x = 0, \tau)$  is calculated. The theoretical predictions according to Eq. (5.27) are presented with the black solid lines. The blue dashed lines  $\Lambda = 1$  in case (ii) and  $\Lambda = \frac{25}{3}$  in case (iv) show the behavior after  $\tau_s > T_J = \frac{1}{\lambda}$ . Standard errors are smaller than the symbol size.

### 5.2.3 $Q$ -, $\Theta$ - and $\Lambda$ -criterion for General Diffusion and Jump-Diffusion Processes

Now, the  $Q$ -,  $\Theta$ - and  $\Lambda$ - criterion are determined for both diffusion and jump-diffusion processes with non-linear drift and multiplicative diffusion term. We numerically integrate Eq. (5.6) with  $D^{(1)}(x) = -x^3$  and  $D^{(2)}(x) = 1 + x^2$ , and Eq. (5.11) for the case with additional jump term with  $\lambda = 100 \text{ s}^{-1}$  and  $\sigma_\xi^2 = 1$  for different integration time step  $\tau_s$ . The result of the non-linear diffusion process is shown in Fig. 5.4 and jump-diffusion process in Fig. 5.5.



**Figure 5.4:**  $Q$ -,  $\Theta$ - and  $\Lambda$ -criterion of diffusion process at  $x = 0$  with  $D^{(1)}(x) = -x^3$  and  $D^{(2)}(x) = 1 + x^2$ , generated for different integration time steps  $\tau_s$ , are plotted in double logarithmic scale. Theoretical predictions according to Eqs. (5.21),(5.26) and (5.27) are presented with the black solid lines. Standard errors are smaller than the symbol size.



**Figure 5.5:**  $Q$ -,  $\Theta$ - and  $\Lambda$ -criterion of jump-diffusion process at  $x = 0$  with  $D^{(1)}(x) = -x^3$ ,  $D^{(2)}(x) = 1 + x^2$ ,  $\lambda = 100 \text{ s}^{-1}$  and  $\sigma_\xi^2 = 1$ , generated for different integration time steps  $\tau_s$ , are plotted in double logarithmic scale. The vertical blue lines show the jump time scale  $T_J = \frac{1}{\lambda}$ . Theoretical predictions according to Eqs. (5.21),(5.26) and (5.27) are presented with the black solid lines. The blue dashed curve  $Q = D^{(2)}(0)\tau_s + \sigma_\xi^2$  and lines  $\Theta = \Lambda = 1$  show the behavior after  $\tau_s > T_J = \frac{1}{\lambda}$ . Standard errors are smaller than the symbol size.

Comparing these results with Fig. 5.1, 5.2 and 5.3, we can observe the similar behavior in the case with linear drift and additive diffusion noise.

After deriving the mathematical criteria to distinguish the diffusive and jumpy behaviors for the simulated stochastic processes, we will discuss how these criteria can be applied dealing with real world data.

### 5.3 Consequences of Downsampling of the Data

In real world, empirical data are given for a fixed sampling time  $\tau_s$  due to experimental constraints. To handle this effect, we have to downsample the numerical data in which a new time step  $\tau_{ds}$  is defined for downsampling time step. For the estimation of the KM conditional moments, it is essential to see in which relation  $\tau_{ds}$  is with the correlation time  $T_C = \frac{1}{\gamma}$  and with the average time scale between the jumps  $T_J = \frac{1}{\lambda}$ .

Influence of downsampling on the diffusion coefficient is first examined. In order to study the downsampled approximation of the diffusion term, it is integrated over a finite downsampled time  $\tau_{ds}$ . In the Itô sense, the integral is generally interpreted as [23, 33]

$$\begin{aligned} \int_t^{t+\tau_{ds}} g(x(s), s) dW_s &= g(x(t), t) \int_t^{t+\tau_{ds}} dW_s \\ &= g(x(t), t) \sum_{k=1}^{N_\tau} (W_k - W_{k-1}) \end{aligned} \quad (5.28)$$

Therefore, the integral of diffusion part in (jump-)diffusion process becomes

$$\begin{aligned} \int_t^{t+\tau_{ds}} \sqrt{D^{(2)}(x)} dW_s &= \sqrt{D^{(2)}(x)} \sum_{k=1}^{N_\tau} \Delta W_k \\ &= \sqrt{D^{(2)}(x)} \tau_s \sum_{k=1}^{N_\tau} \eta_k \end{aligned} \quad (5.29)$$

where  $\eta_k \sim N(0, \sigma_\eta^2 = 1)$  is the standard normal distributed random variable. To estimate the summation  $\sum_{k=1}^{N_\tau} \eta_k$ ,  $N_\tau = \frac{\tau_{ds}}{\tau_s}$  is the number of time steps in the downsampled time step  $\tau_{ds}$ . Let us define the summation as a new effective noise  $\eta_{\text{eff}}$  for diffusion term as,



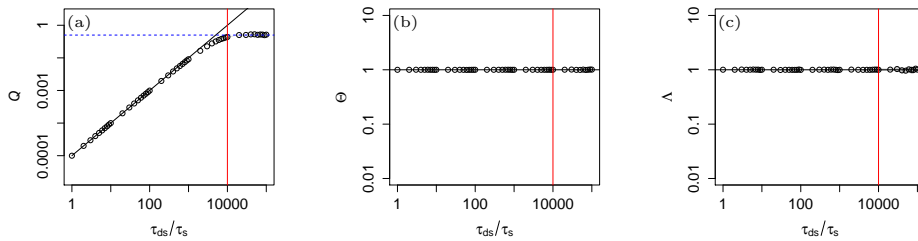
$$\eta_{\text{eff}} = \sum_{k=1}^{N_\tau} \eta_k \quad (5.30)$$

where variance of  $\eta_{\text{eff}}$  will be  $\sigma_{\eta,\text{eff}}^2 = N_\tau \sigma_\eta^2 = N_\tau = \frac{\tau_{\text{ds}}}{\tau_s}$ , (as  $\sigma_\eta^2 = 1$ ), and  $\eta_{\text{eff}}$  tends to a Gaussian random variable. Then, the integral becomes

$$\int_t^{t+\tau_{\text{ds}}} \sqrt{D^{(2)}(x)} dW_s \simeq \sqrt{D^{(2)}(x)} \tau_s \eta_{\text{eff}} = \sqrt{D^{(2)}(x)} \tau_{\text{ds}} \eta \quad (5.31)$$

where  $\eta \sim N(0, 1)$  is the standard normal distributed. Therefore, we can conclude that diffusion coefficient remains unchanged with downsampling and summation of the noises approaches to a Gaussian white noise.

To test it, the OU process with  $\gamma = 1 \text{ s}^{-1}$ ,  $D = 1 \text{ s}^{-1}$  is simulated at the fixed sampling time  $\tau_s = 10^{-4} \text{ s}$ . It is then downsampled with different time step  $\tau_{\text{ds}}$  and the  $Q$ -,  $\Theta$ - and  $\Lambda$ -criterion at  $x = 0$  are determined as shown in Fig. 5.6.



**Figure 5.6:**  $Q$ -,  $\Theta$ - and  $\Lambda$ -criterion of downsampled OU (diffusion) process at  $x = 0$  with  $\gamma = 1 \text{ s}^{-1}$ ,  $D = 1 \text{ s}^{-1}$  are plotted in double logarithmic scale. The vertical red line shows the correlation time scale  $T_C = \frac{1}{\gamma}$ . Theoretical predictions according to Eqs. (5.21),(5.26) and (5.27) are presented with the black solid lines for  $\tau_{\text{ds}} < T_C$ . The blue dashed line in (a) shows  $Q = \sigma_x^2$  after  $\tau_{\text{ds}} > T_C$ . Standard errors are smaller than the symbol size.

We can observe the same diffusive behavior for  $\tau_{\text{ds}} < T_C = \frac{1}{\gamma}$  as described in Eq. (5.21), (5.26) and (5.27). For downsampled time steps larger than correlation length,  $\tau_{\text{ds}} > T_C$ , the results of the criteria show the behavior of uncorrelated Gaussian white noise in which  $Q = \sigma_x^2$  where

$\sigma_x^2 = 0.5$  of OU process  $x$  for this case. The variance  $\sigma_x^2$  can also be calculated with Eq. (2.60) without considering the jump parameters  $\lambda$  and  $\sigma_\xi^2$ . The other criteria  $\Theta = \Lambda = 1$  which is also valid for Gaussian white noise.

Influence of downsampling on the jump amplitude is then examined. The jump term  $\xi dJ_t$  will provide the following stochastic integral

$$\int_t^{t+\tau_{\text{ds}}} \xi dJ_s . \quad (5.32)$$

To deal with this integral, we use the integral relation [33, 81]

$$\int_t^{t+\tau_{\text{ds}}} h(x(s), s) dJ_s \stackrel{\text{ms}}{=} \sum_{k=1}^{J_t} h(x(T_k^-), T_k^-) , \quad (5.33)$$

where  $\stackrel{\text{ms}}{=}$  is Itô mean square equals. The  $T_k^-$  denotes the limit from the left to jump time  $T_k$ . Therefore, the jump integral becomes

$$\int_t^{t+\tau_{\text{ds}}} \xi dJ_s = \sum_{k=1}^{J_t} \xi_k , \quad (5.34)$$

where  $\xi_k \sim N(0, \sigma_\xi^2)$  is the Gaussian distributed random variable with zero mean and variance  $\sigma_\xi^2$ . To estimate the summation  $\sum_{k=1}^{J_t} \xi_k$ , we note that the average time step of jumps in original time series is  $\frac{1}{\lambda}$ , then the average number of jumps becomes  $J_t \simeq \lambda\tau_{\text{ds}}$  in the downsampled time step  $\tau_{\text{ds}}$ . Let us define the summation as a new noise  $\xi_{\text{eff}}$  for jump amplitude as,

$$\xi_{\text{eff}} = \sum_{k=1}^{J_t} \xi_k , \quad (5.35)$$

where variance of  $\xi_{\text{eff}}$  will be  $\sigma_{\xi, \text{eff}}^2 \simeq \lambda\tau_{\text{ds}}\sigma_\xi^2$  and  $\xi_{\text{eff}}$  tends to a Gaussian random variable. Then, we find,

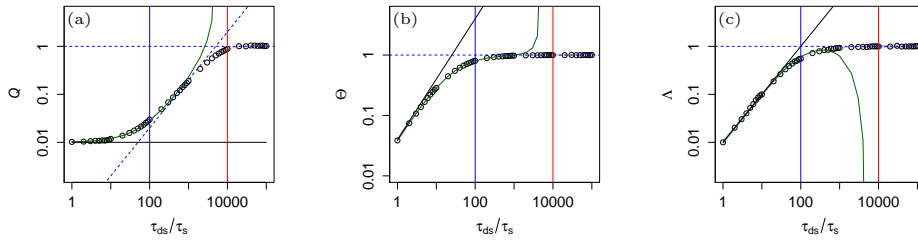
$$\int_t^{t+\tau_{\text{ds}}} \xi \, dJ_s = \xi_{\text{eff}} \simeq \sqrt{\lambda\tau_{\text{ds}}} \xi, \quad (5.36)$$

where  $\xi \sim N(0, \sigma_\xi^2)$  is Gaussian distributed with variance of jump amplitude  $\sigma_\xi^2$ ). Therefore, if we downsample with  $\tau_{\text{ds}} \gg \frac{1}{\lambda}$ ,  $J_t \gg 1$ , then zero-one jump law is no longer fulfilled. However, it can be realized as a jump event with the effective jump amplitude

$$\sigma_{\xi, \text{eff}}^2 = \lambda \sigma_\xi^2 \tau_{\text{ds}}. \quad (5.37)$$

• **Case  $T_J < T_C$ :**

The OU jump-diffusion process with  $\gamma = 1 \text{ s}^{-1}$ ,  $D = 1 \text{ s}^{-1}$ ,  $\lambda = 100 \text{ s}^{-1}$  and  $\sigma_\xi^2 = 1$  at the fixed sampling time  $\tau_s = 10^{-4} \text{ s}$  is generated. It is then downsampled with different time step  $\tau_{\text{ds}}$  and the  $Q$ -,  $\Theta$ - and  $\Lambda$ -criterion are determined as shown in Fig. 5.7.

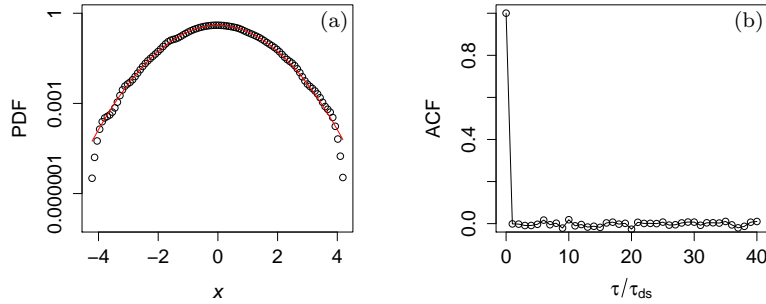


**Figure 5.7:**  $Q$ -,  $\Theta$ - and  $\Lambda$ -criterion of downsampled OU jump-diffusion process at  $x = 0$  with  $\gamma = 1 \text{ s}^{-1}$ ,  $D = 1 \text{ s}^{-1}$ ,  $\lambda = 100 \text{ s}^{-1}$  and  $\sigma_\xi^2 = 1$  are plotted in double logarithmic scale. The vertical blue lines show the jump time scale  $T_J = \frac{1}{\lambda}$  and the vertical red lines the correlation time scale  $T_C = \frac{1}{\gamma}$ . In this case  $T_J < T_C$ . Theoretical predictions according to Eqs. (5.21), (5.26) and (5.27) are presented with the black solid lines. The green solid lines show the approximation up to  $\mathcal{O}(\tau^4)$  of conditional moments described in Eq. (5.16). The blue dashed lines show  $Q = (D + \lambda\sigma_\xi^2)\tau_{\text{ds}}$  for  $T_J < \tau_{\text{ds}} < T_C$  and  $Q = \sigma_x^2$  for  $\tau_s > T_C$ , and  $\Theta = \Lambda = 1$ . Standard errors are smaller than the symbol size.

Here, we can see that the approximations in Eqs. (5.21), (5.26) and (5.27) are still valid for  $\tau_{\text{ds}} \ll T_J$ . (If we consider calculating these criteria up

to  $\mathcal{O}(\tau^4)$  of conditional moments described in Eq. (5.16), we can have a better approximation shown in solid green line in Fig. 5.7.) In the range of  $T_J < \tau_{ds} < T_C$ , the effective value of  $Q$  becomes the sum of effective diffusion and jump terms described in Eqs. (5.31), (5.36) and (5.37) which gives  $Q = (D + \lambda\sigma_\xi^2)\tau_{ds}$ .  $\Theta$  and  $\Lambda$  also become one after a few time steps in this range which apparently show the diffusive behavior. These are shown with the blue dashed lines in Fig. 5.7.

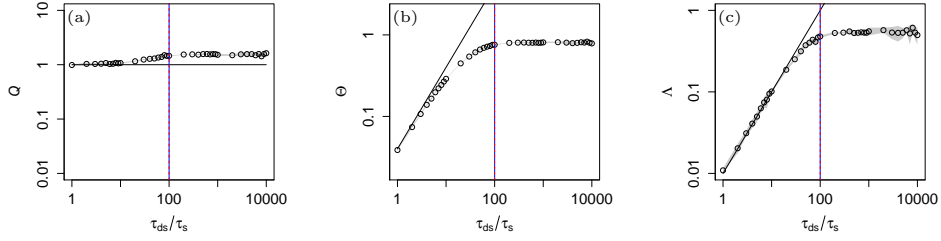
After  $\tau_{ds} > T_C$ , the downsampled process becomes uncorrelated with Gaussian distribution, where  $Q$  becomes the variance of the process  $x$ ,  $Q = \sigma_x^2 = 1$  and  $\Theta = \Lambda = 1$ , as seen in Fig 5.7 right to the vertical red lines. The probability density function (PDF) of the downsampled process  $x$  at  $\tau_{ds} > T_C$  also shows that it is Gaussian-distributed and the autocorrelation function (ACF) shows that it is uncorrelated as well. (see Fig. 5.8). The variance of the process  $x$  can also be calculated analytically with Eq. (2.60) which is  $\sigma_x^2 = 1$  in this case.



**Figure 5.8:** The PDF and ACF of downsampled OU jump-diffusion process at  $\tau > T_C$  in the case of  $T_J < T_C$ . The red solid line shows the PDF of Gaussian distribution with zero mean and the same variance  $\sigma_x^2$ . The PDF is plotted in semi-logarithmic scale.

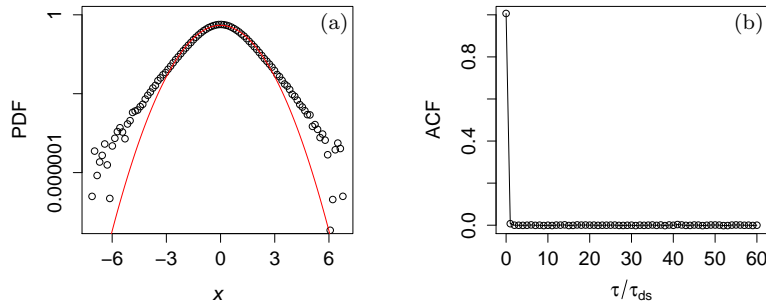
- **Case  $T_J = T_C$ :**

The OU jump-diffusion process with  $\gamma = 100 \text{ s}^{-1}$ ,  $D = 100 \text{ s}^{-1}$ ,  $\lambda = 100 \text{ s}^{-1}$  and  $\sigma_\xi^2 = 1$  at the fixed sampling time  $\tau_s = 10^{-4} \text{ s}$  is generated. It is then downsampled with different time step  $\tau_{ds}$  and the  $Q$ -,  $\Theta$ - and  $\Lambda$ -criterion are determined as shown in Fig. 5.9.



**Figure 5.9:**  $Q$ -,  $\Theta$ - and  $\Lambda$ -criterion of downsampled OU jump-diffusion process at  $x = 0$  with  $\gamma = 100 \text{ s}^{-1}$ ,  $D = 100 \text{ s}^{-1}$ ,  $\lambda = 100 \text{ s}^{-1}$  and  $\sigma_\xi^2 = 1$  are plotted in double logarithmic scale. Theoretical predictions according to Eqs. (5.21),(5.26) and (5.27) are presented with the black solid lines. The vertical blue lines show the jump time scale  $T_J = \frac{1}{\lambda}$  and the vertical red lines the correlation time scale  $T_C = \frac{1}{\gamma}$ . In this case  $T_J = T_C$ . Standard errors of the mean are shown as gray-shaded background.

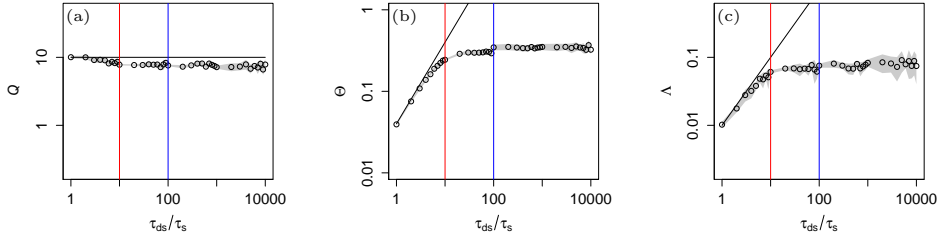
Here, we can see that the approximations in Eqs. (5.21), (5.26) and (5.27) are still valid for  $\tau_{\text{ds}} \ll T_J = T_C$ . For  $\tau_{\text{ds}} > T_J = T_C$ ,  $Q$ ,  $\Theta$  and  $\Lambda$  become constant but do not have the values of Gaussian white noise. The PDF of the downsampled process  $x$  at  $\tau_{\text{ds}} > T_C$  shows that it is non-Gaussian, however, the ACF shows that it is uncorrelated (see Fig. 5.10).



**Figure 5.10:** The PDF and ACF of downsampled OU jump-diffusion process at  $\tau > T_C$  in the case of  $T_J = T_C$ . The red solid line shows the PDF of Gaussian distribution with zero mean and the same variance  $\sigma_x^2$ . The PDF is plotted in semi-logarithmic scale.

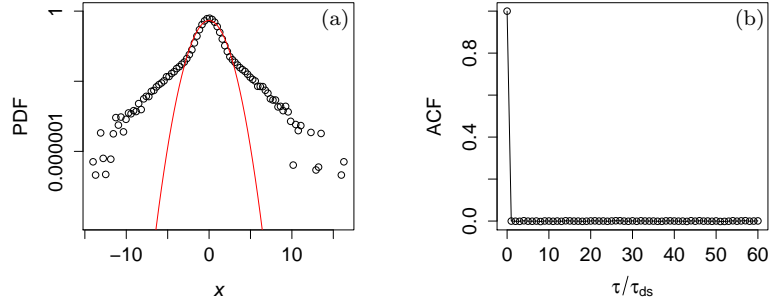
• **Case  $T_J > T_C$ :**

The OU jump-diffusion process with  $\gamma = 1000 \text{ s}^{-1}$ ,  $D = 1000 \text{ s}^{-1}$ ,  $\lambda = 100 \text{ s}^{-1}$  and  $\sigma_\xi^2 = 10$  at the fixed sampling time  $\tau_s = 10^{-4} \text{ s}$  is generated. It is then downsampled with different time step  $\tau_{ds}$  and the  $Q$ -,  $\Theta$ - and  $\Lambda$ -criterion are determined as shown in Fig. 5.11.



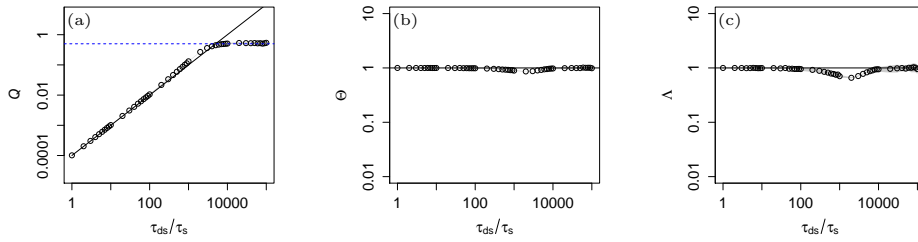
**Figure 5.11:**  $Q$ -,  $\Theta$ - and  $\Lambda$ -criterion of downsampled OU jump-diffusion process at  $x = 0$  with  $\gamma = 1000 \text{ s}^{-1}$ ,  $D = 1000 \text{ s}^{-1}$ ,  $\lambda = 100 \text{ s}^{-1}$  and  $\sigma_\xi^2 = 10$  are plotted in double logarithmic scale. Theoretical predictions according to Eqs. (5.21),(5.26) and (5.27) are presented with the black solid lines. The vertical blue lines show the jump time scale  $T_J = \frac{1}{\lambda}$  and the vertical red lines the correlation time scale  $T_C = \frac{1}{\gamma}$ . In this case  $T_J > T_C$ . Standard errors of the mean are shown as gray-shaded background.

Here, we can see that the approximations in Eqs. (5.21), (5.26) and (5.27) are still valid for  $\tau_{ds} \ll T_C$ . For  $\tau_{ds} > T_J$ ,  $Q$ ,  $\Theta$  and  $\Lambda$  become constant but again do not have the values of Gaussian white noise. The PDF of the downsampled process  $x$  at  $\tau > T_J$  shows that it is non-Gaussian, however, the ACF shows that it is still uncorrelated (see Fig. 5.12).

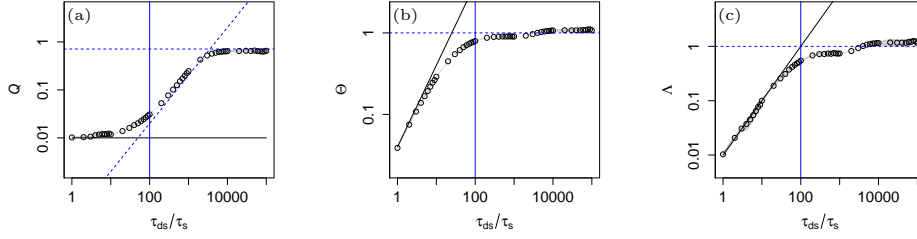


**Figure 5.12:** The PDF and ACF of downsampled OU jump-diffusion process at  $\tau > T_J$  in the case of  $T_J > T_C$ . The red solid line shows the PDF of Gaussian distribution with zero mean and the same variance  $\sigma_x^2$ . The PDF is plotted in semi-logarithmic scale.

Now, the  $Q$ -,  $\Theta$ - and  $\Lambda$ - criterion are determined for both downsampled diffusion and jump-diffusion processes with non-linear drift and multiplicative diffusion terms. We numerically integrate Eq. (5.6) with  $D^{(1)}(x) = -x^3$  and  $D^{(2)}(x) = 1 + x^2$ , and Eq. (5.11) for the case with additional jump term with  $\lambda = 100 \text{ s}^{-1}$  and  $\sigma_\xi^2 = 1$  with a fixed sampling time  $\tau_s = 10^{-4} \text{ s}$ . The results are shown in Fig. 5.13 for diffusion process, and in Fig. 5.14 for jump-diffusion process.



**Figure 5.13:**  $Q$ -,  $\Theta$ - and  $\Lambda$ -criterion of downsampled diffusion process at  $x = 0$  with  $D^{(1)}(x) = -x^3$  and  $D^{(2)}(x) = 1 + x^2$  are plotted in double logarithmic scale. Theoretical predictions according to Eqs. (5.21), (5.26) and (5.27) are presented with the black solid lines. The blue dashed line in (a) shows  $Q = \sigma_x^2$ . Standard errors are smaller than the symbol size.



**Figure 5.14:**  $Q$ -,  $\Theta$ - and  $\Lambda$ -criterion of downsampled OU jump-diffusion process at  $x = 0$  with  $D^{(1)}(x) = -x^3$ ,  $D^{(2)}(x) = 1 + x^2$ ,  $\lambda = 100 \text{ s}^{-1}$  and  $\sigma_\xi^2 = 1$  are plotted in double logarithmic scale. The vertical blue lines show the jump time scale  $T_J = \frac{1}{\lambda}$ . Theoretical predictions according to Eqs. (5.21), (5.26) and (5.27) are presented with the black solid lines. The blue dashed lines show  $Q = (D^{(2)}(0) + \lambda\sigma_\xi^2)\tau_{\text{ds}}$  for  $\tau_{\text{ds}} > T_J$  and  $Q = \sigma_x^2$  (horizontal line), and  $\Theta = \Lambda = 1$ . Standard errors are smaller than the symbol size.

Comparing these results with Fig. 5.6 and 5.7, we can observe the similar behavior in the case with linear drift and additive diffusion noise.

So far for the jump diffusion processes, different cases of between the time scales  $T_J$  and  $T_C$  are analyzed. It remains to study how the magnitude of jump noise with respect to the diffusion noise affect the stochastic process.



## 5.4 Transition from Jumpy to Diffusive Behavior with Changing Jump Amplitude

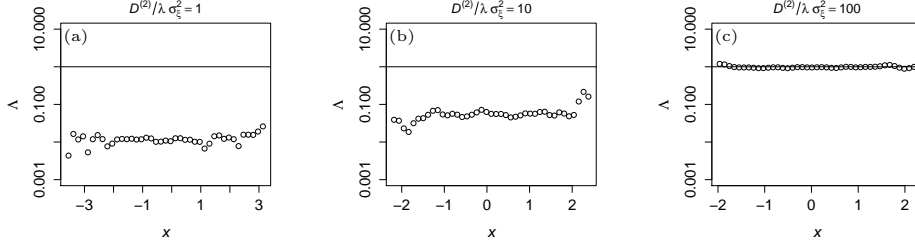
In principle, there are two limits of vanishing jumpy contribution to a stochastic process,  $\lambda \rightarrow 0$  or  $\sigma_\xi^2 \rightarrow 0$ . For realistic scenario to investigate the transition from jumpy to diffusive behavior, we should have a finite  $\lambda$  to make sure that there are significant jump events occurring. Therefore, we consider the fixed  $\lambda$  and investigate the influence of jumps by varying the second important parameter of the jumpy process, namely the jump amplitude  $\sigma_\xi^2$ .

Therefore, the  $\Lambda$ -criterion for OU jump-diffusion process is estimated with different jump amplitudes  $\sigma_\xi^2$ , keeping  $\gamma = 1 \text{ s}^{-1}$ ,  $D = 1 \text{ s}^{-1}$  and  $\lambda = 100 \text{ s}^{-1}$ .  $\sigma_\xi^2$  is varied by several orders of magnitudes ending with small values, corresponding to nearly pure diffusive processes. In particular,  $\sigma_\xi^2$  goes from 10 to  $10^{-7}$ . In order to examine broader aspect between diffusion and jump contribution, the ratio of diffusion and jump noise  $\frac{D}{\lambda\sigma_\xi^2}$  is applied which has been introduced in Chapter 3 and 4.

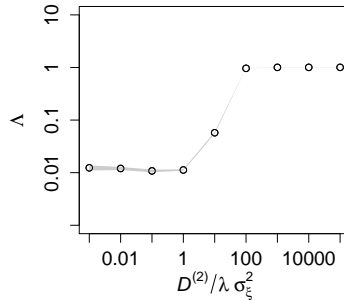
Since the  $\Lambda$ -criterion is associated with the jump rate  $\lambda$  in such a way that  $\Lambda = \lambda\tau_{\text{ds}}$  for  $\tau_{\text{ds}} \ll T_J$  in jump-diffusion process, only  $\Lambda$ -criterion is shown and discussed for this study. Moreover, it is also uniquely constant one for diffusion process for all range of state variable  $x$  and all time scale  $\tau$ . The estimated  $\Lambda = \Lambda(x, \tau = \tau_s)$  are plotted for  $\frac{D}{\lambda\sigma_\xi^2} = \{1, 10, 100\}$  in Fig. 5.15. We can observed that the transition from jump-diffusion to the diffusion process happens at  $\frac{D}{\lambda\sigma_\xi^2} \simeq 10$ .

In order to illustrate more clearly, the mean of  $\Lambda(x, \tau_s)$  in the range of  $x \in [-2\sigma_x, 2\sigma_x]$  is calculated, where  $\sigma_x$  is the standard deviation of  $x$ . Then, the mean of  $\Lambda$  versus  $\frac{D}{\lambda\sigma_\xi^2}$  is plotted in double-logarithmic scale as shown in Fig. 5.16. For  $\frac{D}{\lambda\sigma_\xi^2} \leq 1$ ,  $\Lambda = \lambda\tau_s$  which is clearly a jumpy process. However, we observe a clear transition of jump-diffusion to

diffusion process at  $\frac{D}{\lambda\sigma_\xi^2} \simeq 10$ . From  $\frac{D}{\lambda\sigma_\xi^2} \simeq 100$  onward,  $\Lambda = 1$  and the process behaves like pure diffusion process.



**Figure 5.15:** The function  $\Lambda = \Lambda(x, \tau_s)$  for the ratio of diffusion and jump noise  $\frac{D}{\lambda\sigma_\xi^2} = (1, 10, 100)$  are plotted in semi-logarithmic scale. The OU-jump-diffusion processes are generated with the sampling period  $\tau_s = 10^{-4}$  s,  $\gamma = 1$  s $^{-1}$ ,  $D = 1$  s $^{-1}$  and  $\lambda = 100$  s $^{-1}$  for different  $\sigma_\xi^2$ . Solid black lines show  $\Lambda = 1$ . The transition from jump-diffusion to the diffusion process happens around  $\frac{D}{\lambda\sigma_\xi^2} \simeq 10$ . From  $\frac{D}{\lambda\sigma_\xi^2} \simeq 100$  onward, the process behaves like a pure diffusion process where  $\Lambda = 1$ .



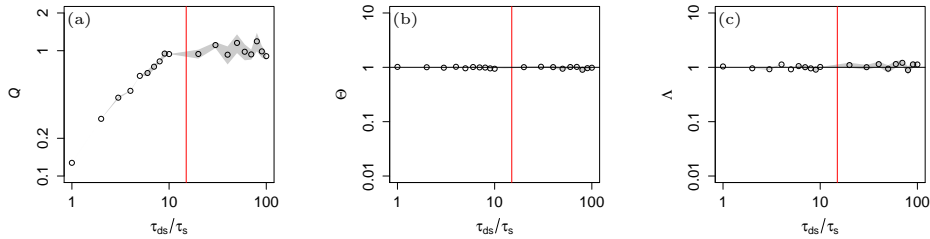
**Figure 5.16:** The mean of  $\Lambda(x, \tau_s)$  in the range of  $x \in [-2\sigma_x, 2\sigma_x]$ , where  $\sigma_x$  is the standard deviation of  $x$  of each OU jump-diffusion process for different ratio of diffusion and jump noise  $\frac{D}{\lambda\sigma_\xi^2}$ , varying from  $10^{-3}$  to  $10^5$  by the order of magnitudes is plotted in double logarithmic scale. Clear transition of jump-diffusion to diffusion process at  $\frac{D}{\lambda\sigma_\xi^2} = 10$  can be observed. From  $\frac{D}{\lambda\sigma_\xi^2} = 100$  onward, the process behaves like a pure diffusion process where  $\Lambda = 1$ . Standard errors are smaller than the symbol size.

Next, we apply these methods on the selected real world empirical data.

## 5.5 Application on Real World Data

The aforementioned criteria were applied on the real world data to investigate whether they are diffusive or jumpy. All the measured data  $x(t)$  with the sampling time  $\tau_s$  are subtracted with their means and divided with their standard deviations, so that the normalized data have zero-means and standard deviations of one. The  $Q$ -,  $\Theta$ - and  $\Lambda$ -criterion are evaluated at  $x = 0$  by downsampling of the given data with downsampling time step  $\tau_{ds}$ . We take five different data sets and separated into three categories.

- (i) In the first category, we analyze are the measurements of the spatial positions of a dielectric bead (polystyrene, diameter- $1\mu\text{m}$ , Bangs Laboratories Inc. USA) trapped in and optical tweezers [58]. The results of  $Q = Q(x = 0, \tau_{ds})$ ,  $\Theta = \Theta(x = 0, \tau_{ds})$  and  $\Lambda = \Lambda(x = 0, \tau_{ds})$  versus  $\tau_{ds}$ , for a single measured data with duration 3 s, and with sample rate 22 kHz are plotted in Fig. 5.17.

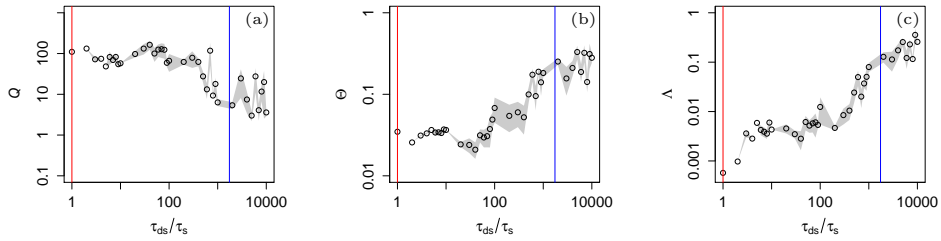


**Figure 5.17:**  $Q$ -,  $\Theta$ - and  $\Lambda$ -criterion of downsampled displacement data of dielectric bead in the optical tweezers are plotted in double logarithmic scale. The vertical red lines the correlation time scale  $T_C = \frac{1}{\gamma}$ . The black solid lines show  $\Theta = \Lambda = 1$ , together with the linear behavior of  $Q$  for  $\tau_{ds} < T_C$ , this process fulfils the necessary conditions of diffusion process. Standard errors of the mean are shown as gray-shaded background and most of them are smaller than the symbol size..

The results from optical tweezers show the behavior of diffusion process. The correlation time  $T_C = \frac{1}{\gamma}$  is first estimated from the drift function which is approximately  $D^{(1)}(x) = -\gamma x$ . Below the correlation time scale  $\tau_{ds} < T_C$ ,  $Q$  has the linear behavior and  $\Theta = \Lambda = 1$ . For  $\tau_{ds} > T_C$ ,  $Q = \sigma_x^2 = 1$  and  $\Theta = \Lambda = 1$  also which show the behavior of uncorrelated Gaussian white noise.

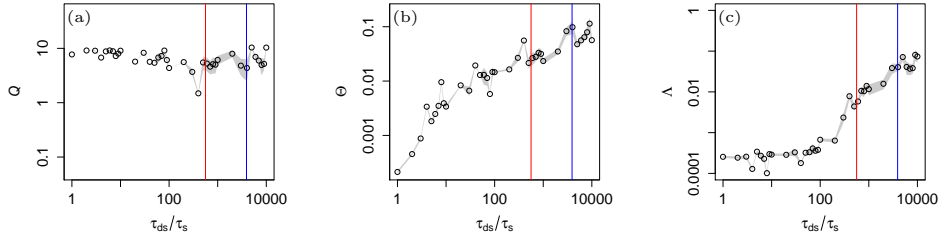
- (ii) In the second category, we analyze DAX stock market index data [27] and epileptic brain data [5].

For the DAX stock market index, the log return is used for the analysis. The results of  $Q = Q(x = 0, \tau_{ds})$ ,  $\Theta = \Theta(x = 0, \tau_{ds})$  and  $\Lambda = \Lambda(x = 0, \tau_{ds})$  versus  $\tau_{ds}$  are plotted in Fig. 5.18.



**Figure 5.18:**  $Q$ -,  $\Theta$ - and  $\Lambda$ -criterion of downsampled log-return DAX data are plotted in double logarithmic scale. The vertical blue lines show the jump time scale  $T_J = \frac{1}{\lambda}$  and the vertical red lines the correlation time scale  $T_C = \frac{1}{\gamma}$ . In this case,  $T_C < T_J$ . Standard errors of the mean are shown as gray-shaded background.

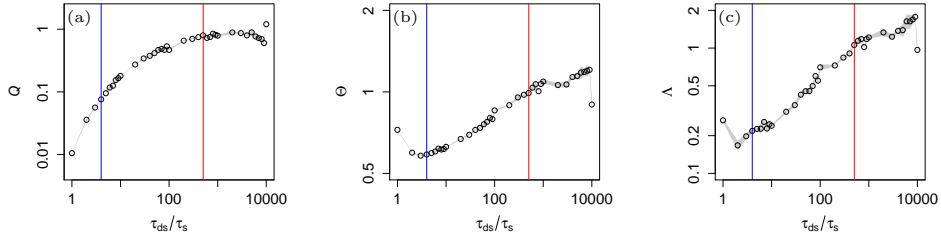
The results of  $Q = Q(x = 0, \tau_{ds})$ ,  $\Theta = \Theta(x = 0, \tau_{ds})$  and  $\Lambda = \Lambda(x = 0, \tau_{ds})$  versus  $\tau_{ds}$  for epileptic brain data are plotted in Fig. 5.19.



**Figure 5.19:**  $Q$ -,  $\Theta$ - and  $\Lambda$ -criterion of downsampled epileptic brain data are plotted in double logarithmic scale. The vertical blue lines show the jump time scale  $T_J = \frac{1}{\lambda}$  and the vertical red lines the correlation time scale  $T_C = \frac{1}{\gamma}$ . In this case,  $T_C < T_J$ . Standard errors of the mean are shown as gray-shaded background and most of them are smaller than the symbol size.

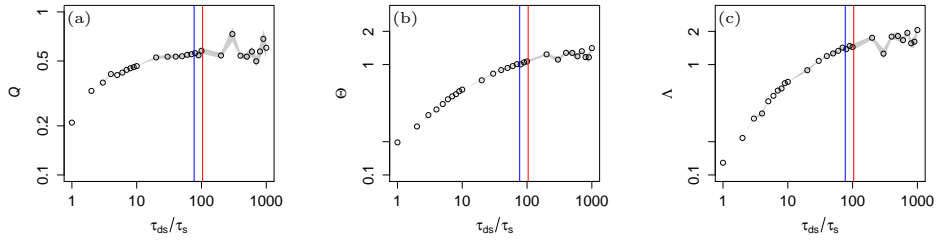
Since  $\Theta$  and  $\Lambda$  are deviated from one, they do not fulfil the necessary conditions for the diffusion process. Here we evaluated again the correlation time  $T_C = \frac{1}{\gamma}$  obtained from the drift function which is approximately  $D^{(1)}(x) = -\gamma x$ . We also estimated  $T_J = \frac{1}{\lambda}$  by Eq. (2.36). Both DAX stock market index data and epileptic brain data show the case of  $T_J > T_C$  which resemble the case of Fig. 5.11. For small  $\tau_{ds}$ ,  $Q$  shows the constant behavior while  $\Theta$  and  $\Lambda$  are not one and also show the functional behavior on  $\tau$ . Thus, one could consider the jump-diffusion process instead to analyze and model these dynamics.

- (iii) In the third category, we look into the data of free jet low temperature helium turbulence [50] with Reynolds number  $Re = 757000$  and solar clear sky index [3]. The results of  $Q = Q(x = 0, \tau_{ds})$ ,  $\Theta = \Theta(x = 0, \tau_{ds})$  and  $\Lambda = \Lambda(x = 0, \tau_{ds})$  versus  $\tau_{ds}$  for the turbulence data are plotted in Fig. 5.20.



**Figure 5.20:**  $Q$ -,  $\Theta$ - and  $\Lambda$ -criterion of downsampled turbulence data are plotted in double logarithmic scale. The vertical blue lines show the jump time scale  $T_J = \frac{1}{\lambda}$  and the vertical red lines the correlation time scale  $T_C = \frac{1}{\gamma}$ . Standard errors are smaller than the symbol size.

The results of  $Q = Q(x = 0, \tau_{ds})$ ,  $\Theta = \Theta(x = 0, \tau_{ds})$  and  $\Lambda = \Lambda(x = 0, \tau_{ds})$  versus  $\tau_{ds}$  for the solar clear sky index are plotted in Fig. 5.21.



**Figure 5.21:**  $Q$ -,  $\Theta$ - and  $\Lambda$ -criterion of downsampled solar clear sky index data are plotted in double logarithmic scale. The vertical blue lines show the jump time scale  $T_J = \frac{1}{\lambda}$  and the vertical red lines the correlation time scale  $T_C = \frac{1}{\gamma}$ . Standard errors are smaller than the symbol size.

As  $\Theta$  and  $\Lambda$  are deviated from one, they do not fulfil the necessary conditions for the diffusion process. Similarly to previous data analysis, we evaluated again the correlation time  $T_C = \frac{1}{\gamma}$  where the drift function is defined by  $D^{(1)}(x) = -\gamma x$ . We also estimated  $T_J = \frac{1}{\lambda}$  by Eq. (2.36). In these two cases,  $T_J < T_C$ . For small  $\tau_{ds}$ , all  $Q$ ,  $\Theta$  and  $\Lambda$  show the functional behavior on  $\tau$ . Thus, one could only conclude that they are not diffusive from the current study. Turbulence is known to

have much more complex nature and it is also not Markovian in time but in scale [23, 70].

In this study, we analyzed mainly on pure diffusion process and jump-diffusion process with Poisson jumps and Gaussian distributed jump sizes. In order to have a broader approach on the real world data, similar rigorous methods should be studied on various known stochastic processes as the future research topics.





# Chapter 6

## Conclusion and Outlooks

The observed fluctuating behavior in real world complex systems may be dissociated into deterministic and stochastic components [23, 81]. The stochastic term can further be separated into Brownian (Wiener process with continuous trajectories) and/or jumps noises (e.g. Poisson jumps, Lévy noise, etc., which produce discontinuous dynamics).

In this thesis, jump-diffusion stochastic dynamics which occur in many real world complex systems are addressed. For discontinuous stochastic jump term, Poisson jumps and Gaussian distributed jump sizes are considered. With these assumptions, the basic analytical results for simple example such as Ornstein-Uhlenbeck process with additional compound Poisson jump term can be derived. The applications in the field of snow physics and the wind energy systems are then introduced. In these applications, the novel concept of the contribution of continuous Brownian and discontinuous jump noise is studied. It can be described by means of the ratio of diffusion and jump parameters,  $\frac{D^{(2)}}{\lambda\sigma_\xi^2}$ , which could give the better insight and understanding in studying the phenomena of continuous and discontinuous stochastic nature in the complex dynamical systems. In our work of snow physics, it can be used to distinguish different snow types. In wind energy systems, it can be an important indicator of different operational regions of a wind turbine.

After investigating the jump-diffusion nature present in these two physical systems, more rigorous analysis and characterization of the methods

are addressed. The study of [45] is extended and the criteria to distinguish continuous and discontinuous processes are introduced and their behaviors at different time scales are discussed. In the most modeling of dynamics of measurements in nature, it is assumed that standard Gaussian white noise-driven Langevin equation, (simply Langevin equation) is sufficient to capture the observed variability. A priori, it is not evident that such assumption is valid.

The analysis of simulating or integrating the stochastic processes at different sampling times are performed. However, for measurements with finite data points, the only plausible way to check this assumption is to use the downsampling the data. Therefore, downsampling of the simulated or empirical data sampled at fixed sampling time with respect to different time scales are determined. In linear jump-diffusion process, there are two unique time scales namely, the correlation time  $T_C = \frac{1}{\gamma}$ , where the linear drift,  $D^{(1)}(x) = -\gamma x$ , describes how fast fluctuations relax to the local mean values of  $x$ , and the average time scale between jumps  $T_J = \frac{1}{\lambda}$ , where  $\lambda$  is the jump rate of Poisson process.

Here, a set of criteria are derived to rule-out that the data can be model with a Langevin equation. which is the main assumptions in the modeling of many phenomena in biology, ecology, complex systems, rheology, neuroscience, Darwin evolution, cosmology, etc. Ruling out the Langevin dynamics opens the possibilities to use other modelings, such as jump-diffusion stochastic dynamics, the generalized Langevin equations, etc.

New aspects of data that are not assimilated from white-noise driven Langevin equations may indicate the presence of abrupt changes in the time series, in which there will be non-trivial physical phenomena that are not able to capture from Langevin equations. This study will pave the way for future research on new phenomena and help us better understand the measured variability of any data set.

---

## References

- [1] L. J. Allen. A primer on stochastic epidemic models: Formulation, numerical simulation, and analysis. *Infectious Disease Modelling*, 2(2):128–142, 2017.
- [2] E. Anahua, S. Barth, and J. Peinke. Markovian power curves for wind turbines. *Wind Energy*, 11(3):219–232, 2008.
- [3] M. Anvari, G. Lohmann, M. Wächter, P. Milan, E. Lorenz, D. Heine-  
mann, M. R. R. Tabar, and J. Peinke. Short term fluctua-  
tions of wind and solar power systems. *New Journal of Physics*,  
18(6):063027, Jun 2016.
- [4] M. Anvari, G. Lohmann, M. Wächter, P. Milan, E. Lorenz, D. Heine-  
mann, M. R. R. Tabar, and J. Peinke. Short term fluctua-  
tions of wind and solar power systems. *New Journal of Physics*,  
18(6):063027, Jun 2016.
- [5] M. Anvari, M. Rahimi Tabar, J. Peinke, and K. Lehnertz. Disen-  
tangling the stochastic behaviour of complex time series. *Scientific  
Reports*, 6:35435, 2016.
- [6] M. Anvari, B. Werther, G. Lohmann, M. Wächter, J. Peinke, and  
H.-P. Beck. Suppressing power output fluctuations of photovoltaic  
power plants. *Solar Energy*, 157:735 – 743, 2017.
- [7] L. Bachelier. Théorie de la spéculation. *Annales scientifiques de  
l'École Normale Supérieure*, 3e série, 17:21–86, 1900.
- [8] L. Bachelier. *Théorie de la spéculation*. Les Grands Classiques  
Gauthier-Villars. [Gauthier-Villars Great Classics]. Éditions Jacques  
Gabay, Sceaux, 1995. Théorie mathématique du jeu. [Mathematical  
theory of games], Reprint of the 1900 original.
- [9] H. Bader and P. Niggli. *Der Schnee und seine Metamorphose:  
Erste Ergebnisse und Anwendungen einer systematischen Unter-  
suchung der alpinen Winterschneedecke. Durchgeführt von der  
Station Weissfluhjoch-Davos der Schweiz. Schnee- und Lawinen-  
forschungskommission 1934-1938*. Kümmerly and Frey, 1939.

- [10] A. Bahraminasab, F. Ghasemi, A. Stefanovska, P. V. E. McClintock, and R. Friedrich. Physics of brain dynamics: Fokker-Planck analysis reveals changes in EEG  $\delta$ – $\theta$  interactions in anaesthesia. *New Journal of Physics*, 11(10):103051, Oct 2009.
- [11] F. M. Bandi and T. H. Nguyen. On the functional estimation of jump-diffusion models. *Journal of Econometrics*, 116(1-2):293–328, 2003.
- [12] F. Black and M. Scholes. The Pricing of Options and Corporate Liabilities. *Journal of Political Economy*, 81(3):637–654, 1973.
- [13] R. Brown. XXVII. A brief account of microscopical observations made in the months of June, July and August 1827, on the particles contained in the pollen of plants; and on the general existence of active molecules in organic and inorganic bodies. *The Philosophical Magazine*, 4(21):161–173, 1828.
- [14] R. Calif and F. G. Schmitt. Multiscaling and joint multiscaling description of the atmospheric wind speed and the aggregate power output from a wind farm. *Nonlinear Processes in Geophysics*, 21(2):379–392, 2014.
- [15] N. Calonne, B. Richter, H. Löwe, C. Cetti, J. Ter Schure, A. Van Herwijnen, C. Fierz, M. Jaggi, and M. Schneebeli. The RHOSSA campaign: Multi-resolution monitoring of the seasonal evolution of the structure and mechanical stability of an alpine snowpack. *The Cryosphere*, 14(6):1829–1848, 2020.
- [16] C. T. Chudley and R. J. Elliott. Neutron Scattering from a Liquid on a Jump Diffusion Model. *Proceedings of the Physical Society*, 77(2):353–361, Feb 1961.
- [17] S. Colbeck, E. Akitaya, R. Armstrong, H. Gubler, J. Lafeuille, K. Lied, D. Mcclung, and E. Morris. The International Classification for Seasonal Snow on the Ground. *The International Commission on Snow and Ice of the International Association of Scientific Hydrology*, 1990.

- 
- [18] J.-M. Courtault, Y. Kabanov, B. Bru, P. Crépel, I. Lebon, and A. Le Marchand. Louis Bachelier on the centenary of “Théorie de la spéculation”. *Math. Finance*, 10(3):341–353, 2000.
- [19] A. Einstein. Investigations on the theory of the brownian movement. *Ann. der Physik*, 1905.
- [20] Z. El kharrazi and S. Saoud. Simulation of COVID-19 epidemic spread using Stochastic Differential Equations with Jump diffusion for SIR Model. In *2021 7th International Conference on Optimization and Applications (ICOA)*, pages 1–4, 2021.
- [21] C. Fierz, R. L. Armstrong, Y. Durand, P. Etchevers, E. Greene, D. M. McClung, K. Nishimura, P. Satyawali, and S. A. Sokratov. *The international classification for seasonal snow on the ground (UNESCO, IHP (International Hydrological Programme)–VII, Technical Documents in Hydrology, No 83; IACS (International Association of Cryospheric Sciences) contribution No 1)*. Jan 2009.
- [22] R. Friedrich, J. Peinke, and C. Renner. How to Quantify Deterministic and Random Influences on the Statistics of the Foreign Exchange Market. *Phys. Rev. Lett.*, 84:5224–5227, May 2000.
- [23] R. Friedrich, J. Peinke, M. Sahimi, and M. Reza Rahimi Tabar. Approaching complexity by stochastic methods: From biological systems to turbulence. *Physics Reports*, 506(5):87–162, 2011.
- [24] C. Gardiner. *Handbook of Stochastic Methods for Physics, Chemistry, and the Natural Sciences*. Proceedings in Life Sciences. Springer-Verlag, 1985.
- [25] F. Ghasemi, M. Sahimi, J. Peinke, R. Friedrich, G. R. Jafari, and M. R. R. Tabar. Markov analysis and Kramers-Moyal expansion of nonstationary stochastic processes with application to the fluctuations in the oil price. *Phys. Rev. E*, 75:060102, Jun 2007.
- [26] F. Ghasemi, M. Sahimi, J. Peinke, and M. R. R. Tabar. Analysis of Non-stationary Data for Heart-rate Fluctuations in Terms of Drift

- and Diffusion Coefficients. *Journal of Biological Physics*, 32(2):117–128, Aug 2006.
- [27] S. Ghashghaie, W. Breymann, J. Peinke, P. Talkner, and Y. Dodge. Turbulent cascades in foreign exchange markets. *Nature*, 381(6585):767–770, Jun 1996.
- [28] P. Glasserman. *Monte Carlo Methods in Financial Engineering*. Springer, New York, NY, USA, 2004.
- [29] J. Gottschall and J. Peinke. Stochastic modelling of a wind turbine's power output with special respect to turbulent dynamics. *Journal of Physics: Conference Series*, 75:012045, Jul 2007.
- [30] J. Gottschall and J. Peinke. How to improve the estimation of power curves for wind turbines. *Environmental Research Letters*, 3(1):015005, Jan 2008.
- [31] H. Haehne, J. Schottler, M. Waechter, J. Peinke, and O. Kamps. The footprint of atmospheric turbulence in power grid frequency measurements. *Europhysics Letters*, 121(3):30001, Mar 2018.
- [32] P. Hagenmuller and T. Pilloix. A New Method for Comparing and Matching Snow Profiles, Application for Profiles Measured by Penetrometers. *Frontiers in Earth Science*, 4(52), May 2016.
- [33] F. Hanson. *Applied Stochastic Processes and Control for Jump Diffusions: Modeling, Analysis, and Computation*. Advances in Design and Control. Society for Industrial and Applied Mathematics, 2007.
- [34] F. B. Hanson. Stochastic Processes and Control for Jump-Diffusions. *SSRN Electronic Journal*, 2007.
- [35] A. M. Hashtroud, E. Mirzahassein, F. Zarei, and M. R. R. Tabar. Jump events in the human heartbeat interval fluctuations. *Journal of Statistical Mechanics: Theory and Experiment*, 2019(8):083213, Aug 2019.
- [36] S. L. Heston. A Closed-Form Solution for Options with Stochastic Volatility with Applications to Bond and Currency Options. *Review of Financial Studies*, 6(2):327–43, 1993.

- 
- [37] P. Huber and E. Ronchetti. *Robust Statistics*. Wiley Series in Probability and Statistics. Wiley, 2011.
- [38] IEC. Wind Turbines Part 1: Design requirements. IEC 61400-1:2005 International Standard, 2005. Technical report, IEC, 2005.
- [39] L. Isserlis. On certain probable errors and correlation coefficients of multiple frequency distributions with skew regression. *Biometrika*, 11(3):185–190, May 1916.
- [40] K. Itô. Stochastic integral. *Proceedings of the Imperial Academy*, 20(8):519 – 524, 1944.
- [41] R. Jarrow and P. Protter. A Short History of Stochastic Integration and Mathematical Finance: The Early Years, 1880-1970. *Lecture Notes-Monograph Series*, 45:75–91, 2004.
- [42] M. Johannes. The Statistical and Economic Role of Jumps in Continuous-Time Interest Rate Models. *The Journal of Finance*, 59(1):227–260, 2004.
- [43] J. B. Johnson and M. Schneebeli. Characterizing the microstructural and micromechanical properties of snow. *Cold Regions Science and Technology*, 30(1):91 – 100, 1999.
- [44] D. Lamouroux and K. Lehnertz. Kernel-based regression of drift and diffusion coefficients of stochastic processes. *Physics Letters A*, 373(39):3507 – 3512, 2009.
- [45] K. Lehnertz, L. Zabawa, and R. Tabar. Characterizing abrupt transitions in stochastic dynamics. *New Journal of Physics*, 20, Nov 2018.
- [46] D. S. Lemons and A. Gythiel. Paul Langevin’s 1908 paper On the Theory of Brownian Motion [Sur la théorie du mouvement brownien, Comptes-rendus de l’Académie des Sciences (Paris) 146, 530-533 (1908)]. *American Journal of Physics*, 65(11):1079–1081, Nov. 1997.
- [47] P. P. Lin, I. Peinke, P. Hagenmuller, M. Wächter, M. R. Rahimi Tabar, and J. Peinke. Stochastic analysis of micro-cone penetration tests in snow. *The Cryosphere*, 16(12):4811–4822, 2022.

- [48] P. G. Lind, I. Herráez, M. Wächter, and J. Peinke. Fatigue Load Estimation through a Simple Stochastic Model. *Energies*, 7(12):8279–8293, 2014.
- [49] H. Löwe and A. van Herwijnen. A Poisson shot noise model for micro-penetration of snow. *Cold Regions Science and Technology*, 70:62–70, 2012.
- [50] P. Manshour, M. R. Rahimi Tabar, and J. Peinke. Fully developed turbulence in the view of horizontal visibility graphs. *Journal of Statistical Mechanics: Theory and Experiment*, 2015(8):P08031, Aug 2015.
- [51] P. Manshour, S. Saberi, M. Sahimi, J. Peinke, A. F. Pacheco, and M. R. Rahimi Tabar. Turbulencelike Behavior of Seismic Time Series. *Phys. Rev. Lett.*, 102:014101, Jan 2009.
- [52] H. Marshall and J. Johnson. Accurate Inversion of High-Resolution Snow Penetrometer Signals for Microstructural and Micromechanical Properties. *Journal of Geophysical Research*, 114, 2009.
- [53] A. McCallum. Cone Penetration Testing (CPT): a valuable tool for investigating polar snow. *Journal of Hydrology (New Zealand)*, 52(2):97–113, 2013.
- [54] R. C. Merton. Option pricing when underlying stock returns are discontinuous. *Journal of Financial Economics*, 3(1):125–144, 1976.
- [55] P. Milan, T. Mücke, A. Morales, M. Wächter, and J. Peinke. Applications of the Langevin Power Curve. In *Proceedings of EWEC 2010*, Warsaw, 2010.
- [56] P. Milan, M. Wächter, and J. Peinke. Turbulent Character of Wind Energy. *Phys. Rev. Lett.*, 110:138701, Mar 2013.
- [57] P. Milan, M. Wächter, and J. Peinke. Stochastic modeling and performance monitoring of wind farm power production. *Journal of Renewable and Sustainable Energy*, 6(3):033119, 2014.



- 
- [58] S. M. Mousavi, S. N. S. Reihani, G. Anvari, M. Anvari, H. G. Alinezhad, and M. R. R. Tabar. Stochastic analysis of time series for the spatial positions of particles trapped in optical tweezers. *Scientific Reports*, 7(1):4832, Jul 2017.
- [59] J.-F. Muzy, R. Baïle, and P. Poggi. Intermittency of surface-layer wind velocity series in the mesoscale range. *Phys. Rev. E*, 81:056308, May 2010.
- [60] T. Mücke, D. Kleinhans, and J. Peinke. Atmospheric turbulence and its influence on the alternating loads on wind turbines. *Wind Energy*, 14(2):301–316, 2011.
- [61] E. A. Nadaraya. On Estimating Regression. *Theory of Probability & Its Applications*, 9(1):141–142, 1964.
- [62] R. F. Pawula. Approximation of the Linear Boltzmann Equation by the Fokker-Planck Equation. *Phys. Rev.*, 162:186–188, Oct 1967.
- [63] I. Peinke, P. Hagenmuller, E. Andò, G. Chambon, F. Flin, and J. Roule. Experimental Study of Cone Penetration in Snow Using X-Ray Tomography. *Frontiers in Earth Science*, 8:63, Mar. 2020.
- [64] I. Peinke, P. Hagenmuller, G. Chambon, and J. Roule. Investigation of snow sintering at microstructural scale from micro-penetration tests. *Cold Regions Science and Technology*, 162:43–55, 2019.
- [65] W. Press, S. Teukolsky, W. Vetterling, and B. Flannery. *Numerical Recipes: The Art of Scientific Computing*. Cambridge University Press, 3 edition, 2007.
- [66] N. Privault. *Stochastic Finance: An Introduction with Market Examples*. Chapman and Hall/CRC Financial Mathematics Series. Taylor & Francis, 2013.
- [67] M. Proksch, H. Löwe, and M. Schneebeli. Density, specific surface area, and correlation length of snow measured by high-resolution penetrometry. *Journal of Geophysical Research: Earth Surface*, 120(2):346–362, 2015.

- [68] J. Prusseit and K. Lehnertz. Stochastic Qualifiers of Epileptic Brain Dynamics. *Phys. Rev. Lett.*, 98:138103, Mar 2007.
- [69] J. Prusseit and K. Lehnertz. Measuring interdependences in dissipative dynamical systems with estimated Fokker-Planck coefficients. *Phys. Rev. E*, 77:041914, Apr 2008.
- [70] C. Renner, J. Peinke, and R. Friedrich. Experimental Indications for Markov Properties of Small-scale Turbulence. *Journal of Fluid Mechanics*, 433:383–409, 2001.
- [71] B. Reuter, M. Proksch, H. Löwe, A. van Herwijnen, and J. Schweizer. Comparing measurements of snow mechanical properties relevant for slab avalanche release. *Journal of Glaciology*, 65(249):55–67, Feb 2019.
- [72] B. Reuter, J. Schweizer, and A. van Herwijnen. A process-based approach to estimate point snow instability. *The Cryosphere*, 9(3):837–847, May 2015.
- [73] P. Rinn, P. G. Lind, M. Wächter, and J. Peinke. The Langevin Approach: An R Package for Modeling Markov Processes. *Journal of Open Research Software*, 4(1), Aug 2016.
- [74] H. Risken. *Fokker-Planck Equation*, pages 63–95. Springer Berlin Heidelberg, Berlin, Heidelberg, 1984.
- [75] H. Risken and T. Frank. *The Fokker-Planck Equation: Methods of Solution and Applications*. Springer Series in Synergetics. Springer Berlin Heidelberg, 1996.
- [76] K. Schmietendorf, J. Peinke, and O. Kamps. The impact of turbulent renewable energy production on power grid stability and quality. *The European Physical Journal B*, 90(11), Nov. 2017.
- [77] M. Schneebeli, C. Pielmeier, and J. B. Johnson. Measuring snow microstructure and hardness using a high resolution penetrometer. *Cold Regions Science and Technology*, 30(1):101 – 114, 1999.

- 
- [78] J. Schweizer, B. Jamieson, and M. Schneebeli. Snow avalanche formation. *Reviews of Geophysics*, 41:1016, Dec 2003.
- [79] R. Stanton. A Nonparametric Model of Term Structure Dynamics and the Market Price of Interest Rate Risk. *The Journal of Finance*, 52(5):1973–2002, 1997.
- [80] R. Stratonovich, R. McDonough, N. McDonough, and R. Bellman. *Conditional Markov Processes and Their Application to the Theory of Optimal Control*. Modern analytic and computational methods in science and mathematics. American Elsevier Publishing Company, 1968.
- [81] M. R. R. Tabar. *Analysis and Data-Based Reconstruction of Complex Nonlinear Dynamical Systems: Using the Methods of Stochastic Processes*. Springer, Cham-Switzerland, 2019.
- [82] P. Tankov. *Financial Modelling with Jump Processes*. Chapman and Hall/CRC Financial Mathematics Series. CRC Press, 2003.
- [83] M. Wächter, H. Heißelmann, M. Hölling, A. Morales, P. Milan, T. Mücke, J. Peinke, N. Reinke, and P. Rinn. The turbulent nature of the atmospheric boundary layer and its impact on the wind energy conversion process. *Journal of Turbulence*, 13:N26, 2012.
- [84] M. Wächter, P. Milan, T. Mücke, and J. Peinke. Power Performance of Wind Energy Converters Characterized as Stochastic Process: Applications of the Langevin Power Curve. *Wind Energy*, 14(6):711–717, 2011.
- [85] G. S. Watson. Smooth Regression Analysis. *Sankhyā: The Indian Journal of Statistics, Series A (1961-2002)*, 26(4):359–372, 1964.
- [86] G. C. Wick. The Evaluation of the Collision Matrix. *Phys. Rev.*, 80:268–272, Oct 1950.
- [87] WindEurope. Wind energy in Europe: 2020 Statistics and the outlook for 2021-2025. <https://windeurope.org/intelligence-platform/product/wind-energy-in-europe-2020-statistics-and-the-outlook-for-2021-2025/>, 2021.



# Curriculum

## Personal Details

Name	Pyei Phyo Lin
Date of Birth	10.08.1989
Place of Birth	Yangon, Myanmar
Nationality	Myanmar

## Tertiary Education

04.2007–03.2010	Diploma <sup>1</sup> in Electrical and Electronic Engineering, Singapore Polytechnic, Singapore
04.2011–12.2013	B.Eng. Engineering Physics, University of Oldenburg
10.2013–03.2016	M.Sc. Engineering Physics, University of Oldenburg

## Work and Scientific Experiences

05.2010–12.2010	Engineering Assistant, Creative Technology Ltd., Singapore
11.2012–02.2016	Tutor, Institute of Physics, University of Oldenburg
07.2013–10.2013	Intern, Fraunhofer – Institute for Wind Energy and Energy System Technology, Bremerhaven, Germany
12.2013–08.2014	Scientific Assistant, Signal Processing Group, Department of Medical Physics and Acoustics, University of Oldenburg

---

<sup>1</sup>Diploma is not equivalent to a German Diplom. It is a qualification of the polytechnics in Singapore, in which the students can enrol after 10 years in school. After completion of the diploma, there is an opportunity to enrol in a higher semester at a university.

## Curriculum

---

- 02.2016–04.2017 Scientific Assistant, Turbulence and Wind Energy and Stochastics Group, ForWind, Institute of Physics, University of Oldenburg
- 09.2016–04.2023 Doctoral Candidate, Turbulence and Wind Energy and Stochastics Group, ForWind, Institute of Physics, University of Oldenburg
- 04.2017–10.2020 Research Fellow, Turbulence and Wind Energy and Stochastics Group, ForWind, Institute of Physics, University of Oldenburg  
**Project:** OWP-Control
- 11.2020–11.2021 Research Fellow, Turbulence and Wind Energy and Stochastics Group, ForWind, Institute of Physics, University of Oldenburg  
**Project:** WEA-Doktor

## List of Publications

P. P. LIN, I. PEINKE, P. HAGENMULLER, M. WÄCHTER,  
M. R. R. TABAR, J. PEINKE:  
Stochastic Analysis of Micro-cone Penetration Tests In Snow,  
The Cryosphere , 16, 4811–4822, 2022.

PPL did the preliminary analysis of the data and simulations, and wrote the main part of the manuscript. PH and IP performed the experiments. JP had the initial idea, and JP, MW and MRRT supervised the work. All the authors interpreted the results, and helped in preparing and editing the manuscript.

P. P. LIN, M. WÄCHTER, M. R. R. TABAR, J. PEINKE:  
Discontinuous Jump Behavior of the Energy Conversion  
in Wind Energy Systems, arXiv preprint arXiv:2301.05553,  
submitted to PRX Energy, 2023.

PPL did the preliminary analysis of the data, and wrote the main part of the manuscript. JP had the initial idea, and JP, MW and MRRT supervised the work. All the authors interpreted the results, and helped in preparing and editing the manuscript.





## Acknowledgement

First of all, I would like to express my sincere gratitude to my supervisors, Prof. Dr. Joachim Peinke, Dr. Matthias Wächter and Prof. Dr. M. Reza Rahimi Tabar, for the insightful comments, encouragements and the guidances which motivated me to widen my research from various perspectives.

My sincere thanks also go to the examiners Prof. Dr. Pedro G. Lind, also for taking the time to review my thesis and Prof. Dr. Kerstin Avila for evaluating the disputation. I am also very grateful to Prof. Dr. Martin Kühn who supported me in the involvement of the research project.

I also would like to thank the coauthors of the publication, Isabel Peinke and Pascal Hagenmuller, as well as Abdulkarim Abdulrazek, Andreas Rott, André Fuchs, Cherif Mihoubi, Christian Behnken, Christian Philipp, Daniela Moreno, Gabriele Centurelli, Gerd Gülker, Hauke Hähne, Hendrik Heißelmann, Ingrid Neunaber, Jarek Puczyłowski, Khaled Yassin, Lars Neuhaus, Lisa Rademacher, Marcel Bock, Martin Wagner, Mehrnaz Anvari, Michael Hölling, Michael Schwarz, Patrick Milan, Philip Rinn, Piyush Singh, Sebastian Ehrich, Shumian Zhao, Thomas Messmer, Tom Wester, Vlaho Petrović, Zeinab Gharibi and all other members and former members of ForWind Oldenburg groups for the helps, supports and helpful discussions.

I would like to express my special thanks to my parents who supported me financially as well as spiritually throughout my education. Besides them, I appreciate all my teachers, tutors, lecturers, supervisors and professors who taught and guided me throughout my whole learning process. Last but not least, I would like to thank my sister, brother, friends and colleagues who cheered me up while working on this thesis.



## Erklärung

Ich erkläre hiermit, dass

- eine Promotion zum Dr. rer. nat. angestrebt wird,
- ich diese Arbeit selbstständig verfasst und nur die angegebenen Hilfsmittel benutzt habe, und dass die Arbeit in Teilen bereits veröffentlicht bzw. zur Veröffentlichung eingereicht wurde,
- diese Dissertation weder in ihrer Gesamtheit noch in Teilen einer anderen Hochschule zur Begutachtung in einem Promotionsverfahren vorgelegen hat.

Oldenburg, den 11. Januar 2023

.....

(Pyei Phyo Lin)