UNIVERSITÀ DEGLI STUDI DI PADOVA

Dipartimento di Fisica e Astronomia “Galileo Galilei”

Master Degree in Physics

Final Dissertation

Self-Similarity and Large Deviations in Two Subordinated Stochastic Processes

Thesis supervisor	Candidate
Prof. Fulvio Baldovin	Anna Guerra

Anno Accademico 2023/2024

Contents

1 On Stochastic Processes **7**
- 1.1 Brownian motion . . . 7
- 1.2 Langevin and Fokker-Planck approach . . . 9
  - 1.2.1 Langevin Equation . . . 9
  - 1.2.2 Fokker-Plank equation . . . 10
- 1.3 Markov Chains . . . 10
  - 1.3.1 Discrete Time . . . 10
  - 1.3.2 Continuous time . . . 12
- 1.4 Characterization of Stochastic Processes . . . 13
  - 1.4.1 Moments . . . 13
  - 1.4.2 Generating functions . . . 15
  - 1.4.3 The Gaussian-Brownian case . . . 16
  - 1.4.4 Anomalous Diffusion . . . 17
- 1.5 Anomalous Scaling and Self-Similarity . . . 17

2 Large Deviation Theory **19**
- 2.1 The large deviation principle . . . 19
  - 2.1.1 Random Walk with Gaussian steps . . . 20
- 2.2 Gärtner-Ellis theorem . . . 20

3 Stochastic diffusivity **22**
- 3.1 Diffusivity Subordinator Processes . . . 22
- 3.2 Diffusing Diffusivity . . . 24
  - 3.2.1 Langevin approach . . . 24
  - 3.2.2 Fokker-Planck approach . . . 26
  - 3.2.3 Cumulants . . . 29
  - 3.2.4 Simulation . . . 33
- 3.3 Two-State Polymer Model . . . 35
  - 3.3.1 Markov Dynamics . . . 35
  - 3.3.2 Cumulants . . . 38
- 3.4 Simulation . . . 41
- 3.5 Discussion . . . 42
  - 3.5.1 Brownian diffusion . . . 43
  - 3.5.2 Long Time Regime and Large Deviations . . . 43
  - 3.5.3 Short Time and Anomalous Scaling . . . 45
  - 3.5.4 Anomalous Scaling and Large Deviation . . . 45

Appendices **49**

A Chapter 1 **50**
- A.1 Einstein Diffusion Coefficient Derivation . . . 50
- A.2 Smoluchowski Diffusion Coefficient Derivation . . . 51

B Chapter 3 **52**
B.1 Ornstein-Uhlenbeck moments with Fokker-Pank approach 52
B.2 $S(t)$ n-th moment for large times 53

Introduction

Bruce has just moved from his hometown to a city where his father found an opportunity for a good job and today is the first day at the new school. He does not know anyone and when the school break starts, he feels a little worried about how to make new friends. At the beginning he does not move much feeling like stopped by an introvert force, but after Tom comes to introduce himself, Bruce starts to become more confident. After some minutes he gets to know Nick, then Leo and then Steve and by the end of the break Bruce walks around the yard happily talking just like all the other integrated students. At the end he only needed a little time to relax to settle in and uniform.
Who is always talkative is Barbara, especially with her friend Sarah who is now on a month holiday. While at the park, Barbara decides to call her to tell all the things she missed out in the time they were not together. Anyway also Sarah has a lot of things to say about her vacation and the two starts a lively conversation and Barbara, as a lot of people do when at the phone, walks around randomly. Anyway her walking velocity changes depending on whether she is listening or talking to her friend. After two hours of phone call, Barbara has done few kilometers but all the most interesting things to say were finished, so they talk calmly. This type of small talks relaxes Barbara and her velocity becomes constant but still with random direction.
Bruce situation has a really different nature with respect to Barbara one, yet an observer who look at them at the beginning or at the end of the interactions could say that they have a lot in common.

As one can intuitively deduce, stochastic diffusivity models are processes for which the diffusion coefficient has a random nature and they are used to describe situations in which one or more assumed condition for the Gaussian-Brownian motion is not satisfied. In this thesis we are going to analyzed two of those processes: the diffusing diffusivity model which looks at the diffusion coefficient as an Ornstein–Uhlenbeck process and the two-states polymer model, that exploits the telegraph model (which is essentially a Markov chain with two possible states) for the diffusion coefficient randomness. Those models are used to describe situations where the microscopical conditions are responsible for the aleatory origin of the diffusion coefficient. In particular the diffusing diffusivity model is related to a dispersion in a not homogeneous medium and the two-state polymer model is associate to a change in time of the diffusive molecules shape.
The presence of another source of randomness calls for the use of particular methods which has to be exploited in order to study those models and indeed we will rely on the properties of subordinated processes, which consider stochastic variable as composition of other random quantity. This technique usually does not permit to obtain for all times the probability or the equation describing the random trajectory as for easier or well-known models, but yet we can gain information by means of other quantities which still characterize uniquely the process and help enlightening its features. In particular the property we are going to look at are Brownian diffusion and the similarity with the Gaussian propagator and for doing it we will focus on the moments which describe how the process diffuses with time. The variance indicates the region in which the dispersion takes place and information about the probability structure, so about the range more favorable for finding the particles, can be obtained from the kurtosis which indicates the similarity with the Gaussian propagator shape. In the case of variance behaving like a Gaussian one, which regards our interested models, the kurtosis is helpful

because it reveals if the analyzed process probability has also a Gaussian form or not. If this does not happen we talk about anomalous scaling which is interesting especially for its relation with self-similarity. The latter confers a strong correspondence between the process and its scaling and affects also the moment structure. On the other hand, the case in which both variance and kurtosis indicate a Gaussian-Brownian diffusion is associated to a regime of large deviation, which regards probability distributions showing an exponential decay. In our interested cases the large deviation is related to the long time regime and the time is indeed the parameter ruling the exponential. Large deviation theory is also useful in virtue of its principles, which can be exploited to obtain useful characteristics of the process.
The fact, as we will see, that the two analyzed models display commons behaviors in the short and long time regimes respectively related to anomalous scaling and large deviation, suggests that they are sharing some properties which are not visible at first sight and so it is interesting to see which are those peculiarities and where they originated from.

Chapter 1

On Stochastic Processes

This chapter aims to provide a general illustration for stochastic processes focusing on the arguments that will be useful for the analysis of the stochastic diffusivity models in Chapter 3. In particular we are going to look at the "father" of stochastic process that is the Gaussian-Brownian motion which will be then a reference for confronting the behaviors and trends in Chapter 3. We will then illustrate the tools which will be needed to analyze the random diffusion coefficient of the diffusing diffusivity model and the two-state polymer model and then there will be introduced the generating functions and moments, which are helpful for complex structure stochastic models. At the end of this chapter there will be described the anomalous scaling, which characterize a type of diffusion which will be encountered.

1.1 Brownian motion

Whoever has to deal with stochastic calculus must be confident with the concept of Brownian motion which plays indeed a central role in probability theory. Nowadays applied in multiple fields, this motion origins from the description of the movement of a particle in a fluid and is named after Robert Brown, a Scottish botanist who first in 1827 reported the observation of particles of 10^{-3} cm size moving in water while studying plants pollen [10]. Even if Brown was not able to justify the phenomenon, his discovery had a significant relevance since for the first time random motion was attributed to non-living particles. The explanation of this motion had to wait 80 years, when Einstein in one of the Annus Mirabilis Papers [7] found the cause in the collision between analyzed molecules, called *Brownian particles*, and the surrounding fluid ones. Given the large mass difference between the two molecules types, the single hit cannot affect the trajectory of the bigger particle (the pollen in the Brown case) but the collision frequency is too high to be neglected and is indeed the origin of the motion. Anyway a deterministic treatment is not possible since it would require to consider a number of equation of the order of the Avogadro number, so in order to study the system Einstein relied on a probabilistic approach. Exploiting the number of particles conservation, he was able to derive the expression [7], also known as the *Fick's second law*, for the Brownian particles probability density function $p(\vec{x}, t)$:

$$\frac{\partial p(\vec{x}, t)}{\partial t} = D \Delta p(\vec{x}, t), \tag{1.1.1}$$

where Δ is the Laplacian and D quantify the amount of dispersion and is called *diffusion coefficient.* The solution of the previous equation for natural boundary conditions results to be a normal distribution centered in $\vec{x}_0$ and is known as *Gaussian propagator*:

$$p(\vec{x}, t|\vec{x}_0, 0) = \frac{1}{(4\pi Dt)^{\frac{d}{2}}} \exp\left\{-\frac{(\vec{x}-\vec{x}_0)^2}{4Dt}\right\}, \tag{1.1.2}$$

where d is the spatial dimension in which the diffusion is taking place. If the diffusion coefficient does not depend on the coordinates, then the mean and variance can be easily recognized from the normal distribution structure of (1.1.2):

$$\langle \vec{x}(t) \rangle = \vec{x}_0, \qquad \langle (\vec{x}(t) - \vec{x}_0)^2 \rangle = 2dDt. \tag{1.1.3}$$

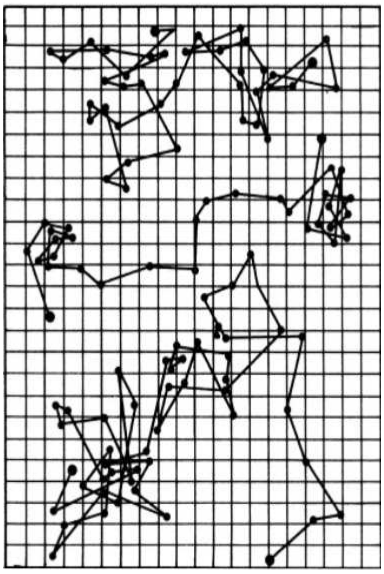

Figure 1.1: Brownian particle motion observed by Perrin [36] where the dots corresponds to the particle position after $30s$

The previous results can be interpreted supposing that a given amount of macromolecules are released in a fluid, then the concentration will always be higher in the initial deposition point while the particles spread throughout the medium will increase linearly in time and proportionally to the diffusion coefficient. The latter is the focus of the other enlightening part of Einstein's paper where he pointed out the relation between D and the temperature T of the region where the analyzed particle is moving [7]:

$$D = \mu k_B T, \tag{1.1.4}$$

with μ indicating the particle mobility and k_B the Boltzmann constant. Eq. (1.1.4) is known *Einstein-Smoluchowski relation* since both the physicists[1] released a formula for the diffusion coefficient within almost the same year with two different approaches (illustrated in Appendix A.1). Einstein [7] followed a thermodynamic approach in the case of thermal equilibrium, while Smoluchowski arrived at the result considering the diffusing particles undergoing an unbiased random walk model [8] . Denoting the random walk length and time step respectively with l and Δt, the comparison between the two derivations [9] in one dimension relates the macroscopic quantities with microscopical characteristics

$$k_B T = \frac{l^2}{2\Delta t}. \tag{1.1.5}$$

It is worth to underline that both the approaches are valid under the assumption that all the collisions between the medium and the diffusive particle can be considered equal and do not change in time. Anyway there are situations in which those hypothesis underlying a constant diffusion coefficient are not satisfied anymore and this will be the case of the processes treated in Chapter 3.

[1] Smoluchowski was studying the Brownian motion in the same period Einstein was and indeed he also published his own derivation of the Fick's second law

1.2 Langevin and Fokker-Planck approach

We are now going to illustrate two of the most common and famous approaches used when treating stochastic processes, which are the *Langevin equation* [3] and the *Fokker-Plank equation* [3].
Those equations permit to enlarge the Brownian motion treated in the previous section to situations in which an external force is exerted on the stochastic diffusing particles and they will be exploited in Section 3.2 when treating the Ornstein–Uhlenbeck process, that concerns the case of an harmonic oscillator performing Brownian motion. In particular we will exploit the different perspectives the two equations provide, indeed through the Fokker-Planck approach the process analysis is based on its probability, while the Langevin equation sets the focus on the random variable trajectory.

1.2.1 Langevin Equation

Few years after the Einstein's revolutionary paper on the random motion, the French physicist Langevin proposed to describe the Brownian particle using a new approach which focuses no more on the probability as done until then, but on the particle trajectory treating its randomness through a stochastic differential equation.
Considering a particle immerse in a fluid, the force exerted on the particle can be split in two parts:

- the viscous drag acting in the opposite velocity $\vec{v}$ direction originating from the macroscopic hydrodynamics character
$$\vec{F}_d = -\frac{1}{\gamma}\vec{v}, \tag{1.2.1}$$
where γ is the friction
- a microscopic force $\vec{\eta}(t)$ due to the collisions between the analyzed particle and the fluid molecules.

The issue stands now in finding a proper formulation for $\vec{\eta}(t)$. For the very same reason mentioned in Section 1.1, the hits can not be treated in a deterministic way, so $\vec{\eta}(t)$ must have a stochastic nature conferring the random motion to the particle. Since the impacts has no preferred direction and it is reasonable to consider all of them acting rapidly with the same strength , $\vec{\eta}(t)$ is required to satisfy the following properties:

$$\begin{cases} \langle \eta_i(t) \rangle = 0 \\ \langle \eta_i(t)\eta_j(t') \rangle = 2D\delta_{ij}\delta(t-t') \end{cases} \tag{1.2.2}$$

where the mean is taken with respect to the $\vec{\eta}(t)$ realizations and D is the diffusion coefficient which is proportional to the strength of the collision force. The (1.2.2) relations bring us to identify

$$\vec{\eta}(t) = \sigma\frac{d\vec{W}(t)}{dt}, \tag{1.2.3}$$

where $W_i(t)$ is a Wiener process with $\langle dW_i^2(t) \rangle = dt^2$. The 2-nd Newton's law for the Brownian particle results then:

$$m\frac{d^2\vec{x}}{dt^2} = -\frac{1}{\gamma}\frac{d\vec{x}}{dt} + \sqrt{2D}\frac{d\vec{W}(t)}{dt}, \tag{1.2.4}$$

which divided by the mass m and written in terms of the velocity $\vec{v}$ returns the *Langevin equation* for the velocity:

$$\frac{d\vec{v}}{dt} = -\frac{1}{m\gamma}\vec{v} + \frac{\sqrt{2D}}{m}\frac{d\vec{W}(t)}{dt}. \tag{1.2.5}$$

In the low Reynolds number regime the ratio between the mass and the friction is negligible, hence when considering eq. (1.2.4) the inertial effect encoded in the acceleration term is vanishing leading the particle to have constant velocity. Including the possibility of having also an external force, eq. (1.2.5) becomes:

$$0 = -\frac{1}{\gamma m}\frac{d\vec{x}}{dt} + \vec{F}_{ext}(\vec{x})dt + \frac{\sqrt{2D}}{m}\frac{d\vec{W}(t)}{dt}, \tag{1.2.6}$$

where we returned to express the velocity as derivative of the position w.r.t. time. Eq. (1.2.6) can be rewritten as

$$d\vec{x} = m\gamma\vec{F}_{ext}(\vec{x})dt + \sqrt{2D}\gamma d\vec{W}(t), \tag{1.2.7}$$

which corresponds to the *Overdamped Langevin equation* [3].

1.2.2 Fokker-Plank equation

The Fokker-Plank equation describes the evolution in time of the probability density function particle and can be considered the extension of eq. (1.1.1) to the case where the particle is subjected to the action of external forces and has a constant velocity. The Fokker-Plank equation for the spatial coordinate [3] can be obtained from the overdamped Langevin equation (1.2.7) confirming the equivalence of the two approaches. In one dimension it takes the form:

$$\frac{\partial}{\partial t}p(x,t|x_0,0) = -\frac{\partial}{\partial x}\left(\frac{F(x)}{\gamma m}p(x,t|x_0,0)\right) + \frac{\sigma^2}{2\gamma^2 m^2}\frac{\partial^2}{\partial x^2}p(x,t|x_0,0), \tag{1.2.8}$$

where $F(x)$ is the external force responsible of the velocity drift and γ, m and σ correspond the same quantities used in previous section.

1.3 Markov Chains

In order to study the diffusion coefficient randomness of the two-state polymer model we are going to do a briefly recap on the Markov chains and their properties which will be successively useful.
Markov[2] introduced the so called *Markov chains* [4] [11] [12] in order to study processes whose future outcome depends only the present state and which take value over a numerable set of possible states. Given the probabilities of changing states and the initial distribution, Markov chains enable us to examine the probability evolution in time exploiting the factorization of the probability evolution into terms which relates the probability between two consecutive states.

1.3.1 Discrete Time

Consider a discrete Markov process $\{\omega_i\}_{i=0}^{n}$ indexed by i-th step in time which can take value over a set of l possible states $\Omega = \{s_i\}_{i=1}^{l}$. The probability of flipping or remaining in a state within a time step of duration Δt is called *transition probability*, indicated as $\pi(s_i|s_j)$ with s_i and s_j initial and final state respectively. After n time steps the probability that the process follows the sequence $\omega = \{\omega_n, \omega_{n-1}, \ldots, \omega_0\}$ is given by:

$$\mathbb{P}(\omega = \{\omega_n, \omega_{n-1}, \ldots, \omega_0\}) = \pi(\omega_n|\omega_{n-1})\pi(\omega_{n-1}|\omega_{n-2})\ldots\pi(\omega_1|\omega_0)\rho(\omega_0), \tag{1.3.1}$$

where $\rho(\omega_0)$ is the probability of starting with ω_0. Usually one is more interested on the final state than on the sequence, so if we now focus only on ω_n its probability is given summing (1.3.1) over all the possible paths (i.e. all the possible ω_j for $0 \leq j \leq n-1$)

[2] Andrey Markov (1856-1922)

$$\mathbb{P}(\omega_n|\rho(\omega_0)) = \sum_{\omega_n\in\Omega}\sum_{\omega_{n-1}\in\Omega}\cdots\sum_{\omega_1\in\Omega}\sum_{\omega_0\in\Omega}\pi(\omega_n|\omega_{n-1})\pi(\omega_{n-1}|\omega_{n-2})\ldots\pi(\omega_1|\omega_0)\rho(\omega_0). \tag{1.3.2}$$

If transition probabilities do not change with time, the process can be analyzed rewriting eq. (1.3.2) by means of matrices and bra-ket notation. More precisely, let us introduce the *stochastic matrix* Π whose entries Π_{ij} correspond to $\pi(s_i|s_j)$ (written as $\pi(i|j)$ for simplicity):

$$\Pi = \begin{bmatrix} \pi(1|1) & \pi(2|1) & \ldots & \pi(l|1) \\ \pi(1|2) & \pi(2|2) & \ldots & \pi(l|2) \\ \vdots & \vdots & \ddots & \vdots \\ \pi(1|l) & \pi(2|l) & \ldots & \pi(l|l) \end{bmatrix} = \Delta t \begin{bmatrix} r_{11} & r_{21} & \ldots & r_{l1} \\ r_{12} & r_{22} & \ldots & r_{l2} \\ \vdots & \vdots & \ddots & \vdots \\ r_{1l} & r_{2l} & \ldots & r_{ll} \end{bmatrix}, \tag{1.3.3}$$

where r_{ij} are the $\pi(i|j)$ *transition rate*. It can be proven [3] that because of probability conservation, the sum over the Π columns has to return 1:

$$\sum_{i=0}^{l} \pi(j|i) = 1 \qquad \forall j. \tag{1.3.4}$$

Let us now define the *probability ket* as:

$$|\rho_n\rangle = \begin{bmatrix} \rho_{n,1} \\ \vdots \\ \rho_{n,l} \end{bmatrix}, \tag{1.3.5}$$

with $\rho_{n,i}$ indicating the probability to find the process in state s_i after n steps. With the formalism just introduced the ket describing the probabilities after one step is obtained multiplying $|\rho_0\rangle$, i.e. the ket describing the initial distribution, by the stochastic matrix:

$$|\rho_1\rangle = \Pi\,|\rho_0\rangle\,. \tag{1.3.6}$$

Iterating the reasoning, the ket related to the process after $n\Delta t$ is obtained applying n times the stochastic matrix to $|\rho_0\rangle$ permitting to express (1.3.2) with ket notation:

$$|\rho_n\rangle = \Pi^n\,|\rho_0\rangle\,. \tag{1.3.7}$$

Notice that defining the basis $\{|s_i\rangle\}_{i=1}^{l}$ with

$$|s_1\rangle = \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \qquad |s_2\rangle = \begin{bmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{bmatrix}, \qquad \ldots, \qquad |s_l\rangle = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{bmatrix}, \tag{1.3.8}$$

each $\rho_{n,i}$ is the $|\rho_n\rangle$ projection on $|s_i\rangle$, hence corresponding to the contraction with the bra $\langle s_i|$.
Because of probability conservation, the sum over the $|\rho_n\rangle$ components has to be equal to 1 for all n and it is translated to the bra-ket notation through:

$$\langle 1|\rho_n\rangle = \sum_{i=0}^{l} \rho_{n,i} = 1 \qquad \text{with} \quad \langle 1| = \left[1, 1, \ldots 1\right]. \tag{1.3.9}$$

For reason that will be clear in a while, it is usually preferred to rewrite Π by means of its diagonal matrix. If Π is diagonalizable, it exists an invertible matrix D such that:

$$\Pi = D\Pi_D D^{-1}, \tag{1.3.10}$$

where Π_D is the diagonal matrix whose entrances are Π eigenvalues. Inserting (1.3.10) in (1.3.7) returns:

$$|\rho_n\rangle = D{\Pi_D}^n D^{-1} |\rho_0\rangle . \tag{1.3.11}$$

With the upper writing the evaluation of probability ket evolution in time is easier, since the n-th power acts directly on the Π eigenvalues. Moreover, in this way it is possible to obtain the probability ket in the long time regime. Indeed if Π is positive, the largest eigenvalue corresponds to 1 and has multeplicity 1 because of the Perron-Frobenius theorem [12], which holds for non-negative and regular matrices. Π_D has then the form:

$$\Pi_D = \begin{bmatrix} 1 & 0 & \dots & 0 \\ 0 & \lambda_1 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \lambda_{l-1} \end{bmatrix}, \qquad \text{with} \quad \lambda_i < 1 \quad \text{for} \quad 1 \leq i \leq l-1. \tag{1.3.12}$$

In the long time limit (i.e. $n \to \infty$) only the unit eigenvalue survives, leading to:

$$\lim_{n\to\infty} \Pi_D = \begin{bmatrix} 1 & 0 & \dots & 0 \\ 0 & 0 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 0 \end{bmatrix}, \tag{1.3.13}$$

which inserted in eq. (1.3.10) gives the long time regime probability ket. The latter results proportional to the eigenvector $|v_1\rangle$ related to the unit eigenvalue. Indeed decomposing the initial ket in the eigenvectors basis $\{v_i\}_{i=0}^{l-1}$:

$$|\rho_0\rangle = \sum_{i=1}^{l} \tilde{\rho}_{0,i} |v_i\rangle , \tag{1.3.14}$$

it is possible to rewrite $|\rho_n\rangle$ as

$$|\rho_n\rangle = \sum_{i=1}^{l} \tilde{\rho}_{0,i}\lambda_i^n |v_i\rangle . \tag{1.3.15}$$

Then, since all the eigenvalues are strictly smaller than 1 except the one corresponding to $|v_0\rangle$, when taking the limit $n \to \infty$ on (1.3.15) only the $|v_0\rangle$ term survives.

1.3.2 Continuous time

Previously the time was indexed by the number of steps done from the referenced starting time whose probability density was ρ_0. Now we want to move from a discrete to a continuous time representation. The procedure is the common one: the time between two consecutive steps is labeled with Δt (as before) for then moving to the continuous putting $\Delta t \to 0$. First of all, it is convenient to rewrite the stochastic matrix with the time parameter embedded in an exponential function:

$$\Pi_{\Delta t} = e^{G\Delta t} = \sum_{n=0}^{\infty} \left(\frac{G\Delta t}{n!} \right)^n . \tag{1.3.16}$$

G is called *generating matrix* and or sufficiently small Δt, the series can be truncated at first order:

$$\Pi_{\Delta t} = e^{G\Delta t} = \mathbb{1} + G\Delta t + O(\Delta t^2). \tag{1.3.17}$$

By construction (1.3.16), $\Pi_{\Delta t}$ and G share the same eigenvectors while the eigenvalues are related by:

$$\lambda_i = e^{\gamma_i \Delta t} \qquad \Longleftrightarrow \qquad \gamma_i = \frac{1}{\Delta t} \ln \Big(\lambda_i \Big). \tag{1.3.18}$$

Notice that since $\lambda_i \leq 1$, all γ_i are negative except the one $\gamma_1 = 0$ related to $\lambda_1 = 1$.
The exponential notation in (1.3.16) permits to express the process in the continuous time representation. Indeed writing the probability ket as done in (1.3.15) and moving the time parameter from n to:

$$\lim_{\Delta t \to 0} n\Delta t = t, \tag{1.3.19}$$

it is possible to express the probability evolution in time with the ket:

$$|\rho_t\rangle = e^{Gt} |\rho_0\rangle = \sum_{i=1}^{l} \tilde{\rho}_{0,i} e^{\gamma_i t} |v_i\rangle . \tag{1.3.20}$$

1.4 Characterization of Stochastic Processes

As we already mentioned, many stochastic processes have a complex structure and the possibility of obtaining the probability density function is not obvious. However we can still gain information relying on other types of characterizations such as moments and cumulants or their generating functions [1] [13] [14].
In particular we are going to see in Chapter 3 that only two central moments and their relation with the Gaussian case helps obtaining information about the process without knowing the probability.

1.4.1 Moments

Among the most useful quantities to describe a stochastic processes, there are *moments*. The *n-th moment* of the stochastic process X_t is defined as:

$$\mu_n(t) := \langle {X_t}^n \rangle = \int_{\mathbb{R}} p(x,t) x^n dx. \tag{1.4.1}$$

Evaluating moments is a main purpose in probability theory because they characterize the process and helps identifying distributions. They are also used to confront between probability distributions and usually the Gaussian distribution is taken as main reference.
The most common moment is the first one which correspond to the mean while higher moments are related to tail behavior. Sometimes rather than evaluating moments as in (1.4.1), it is preferred to compute them with respect to the μ_1:

$$\nu_n(t) := \langle (X_t - \mu_1)^n \rangle = \int_{\mathbb{R}} p(x,t)(x - \mu_1(t))^n dx. \tag{1.4.2}$$

Eq. (1.4.2) is named *n-th central moment* and it is trivial to see that ν_1 is zero whereas ν_2 is the variance (frequently denoted with σ^2) which measures the dispersion around the mean. Normalizing eq. (1.4.2) with respect to σ, it is obtained the so called *standardized moments*:

$$\bar{\nu}_n(t) := \frac{\nu_n(t)}{\nu_2(t)^{\frac{n}{2}}}. \tag{1.4.3}$$

The latter is dimensionless and it is useful if one is interested in scaling invariant quantities. The most used standardized moments are *skew* and *kurtosis* [15], which are helpful indicators permitting to describe and interpret the probability shape. The skew corresponds to the third standardized moment and it quantifies the probability distribution asymmetry with respect to the mean. In particular if $\bar{\nu}_3(t)$ is positive the probability density function is respectively concentrated to the left of the mean with a longer right tail and vice versa if the skew is negative while $\bar{\nu}_3$ is null if the probability is symmetric.
Since the kurtosis will play a key role in the analysis reported in Chapter 3, let us deepen it properly. The *kurtosis* is the 4-th standardized moment, so it relates the 4-th central moment ν_4 and variance ν_2^2:

$$\kappa_X(t) = \frac{\nu_4(t)}{\nu_2(t)^2} = \frac{\langle (X_t - \mu_1(t))^4 \rangle}{\langle (X_t - \mu_1(t))^2 \rangle^2}. \tag{1.4.4}$$

Notice that by definition κ_X is always positive. From the kurtosis it is possible to gain information about the probability profile around tails and peak (usually referred as *tailedness* and *peakness*). Indeed confronting κ_X with the Gaussian kurtosis, which is 3, we have that X_t probability subtends a larger area (with respect to the normal distribution) nearby the peak and in the tails if κ_X is greater than 3, while the probability density distribution is lower and concentrated in a region of variance length if $\kappa_X < 3$. An intuitive explanation for that can be given considering processes with the same variance (indeed the kurtosis is standardized) and recalling that the 4-th central moment has an inverse proportionality to the peak height.
Usually it is preferred to use $\kappa_X - 3$, called *excess kurtosis* to get an immediate comparison with the Normal distribution. In this sense $\kappa_X - 3$ is a "Gaussianity" indicator and depending on whether its value is positive, negative or null 0 the X_t probability density function is called respectively leptokurtic, platykurtic or mesokurtic. Fig. 1.2 displays the distribution shapes of the three cases: the mesokurtic probability density is similar to the bell-shaped Gaussian curve, in the leptokurtic case the prominent peak favors events in a restricted range around the mean and on the tail while the platycurtic distribution is much compact (indeed the uniform probability has $\kappa_X < 3$) accumulating the probability for tails events.

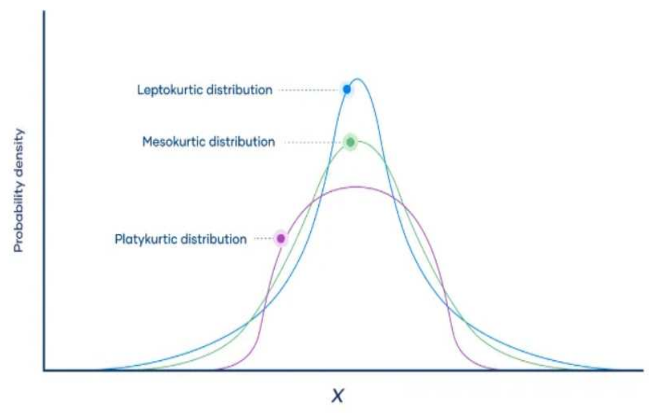

Figure 1.2: Different types of probability density function shapes depending on excess kurtosis

1.4.2 Generating functions

Moments can be collected in an elegant way by means of generating functions, whose general form is the expectation value of an exponential function containing the random variable in its argument. One of the most common is the *moment generating function*, which for a stochastic process X_t is defined as:

$$\phi(k,t) := \langle e^{k \cdot X_t} \rangle = \sum_{j=0}^{\infty} \frac{\langle (kX_t)^j \rangle}{j!}. \tag{1.4.5}$$

From the sum above it is possible to see that $\phi(k,t)$ contains all the X_t moments which can be gained from (1.4.5) differentiating with respect to k and then setting $k = 0$:

$$\langle X_t{}^n \rangle = \left(\frac{\partial}{\partial k} \right)^n \phi(k,t) \bigg|_{k=0}. \tag{1.4.6}$$

So from $\phi(k,t)$ it is possible to evaluate all the moments and this fully characterizes the process. A mathematical explanation for this uniquely characterization can be given in the continuous case, in which the moment generating function corresponds to the Laplace transform of the X_t probability density function which (as well as its anti-transformed) is unique.
The moment generating function in stochastic theory is the analogue of the statistical mechanics partition function [6]. Indeed they both enable us to describe uniquely the systems and the interested mean values (averages of X_T or of the energy) can be extracted differentiating with respect to a parameter (the inverse temperature in the statistical mechanics case).
It is also used to express eq. (1.4.5) with k purely imaginary so by means of Fourier transform. In this case we talk about *characteristic function* and the analogues of eq. (1.4.5) and (1.4.6) are:

$$\tilde{\phi}(k,t) := \langle e^{isX_t} \rangle = \sum_{j=0}^{\infty} \frac{\langle (isX_t)^j \rangle}{j!}, \qquad \langle X_t{}^n \rangle = \left(i \frac{\partial}{\partial k} \right)^n \phi(k,t) \bigg|_{s=0} \qquad \text{with} \quad s \in \mathbb{R} \tag{1.4.7}$$

Another common function useful to characterize a process is the so-called *cumulant generating function*:

$$\Phi(k,t) := \ln \langle e^{kX_t} \rangle = \sum_{j=0}^{\infty} \frac{k^j}{j!} \sigma_j, \tag{1.4.8}$$

where σ_j is known j-th *cumulant*, which represents an alternative to the moment characterization. The firsts cumulants are:

$$\sigma_1 = \langle X \rangle, \tag{1.4.9}$$
$$\sigma_2 = \langle X^2 \rangle - \langle X \rangle^2, \tag{1.4.10}$$
$$\sigma_3 = \langle X^3 \rangle - 3\langle X^2 \rangle \langle X \rangle + 2\langle X \rangle^3, \tag{1.4.11}$$
$$\sigma_4 = \langle X^4 \rangle - 4\langle X^3 \rangle \langle X \rangle - 3\langle X^2 \rangle^2 + 12\langle X^2 \rangle \langle X \rangle^2 - 6\langle X \rangle^4. \tag{1.4.12}$$

Sometimes it is preferred the use of cumulants over moments because the firsts σ_i have a more immediate confront with the probability distribution structure. Indeed from the previous equations it is possible to recognize the mean in σ_1, the variance in σ_2 and the 3-rd central moment in σ_3, while higher order cumulants do not correspond exactly to moments or central moments but they are a combination of them with highest moment order corresponding to their order.

Note that, since cumulant generating function is defined as the logarithm of the partition function correspondent $\phi(k,t)$, it is natural to associate it with the equivalent of the statistical mechanics free energy. As for the latter case, it is possible to prove [13] that the cumulant generating function is a convex and infinitely differentiable function and its existence is guaranteed if and only if the probability density function in which it is averaged has exponential decaying tails.

1.4.3 The Gaussian-Brownian case

Since most of the times when dealing with complex stochastic processes their characteristics are compared with the Gaussian-Brownian motion ones (Section 1.1), it is worth to illustrate the latter quantities mainly exploited for confronting. In particular we are going to focus on the moments and the generating functions which will be involved in the analysis of the subordinator processes in Chapter 3.
Already reported in Section 1.1, the Gaussian-Brownian variance can be evaluated also from the following formula [16] which permits to estimate quickly the $n-$th central moment for processes distributed according to the Gaussian propagator (1.1.2):

$$\nu_n^{GB}(t) = \langle (X_t - x_0)^n \rangle = \begin{cases} 0 & \text{if} \quad n \quad \text{is odd} \\ (2dDt)^{\frac{n}{2}} n!! & \text{if} \quad n \quad \text{is even,} \end{cases} \tag{1.4.13}$$

where the double factorial $n!!$ indicates the product of the integers from 1 to $(n-1)$ with the same parity. While the odd central moments are vanishing because of the Gaussian symmetry with respect to its mean, eq. (1.4.13) shows that in the even n case the $\nu_n^{GB}(t)$ dependence on the time goes like $t^{\frac{n}{2}}$. Because of this we have that the standardized moments (1.4.3) are time-independent:

$$\bar{\nu}_{2n}^{GB}(t) = \frac{(2Dt)^n n!!}{(2Dt)^n} = n!!, \tag{1.4.14}$$

which in the case $n = 4$ returned the kurtosis $\kappa_X = 3$ as anticipated in Section 1.4.
Looking now at the generating function in the 1-dimensional case, the Gaussian-Brownian motion moment generating function can be evaluate inserting eq. (1.1.2) (with $d = 1$) in (1.4.5):

$$\begin{aligned} \phi(k,t)^{GB} &= \int_{-\infty}^{\infty} dx \frac{1}{(4\pi Dt)^{\frac{1}{2}}} \exp\Big\{ -\frac{(x-x_0)^2}{4Dt} + kx \Big\} \\ &= e^{kx_0 + Dtk^2} \int_{-\infty}^{\infty} dy \frac{1}{(4\pi Dt)^{\frac{1}{2}}} \exp\Bigg\{ -\frac{(y - \frac{k}{2dt})^2}{4Dt} \Bigg\} \\ &= e^{x_0 k + Dtk^2}, \end{aligned} \tag{1.4.15}$$

where from the first to the second line it was made the change of variable $y = x - x_0$ and the quantity e^{kDt} was added and subtracted in order to complete the square which makes the integral to correspond to 1. Recalling eq. (1.4.8), the cumulant generating function is easily obtained taking (1.4.15) logarithm:

$$\Phi(k,t)^{GB}(t) = x_0 k + Dtk^2. \tag{1.4.16}$$

Since the previous equation does not contain terms like k^α with $\alpha \geq 3$, it follows that the only non-vanishing cumulants are the mean and the variance.

1.4.4 Anomalous Diffusion

As seen in section 1.1, the mean square displacement derived from the Gaussian propagator displays a linear time dependence. This will be the also an important feature of the analyzed stochastic diffusivity models and for completeness we mention the categorization used for the diffusion depending on the variance time dependence.
When the variance does not exhibits a linear dependence on time we talk of *anomalous diffusion*. In particular it is possible to subdivide diffusion processes through their mean squared displacement power law of time $\langle \Delta x^2 \rangle \propto t^{2a}$ with the following categorization:

$$\begin{cases} a < \frac{1}{2}: & \text{subdiffusion,} \\ a = \frac{1}{2}: & \text{Fickian diffusion,} \\ a > \frac{1}{2}: & \text{superdiffusion.} \end{cases}$$

An example of superdiffusion was already pointed out in 1926 by Lewis Fries Richardson [17], who observed a variance proportional to the time cubed analyzing particle dispersion with turbulent flow. Many models were made out in order to study and characterize the different types of diffusion [36] and their features and they can be basically distinguished through the related probability density function which is either Gaussian-like or not.

1.5 Anomalous Scaling and Self-Similarity

The variance linearity in time is a characteristic feature of the milestone Brownian motion which indeed gives the name to this type of diffusion. Anyway even if the Brownian motion is known to describe a countless number of phenomena, in many cases [36] [37] [38], and also in the processes studied in Chapter 3, the variance is observed to grow linearly in time but the associated probability is not a Gaussian curve as in the Brownian case (1.1.2). When this occurs we talk about *anomalous scaling* [36] and in the majority of the times the probability density function is found to have the form:

$$p_{X_t}(x,t) \sim \frac{1}{t^\nu} f\left(\frac{x}{t^\nu}\right), \tag{1.5.1}$$

where ν is the *scaling exponent* and f is called *scaling function*. The anomalous scaling is strictly correlated to the concept of *self-similarity*. Adopting the definition used in [19], a stochastic process Y_t is *self-similar* [19] [22] [23] if $\forall a \geq 0$ there exists b such that:

$$\big(Y_{at}, t \geq 0\big) \stackrel{d}{=} \big(bY_t, t \geq 0\big), \tag{1.5.2}$$

where $=^d$ indicates that the processes are identically distributed. In the case of probability distributions of kind (1.5.1), self-similarity can be easily checked scaling X_t as:

$$\bar{X}_t = a^\nu X_t, \tag{1.5.3}$$

then the latter probability density function is given by:

$$\begin{aligned} p_{\bar{X}_t}(\bar{x},t) &= \int_{\mathbb{R}} dx \delta(\bar{x} - a^\nu x_t) p_{X_t}(x,t) \\ &= \frac{1}{t^\nu a^\nu} f\big(\frac{\bar{x}}{t^\nu a^\nu}\big). \end{aligned} \tag{1.5.4}$$

Eq. (1.5.4) corresponds also to the probability of the initial process but indexed by at, so the following processes share the same distribution:

$$\big(X_{at}, t \geq 0\big) \stackrel{d}{=} \big(a^{\nu} X_t, t \geq 0\big). \tag{1.5.5}$$

Processes satisfying eq. (1.5.6), either with probability structure of type (1.5.1) or not, exhibit anomalous scaling. Indeed choosing $a = \frac{1}{t^{\nu}}$ and exploiting self-similarity (1.5.6), the $n-$th X_t moment grows in time as:

$$\langle X_t^n \rangle = t^n \langle \bar{X}_1 \rangle. \tag{1.5.6}$$

It follows that either all the X_t moments are null or none is. For successive purpose, it is useful to point out that thanks to eq. (1.5.6) in case of self-similarity (1.5.6) with zero mean the standardized moments, including then the kurtosis excess, are constant.

Chapter 2

Large Deviation Theory

The ground basis of large deviation theory was set by Cramérwho in his article "On a new limit theorem in probability theory" [29] pointed out a theorem which permits, under certain conditions, to relate the probability function with an exponential decay. The paper did not attract interest at that time and a development of large deviation theory had to wait 40 years for Friedlin and Wentzell [24] and for Donsker and Varadhan [25]. The aim of this theory is to study probabilities which exhibit an exponential asymptotic behavior and trying to give criteria on whether this trend should be expected or not. The interest in this type of probability is related to the exponential argument dependence on a parameter whose growth make the probability to be peaked around a point which corresponds to the most favorable event and can be determined through the saddle point approximation.
For our purposes, large deviation theory is useful to study the asymptotic time regime trend in Chapter 3 where indeed the exponential decay is ruled by the time.

2.1 The large deviation principle

Consider a discrete random variable X_n indexed by the integer n and let us indicate with $P(X_n \in C)$ the probability of X_n to take value on the set C. $P(X_n \in C)$ is said to satisfy the *large deviation principle* if the following limit exists:

$$\lim_{n\to\infty} -\frac{1}{n} \ln P(X_n \in C) = I_C. \tag{2.1.1}$$

The existence of the upper limit implies not only an exponential behavior of $P(X_n \in C)$ for large n, but also a linear growth of the exponential argument, since a different dependence would make I_C, called *rate*, diverge or vanish. Notice that the large deviation principle (2.1.1) is equivalent to say that $P(X_n \in C) \approx e^{-nI_C}$ for large n, so in order to have a finite probability I_C must be positive.
The limit (2.1.1) is extended to the continuous case considering the interval $C = [x, x + dx]$ with dx infinitesimal displacement, then the probability $P(\tilde{X}_n, [x, x + dx])$ for $\tilde{X}_n$ continuous random variable can be written through the probability density function $p(\tilde{X}_n = x)$:

$$P(\tilde{X}_n, [x, x + dx]) = p(\tilde{X}_n = x)dx, \tag{2.1.2}$$

Inserting now (2.1.2) in (2.1.1) returns the large deviation principle for the continuous:

$$\lim_{n\to\infty} -\frac{1}{n} \ln p(\tilde{X}_n = x)dx = J(x) - \lim_{n\to\infty} \frac{1}{n} \ln dx = J(x). \tag{2.1.3}$$

$J(x)$ is called *rate function* and is always positive for the same reason discussed for I_C. The notation $P(X_n \in dx) \asymp e^{-nJ(x)}dx$ (where "$\asymp$" refers to the asymptotic behavior)is used to indicate the probability obeying the large deviation principle (2.1.3).

The limits (2.1.1) and (2.1.3) contain the basic underlying concept of large deviation theory that consists in finding good criteria (i.e. large deviation principles) in order to distinguish among probabilities which do and do not exhibit an asymptotic exponential behavior. The second main goal is the evaluation of the rate/function rate which is controlling the distribution in the large n regime where the most favorable events correspond to small values of I_C (or minima of $J(x)$).

2.1.1 Random Walk with Gaussian steps

The random walk with Gaussian steps is a simple but instructive example of large deviations.
Consider a 1-dimension random walker taking steps that are independent and identically Gaussian distributed with mean μ and variance σ^2. Assuming that he starts from the axis origin, his coordinates after n steps is a sum of Gaussian random variables:

$$S_n = \sum_{i=0}^{n} Y_i, \tag{2.1.4}$$

where Y_i corresponds to the i-th step. Because of the properties of Gaussian distributions the S_n probability corresponds:

$$p(S_n = s) = \frac{1}{2\pi\sigma} e^{-\frac{n(s-\mu)^2}{2\sigma^2}}. \tag{2.1.5}$$

Taking now the limit (2.1.3) for $p(S_n = s)$ we find the S_n rate function:

$$J(s) = \frac{(s-\mu)^2}{2\sigma^2}, \tag{2.1.6}$$

which indicates that more the walker moves, more it is probable to find it at μ position. This is confirmed by Fig. [2.1] where it is shown the growth of $p(S_n = s)$ peak around the mean, which is also the $J(s)$ global minimum, as increasing the number of steps.

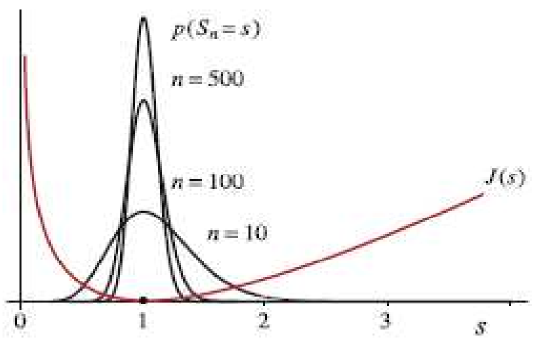

Figure 2.1: $J(s)$ and $p(S_n = s)$ variation as increasing the number of steps n for $\mu = \sigma = 1$

2.2 Gärtner-Ellis theorem

In most cases the evaluation of the rate function is not simple as done for the Random Walk with Gaussian steps, so we need techniques such as the Gärtner-Ellis theorem [27] [30] which enable us to determine it.
Consider a real random variable X_n indexed by positive integer n and distributed according to $P(X_n \in dx)$ and define the *scaled cumulant generating function* as:

$$\lambda(k) := \lim_{n\to\infty} \frac{1}{n} \ln\langle e^{nkX_n}\rangle, \tag{2.2.1}$$

where the average is meant over the possible X_n realizations. If $\lambda(k)$ exists and differentiable for all $k \in \mathbb{R}$, then the Gärtner-Ellis theorem guarantees that $P(X_n \in dx)$ obeys a large deviation principle with rate function:

$$J(x) = \sup_{k \in \mathbb{R}} \{kx - \lambda(k)\}, \tag{2.2.2}$$

which corresponds to the *Legendre-Fenchel transform* [26] of $\lambda(k)$.
Let us give a general argument for the derivation of the theorem. If A_n satisfy a large deviation principle then for large n the following holds:

$$\langle e^{nkX_n} \rangle \asymp \int_{-\infty}^{\infty} e^{n(kx - J(x))} dx \tag{2.2.3}$$

$$\asymp \exp\bigg(-k \sup_{x \in \mathbb{R}} \{kx - J(x)\} \bigg), \tag{2.2.4}$$

where from the first to the second row the saddle-point approximation method [31] was applied. Inserting (2.2.3) in (2.2.1) the scaled cumulant generating function becomes:

$$\lambda(k) = \sup_{x \in \mathbb{R}} \{kx - J(x)\}. \tag{2.2.5}$$

Eq. (2.2.2) is then obtained exploiting the Legendre-Fenchel transform ensured by $\lambda(k)$ differentiability.
The rate functions derived by the Gärtner-Ellis theorem respects the required positivity. Indeed from the definition (2.2.1) we have that $\lambda(0) = 0$ which matched eq. (2.2.5) evaluate at $k = 0$ returns:

$$\lambda(0) = \sup_{x \in \mathbb{R}} \{-J(x)\} = \inf_{x \in \mathbb{R}} \{J(x)\} = 0. \tag{2.2.6}$$

So the theorem guarantees not only a positive rate function, but also the existence of its global minimum at 0. The possibility of more minima is rejected as consequence of $J(x)$ gained by a Legendre-Fenchel transform which is always convex. Actually exploiting $\lambda(k)$ differentiability, it is possible to show [27] that in this case the rate function is *strictly convex* (i.e. does not contain linear terms in x).
The Gärtner-Ellis theorem is then a powerful tool to study large deviations even if the condition on the existence and differentiability of the scaled cumulant generating function sets a limit for its application.

Chapter 3

Stochastic diffusivity

In Chapters 1 and 2 the ground basis for studying stochastic diffusivity processes were settled and now it is finally come the time to start the analysis. We are going to study these processes, also referred as *superstatistical* [36] for the randomness of their transport parameter, by means of subordinator processes [1] [33] [34], which consider stochastic variables as function of other random quantities in turn. In particular we are going to focus on processes whose corresponding Langevin equation has the form:

$$dX(t) = \sqrt{2D(t)}dW(dt), \tag{3.0.1}$$

where $W(dt)$ is again a Wiener process and the diffusion coefficient $D(t)$ is a stochastic process. These models are used to describe diffusions for which one or more Einstein's conditions for exhibiting Brownian motion are not satisfied. Specifically for the processes analyzed in the following, the assumption of equal collisions fails.
However since the composition (3.0.1) cannot be managed as in the deterministic case, accurate methods and theorems have to be exploited in order to analyze such processes and deduce some interesting features.
In the following we will try to give a general idea about subordinator processes involving stochastic diffusivity and how to deal with them, for then studying in detail two particular cases: firstly we are going to analyze the diffusing diffusivity model [37] [38] [39] which considers the diffusion coefficient to be the square of a Ornstein-Uhlenbeck process, while for the second process studied the diffusivity is associated to a telegraph model.

3.1 Diffusivity Subordinator Processes

Let us consider (3.0.1), using the *subordinator process* defined as

$$dS \equiv 2D(t)dt \qquad \text{or equivalently} \qquad S(t) \equiv \int_0^t 2D(t')dt', \tag{3.1.1}$$

it is possible [1] to encode the time dependence of the diffusing coefficient on dS, so then the coordinate infinitesimal displacement can be written as:

$$dS = D(t)dt \quad \Rightarrow \quad dX(t) = dW(dS) \tag{3.1.2}$$

With this parameterization the $X(t)$ probability density function $p_X(x, t|x_0, 0)$ can be gained through the *subordination formula* [1]:

$$p_X(x,t|x_0,0) = \int_0^\infty p_S(s,t)G_{BG}(x,s|x_0)ds \qquad \text{with} \quad G_{BG}(x,s|x_0) = \frac{e^{\frac{-(x-x_0)^2}{2s}}}{\sqrt{2\pi s}} \tag{3.1.3}$$

where the integration boundaries are due to the fact that diffusivity is positive-defined and the time range considered starts from 0.
From eq. (3.1.3) it is possible to express the $X(t)$ moments as function of the $S(t)$ ones. Indeed exploiting the probability density function gained by the subordination formula (3.1.3) the n-th moment with respect to x_0 becomes:

$$\begin{aligned}
\langle (X(t)-x_0)^n\rangle &= \int_{-\infty}^\infty dx p_X(x,t|x_0)(x-x_0)^n \\
&= \int_{-\infty}^\infty dx \int_0^\infty ds G_{BG}(x,s|x_0)(x-x_0)^n p_S(s,t) \\
&= \int_{-\infty}^\infty dx \int_0^\infty ds \frac{e^{\frac{-(x-x_0)^2}{2s}}}{\sqrt{2\pi s}}(x-x_0)^n p_S(s,t) \\
&= \int_{-\infty}^\infty dx' x^n \frac{e^{-x'^2}}{\sqrt{2\pi}} \int_0^\infty ds p_S(s,t) s^{\frac{n}{2}},
\end{aligned} \tag{3.1.4}$$

where from the 2-nd to the 3-rd line the $x' = \frac{x-x_0}{\sqrt{2s}}$ change of variable is done. In eq. (3.1.4) the integration over dx corresponds to the n-th moment G_n of Gaussian distribution with unit variance and zero mean, while in the other integral the $\langle s^{\frac{n}{2}}\rangle$ is recognized. Eq. (3.1.4) becomes then:

$$\langle (X(t)-x_0)^n\rangle = G_n \langle S(t)^{\frac{n}{2}}\rangle. \tag{3.1.5}$$

The equation provides a powerful tool because it permits to derive the interested $X(t)$ moments directly from $\langle S(t)^{\frac{n}{2}}\rangle$ without necessarily knowing $p_X(x,t|x_0,0)$, whose evaluation through (3.1.3) is not always possible. Since $G_n = 0$ for n odd, also (3.1.5) vanishes for any choice of x_0. Thus, processes that can be parameterized by (3.1.2) are symmetric with respect to their initial position, which coincides with their mean, while their variance corresponds to the $S(t)$ mean and (recalling (1.4.4)) the kurtosis can be rewritten as:

$$\kappa_X = 3\frac{\langle S(t)^2\rangle}{\langle S(t)\rangle^2}, \tag{3.1.6}$$

or equivalently:

$$\kappa_X - 3 = 3\frac{\langle S(t)^2\rangle - \langle S(t)\rangle^2}{\langle S(t)\rangle^2}. \tag{3.1.7}$$

The $S(t)$ firsts two moments can be obtained from the 1-st and 2-nd $D(t)$ moments recalling (3.1.1). Indeed since the time considered is not aleatory, when computing $\langle S(t)^n\rangle$ through eq. (3.1.1) the mean operator comes inside the n integrals and acts on the $D(t_i)$ products:

$$\langle S(t)\rangle = 2\int_0^t \langle D(t_1)|D_0\rangle dt_1, \qquad \langle S(t)^2\rangle = 4\int_0^t dt_2 \int_0^t dt_1 \langle D(t_1)D(t_2)|D_0\rangle. \tag{3.1.8}$$

3.2 Diffusing Diffusivity

Diffusing diffusivity is the model in which the diffusion coefficient of eq. (3.0.1) is distributed as $D(t) = \vec{Y}(t)^2$ where $\vec{Y}(t)$ is an Ornstein-Uhlenbeck process. This model is proper to describes particles dispersing in an heterogeneous medium [37] [38], where the diffusion nature changes depending on whether we consider short or long space (and time) scales. Indeed if a particle is dispersed in an heterogeneous fluid, for short distances it will feel different collisions depending on which type of medium molecules are hitting, while for long space scale the heterogeneity is not more visible and the motion assumes Brownian characteristics. For this reason the Ornstein-Uhlenbeck process is found to give a good description of such diffusion. Indeed recalling its differential equation

$$d\vec{Y}(t) = -\frac{\vec{Y}(t)}{\tau}dt + \sigma d\vec{W}(t), \tag{3.2.1}$$

we see that it is possible to calibrate the deviation from the Brownian motion adjusting the τ timescale parameter.
In order to manage this sort of matryoshka of stochasticity, it is convenient to start studying $Y(t)$ to obtain $D(t)$ information such as $\langle D(t)^2\rangle$ and $\langle D(t)\rangle^2$ which will then be relevant for discussing the starting interested process $X(t)$.

3.2.1 Langevin approach

The Langevin approach focuses on the trajectory of the particle itself which is obtained solving Eq. (3.2.1) [1]:

$$\vec{Y}(t) = \vec{Y}_0 e^{\frac{-t}{\tau}} + \int_0^{\infty} e^{\frac{t'-t}{\tau}} \sigma \vec{\eta}(t')dt' \tag{3.2.2}$$

where $\eta_i(t) =$ "$\frac{dW_i(t)}{dt}$" is the Gaussian white noise (already mentioned in section 1.2.1) satisfying:

$$\begin{cases} \langle \vec{\eta}_(t)\rangle = 0, \\ \langle \vec{\eta}(t_1)\cdot\vec{\eta}(t_2)\rangle = n\delta(t_1 - t_2). \end{cases} \tag{3.2.3}$$

where n is the dimension of the Ornstein-Uhlenbeck process considered. Exploiting those properties $D(t)$ mean and autocorrelation given that the process started in $x = 0$ at $t = 0$ (so supposing $p_D(t_0 = 0) = \delta(D - D_0)$) are evaluated:

$$\begin{aligned}
\langle D(t)|D_0\rangle = \langle \vec{Y}(t)^2|\mathrm{Y}_0\rangle &= \langle \left(\vec{Y}_0 e^{\frac{-t}{\tau}} + \int_0^t e^{\frac{t'-t}{\tau}}\sigma\vec{\eta}(t')dt'\right)\left(\vec{Y}_0 e^{\frac{-t}{\tau}} + \int_0^t e^{\frac{t''-t}{\tau}}\sigma\vec{\eta}(t'')dt''\right)\rangle \\
&= \vec{Y}_0^2 e^{\frac{-2t}{\tau}} + 2\vec{Y}_0 e^{\frac{-t}{\tau}}\int_0^t e^{\frac{t'-t}{\tau}}\sigma\langle\vec{\eta}(t')\rangle dt' + \int_0^t\int_0^t e^{\frac{t'+t''-2t}{\tau}}\sigma^2\langle\vec{\eta}(t')\vec{\eta}(t'')\rangle dt'dt'' \\
&= \vec{Y}_0^2 e^{\frac{-2t}{\tau}} + \int_0^t e^{\frac{2t'-t}{\tau}}\sigma^2 dt' \\
&= \vec{Y}_0^2 e^{\frac{-2t}{\tau}} - \frac{n\sigma^2\tau}{2}e^{\frac{-2t}{\tau}} + \frac{n\sigma^2\tau}{2}.
\end{aligned} \tag{3.2.4}$$

Substituting $\vec{Y}_0^2 = D_0$ the it follows:

$$\langle D(t)|D_0\rangle = D_0 e^{\frac{-2t}{\tau}} + \frac{n\sigma^2\tau}{2}(1 - e^{\frac{-2t}{\tau}}). \tag{3.2.5}$$

Note that if the initial distribution is at the equilibrium, then the eq. (3.2.5) does not depend on time and this occurs if for any time

$$\langle D\rangle^* = \frac{n\sigma^2\tau}{2}. \tag{3.2.6}$$

For the autocorrelation the counts are a little bit longer:

$$\begin{aligned}\langle D(t_1)D(t_2)|D_0\rangle &= \langle \vec{Y}(t_1)^2\vec{Y}(t_2)^2|\vec{Y}_0\rangle\\ &= \langle\Big(\vec{Y}_0^2e^{\frac{-2t_1}{\tau}} + 2\vec{Y}_0e^{\frac{-t_1}{\tau}}\int_0^{t_1} e^{\frac{t'-t_1}{\tau}}\sigma\vec{\eta}(t')dt' + \Big(\int_0^{t_1} e^{\frac{t'-t_1}{\tau}}\sigma^2\vec{\eta}(t')dt'\Big)^2\Big)\cdot\\ &\cdot\Big(\vec{Y}_0^2e^{\frac{-2t_2}{\tau}} + 2\vec{Y}_0e^{\frac{-t_2}{\tau}}\int_0^{t_2} e^{\frac{t''-t_2}{\tau}}\sigma\vec{\eta}(t'')dt'' + \Big(\int_0^{t_2} e^{\frac{t''-t_2}{\tau}}\sigma^2\vec{\eta}(t'')dt''\Big)^2\Big)\rangle. \end{aligned}\tag{3.2.7}$$

Eq. (3.2.7) contains terms of the form $\langle\vec{\eta}(t')^n\cdot\vec{\eta}(t'')^m\rangle$ with $n, m = 0, 1, 2$ but a huge simplification comes from remembering that the mean of an odd function on Gaussian distributed variables is null:

$$\langle\vec{\eta}(t')^2\vec{\eta}(t'')\rangle = \langle\vec{\eta}(t')\vec{\eta}(t'')^2\rangle = 0 \tag{3.2.8}$$

which combined with (3.2.3) deletes the terms with $n + m = 1, 3$ so then (3.2.7) becomes:

$$\begin{aligned}\langle\vec{Y}(t_1)^2\vec{Y}(t_2)^2|\vec{Y}_0\rangle = &\vec{Y}_0^4e^{-2\frac{t_2+t_1}{\tau}} + \langle\vec{Y}(t_1)^2|\vec{Y}_0\rangle\vec{Y}_0^2e^{\frac{-2t_2}{\tau}} + \langle\vec{Y}(t_2)^2|\vec{Y}_0\rangle\vec{Y}_0^2e^{\frac{-2t_1}{\tau}}+\\ &+ 4e^{-\frac{t_2+t_1}{\tau}}\langle\Big(\int_0^{t_1} e^{\frac{t'-t_1}{\tau}}\sigma\vec{Y}_0\cdot\vec{\eta}(t')dt'\Big)\Big(\int_0^{t_2} e^{\frac{t''-t_2}{\tau}}\sigma\vec{Y}_0\cdot\vec{\eta}(t'')dt''\Big)\rangle+\\ &+\langle\Big(\int_0^{t_1} e^{\frac{t'-t_1}{\tau}}\sigma^2\vec{\eta}(t')dt'\Big)^2\Big(\int_0^{t_2} e^{\frac{t''-t_2}{\tau}}\sigma^2\vec{\eta}(t'')dt''\Big)^2\rangle. \end{aligned}\tag{3.2.9}$$

$$\tag{3.2.10}$$

In order to treat the last term in (3.2.9) it is useful to exploit the Wick's probability theorem which states that given a set of Gaussian distributed variable $\{x_1, ..., x_n\}$ the following holds:

$$\langle x_{i_1}...x_{i_n}\rangle = \sum_{p\in P_n}\langle x_{p_1}x_{p_2}\rangle...\langle x_{p_{n-1}}x_{p_n}\rangle, \tag{3.2.11}$$

where the sum is over all the possible permutations P_n with n elements. Applying this theorem the our case leads:

$$\begin{aligned}\langle\vec{\eta}(t_1')\vec{\eta}(t_2')\vec{\eta}(t_1'')\vec{\eta}(t_2'')\rangle = &n\big(\delta(t_1'-t_2')\delta(t_1''-t_2'') + \delta(t_1'-t_1'')\delta(t_2'-t_1'') + \delta(t_1'-t_2'')\delta(t_2'-t_1')\big)+\\ &+ n(n-1)\delta(t_1'-t_2')\delta(t_1''-t_2''). \end{aligned}\tag{3.2.12}$$

Exploiting the upper result in eq. (3.2.9) yields (assuming $t_2 > t_1$):

$$\begin{aligned}
\langle \vec{Y}(t_1)^2\vec{Y}(t_2)^2|\vec{Y}_0\rangle &= \vec{Y}_0^4 e^{-2\frac{t_2+t_1}{\tau}} + \langle \vec{Y}(t_1)^2|\vec{Y}_0\rangle \vec{Y}_0^2 e^{\frac{-2t_2}{\tau}} + \langle \vec{Y}(t_2)^2|\vec{Y}_0\rangle \vec{Y}_0^2 e^{\frac{-2t_1}{\tau}} + 4\vec{Y}_0^2 e^{-2\frac{t_2+t_1}{\tau}} \int_0^{t_1} e^{\frac{t'}{\tau}}\sigma dt' \\
&+ \sigma^4 e^{-2\frac{t_2+t_1}{\tau}} \left(\int_0^{t_1} dt' \int_0^{t_2} dt''(ne^{-2\frac{t'+t''}{\tau}} + n(n-1)e^{-2\frac{t'+t''}{\tau}}) + 2\int_0^{t_1} dt' \int_0^{t_1} dt'' ne^{-2\frac{t'+t''}{\tau}} \right) \\
&= \vec{Y}_0^4 e^{-2\frac{t_2+t_1}{\tau}} + \frac{1}{2}n\sigma^2\tau\vec{Y}_0^2 e^{-2\frac{t_1}{\tau}}(1-e^{-2\frac{t_2}{\tau}}) + \frac{1}{2}n\sigma^2\tau {Y_0}^2 e^{-2\frac{t_2}{\tau}}(1-e^{-2\frac{t_1}{\tau}}) + \\
&+ 2\sigma^2\tau {Y_0}^2 e^{-2\frac{t_2}{\tau}}(1-e^{-2\frac{t_1}{\tau}}) + \frac{1}{2}n\sigma^4\tau^2(e^{-2\frac{t_2-t_1}{\tau}} - 2e^{-2\frac{t_2}{\tau}} + e^{-2\frac{t_2+t_1}{\tau}}) + \\
&+ \frac{1}{4}n^2\sigma^4\tau^2(1-e^{-2\frac{t_1}{\tau}} - e^{\frac{-2t_2}{\tau}} + e^{-2\frac{t_1+t_2}{\tau}}).
\end{aligned} \tag{3.2.13}$$

The latter equation rewritten in terms of $D(t)$ gives the interested autocorrelation:

$$\begin{aligned}
\langle D(t_1)D(t_2)|D_0\rangle &= \frac{1}{4}n^2\sigma^4\tau^2 + \left(\frac{1}{2}D_0 n\sigma^2\tau - \frac{1}{4}n^2\sigma^4\tau^2\right)e^{-2\frac{t_1}{\tau}} + \left(\frac{1}{2}(n+4)D_0\sigma^2\tau - \frac{1}{4}(n^2+4n)\sigma^4\tau^2\right)e^{-2\frac{t_2}{\tau}} + \\
&+ \left({D_0}^2 - (n+2)D_0\sigma^2\tau + \frac{1}{4}(n^2+2n)\sigma^4\tau^2\right)e^{-2\frac{t_2+t_1}{\tau}} + \frac{1}{2}n\sigma^4\tau^2 e^{-2\frac{t_2-t_1}{\tau}}.
\end{aligned} \tag{3.2.14}$$

The 2-nd moment comes straightforwardly putting $t_1 = t_2 = t$ in (3.2.8):

$$\begin{aligned}
\langle D(t)^2|D_0\rangle =& \frac{1}{4}(n^2+2n)\sigma^4\tau^2 + \left((n+2)D_0\sigma^2\tau - \frac{1}{2}(n^2+2n)\sigma^4\tau^2\right)e^{-2\frac{t}{\tau}} + \\
&+ \left({D_0}^2 - (n+2)D_0\sigma^2\tau + \frac{1}{4}(n^2+2n)\sigma^4\tau^2\right)e^{-4\frac{t}{\tau}}.
\end{aligned} \tag{3.2.15}$$

Following the same reasoning which led to (3.2.6), eq.s (3.2.14) and (3.2.15) in case of equilibrium initial condition becomes:

$$\langle D(t_1)D(t_2)\rangle^* = \frac{1}{4}n^2\sigma^4\tau^2 + \frac{1}{2}n\sigma^4\tau^2 e^{-2\frac{t_2-t_1}{\tau}}, \tag{3.2.16}$$

$$\langle D^2\rangle^* = \frac{1}{4}(n^2+2n)\sigma^4\tau^2, \tag{3.2.17}$$

where in the case of autocorrelation (3.2.16) the term related to the difference between the two times considered is allowed.

3.2.2 Fokker-Planck approach

As the procedure leading to the Fokker-Plank equation does, this method focuses on the probability density function. For the Ornstein-Uhlenbeck process in n dimensions the probability density function to be in Y at t given that the process started in Y_0 at $t_0 = 0$ is [1]:

$$p_{\vec{Y}}(\vec{Y}, t|\vec{Y}_0, 0) = \prod_{i=1}^{n} p_{Y_i}(Y_i, t|Y_{0,i}, 0) \tag{3.2.18}$$

with

$$p_{Y_i}(Y_i,t|Y_{0,i},0)=\frac{1}{\sqrt{\pi\sigma^2\tau(1-e^{\frac{-2t}{\tau}})}}e^{\frac{-(Y_i-Y_{0,i}e^{\frac{-t}{\tau}})^2}{\sigma^2\tau(1-e^{\frac{-2t}{\tau}})}}. \tag{3.2.19}$$

The previous equation permits to gain $p_D(D,t|D_0,0)$ through:

$$p_D(D,t|D_0,0)=\int_{-\infty}^{\infty}dy\delta(D-y^2)p_Y(y,t|y_0,0). \tag{3.2.20}$$

If the Ornstein-Uhlenbeck process considered has dimension $n=1$, the $D(t)$ results:

$$p_D(D,t|D_0,0)=\frac{1}{\sqrt{4\pi\sigma^2\tau D(1-e^{\frac{-2t}{\tau}})}}(e^{\frac{-(\sqrt{D}-\sqrt{D_0}e^{\frac{-t}{\tau}})^2}{\sigma^2\tau(1-e^{\frac{-2t}{\tau}})}}+e^{\frac{-(\sqrt{D}+\sqrt{D_0}e^{\frac{-t}{\tau}})^2}{\sigma^2\tau(1-e^{\frac{-2t}{\tau}})}}). \tag{3.2.21}$$

However already in $n=1$ it possible to see that (3.2.19) is easier to manage than (3.2.21), so it is convenient to evaluate the $Y(t)$ moments and autocorrelation for then reverting to the $D(t)$ ones. First of all, notice that every dimensional component is independent from another, so we can factorize $\langle\vec{Y}(t)^2|Y_0\rangle$ and $\langle\vec{Y}(t_1)^2\vec{Y}(t_2)^2|Y_0\rangle$ as:

$$\langle\vec{Y}(t)^2|\vec{Y_0}\rangle=\sum_{i=1}^{n}\langle Y_i(t)^2|Y_{0,i}\rangle, \tag{3.2.22}$$

$$\langle\vec{Y}(t_1)^2\vec{Y}(t_2)^2|\vec{Y_0}\rangle=\sum_{i\neq j}^{n}\langle Y_i(t_1)^2|Y_{0,i}\rangle\langle Y_j(t_2)^2|Y_{0,j}\rangle+\sum_{i=1}^{n}\langle Y_i(t_1)^2Y_i(t_2)^2|Y_{0,i}\rangle. \tag{3.2.23}$$

Therefore we can evaluate the 1-dimension $Y(t)$ moments for then inserting the results in (3.2.22) and (3.2.23)

$$\langle Y(t)^2|Y_0\rangle=\int_{-\infty}^{\infty}Y^2\frac{1}{\sqrt{\pi\sigma^2\tau(1-e^{\frac{-2t}{\tau}})}}e^{\frac{-(Y-Y_0e^{\frac{-t}{\tau}})^2}{\sigma^2\tau(1-e^{\frac{-2t}{\tau}})}}dY, \tag{3.2.24}$$

$$\begin{aligned}\langle Y(t_1)^2Y(t_2)^2|Y_0\rangle=\int_{-\infty}^{\infty}dY_1\int_{-\infty}^{\infty}dY_2{Y_1}^2{Y_2}^2\frac{1}{\sqrt{\pi\sigma^2\tau(1-e^{\frac{-2t_1}{\tau}})}}e^{\frac{-(Y_1-Y_0e^{\frac{-t_1}{\tau}})^2}{\sigma^2\tau(1-e^{\frac{-2t_1}{\tau}})}}\cdot\\ \cdot\frac{1}{\sqrt{\pi\sigma^2\tau(1-e^{-2\frac{t_2-t_1}{\tau}})}}e^{\frac{-(Y_2-Y_1e^{-\frac{t_2-t_1}{\tau}})^2}{\sigma^2\tau(1-e^{-2\frac{t_2-t_1}{\tau}})}},\end{aligned} \tag{3.2.25}$$

where in eq. (3.2.25) it is chosen $t_2>t_1$. In order to evaluate the integrals above, let us perform the following change of variables (another possible way is illustrated in Appendix B.1):

$$\begin{cases}z=\frac{-(Y_1-Y_0e^{\frac{-t_1}{\tau}})}{\sqrt{\sigma^2\tau(1-e^{\frac{-2t_1}{\tau}})}},\\ q=\frac{-(Y_2-Y_1e^{-\frac{t_2-t_1}{\tau}})}{\sqrt{\sigma^2\tau(1-e^{-2\frac{t_2-t_1}{\tau}})}},\end{cases} \tag{3.2.26}$$

where the change in q is involved in the derivation of the autocorrelation. $\langle Y(t_1)^2|Y_0\rangle$ and $\langle Y(t_1)^2Y(t_2)^2|Y_0\rangle$ are then obtained from an integration by parts recalling that the 2-nd and 4-th moments of Gaussian distribution with unit variance and zero mean are respectively $G_2 = 1$ and $G_4 = 3$ while $G_n = 0$ for n odd :

$$\begin{aligned}\langle Y(t_1)^2|Y_0\rangle &= \int_{-\infty}^{+\infty}\frac{e^{-z^2}}{\sqrt{\pi}}\Big(\sigma^2\tau(1-e^{\frac{-2t_1}{\tau}})z^2 + {Y_0}^2e^{-2\frac{t_1}{\tau}}+ \\ &+ 2\sqrt{\sigma^2\tau(1-e^{\frac{-2t_1}{\tau}})}Y_0e^{-\frac{t_1}{\tau}}z\Big)dz &(3.2.27)\\ &= {Y_0}^2e^{\frac{-2t_1}{\tau}} - \frac{\sigma^2\tau}{2}e^{\frac{-2t_1}{\tau}} + \frac{\sigma^2\tau}{2}, &(3.2.28)\end{aligned}$$

$$\begin{aligned}\langle Y(t_1)^2Y(t_2)^2|Y_0\rangle &= \int_{-\infty}^{+\infty}dq\int_{-\infty}^{+\infty}dz\frac{e^{-z^2}e^{-q^2}}{\pi}\Big(\sqrt{\sigma^2\tau(1-e^{\frac{-2t_1}{\tau}})}z + Y_0e^{-\frac{t_1}{\tau}}\Big)^2\cdot \\ &\cdot\Big(\sqrt{\sigma^2\tau(1-e^{-2\frac{t_2-t_1}{\tau}})}q + \big(\sqrt{\sigma^2\tau(1-e^{\frac{-2t_1}{\tau}})}z + Y_0e^{-\frac{t_1}{\tau}}\big)e^{-\frac{t_2-t_1}{\tau}}\Big)^2 \\ &= \int_{-\infty}^{+\infty}dq\int_{-\infty}^{+\infty}dz\frac{e^{-z^2}e^{-q^2}}{\pi}\Big(\sigma^4\tau^2(1-e^{\frac{-2t_1}{\tau}})(1-e^{-2\frac{t_2-t_1}{\tau}})z^2q^2+ \\ &+ \sigma^4\tau^2(1-e^{\frac{-2t_1}{\tau}})^2e^{-2\frac{t_2-t_1}{\tau}}z^4 + \sigma^2\tau(1-e^{\frac{-2t_1}{\tau}}){Y_0}^2e^{-2\frac{t_2}{\tau}}z^2+ \\ &+ \sigma^2\tau{Y_0}^2(1-e^{-2\frac{t_2-t_1}{\tau}})e^{-2\frac{t_1}{\tau}}q^2 + {Y_0}^4e^{-2\frac{t_2+t_1}{\tau}}+ \\ &+ \sigma^2\tau(1-e^{\frac{-2t_1}{\tau}}){Y_0}^2e^{-2\frac{t_2}{\tau}}z^2 + 4\sigma^2\tau(1-e^{\frac{-2t_1}{\tau}}){Y_0}^2e^{-2\frac{t_2}{\tau}}z^2\Big) &(3.2.29)\\ &= {Y_0}^4e^{-2\frac{t_2+t_1}{\tau}} + \frac{1}{2}\sigma^2\tau{Y_0}^2e^{-2\frac{t_1}{\tau}}(1-e^{-2\frac{t_2}{\tau}}) + \frac{1}{2}\sigma^2\tau{Y_0}^2e^{-2\frac{t_2}{\tau}}(1-e^{-2\frac{t_1}{\tau}})+ \\ &+ \frac{1}{4}\sigma^4\tau^2(1-e^{-2\frac{t_1}{\tau}})(1-e^{-2\frac{t_2}{\tau}}) + \frac{1}{2}\sigma^4\tau^2e^{-2\frac{t_2}{\tau}}(e^{-\frac{t_1}{\tau}}+e^{\frac{t_1}{\tau}})^2+ \\ &+ 2\sigma^2\tau{Y_0}^2e^{-2\frac{t_2}{\tau}}(1-e^{-2\frac{t_1}{\tau}}). &(3.2.30)\end{aligned}$$

Inserting the upper results in (3.2.22) and (3.2.23) it is find the same autocorrelation and 2-nd moment obtained through the Langevin approach in (3.2.4) and (3.2.13).
From the probability distribution (3.2.21) it is possible to obtain information about the stationary regimes. Considering for simplicity the $Y(t)$ 1 dimensional case whose $D(t)$ probability density function is reported in (3.2.21), in order to have stationarity it is needed:

$$\partial_t p_D(D,t|D_0,0) = 0. \tag{3.2.31}$$

Instead of computing the upper equation as it is written, it is better to exploit the fact that the time dependence of $p_D(D,t|D_0,0)$ appears with a factor $e^{\frac{-t}{\tau}}$:

$$\frac{\partial\big(e^{\frac{-t}{\tau}}\big)}{\partial t}\frac{1}{\partial\big(e^{\frac{-t}{\tau}}\big)}p_D(D,t|D_0,0) = -\frac{1}{\tau}e^{\frac{-t}{\tau}}\frac{\partial p_D(D,t|D_0,0)}{\partial\big(e^{\frac{-t}{\tau}}\big)}. \tag{3.2.32}$$

Since partial derivative on the r.h.s of (3.2.32) is never vanishing, it follows stationarity holds for $\tau\to\infty$ (or $\frac{t}{\tau}\to 0$) and $t\to\infty$, so for the short and long time regimes. That can also be demonstrated looking at the $p_D(D,t|D_0,0)$ (taking for simplicity $n=1$):

- for the short time regime:

$$\begin{cases} e^{\frac{-t}{\tau}} \approx 1, \\ \tau(1 - e^{\frac{-2t}{\tau}}) \approx 2t, \end{cases} \tag{3.2.33}$$

where in the second line it is used:

$$e^{\frac{-2t}{\tau}} = \sum_{j=0}^{\infty} \frac{1}{j!}\Big(\frac{-2t}{\tau}\Big)^j = 1 - \frac{2t}{\tau} + O\Big(\Big(\frac{t}{\tau}\Big)^2\Big). \tag{3.2.34}$$

Inserting the (3.2.33) in (3.2.21) we find:

$$p_D{}^{ST}(D, t|D_0, 0) = \frac{1}{\sqrt{8\pi\sigma^2 tD}}\big(e^{\frac{-(\sqrt{D}-\sqrt{D_0})^2}{2\sigma^2 t}} + e^{\frac{-(\sqrt{D}+\sqrt{D_0})^2}{2\sigma^2 t}}\big). \tag{3.2.35}$$

The latter equation could seem to depend on time at first sight but recalling the following limit:

$$\lim_{\varepsilon \to 0^+} \frac{1}{\sqrt{2\pi\varepsilon^2}} e^{-\frac{1}{2}\frac{x^2}{\varepsilon^2}} = \delta(x), \tag{3.2.36}$$

the probability density function becomes $p_D{}^{ST}(D) = \frac{1}{2}(\delta(D-D_0)+\delta(D+D_0))$ when taking the $t \to 0$. The latter has to be corrected since the second Dirac delta (related to $D = -D_0$) has no physical meaning because the diffusivity cannot be negative. Finally the short time probability density function results to be:

$$p_D{}^{ST}(D) = \delta(D - D_0). \tag{3.2.37}$$

Thus for $t \ll \tau$ the diffusivity can be considered a constant with the same initial value D_0.

- for the long time regime:

$$e^{\frac{-t}{\tau}} \approx 0, \tag{3.2.38}$$

which inserted in (3.2.21) removes the time dependence leading directly to the long time probability density function:

$$p_D{}^{LT}(D) = \frac{1}{\sqrt{\pi D\sigma^2\tau}} e^{-\frac{D}{\sigma^2\tau}}, \tag{3.2.39}$$

where the initial value D_0 is completely forgotten.

3.2.3 Cumulants

In the case of equilibrium initial condition, the first $S(t)$ mean evaluation comes straightforwardly since the diffusivity mean (3.2.6) does not depend on time:

$$\langle S(t)\rangle^* = n\sigma^2\tau t. \tag{3.2.40}$$

The upper equation corresponds to the $X(t)$ variance and it demonstrate that the $X(t)$ undergoes Fickian diffusion for any time.
For the $X(t)$ variance, recalling eq. (3.2.16), it is found:

$$\langle S(t)^2\rangle^* = 8\int_0^t dt_2 \int_0^{t_2} dt_1 \langle D(t_1)D(t_2)\rangle^* \tag{3.2.41}$$
$$= 8\int_0^t dt_2 \int_0^{t_2} dt_1 \left(\frac{1}{4}n^2\sigma^4\tau^2 + \frac{1}{2}n\sigma^4\tau^2 e^{-2\frac{t_2-t_1}{\tau}}\right)$$
$$= n^2\sigma^4\tau^2 t^2 + n\sigma^4\tau^2(2\tau t + \tau^2 e^{-2\frac{t}{\tau}} - \tau^2) \tag{3.2.42}$$

where in (3.2.41) it was exploited the fact that the $dt_1 dt_2$ integration over a square of size $[0,t]\times[0,t]$ is equivalent to 2 times the integration over the isosceles triangle dividing the square in two equal part. From the results gained in eq.s (3.2.40) and (3.2.42), recalling (3.1.7), the reduced excess kurtosis is gained:

$$\kappa_X{}^* - 3 = 3\frac{2\tau t + \tau^2 e^{-2\frac{t}{\tau}} - \tau^2}{nt^2} \tag{3.2.43}$$

Due to the time squared dependence on the denominator, the upper equation vanishes in the long time regime, while the short time kurtosis is obtained expanding the exponential term up to the second order in t:

$$(\kappa_X{}^* - 3)_{ST} = \frac{6}{n}, \qquad (\kappa_X{}^* - 3)_{LT} = \frac{6}{n\frac{t}{\tau}} \tag{3.2.44}$$

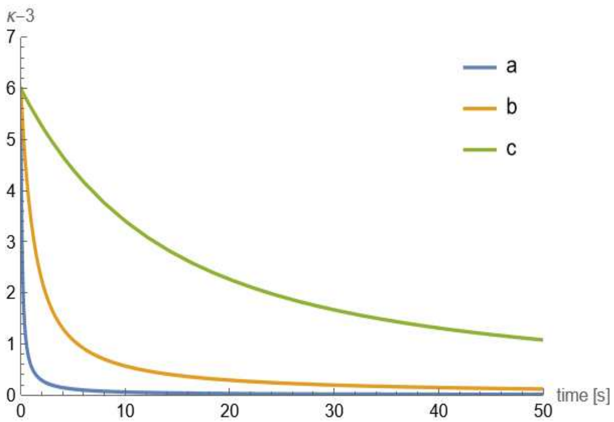

Figure 3.1: $\kappa_X{}^* - 3$ at varying of the proper time for $n = 1$: a: $\tau = 0.1s$; b: $\tau = 1s$; c: $\tau = 10s$.

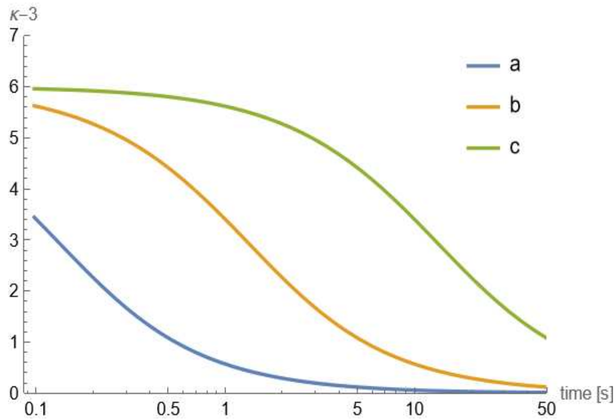

Figure 3.2: $\kappa_X{}^* - 3$ at varying of the proper time for $n = 1$ with logarithmic time axis: a: $\tau = 0.1s$; b: $\tau = 1s$; c: $\tau = 10s$.

From the figure above it is possible to see that for any choice of τ the kurtosis behaves similarly: at long times it vanishes while for short times it converge to the same point (as reported in (3.2.44)). Fig.s [3.1][3.2] draw an initial leptokurtic probability distribution, which modifies into a mesokurtic curve as increasing the time and the velocity at which this change in shape happens is determined by the τ parameter. Thus, since the diffusion is found to be Fickian for all the times (3.2.40), the kurtosis vanishing for large times indicates a large deviation regime, while the constant trend for short time (Fig. [3.2]) suggests an anomalous scaling behavior.
Looking now at the general case without equilibrium initial condition, the $X(t)$ variance is gained inserting eq. (3.2.5) in the first equation of (3.1.8):

$$\begin{aligned}\langle X(t)^2\rangle &= 2\int_0^t \left[D_0 e^{\frac{-2t}{\tau}} + \frac{n\sigma^2\tau}{2}(1-e^{\frac{-2t}{\tau}})\right]dt_1 \\ &= \big(\frac{1}{2}n\sigma^2\tau^2 - \tau D_0\big)e^{-2\frac{t}{\tau}} + n\sigma^2\tau t - \frac{1}{2}n\sigma^2\tau^2 + D_0\tau. \end{aligned} \tag{3.2.45}$$

The variance behavior for long times is easily evaluated form (3.2.45) putting to 1 the exponential factor, while the latter expanded at first order in t gives the short time behavior:

$$\begin{cases}\langle X(t)^2\rangle_{ST} = 2D_0 t, \\ \langle X(t)^2\rangle_{LT} = n\sigma^2\tau t - \frac{1}{2}n\sigma^2\tau^2 + D_0\tau.\end{cases} \tag{3.2.46}$$

The diffusion results then Fickian for both regimes, but with different coefficients.
As done previously, in order to obtain the kurtosis the $\langle S(t)^2\rangle$ is computed inserting eq. (3.2.14) in the second equation of (3.1.8):

$$\begin{aligned}\langle S(t)^2\rangle = 8\int_0^t dt_2 \int_0^{t_2} dt_1 \Bigg(&\Big(\frac{1}{2}D_0 n\sigma^2\tau - \frac{1}{4}n^2\sigma^4\tau^2\Big)e^{-2\frac{t_1}{\tau}} + \Big(\frac{1}{2}(n+4)D_0\sigma^2\tau - \frac{1}{4}(n^2+4n)\sigma^4\tau^2\Big)e^{-2\frac{t_2}{\tau}} + \\ &+ \Big(D_0{}^2 - (n+2)D_0\sigma^2\tau + \frac{1}{4}(n^2+2n)\sigma^4\tau^2\Big)e^{-2\frac{t_2+t_1}{\tau}} + \frac{1}{2}n\sigma^4\tau^2 e^{-2\frac{t_2-t_1}{\tau}} + \frac{1}{4}n^2\sigma^4\tau^2\Bigg) \\ = n^2\sigma^4\tau^2 t^2 + \big(&2n\sigma^2\tau^2 D_0 + (2n-n^2)\sigma^4\tau^3\big)t + \big(\frac{1}{2}(4n-n^2)\sigma^4\tau^4 + 2n\sigma^2\tau^3 D_0 - 2\tau^2 D_0{}^2\big)e^{-2\frac{t}{\tau}} + \\ \big((4n+n^2)&\sigma^4\tau^3 - (2n+8)\sigma^2\tau^2 D_0\big)e^{-2\frac{t}{\tau}}t + \big(\frac{1}{4}(2n+n^2)\sigma^4\tau^4 - (n+2)\sigma^2\tau^3 D_0 + \tau^2 D_0{}^2\big)e^{-4\frac{t}{\tau}} + \\ + \frac{1}{4}(n^2&-10)\sigma^4\tau^4 + (2-n)\sigma^2\tau^3 D_0 + \tau^2 D_0{}^2, \end{aligned} \tag{3.2.47}$$

which, combined with eq. (3.2.45) and again using eq. (3.1.7), leads to the reduced excess kurtosis:

$$\begin{aligned}\kappa_X - 3 = &3\Big[2n\sigma^4\tau^3 t + 2n\sigma^4\tau^4 e^{-2\frac{t}{\tau}} + \big(4n\sigma^4\tau^3 - 8D_0\sigma^2\tau^2\big)e^{-2\frac{t}{\tau}}t + \\ &+ \big(\frac{1}{2}n\sigma^4\tau^4 - 2\sigma^2\tau^3 D_0\big)e^{-4\frac{t}{\tau}} - \frac{5}{2}n\sigma^4\tau^4 + 2\sigma^2\tau^3 D_0\Big] \cdot \\ &\cdot \Big[\big(\frac{1}{2}n\sigma^2\tau^2 - \tau D_0\big)e^{-2\frac{t}{\tau}} + n\sigma^2\tau t - \frac{1}{2}n\sigma^2\tau^2 + D_0\tau\Big]^{-2}. \end{aligned} \tag{3.2.48}$$

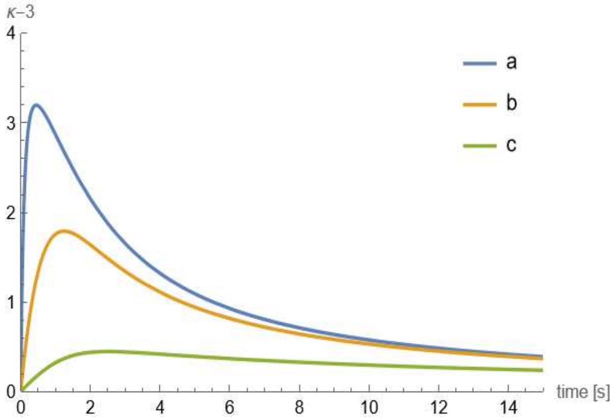

Figure 3.3: Reduced kurtosis excess at varying of D:
a: $\{\sigma = 1\frac{cm}{s}, \tau = 1s, D_0 = 0,1\frac{cm^2}{s}\}$;
b:$\{\sigma = 1\frac{cm}{s}, \tau = 1s, D_0 = 1\frac{cm^2}{s}\}$;
c:$\{\sigma = 1\frac{cm}{s}, \tau = 1s, D_0 = 10\frac{cm^2}{s}\}$.

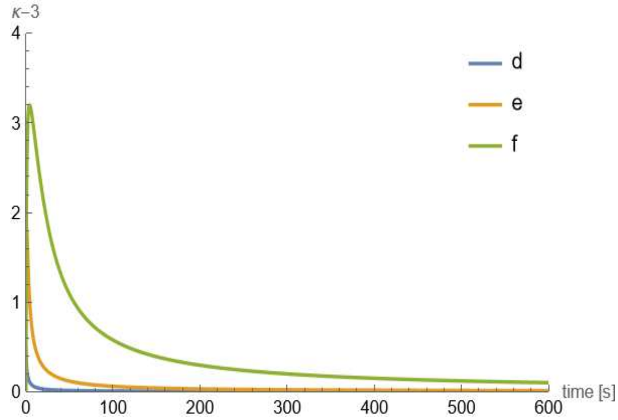

Figure 3.4: Reduced kurtosis excess at varying of τ:
d: $\{\sigma = 1\frac{cm}{s}, \tau = 0,1s, D_0 = 1\frac{cm^2}{s}\}$;
e:$\{\sigma = 1\frac{cm}{s}, \tau = 1s, D_0 = 1\frac{cm^2}{s}\}$;
f :$\{\sigma = 1\frac{cm}{s}, \tau = 10s, D_0 = 1\frac{cm^2}{s}\}$.

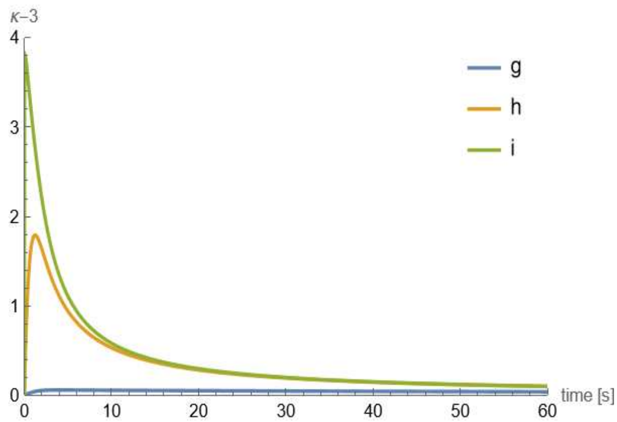

Figure 3.5: Reduced kurtosis excess at varying of σ:
g: $\{\sigma = 0,1\frac{cm}{s}, \tau = 1s, D_0 = 1\frac{cm^2}{s}\}$;
h: $\{\sigma = 1\frac{cm}{s}, \tau = 1s, D_0 = 1\frac{cm^2}{s}\}$;
i: $\{\sigma = 10\frac{cm}{s}, \tau = 1s, D_0 = 1\frac{cm^2}{s}\}$.

In the above figure the reduced excess kurtosis is plotted for processes with scales around seconds and centimeters. The behaviors for different parameters are dissimilar for short times while they all tend to vanish as increasing the time. The long time trend could be guessed also from (3.2.48) where for $t \to \infty$ the $e^{-\frac{2t}{\tau}}$ and $e^{-\frac{4t}{\tau}}$ terms cancel out bringing to the reduced excess kurtosis:

$$(\kappa_X - 3)_{LT} = 6\frac{\tau}{t} \xrightarrow{\frac{t}{\tau}\to\infty} 0. \tag{3.2.49}$$

Therefore the $X(t)$ probability density function for long times is mesokurtic independently from the choice of the parameters. A similarity with the Gaussian distribution is thus expected and the subordinator formula (3.1.3) is exploited to investigate on. In order to apply this formula it is necessary to know $p_S(s,t)$. The ergodicity of $D(t)$ process comes in help: it guarantees that the long time average $\langle D\rangle_{LT}$ is the same that the ensemble one (evaluated with (3.2.39))

$$\frac{1}{t}\int_0^t D(t')dt' = \langle D\rangle_{LT} = \frac{n\sigma^2\tau}{2}, \tag{3.2.50}$$

where in the last line it is used (3.2.5) for $t \to \infty$. Recalling (3.1.1) the subordinator becomes:

$$s(t) = n\sigma^2\tau t. \tag{3.2.51}$$

The upper equation implies that for long times s is no more an aleatory variable but its evolution is completely deterministic on t. Eq. (3.2.51) is achieved also starting from a Large Deviation principle [27] [37] for $\frac{s}{t} \equiv \bar{s}$ that is the existence of the following limit:

$$\lim_{t\to\infty} -\frac{1}{t}\ln\left(p_{\bar{s}}(\bar{s},t)\right) \propto \bar{s} - n\sigma^2\tau. \tag{3.2.52}$$

The latter ensures that $p_{\bar{s}}(\bar{s},t)$ peak around $\sigma^2\tau$ raises as increasing t so then for $t \to \infty$:

$$p_{\bar{s}}(\bar{s},t) = \delta(\bar{s} - n\sigma^2\tau) \qquad \xrightarrow{(3.2.20)} \qquad p_S(s,t) = \delta(s - n\sigma^2\tau t). \tag{3.2.53}$$

Inserting $p_S(s,t)$ just obtained in the (3.1.3) returns:

$$p_X(x,t|x_0,0)_{LT} = \frac{e^{\frac{-(x-x_0)^2}{2n\sigma^2\tau t}}}{\sqrt{2\pi n\sigma^2\tau t}}. \tag{3.2.54}$$

Therefore $p_X(x,t|x_0,0)^{LT}$ is not only mesokurtic, as indicated by the kurtosis, but it is a Gaussian with x_0 mean and $\sigma^2\tau t$ variance. Eq. (3.2.54) corresponds to the unbiased Brownian distribution (1.1.2) with diffusive coefficient $\frac{n\sigma^2\tau}{2}$ and obeys normal diffusion which broadens $X(t)$ probability density function as increasing time because of its variance linear dependence on t.
The short time regime is more complicated. Indeed putting simply $t = 0$ in (3.2.48) gives the indeterminate form $\frac{0}{0}$ so it is needed to go to higher orders in t to find a better expression. Expanding the exponential and neglecting $O(t^3)$ the time dependent factors become:

$$\begin{cases} e^{-\frac{2t}{\tau}} = 1 - \frac{2t}{\tau} + \frac{t^2}{\tau^2} - \frac{4}{3}\frac{t^3}{\tau^3} + O\left(\frac{t^3}{\tau^3}\right), \\ e^{-\frac{4t}{\tau}} = 1 - \frac{4t}{\tau} + 8\frac{t^2}{\tau^2} - \frac{32}{3}\frac{t^3}{\tau^3} + O\left(\frac{t^3}{\tau^3}\right), \\ \frac{t}{\tau}e^{-\frac{2t}{\tau}} = t - 2\frac{t^2}{\tau^2} + 2\frac{t^3}{\tau^3} + O\left(\frac{t^3}{\tau^3}\right), \end{cases} \tag{3.2.55}$$

which inserted in (3.2.48) give the reduced kurtosis excess for short times:

$$(\kappa_X - 3)_{ST} = 4\frac{2\sigma^2\tau}{D_0}\frac{t}{\tau} \xrightarrow{\frac{t}{\tau}\to 0} 0. \tag{3.2.56}$$

The above result indicates that the short time $X(t)$ excess kurtosis is vanishing (i.e. the correspondent probability distribution is mesokurtic) for any choice of parameters.

3.2.4 Simulation

Since the diffusing diffusivity model has a complex stochastic nature, it is useful to simulate the process in order to have a better comprehension. We analyze the case in which the $X(t)$ known initial position is set to coincide with the axis origin and the parameters of the Ornstein-Ulhenbeck equation (3.2.1) are for simplicity $\sigma = 1cms^{-1}$, $\tau = 1s$ and $Y_0 = 1cms^{-\frac{1}{2}}$ (i.e. $D_0 = 1cm^2s^{-1}$) for the case in which the process does not start at the equilibrium, while in the situation of equilibrium initial conditions the stationary probability density function is taken from [38]. Simulating the process several times, it is possible to extract the variance and kurtosis which can be confronted with the results gained in the previous section.

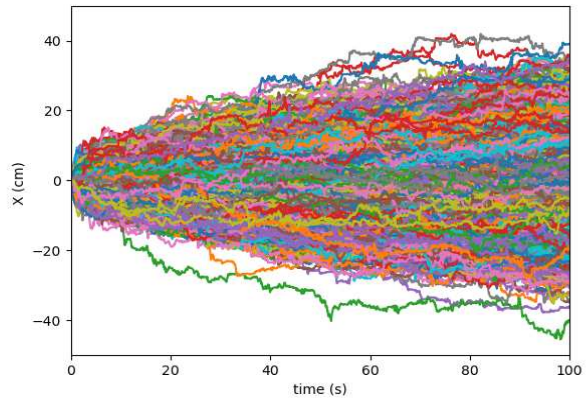

Figure 3.6: 100000 simulations of the diffusing diffusivity model for the Ornstein-Uhlenbeck process with
$D_0 = 1cm^2s^{-1}$, $\sigma = 1cms^{-1}$ and $\tau = 1s$.

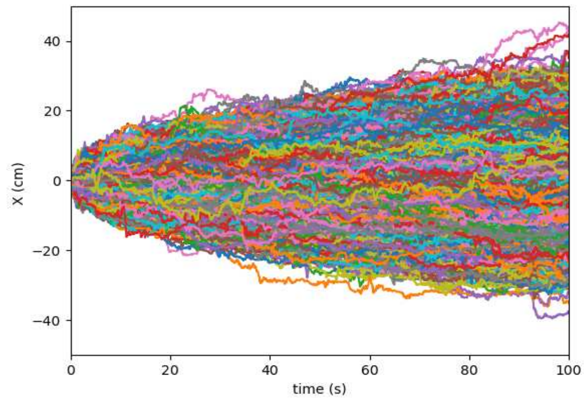

Figure 3.7: 100000 simulations of the diffusing diffusivity for the Ornstein-Uhlenbeck process with $\sigma = 1cms^{-1}$ and $\tau = 1s$ with equilibrium initial conditions.

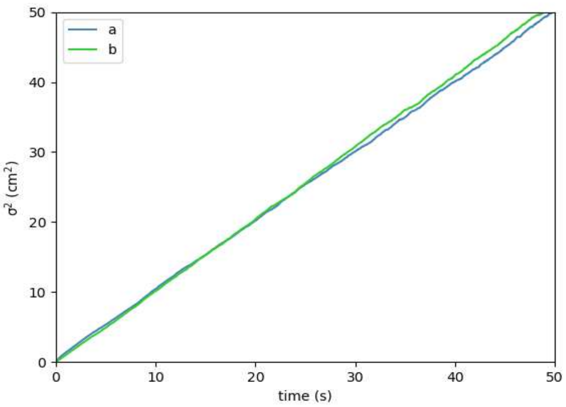

Figure 3.8: Variance computed from the paths displayed in the graphs above with a and b referring respectively to Fig.[3.6] and Fig.[3.7].

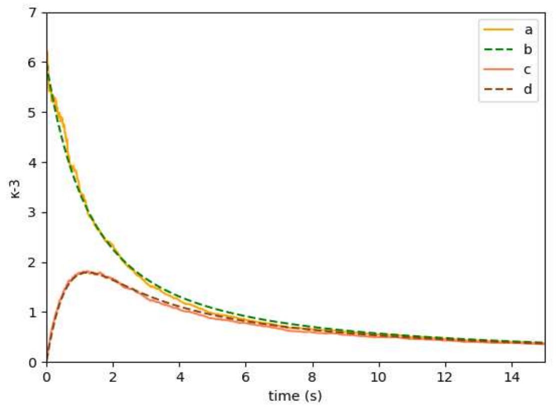

Figure 3.9: Reduced kurtosis excess computed from the paths displayed in the graphs above (solid lines) and from the result gain in Section 3.2.3 (dashed lines). In particular c, d and e,f refer respectively to the cases in which the Ornstein-Ulhenbeck process starts at the equilibrium and when it does not.

Fig.s [3.6] and [3.7] display how the diffusing diffusivity model is expected to disperse with time and they exhibit a position distribution symmetric with respect to the axes origin, thus suggesting a zero mean for every time as already eq. (3.1.5) predicted. The difference between the two graphs is the dispersion widths, which become more similar as increasing the time.
In Fig. [3.9] it possible to see that the reduced excess kurtosis from the simulations are in accordance with what found in Section 3.2.3. As it was expected, the diffusing coefficient initial equilibrium kurtosis differs from the non-equilibrium one in the short time range, while the curves tend to overlap for large times.
Both the simulated situations show the variance to seem linearly dependent on time as indeed the previous section pointed out. To test this dependence, linear fits are done in the two different regimes[1].

[1]for the fit on short time new simulations (as many as before) are done with a better position precision.

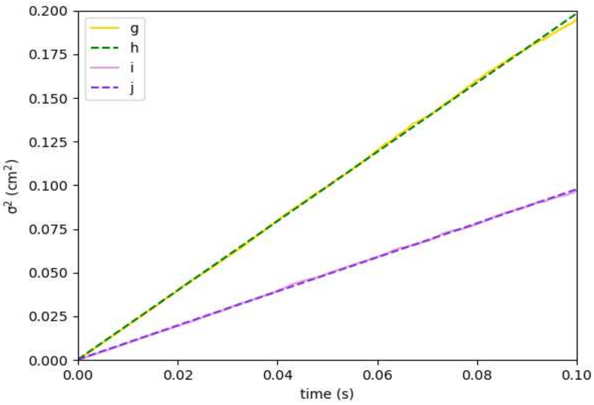

Figure 3.10: Linear fit on the variance for the short time regime with h and j referring respectively to Fig.[3.6]and Fig.[3.7].

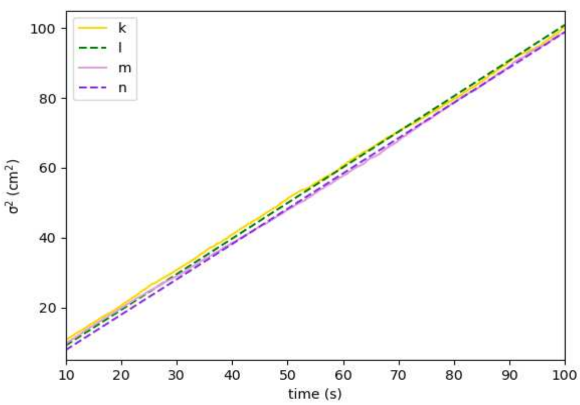

Figure 3.11: Linear fit on the variance for the long time regime with l and n referring respectively to Fig.[3.6]and Fig.[3.7].

$$h(t) = t, \qquad j(t) = 0,5 + t; \qquad\qquad l(t) = n(t) = t. \tag{3.2.57}$$

The fits in the figures above confirm the linear dependence already found in the previous section. In the short time regime, the two different cases show different coefficients (in accordance with eq.s (3.2.40) and (3.2.46)), while in the long time limit the initial condition is forgotten and the curves overlap.

3.3 Two-State Polymer Model

The *two-state polymer model* is a model for describing polymers which shift configurations among two possible ones. The two shapes have different diffusion coefficient and their changing in time affects the way the polymer spreads in the medium. Indeed the dispersion dependence of the diffusive particle structure can be evinced rewritten the Einstein-Smoluchowski (1.1.4) relation through the *Einstein-Stokes relation* [1]:

$$D = \frac{k_B T}{6\pi\eta r} \tag{3.3.1}$$

where η is the viscosity and r the particle radius. A common situation in which a polymers change shape is when it passes from a stretched to a relaxed configuration. Given the different diffusion coefficients and the probability of flipping between them, the situation can be treated by means of the Markov chains. In particular, since the possible configurations are two, we will use the so-called *Telegraph Model*, whose named is due to its historical origin (indeed it was designed to described the telegraph signal in order to get information about the noise).
In the following it is used the same approach exploited in the diffusing diffusivity model, i.e. we start from the diffusivity randomness analysis for then exploit the obtained results to the subordinator parameterization.

3.3.1 Markov Dynamics

The stochasticity of diffusivity having two possible states is well described by Markov chains. Denoting as s_1 and s_2 the diffusivity states corresponding respectively to D_1 and D_2, the general forms of stochastic matrix and probability ket in this case are:

$$\Pi = \begin{bmatrix} \pi(1|1) & \pi(2|1) \\ \pi(1|2) & \pi(2|2) \end{bmatrix} = \Delta t \begin{bmatrix} r_{11} & r_{12} \\ r_{21} & r_{22} \end{bmatrix} = \Delta t \begin{bmatrix} r_{11} & \frac{1}{\Delta t} - r_{22} \\ \frac{1}{\Delta t} - r_{11} & r_{22} \end{bmatrix}, \tag{3.3.2}$$

$$|\rho_n\rangle = \begin{bmatrix} \rho_{n,1} \\ \rho_{n,2} \end{bmatrix}, \tag{3.3.3}$$

where again, as in section 1.3, we used $\pi(i|j)$ to indicate the $\pi(s_i|s_j)$ and in the last passage of (3.3.2) we exploited (1.3.4). The Π eigenvalues and eigenvectors are found to be:

$$\lambda_1 = 1, \qquad \lambda_2 = (r_{11} + r_{22})\Delta t - 1, \tag{3.3.4}$$

$$|v_1\rangle = \frac{1}{1+\Lambda}\begin{bmatrix} \Lambda \\ 1 \end{bmatrix}, \qquad |v_2\rangle = \frac{1}{\sqrt{2}}\begin{bmatrix} 1 \\ -1 \end{bmatrix} \qquad \text{with} \quad \Lambda = \frac{1 - r_{22}\Delta t}{1 - r_{11}\Delta t} = \frac{r_{21}}{r_{12}}. \tag{3.3.5}$$

As mentioned in section 1.3, the stochastic matrix biggest eigenvalue corresponds to 1 and its multeplicity is 1 providing that Π is diagonalizable. For $\lambda_1 = \lambda_2 = 1$ the stochastic matrix has null determinant so its diagonalization is not possible. This occurs if and only if $\pi_{11} = \pi_{22} = 1$ $(0 \leq \pi_{11}, \pi_{22} \leq 1)$ bringing to $\Pi = \mathbb{1}_{2\times 2}$ already diagonal. This situation describes a stochastic process no longer varying in time since in this case the probability distribution is the same for every time (i.e. $|\rho_n\rangle = |\rho_0\rangle \ \forall n$). The diagonalization of Π leads to the following results:

$$\Pi_D = \begin{bmatrix} \lambda_0 & 0 \\ 0 & \lambda_1 \end{bmatrix}, \tag{3.3.6}$$

$$D = \begin{bmatrix} \Lambda & 1 \\ 1 & -1 \end{bmatrix} \qquad D^{-1} = \frac{1}{\Lambda + 1}\begin{bmatrix} 1 & 1 \\ 1 & -\Lambda \end{bmatrix}. \tag{3.3.7}$$

Since

$$\lim_{n\to\infty} {\Pi_D}^n = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \tag{3.3.8}$$

recalling eq. (1.3.11), the long time regime probability ket is found to be:

$$\lim_{n\to\infty} |\rho_n\rangle = \frac{1 - r_{11}\Delta t}{2 - (r_{11} + r_{22})\Delta t}\begin{bmatrix} \Lambda \\ 1 \end{bmatrix} = \frac{1}{1+\Lambda}\begin{bmatrix} \Lambda \\ 1 \end{bmatrix}, \tag{3.3.9}$$

which exactly corresponds to the eigenvector related the unit eigenvalue as predicted in section 1.3. If we now move to the continuous time representation, the eigenvalues of the generating matrix G related to Π are:

$$\gamma_1 = 0 \qquad \gamma_2 = -(r_{11} + r_{22}), \tag{3.3.10}$$

where for γ_2 evaluation it is used the limit:

$$\lim_{x\to 0} \frac{1}{x} \ln\Big(ax - 1\Big) = -ax. \tag{3.3.11}$$

Decomposing the initial probability ket with G eigenvectors basis

$$|\rho_0\rangle = \tilde{\rho}_{0,1}\,|v_1\rangle + \tilde{\rho}_{0,2}\,|v_2\rangle \tag{3.3.12}$$

the probability evolution in the continuous time is described by the ket:

$$|\rho_t\rangle = e^{Gt}\,|\rho_0\rangle = \tilde{\rho}_{0,1}\,|v_1\rangle + \tilde{\rho}_{0,2}e^{\gamma_2 t}\,|v_2\rangle\,. \tag{3.3.13}$$

Notice that in the long time regime $|\rho_t\rangle$ tends to $|v_1\rangle$ both in the continuous and discrete representation and $|v_1\rangle$ corresponds to the stationary distribution, i.e. if

$$|\rho_0\rangle = |v_1\rangle \qquad \Longrightarrow \qquad |\rho_t\rangle = |v_1\rangle \quad \forall t. \tag{3.3.14}$$

This means that the diffusion coefficient mean for long times tends to the the stationary average $\langle D\rangle^*$ which is [3]:

$$\langle D\rangle^{LT} = \langle D\rangle^* = \frac{\Lambda}{1+\Lambda}D_1 + \frac{1}{1+\Lambda}D_2 = \frac{(D_1 r_{21} + D_2 r_{12})}{(r_{12}+r_{21})}t \tag{3.3.15}$$

Recalling now that the probability conserving condition imposes that for any time the probability ket contraction with $\langle 1| = [1,1]$ has to return 1, the following difference has to be null:

$$\langle 1|\rho_0\rangle - \langle 1|\rho_t\rangle = \sum_{i=1,2} \tilde{\rho}_{0,i}(1-e^{\gamma_i t})\langle 1|v_i\rangle = \tilde{\rho}_{0,2}(1-e^{\gamma_2 t})\langle 1|v_2\rangle = 0. \tag{3.3.16}$$

Recalling (3.3.5), the upper relation is found to be always satisfied. Notice that the $\langle 1|\rho_n\rangle = 1$ condition implies $\tilde{\rho}_{0,1} = 1$. Anyway this does not constrain the initial probability since we are still free in the $\tilde{\rho}_{0,2}$ choice and $|v_2\rangle$ is the one regularizing the difference between the two states initial probabilities. Indeed the only restriction for $\tilde{\rho}_{0,2}$ is given by the probabilities of starting in s_1 or s_2 expressed in $\{|v_1\rangle, |v_2\rangle\}$ basis:

$$\rho^0{}_1 = \frac{\Lambda}{1+\Lambda} + \frac{\tilde{\rho}_{0,2}}{\sqrt{2}}, \tag{3.3.17}$$

$$\rho^0{}_2 = \frac{1}{1+\Lambda} + \frac{\tilde{\rho}_{0,2}}{\sqrt{2}}, \tag{3.3.18}$$

whose condition $|\rho_{0,1} - \rho_{0,2}| \leq 1$ imposes

$$-\sqrt{2}\frac{\Lambda}{\Lambda+1} \leq \tilde{\rho}_{0,2} \leq \sqrt{2}\frac{1}{\Lambda+1}. \tag{3.3.19}$$

Let us consider the case in which $r_{11} = r_{22} \equiv r$, which corresponds to the stochastic matrix:

$$\Pi_{\Delta t} = \begin{bmatrix} r\Delta t & 1-r\Delta t \\ 1-r\Delta t & r\Delta t \end{bmatrix}. \tag{3.3.20}$$

The eigenvalues and eigenvectors are then:

$$\lambda_0 = 1, \qquad \lambda_1 = 2r\Delta t - 1 \quad \Longleftrightarrow \quad \gamma_0 = 0, \qquad \gamma_1 = \frac{1}{\Delta t}\ln\Big(2r\Delta t - 1\Big), \tag{3.3.21}$$

$$|v_1\rangle = \frac{1}{\sqrt{2}}\begin{bmatrix}1\\1\end{bmatrix}, \qquad |v_2\rangle = \frac{1}{\sqrt{2}}\begin{bmatrix}1\\-1\end{bmatrix}. \tag{3.3.22}$$

In this case the only parameter controlling the Markov chain is the transition rate r. In particular (3.3.20) shows that the transition rates related to remaining in a state or changing it are the same whether we consider state 1 or 2.
As done in the previous section, the probability of founding the diffusivity in state 1 or 2 at time t is

$$\mathbb{P}(D(t) = s_{1,2}|\rho_0) = \langle s_{1,2}|e^{Gt}|\rho_0\rangle = \sum_{i=\{0,1\}} \tilde{\rho^0}_i e^{\gamma_i t}\langle s_{1,2}|v_i\rangle = \frac{1}{\sqrt{2}}\langle s_{1,2}|v_1\rangle + \tilde{\rho^0}_2 e^{\gamma_2 t}\langle s_{1,2}|v_2\rangle \tag{3.3.23}$$

Since $e^{\gamma_2} \leq 1$, for large times the $\langle s_{1,2}|v_2\rangle$ term in (3.3.23) vanishes and the $D(t)$ process reduces to the unbiased random walk:

$$\lim_{t\to\infty} \mathbb{P}(D(t) = s_1|\rho_0) = \langle s_1|v_1\rangle = \frac{1}{2}, \tag{3.3.24}$$

$$\lim_{t\to\infty} \mathbb{P}(D(t) = s_2|\rho_0) = \langle s_2|v_1\rangle = \frac{1}{2}. \tag{3.3.25}$$

3.3.2 Cumulants

Now that the diffusivity has been analyzed, the $X(t)$, in particular the excess kurtosis, process can be studied. The procedure is the same employed in the diffusing diffusivity model: the $X(t)$ kurtosis is derived by means of the subordinator $S(t)$ through the relation (3.1.7). For this derivation we will start again with discrete time for then moving to the continuous sending $\Delta t \to 0$. The integral in the subordinator definition becomes a sum when discretizing the time:

$$S(t) = 2\sum_{i=1}^{n} D(\omega_i)\Delta t \qquad \text{with } t = n\Delta t. \tag{3.3.26}$$

Recalling that in our case $\omega_i = \{s_1, s_2\}$. The moment generating function for the $S(t)$ process is:

$$M_{S,\Delta t}(k,t) = \langle e^{kS(t)}\rangle \tag{3.3.27}$$

$$= \sum_{\omega_0=\{s_1,s_2\}} \sum_{\omega_1=\{s_1,s_2\}} \cdots \sum_{\omega_n=\{s_1,s_2\}} e^{2k\sum_{i=0}^{n}\omega_i\Delta t} \times$$
$$\times\, \pi_{\Delta t}(\omega_n|\omega_{n-1})\pi_{\Delta t}(\omega_{n-1}|\omega_{n-2})\ldots\pi_{\Delta t}(\omega_1|\omega_0)\rho(\omega_0). \tag{3.3.28}$$

In order to get a manageable expression for $M_{S,\Delta t}(k,t)$ we introduce the so called *titled transition matrix* [41]:

$$(\hat{\Pi}_{\Delta t})_{\omega_i,\omega_{i-1}} = \pi_{\Delta t}(\omega_i|\omega_{i-1})e^{2kD(\omega_i)\Delta t}, \tag{3.3.29}$$

which can be factorized with two exponentials:

$$\hat{\Pi}_{\Delta t} = e^{G\Delta t}e^{2kD\Delta t} \qquad \text{with} \quad D = \begin{bmatrix} D_1 & 0 \\ 0 & D_2 \end{bmatrix}. \tag{3.3.30}$$

The matrix notation, as in the previous section, incorporates the $\{\omega_i\}$ sums appearing in (3.3.28) so the generating function becomes:

$$M_{S,\Delta t}(k,t) = \sum_{\omega_0=\{s_1,s_2\}} \sum_{\omega_n=\{s_1,s_2\}} e^{2kD(\omega_0)}(\hat{\Pi}_{\Delta t})^n{}_{\omega_n,\omega_0}\rho(\omega_0). \tag{3.3.31}$$

We are now almost ready for taking the continuum limit and the last thing to do is to work on $\hat{\Pi}$. For sufficiently small Δt the titled transition matrix can be expanded as:

$$\hat{\Pi}_{\Delta t} = \Big(\mathbb{1} + G\Delta t + O(\Delta t^2)\Big)\Big(\mathbb{1} + 2kD\Delta t + O(\Delta t^2)\Big) = \mathbb{1} + \Big(G + 2kD\Big)\Delta t + \Big(O(\Delta t^2)\Big). \tag{3.3.32}$$

Remembering that $n = \frac{t}{\Delta t}$ and exploiting the limit

$$\lim_{\Delta t \to \infty} \Big(\mathbb{1} + \Big(G + 2kD \Big) \Delta t + \Big(O(\Delta t^2) \Big) \Big)^{\frac{t}{\Delta t}} = e^{(G+2kD)t}, \tag{3.3.33}$$

the continuous time representation for the generating function is obtained:

$$\begin{aligned} M_S(k,t) &= \lim_{\Delta t \to \infty} \sum_{\omega_0 = \{s_1, s_2\}} \sum_{\omega_L = \{s_1, s_2\}} e^{(G+2kD)t} \rho(\omega_0) = \\ &= \Big\langle 1 \Big| e^{(G+2kD)t} \Big| \rho_0 \Big\rangle . \end{aligned} \tag{3.3.34}$$

In order to be able to derive $M_S(k,t)$ for getting $\langle S(t) \rangle$ and $\langle S(t)^2 \rangle$, it is convenient to rewrite (3.3.34) using $R \equiv G + 2kD$ eigenvalues and eigenvectors. The G matrix is extracted expanding at first order $e^{G\Delta t}$ and matching it with (3.3.4) for then sending $\Delta t \to 0$:

$$G = \begin{bmatrix} -r_{12} & r_{21} \\ r_{12} & -r_{21} \end{bmatrix} \qquad \text{with} \quad c = \lim_{\Delta t \to 0} \frac{1-\pi}{\Delta t}. \tag{3.3.35}$$

R, its eigenvalues and eigenvectors are then:

$$R = \begin{bmatrix} 2kD_1 - r_{12} & r_{21} \\ r_{12} & 2kD_2 - r_{21} \end{bmatrix}, \tag{3.3.36}$$

$$\eta_1(k) = \frac{1}{2}(2k(D_1 - D_2)) - r_{12} - r_{21} - q, \tag{3.3.37}$$

$$\eta_2(k) = \frac{1}{2}(2k(D_1 - D_2)) - r_{12} - r_{21} + q, \tag{3.3.38}$$

$$|n_1(k)\rangle = \begin{bmatrix} -\frac{1}{2r_{12}}(2k(D_1 - D_2)) - r_{12} - r_{21} - q \\ 1 \end{bmatrix}, \tag{3.3.39}$$

$$|n_2(k)\rangle = \begin{bmatrix} \frac{1}{2r_{12}}(2k(D_1 - D_2)) - r_{12} - r_{21} + q \\ -1 \end{bmatrix}, \tag{3.3.40}$$

with

$$q = \sqrt{(r_{12} + r_{21} - 2k(D_1 - D_2))^2 - 4(4k^2 D_1 D_2 - 2kD_2 r_{12} - 2kD_1 r_{21})}. \tag{3.3.41}$$

Since $R|_{k=0} = G$ (meaning also $\eta_i|_{k=0} = \gamma_i$ and $n_i|_{k=0} = v_i$) and recalling that for the evaluation of $S(t)$ moments $k = 0$ will be set, it is reasonable to decompose ρ_0 as:

$$|\rho_0\rangle = \frac{1}{\sqrt{2}||n_1(k)||} |n_1(k)\rangle + \frac{\tilde{\rho}_{0,2}}{||n_2(k)||} |n_2(k)\rangle, \tag{3.3.42}$$

where it is used $\tilde{\rho}_{0,1} = \frac{1}{\sqrt{2}}$. Inserting the latter in eq. (3.3.34) the $M_S(k,t)$ appropriate expression for deriving $\langle S(t) \rangle$ and $\langle S(t)^2 \rangle$ is obtained:

$$\begin{aligned} M_S(k,t) = &\frac{1}{\sqrt{2}} e^{t(\frac{1}{2}(2k(D_1 - D_2)) - r_{12} - r_{21} - q)} \left(\frac{(\frac{1}{2}(-2k(D_1 - D_2)) - r_{12} - r_{21} - q) + 1}{\sqrt{(\frac{1}{2}(-2k(D_1 - D_2)) - r_{12} - r_{21} - q) + 1}} \right) + \\ &+ \frac{1}{\sqrt{2}} e^{t(\frac{1}{2}(2k(D_1 - D_2)) - r_{12} - r_{21} + q)} \left(\frac{(\frac{1}{2}(2k(D_1 - D_2)) - r_{12} - r_{21} + q) + 1}{\sqrt{(\frac{1}{2}(2k(D_1 - D_2)) - r_{12} - r_{21} + q) + 1}} \right), \end{aligned} \tag{3.3.43}$$

$$\langle S(t)\rangle = \frac{dM_S(k,t)}{dk}\bigg|_{k=0} = -\frac{1}{r_{12}+r_{21}}\left(D_2 - D_1 + \frac{(r_{21}-r_{12})(D_2-D_1)}{r_{12}+r_{21}}\right)+ \\ + \tilde{\rho}_{0,2}e^{-t(r_{12}+r_{21})}\frac{\sqrt{2}}{2r_{12}}\left(D_2 - D_1 + \frac{(r_{12}-r_{21})(D_2-D_1)}{r_{12}+r_{21}}\right)+ \\ 2t\frac{(D_1 r_{21} + D_2 r_{12})}{(r_{12}+r_{21})}, \tag{3.3.44}$$

$$\langle S(t)^2\rangle = \frac{d^2M_S(k,t)}{dk^2}\bigg|_{k=0} = 4\sqrt{2}e^{-t(r_{12}+r_{21})}\tilde{\rho}_{0,2}\frac{r_{21}(D_1-D_2)^2}{(r_{12}+r_{21})^3} + 8\frac{r_{21}r_{12}(D_1-D_2)^2}{(r_{12}+r_{21})^4}+ \\ -4\sqrt{2}e^{-t(r_{12}+r_{21})}t\tilde{\rho}_{0,2}\frac{(r_{12}D_1 + r_{21}D_2)(D_1-D_2)}{(r_{12}+r_{21})^2} + 8t\frac{r_{12}r_{21}(D_1-D_2)^2}{(r_{12}+r_{21})^3}+ \\ +4t^2\frac{(D_1r_{21}+D_2r_{12})^2}{(r_{12}+r_{21})^2}. \tag{3.3.45}$$

Again due to eq. (3.1.5), the $S(t)$ mean correspond to $X(t)$ variance, whose time dependence is present with an exponential decay and a linear term, indicating the $X(t)$ diffusion to become more Fickian as increasing time:

$$\langle X(t)^2\rangle_{LT} = 2\frac{(D_1r_{21}+D_2r_{12})}{(r_{12}+r_{21})}t = 2\langle D\rangle^* t. \tag{3.3.46}$$

Notice that for the stationary case (3.3.14) $\tilde{\rho}_{0,2}$ vanishes so the exponential term in eq. (3.3.44) drops out and the $X(t)$ process undergoes Brownian diffusion for all times.
Substituting the results above in (3.1.7) the reduced excess kurtosis is obtained. Given the complex structure of the above results, the kurtosis evaluation is done with Mathematica [46]. Anyway it is possible to see at first sight that the kurtosis numerator does not contain quadratic terms in t, since the dependence on t^2 of (3.3.45) is the square root of (3.3.44) dependence on t. Note that for $D_1 = D_2$ (i.e. no shifting in the diffusion coefficient and so in the polymer shape) both the majority of the terms in eq.s (3.3.44) and (3.3.45) and the X_t variance and kurtosis correspond to the Gaussian-Brownian ones as expected.
Since there are a lot of parameters at play, it is interesting to look at how the kurtosis changes in function of them.

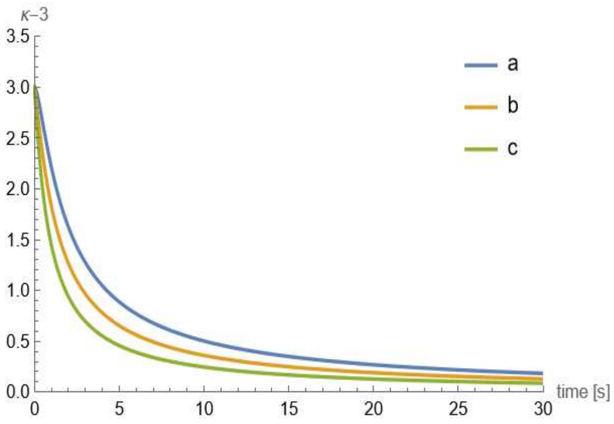

Figure 3.12: Excess kurtosis at varying of diffusing coefficient:
a: $\{D_1 = 5\cdot 10^{-9}m^2s^{-1},\ D_2 = 10^{-9}m^2s^{-1},\ r_{12} = 0.5s^{-1},\ r_{21} = 0.5s^{-1}, {\rho^0}_1 = {\rho^0}_2 = 0.5\}$;
b: $\{D_1 = 10^{-9}m^2s^{-1},\ D_2 = 10^{-10}m^2s^{-1},\ r_{12} = 0.5s^{-1},\ r_{21} = 0.5s^{-1}, {\rho^0}_1 = {\rho^0}_2 = 0.5\}$;
c :$\{D_1 = 5\cdot 10^{-9}m^2s^{-1},\ D_2 = 10^{-11}m^2s^{-1},\ r_{12} = 0.5s^{-1},\ r_{21} = 0.5s^{-1}, {\rho^0}_1 = {\rho^0}_2 = 0.5\}$.

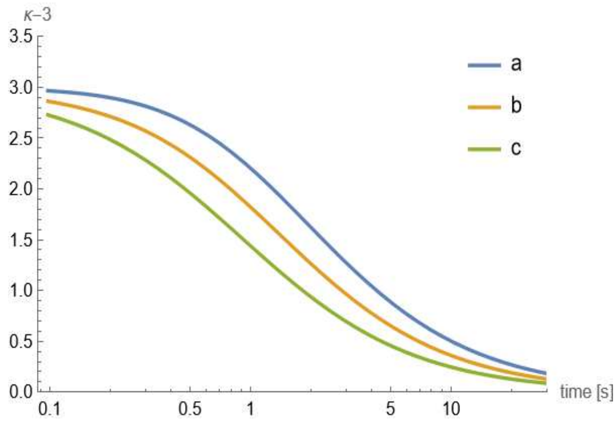

Figure 3.13: Excess kurtosis at varying of diffusing coefficient for the log scale in the time axis:
a: $\{D_1 = 5\cdot 10^{-9}m^2s^{-1},\ D_2 = 10^{-9}m^2s^{-1},\ r_{12} = 0.5s^{-1},\ r_{21} = 0.5s^{-1}, {\rho^0}_1 = {\rho^0}_2 = 0.5\}$;
b: $\{D_1 = 10^{-9}m^2s^{-1},\ D_2 = 10^{-10}m^2s^{-1},\ r_{12} = 0.5s^{-1},\ r_{21} = 0.5s^{-1}, {\rho^0}_1 = {\rho^0}_2 = 0.5\}$;
c :$\{D_1 = 5\cdot 10^{-9}m^2s^{-1},\ D_2 = 10^{-11}m^2s^{-1},\ r_{12} = 0.5s^{-1},\ r_{21} = 0.5s^{-1}, {\rho^0}_1 = {\rho^0}_2 = 0.5\}$.

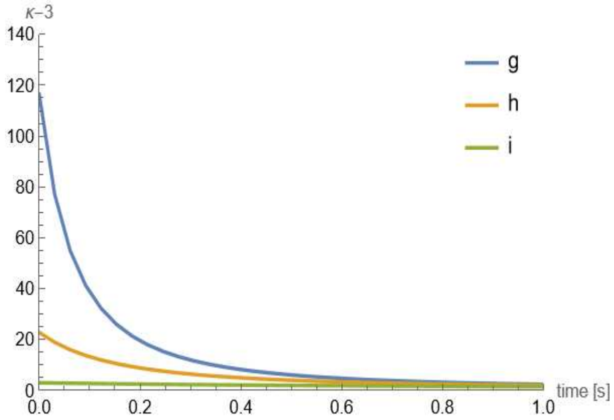

Figure 3.14: Excess kurtosis at varying of initial probabilities:
g: $\{D_1 = 5 \cdot 10^{-9} m^2 s^{-1}$, $D_2 = 10^{-9} m^2 s^{-1}$, $r_{12} = 0.5 s^{-1}$, $r_{21} = 0.5 s^{-1}$,$\rho^0{}_1 = \frac{\sqrt{2}+2}{6}$, $\rho^0{}_2 = \frac{-\sqrt{2}+3}{6}\}$;
h: $\{D_1 = 5 \cdot 10^{-9} m^2 s^{-1}$, $D_2 = 10^{-9} m^2 s^{-1}$, $r_{12} = 0.5 s^{-1}$, $r_{21} = 0.5 s^{-1}$,$\rho^0{}_1 = \frac{\sqrt{2}+1}{2\sqrt{2}}$, $\rho^0{}_2 = \frac{\sqrt{2}-1}{2\sqrt{2}}\}$;
i: $\{D_1 = 5 \cdot 10^{-9} m^2 s^{-1}$, $D_2 = 10^{-9} m^2 s^{-1}$, $r_{12} = 0.5 s^{-1}$, $r_{21} = 0.5 s^{-1}$,$\rho^0{}_1 = \rho^0{}_2 = 0.5\}$.

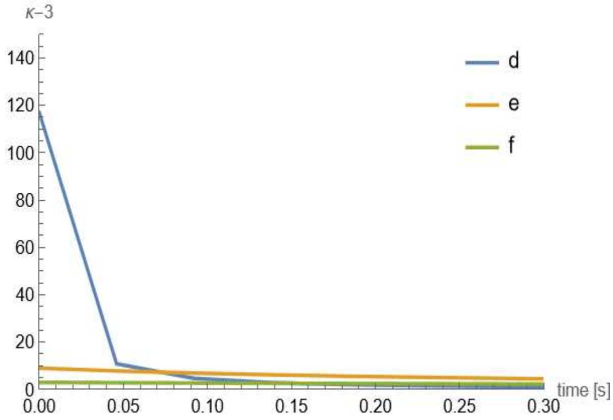

Figure 3.15: Excess kurtosis at varying of transition rates:
d: $\{D_1 = 5 \cdot 10^{-9} m^2 s^{-1}$, $D_2 = 10^{-9} m^2 s^{-1}$, $r_{12} = 10 s^{-1}$, $r_{21} = 0.5 s^{-1}$,$\rho^0{}_1 = \rho^0{}_2 = 0.5\}$;
e: $\{D_1 = 5 \cdot 10^{-9} m^2 s^{-1}$, $D_2 = 10^{-9} m^2 s^{-1}$, $r_{12} = 0.5 s^{-1}$, $r_{21} = 0.5 s^{-1}$,$\rho^0{}_1 = \rho^0{}_2 = 0.5\}$;
f: $\{D_1 = 5 \cdot 10^{-9} m^2 s^{-1}$, $D_2 = 10^{-9} m^2 s^{-1}$, $r_{12} = 0.5 s^{-1}$, $r_{21} = 0.5 s^{-1}$,$\rho^0{}_1 = \rho^0{}_2 = 0.5\}$.

The above figures display similar kurtosis form for every parameter choice. However the latter affects the peak height and broadness. In Fig. [3.12] the curves are observed to widen as the difference between D_1 and D_2 increases and Fig. [3.15] shows the peak to rise when $\tilde{\rho}_{0,2}$ grows. The peak broadness is also affected by r_{12} and r_{12}, indeed the process is expected to accelerate as the transition rates increase. Note that both in Fig. [3.12] and Fig. [3.14] $\rho_{0,1} = \rho_{0,2} = 0.5$ was set and the same curves would have been obtained if D_1 and D_2 values were exchanged one another. This is not surprising since when studying the telegraph model no preference was made for any state. Indeed in eq. (3.3.44) and eq. (3.3.45) the only term not invariant under D_1 and D_2 interchange contains $\tilde{\rho}_{0,2}$, which regulates the initial probability distribution, and $e^{-t(r_{12}+r_{21})}$ so it is going to fade as the time increases coherently with the fact that for long times the initial distribution is forgotten.
Moreover, from Fig. [3.13] it is possible to see that the stationary distribution (which corresponds to $\rho^0{}_1 = \rho^0{}_2 = 0.5$ with the used parameters) has a remarkable constant kurtosis trend with with respect the other ones and this suggests to search for an anomalous scaling.

3.4 Simulation

As done for the diffusing diffusivity model, the results found in the previous section are checked through a simulation taking into account the stationary case (3.3.14) assuming that the diffusion process starts at the axis origin.

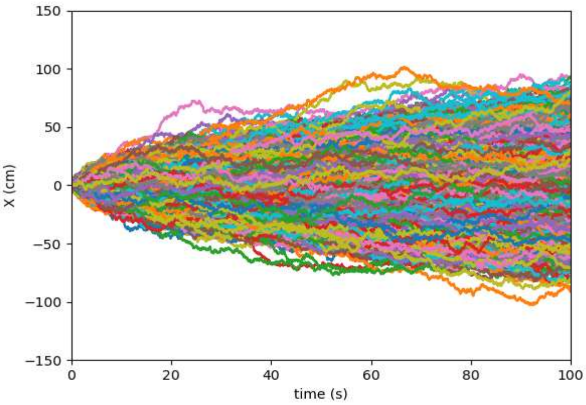

Figure 3.16: 100000 simulations for the two state polymer model with $\{D_1 = 5 \cdot 10^{-4} m^2 s^{-1}$, $D_2 = 10^{-4} m^2 s^{-1}$, $r_{12} = 50 s^{-1}$, $r_{21} = 50 s^{-1}$,${\rho^0}_1 = {\rho^0}_2 = 0.5\}$.

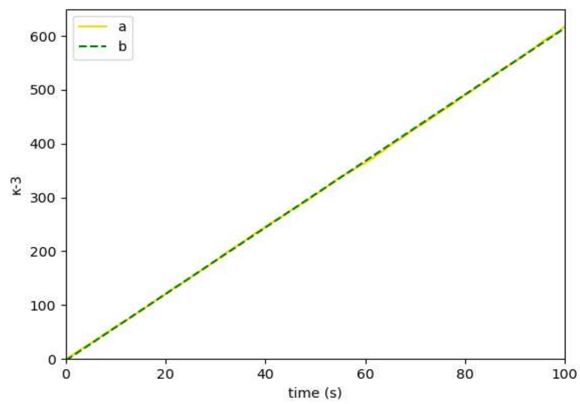

Figure 3.17: Variance associated to the simulation of Fig. [3.16] (a) and linear fit on it (b).

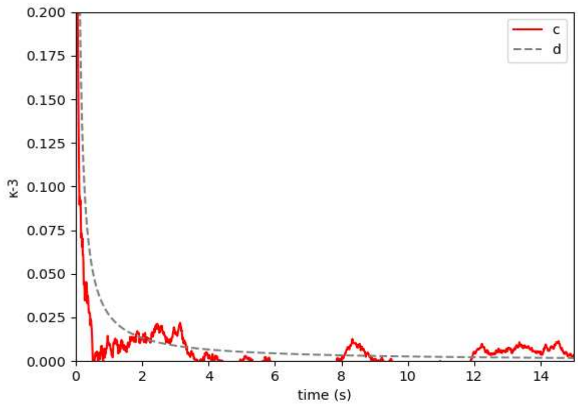

Figure 3.18: Kurtosis excess obtained from the simulation of Fig. [3.16] (c) and from the results of Section 3.3.2 .

$$b(t) = 6t. \tag{3.4.1}$$

Also in this case the simulated diffusion is symmetrical with respect to the initial position, in according with the fact that processes which admits subordinator parameterization have odd central moments vanishing (3.1.5).

As expected from the previous section, the variance show a linear dependence on time at first sight and indeed the fit slope corresponds to the one found inserting the parameters used for the simulations in eq. (3.3.44). In (3.4.1) there is not present a constant term predicted by (3.3.44) of order 10^{-2} due to the precision of the simulation. The latter affects also the kurtosis computation in Fig. [3.18] where, even if it is possible to see the simulation kurtosis following the trend delineated by the results of Section (3.3.2), it is suggested an increased number of simulations which we cannot done because of computational impossibility.

3.5 Discussion

Even though the diffusing diffusivity and the two-state polymer models describes two different diffusing processes, they share some interesting features which comes from common origin. The fact that the two processes comes from the same subordination structure already indicates that both inherit some general features. Also the cumulants studied in sections 3.2.3 and 3.3.2 exhibit a similar behavior. It is thus interesting to investigate on how and why shared characteristics come from and what is their meaning and relation.

3.5.1 Brownian diffusion

The first common feature to highlight is the predominance of the Brownian diffusion in both the process, since both the variances (3.2.45) and (3.3.44) contains a decaying exponential term and a linear one in t. Moreover, in case of initial stationary conditions the exponential term vanishes for both the processes, meaning that the Brownian or Fickian diffusion occurs for all times. Their Brownian motion is also suggested by their kurtosis (Fig.s[3.3-3.5] and Fig.s[3.12-3.14]) which indicate mesokurtic probability density functions for long times. Anyway, even without knowing the probability density function, we can exclude the possibility of Gaussian-like probabilities for all times, since there is a range in Fig.s[10][11] and Fig.[3.1] where the kurtosis indicate leptokurtic probability distributions. Those are then situations of *Brownian yet Non-Gaussian diffusion* [20] [21] [36] [37] [38] [39].

3.5.2 Long Time Regime and Large Deviations

For the long time regime the stationary of the asymptotic diffusion coefficient probability density function makes X_t probability tends to Gaussian propagator. In order to prove this, let us start analyzing the $\langle S(t)^2\rangle$. Managing the integrals domain as already done in eq.s (3.2.42) and (3.2.47), we have that the $S(t)$ 2-nd can be written as:

$$\begin{aligned}\langle S(t)^2\rangle &= 8\int_0^t dt_2\int_0^{t_2} dt_1\langle D(t_2)D(t_1)\rangle \\ &= 8\int_0^t dt_2\int_0^{t_2} dt_1\int_0^\infty dD_2\int_0^\infty dD_1 p(D_2,t_2|D_1,t_1)p(D_1,t_1|D_0,0)D_1D_2, \end{aligned}\tag{3.5.1}$$

where in (3.5.1) it was exploited the Markovian nature belonging both to the Ornstein–Uhlenbeck process and telegraph model. Let us consider now a time t^* large enough to permit to consider

$$p(D,t|D_0,0)\approx p^*(D) = p_D(D)_{LT} \qquad \forall t\geq t^*, \tag{3.5.2}$$

where $p_D(D)_{LT}$ refers to eq. (3.2.39) and $|v_1\rangle$ in (3.3.5) respectively for the diffusing diffusivity model and the two state polymer model. Exploiting (3.5.2), the integral (3.5.1) can be split as:

$$\langle S(t)^2\rangle = 8\int_0^{t^*} dt_2\int_0^{t_2} dt_1\int_0^\infty dD_2\int_0^\infty dD_1 p(D_2,t_2|D_1,t_1)p(D_1,t_1|D_0,0)D_2D_1+ \tag{3.5.3}$$

$$+8\int_{2t^*}^t dt_2\int_0^{t^*} dt_1\int_0^\infty dD_2\int_0^\infty dD_1 p(D_2,t_2|D_1,t_1)p(D_1,t_1|D_0,0)D_2D_1+ \tag{3.5.4}$$

$$+8\int_{2t^*}^t dt_2\int_{t_2-t^*}^{t_2} dt_1\int_0^\infty dD_2\int_0^\infty dD_1 p(D_2,t_2|D_1,t_1)p(D_1,t_1|D_0,0)D_2D_1+ \tag{3.5.5}$$

$$+8\int_{2t^*}^t dt_2\int_{t^*}^{t_2-t^*} dt_1\int_0^\infty dD_2\int_0^\infty dD_1 p_D(D_1)_{LT}p_D(D_2)_{LT}D_2D_1 \tag{3.5.6}$$

Notice that among the previous integrals the only one having a quadratic term in time is (3.5.6), which corresponds to:

$$8\int_{2t^*}^{t} dt_2 \int_{t^*}^{t_2-t^*} (\langle D\rangle_{LT})^2 = \big(2\langle D\rangle_{LT}(t-t^*)\big)^2 \tag{3.5.7}$$

The extension of the reasoning just done to the general n-th moment case (explained in detail in Appendix B.2) leads to:

$$\langle S(t)^n\rangle = C(t,t^*) + \big(2\langle D\rangle_{LT}(t-t^*)\big)^n, \tag{3.5.8}$$

where $C(t,t^*)$ is a t polynomial of order smaller than n. Sending now t to infinity we have that:

$$\lim_{t\to\infty}\langle\Big(\frac{S(t)}{t}\Big)^n\rangle = \big(2\langle D\rangle_{LT}\big)^n. \tag{3.5.9}$$

As we have said in section 1.4, the knowledge of all moments fully characterizes a distribution and since eq. (3.5.9) shows that $\bar{s}(t) = \frac{S(t)}{t}$ has the same moments of $2\langle D\rangle_{LT}$ in the long time limit they are identically distributed:

$$\big(\frac{S(t)}{t}, t\to\infty\big) \overset{d}{=} \big(2\langle D\rangle_{LT}, t\to\infty\big). \tag{3.5.10}$$

Anyway $\langle D\rangle_{LT}$ is constant, which means that $\bar{s}$ is no more aleatory for large times and its probability density function corresponds then to a Dirac delta:

$$p_{\bar{s}}(\bar{s})_{LT} = \delta(\bar{s} - 2\langle D\rangle_{LT}), \tag{3.5.11}$$

which rewritten in terms of $S(t)$ becomes:

$$p_S(s,t)^{LT} = \delta(s - 2\langle D\rangle_{LT}t). \tag{3.5.12}$$

Inserting now 3.5.12 in the subordination formula (3.1.3) returns the $X(t)$ probability density function:

$$p_X(x,t|x_0)_{LT} = \int_0^\infty ds\frac{e^{\frac{-(x-x_0)^2}{2s}}}{\sqrt{2\pi s}}\delta(s-2\langle D\rangle t) = \frac{1}{\sqrt{4\pi\langle D\rangle_{LT}t}}\exp\Big\{\frac{-(x-x_0)^2}{4\langle D\rangle_{LT}t}\Big\}. \tag{3.5.13}$$

The probability densities above is indeed a Gaussian curve as suggested by the kurtosis in Sections 3.2.3 and 3.3.2 and its variance corresponds to the ones found in the previous Sections (eq.s (3.2.46) and (3.3.46)). Moreover, confronting eq.s (3.5.13) and (1.1.2) we can see that in the long time regime these models diffusions behave like a Gaussian-Brownian motion with diffusion coefficient $\langle D\rangle_{LT}$. Eq. (3.5.13) is also in agreement with the X_t long time regime probability density found in (3.2.54) in the diffusing diffusivity case. It is relevant to stress that the recovery of the Brownian motion is due to the asymptotic stationary of the D probability density function which permits to factorize in n independent terms the average on the n Ds.
In eq. (3.5.9) it is possible to see the structure of large deviation (2.1.5) where t^{-1} is the asymptotic parameter and the rate is:

$$J(x) = \frac{(x-x_0)^2}{4\langle D\rangle_{LT}}. \tag{3.5.14}$$

We thus expect the $X(t)$ probability to be peaked in the initial position x_0 for long times. This is indeed in agreement with (3.5.13) since, recalling eq. (3.2.36):

$$\lim_{t\to\infty} p_X(x,t|x_0)_{LT} = \delta(x-x_0). \tag{3.5.15}$$

3.5.3 Short Time and Anomalous Scaling

In the case of diffusing diffusivity model with the diffusion coefficient starting from initial equilibrium we have found that the kurtosis in Fig.[3.12] exhibits a constant trend in the short time regime, which indicates a leptokurtic distribution, but the variance corresponds to a Fickian diffusion. A very same behavior is present in the two-state polymer model in the case in which the diffusion coefficient starts with the stationary probability. We have then that both the models display anomalous scaling for short times.
Even though the initial D probability density functions are different for the two analyzed models, it is interesting to see if their probability structures have something in common in this regime and if they corresponds the probability form of (1.5.1). Since for the diffusing diffusivity model we already prove that the short time regime D does not depend on time (eq. (3.2.37)) and the same is valid also for the two-state polymer model given that D starts with the stationary ket (3.3.14), the subordinator (3.1.1) has a linear time dependence:

$$S(t) = 2Dt, \tag{3.5.16}$$

which means that, indicating with $p_{D^*}(D)$ the D probability density for which the anomalous scaling is observed, the $S(t)$ probability has the form:

$$p_S(s,t)_{ST} = \int_0^\infty dD\delta(s-2Dt)p_{D^*}(D) = \frac{1}{2t}p_{D^*}(\frac{s}{2t}), \tag{3.5.17}$$

or equivalently:

$$p_{\bar{S}}(\bar{s},t)_{ST} = \frac{1}{2}p_{D^*}(\frac{\bar{s}}{2}), \tag{3.5.18}$$

where with . Eq. (3.5.20) has thus the same structure encountered in eq. (1.5.1) and because of this, recalling 1.5, $S(t)$ inherits self-similarity.

3.5.4 Anomalous Scaling and Large Deviation

For the diffusing diffusivity model for long times it is possible to prove [37] that, thanks to a Large Deviation principle [27], the subordinator probability distribution is an exponential decay:

$$p_S(s,t) \asymp e^{-t^\alpha J(\frac{s}{t})} \qquad \text{with} \quad \alpha > 0. \tag{3.5.19}$$

On the other hand, if the model also displays anomalous scaling for short times with $\bar{S}(t)$ probability density function of the type (1.5.1), then eq. (3.5.20) can be generalize with:

$$p_{\bar{S}}(\bar{s},t)_{ST} = \frac{1}{t^{\nu-1}}f(\frac{\bar{s}}{t^{\nu-1}}). \tag{3.5.20}$$

In order for the two probability (3.5.20) and (3.5.20) to be compatible we must have:

$$f\left(\frac{\bar{s}}{t^{\nu-1}}\right) = \exp\left\{-\left(\frac{\bar{s}}{t^{\nu-1}}\right)^\gamma\right\}, \tag{3.5.21}$$

with the exponents being related by:

$$\alpha = \gamma(1 - \nu). \tag{3.5.22}$$

Conclusions

In Chapter 1 we have started illustrating the tools which were then exploited to analyzed the stochastic diffusivity processes in Chapter 3. Since a crucial part in the stochastic diffusivity models analysis regards the confront with the Gaussian-Brownian motion, the latter was described with its main features. Then we briefly illustrated the Langevin equation with the Fokker-Planck one and the Markov chains which were then used to analyzed the diffusion coefficient randomness respectively for the diffusing diffusivity model and for the two state polymer model. It was then introduced an alternative to the probability distribution for the stochastic process characterization which consists in the generating functions and the moments. In particular it was highlighted the roles, successively exploited, of the kurtosis as Gaussianity indicator and of the variance time dependence which is used to subdivide the diffusion types. Then we moved to study two important ingredients for the Chapter 3 analysis which are the anomalous scaling with its relation with self-similarity and the large deviation theory (Chapter 2) used to analyze probability distributions exhibiting an exponential decay.
Finally we arrived at the core of thesis in Chapter 3 where it was firstly illustrated the subordinator approach used for treating stochastic diffusivity models. The first analyzed model was the diffusing diffusivity one for which, after studying the diffusion coefficient randomness, the variance and the kurtosis were evaluated. Here it was found that the diffusion is Fickian both for long and short times and the probability density distribution tends to a mesokurtic curve for large times independently on which initial probability is chosen for the diffusion coefficient. The study of the probability distribution revealed that the long time it has the form of a Gaussian propagator where the diffusion coefficient is substituted with its long time average (which is time independent) in according with what was already found [38]. The large deviation structure was so confirmed in the asymptotic regime for any chosen condition or parameters. What it was belonging to a specific type of constrain is the anomalous scaling regime, which was exhibited only for short times in the initial equilibrium condition of the diffusion coefficient distribution.
We then moved to two state polymer model whose analysis followed the quite same procedure adopted for the diffusing diffusivity model considering that in this case the diffusion coefficient was studied by means of the telegraph model. Again here it was found Brownian diffusion for short and long times and the same long time regime behavior encountered in the diffusing diffusivity model. Also in the two polymer state it was observed a self similarity trend for short times if starting with the stationary diffusion coefficient distribution.
The analogies between the two analyzed models suggested that they are sharing some particular features. Since the two processes has in common a stationary diffusion coefficient probability density function for long times, that was seen as a good starting point to investigate. Indeed looking at the subordinator moments in the long time regime, the subordinator was found to "lose" his aleatory nature, or better to not inherit it from the diffusion coefficient. Through the subordination formula this leads to the interested process long time probability to be a Gaussian propagator with the the diffusion coefficient substituted with its long time average. Even if this was already found for the diffusing diffusivity model, the latter result has a more general nature because it extends to all the stochastic diffusivity models having the diffusion coefficient distribution whose long time limit is time independent. We also point out that what just argued returns the same probability density function that would be found using the Cohen-Beck formula [44] with the long time limit diffusion coefficient probability.
Looking now at the anomalous scaling regime, it was prove that again the independence of the diffusion

coefficient on time can be exploited but this time it does not lead to a Gaussian distribution, but to a self similarity regime.
Finally, it was observed that the Fisher relation is respected for the analyzed models.
We would like now to mention an interesting point that seemed interesting for further investigations which regards the stationary distribution of the diffusion coefficient in the two polymer state model. This stationary seemed to be similar to a sum of independent and identically distributed variables. Indeed discretizing the process we have that

$$dX_i = \sqrt{2D_i}dW_i \tag{3.5.23}$$

The discretized X_t can then be seen as sum of i.i.d. variables since both D_i and W_i does not depends on previous outcomes. This is interesting because it is related to the concept of exchangiability [45] which was indeed already found to be related with anomalous scaling [40].

Appendices

Appendix A

Chapter 1

A.1 Einstein Diffusion Coefficient Derivation

Considering the diffusing particles as molecules suspended in a fluid, we can look at them as a thermodynamic system whose macroscopic quantities are influenced by the fluid. Assuming that there is present a stationary external force f acting on each diffusive particle and working for simplicity in one dimension, the variation of the internal energy U and entropy S can be written as:

$$\delta U = -\int_{\mathbb{R}} dx f c \delta x, \qquad \delta S = -\int_{\mathbb{R}} dx k_B \frac{\partial c}{\partial x} \delta x, \tag{A.1.1}$$

where k_B is the Boltzmann constant and c is the macromolecule density. In equilibrium conditions the variation of the free energy vanishes:

$$\delta F = \delta E - T\delta S = 0. \tag{A.1.2}$$

Combining the above equation with (A.1.1) leads to:

$$fc - Tk_B \frac{\partial c}{\partial x} = 0. \tag{A.1.3}$$

On the other hand, thermodynamic equilibrium implies that the total number of macromolecules passing through a point per unit time (i.e. the flux in one dimension) has to be null. The terms generating flux are:

- the force f which imparts to each particle a velocity $\frac{\mu}{f}$, where μ is the mobility, originating then a flux $\frac{c\mu}{f}$
- the collisions between particles causing a diffusion flux $D\frac{\partial c}{\partial x}$.

In order to compensate each other, the two flux components have to act in opposite directions:

$$\frac{c\mu}{f} - D\frac{\partial c}{\partial x} = 0. \tag{A.1.4}$$

Comparing now eq.s (A.1.3) and (A.1.4) it is possible to obtain the relation between the diffusion coefficient with mobility and temperature corresponding to eq. (1.1.4).

A.2 Smoluchowski Diffusion Coefficient Derivation

The Smoluchowski approach to Brownian motion was different with respect to the Einstein one. Indeed he looked at the process focusing on a single diffusive particle and analyzed its motion by means of the unbiased random walk model [3]. Since the collisions have no preferred direction, it is reasonable to consider their effect on a particle as a random shift whose direction is equally probable for any coordinate axis. In particular, again in one dimension, we can consider the particle making a step of length l to the right or to the left with equal probability (i.e. $\frac{1}{2}$) within every time interval of duration Δt. Denoting with n the number of step done and identifying the axis origin with the initial position, it is possible to show that the probability of finding the particle in coordinate il after $n\Delta t$ corresponds to:

$$p(il, n\Delta t) = \left(\frac{1}{2}\right)^n \binom{n}{\frac{n+i}{2}}, \tag{A.2.1}$$

which for large n can be written as:

$$p(il, t) = \left(\frac{2}{\pi n}\right)^{\frac{1}{2}} e^{\frac{i^2}{2n}}. \tag{A.2.2}$$

From the above probability the mean and variance are gained [8] resulting:

$$\langle x \rangle = 0, \qquad \langle x^2 \rangle = \frac{l^2}{\Delta t} t \qquad \text{with} \quad t = N\Delta t. \tag{A.2.3}$$

The obtained variance combined with (1.1.4) returns (1.1.5).

Appendix B

Chapter 3

B.1 Ornstein-Uhlenbeck moments with Fokker-Pank approach

An other possible way for evaluating $\langle Y(t_1)^2|Y_0\rangle$ and $\langle Y(t_1)^2 Y(t_2)^2|Y_0\rangle$ employs the Gamma function. Since in (3.2.27) and (3.2.29) the integrated function is even the boundaries can be rewritten as:

$$\int_{-\infty}^{+\infty} \rightarrow 2\int_{0}^{+\infty}. \tag{B.1.1}$$

Through another change of variable $r = z^2$ (and $s = q^2$) the interested quantities are found:

$$\begin{aligned}
\langle Y(t_1)^2|Y_0\rangle &= 2\int_0^{+\infty} \frac{e^{-r}}{\sqrt{\pi}}\frac{1}{2\sqrt{r}}\Big(\sigma^2\tau(1-e^{\frac{-2t_1}{\tau}})r + {Y_0}^2 e^{-2\frac{t_1}{\tau}}\Big)dr \\
&= \frac{1}{\sqrt{\pi}}\sigma^2\tau(1-e^{\frac{-2t_1}{\tau}})\Gamma\Big(\frac{3}{2}\Big) + \frac{1}{\sqrt{\pi}}{Y_0}^2 e^{-2\frac{t_1}{\tau}}\Gamma\Big(\frac{1}{2}\Big) = \\
&= {Y_0}^2 e^{\frac{-2t_1}{\tau}} - \frac{\sigma^2\tau}{2}e^{\frac{-2t_1}{\tau}} + \frac{\sigma^2\tau}{2},
\end{aligned} \tag{B.1.2}$$

$$\begin{aligned}
\langle Y(t_1)^2 Y(t_2)^2|Y_0\rangle = 4\int_0^{+\infty} dr \int_0^{+\infty} ds \frac{e^{-r}e^{-s}}{\pi}\frac{1}{4\sqrt{rs}}\Big(&\sigma^4\tau^2(1-e^{\frac{-2t_1}{\tau}})(1-e^{-2\frac{t_2-t_1}{\tau}})rs + \\
&+ \sigma^2\tau(1-e^{\frac{-2t_1}{\tau}})e^{-2\frac{t_2-t_1}{\tau}}r^2 + \sigma^2\tau{Y_0}^2(1-e^{-2\frac{t_2-t_1}{\tau}})e^{-2\frac{t_1}{\tau}}s + {Y_0}^4 e^{-2\frac{t_2-t_1}{\tau}} + \\
&+ \big(2\sigma^2\tau(1-e^{\frac{-2t_1}{\tau}}){Y_0}^2 e^{-2\frac{t_2}{\tau}} + 4\sigma^2\tau(1-e^{\frac{-2t_1}{\tau}}){Y_0}^2 e^{-2\frac{t_2}{\tau}}\big)r\Big) \\
= \frac{1}{\pi}\sigma^4\tau^2(1-e^{\frac{-2t_1}{\tau}})&(1-e^{-2\frac{t_2-\frac{1}{\pi}t_1}{\tau}})\Gamma\Big(\frac{3}{2}\Big)\Gamma\Big(\frac{3}{2}\Big) + \frac{1}{\pi}\sigma^2\tau(1-e^{\frac{-2t_1}{\tau}}) + \\
&+ e^{-2\frac{t_2-t_1}{\tau}}\Gamma\Big(\frac{5}{2}\Big)\Gamma\Big(\frac{1}{2}\Big) + \frac{1}{\pi}\sigma^2\tau{Y_0}^2(1-e^{-2\frac{t_2-t_1}{\tau}})e^{-2\frac{t_1}{\tau}}\Gamma\Big(\frac{3}{2}\Big)\Gamma\Big(\frac{1}{2}\Big) + \\
&+ \frac{1}{\pi}{Y_0}^4 e^{-2\frac{t_2-t_1}{\tau}}\Gamma\Big(\frac{1}{2}\Big)\Gamma\Big(\frac{1}{2}\Big) + \frac{1}{\pi}\big(2\sigma^2\tau(1-e^{\frac{-2t_1}{\tau}}){Y_0}^2 e^{-2\frac{t_2}{\tau}} + \\
&+ 4\sigma^2\tau(1-e^{\frac{-2t_1}{\tau}}){Y_0}^2 e^{-2\frac{t_2}{\tau}}\big)\Gamma\Big(\frac{3}{2}\Big)\Gamma\Big(\frac{1}{2}\Big) \\
= {Y_0}^4 e^{-2\frac{t_2+t_1}{\tau}} &+ \frac{1}{2}\sigma^2\tau{Y_0}^2 e^{-2\frac{t_1}{\tau}}(1-e^{-2\frac{t_2}{\tau}}) + \frac{1}{2}\sigma^2\tau{Y_0}^2 e^{-2\frac{t_2}{\tau}}(1-e^{-2\frac{t_1}{\tau}}) + \\
&+ \frac{1}{4}\sigma^4\tau^2(1-e^{-2\frac{t_1}{\tau}})(1-e^{-2\frac{t_2}{\tau}}) + \frac{1}{2}\sigma^4\tau^2 e^{-2\frac{t_2}{\tau}}(e^{-\frac{t_1}{\tau}} + e^{\frac{t_1}{\tau}})^2 + \\
&+ 2\sigma^2\tau{Y_0}^2 e^{-2\frac{t_2}{\tau}}(1-e^{-2\frac{t_1}{\tau}}),
\end{aligned} \tag{B.1.3}$$

where in the last passages of the above equations it is used:

$$\begin{cases} \Gamma\left(\frac{1}{2}\right) = \sqrt{\pi}, \\ \Gamma\left(\frac{3}{2}\right) = \frac{1}{2}\sqrt{\pi}, \\ \Gamma\left(\frac{5}{2}\right) = \frac{3}{4}\sqrt{\pi}. \end{cases} \tag{B.1.4}$$

B.2 $S(t)$ n-th moment for large times

The generalization of the reasoning done for the 2-nd $S(t)$ moment in Section 3.5 to the n-th moment starts by dividing the n integral domain as done for $\langle S(t)^2\rangle$ evaluation (i.e. $t_n > t_{n-1} > \cdots > t_1$) and recalling that this splitting is payed with a $n!$:

$$\langle S(t)^n\rangle = 2^n n! \int_0^t dt_n \int_0^{t_n} dt_{n-1} \cdots \int_0^{t_2} dt_1 \langle D(t_n)D(t_{n-1})\dots D(t_1)\rangle \tag{B.2.1}$$

Using now t^* as done in Section 3.5.2, in order to get a term which goes like t^n (which will be the dominant one in the long time limit) we must have integrals of the type

$$\int_{t^*}^{t_{i+1}-t^*} dt_i \tag{B.2.2}$$

for any i from 1 to n. Such integral comes out of (B.2.1) only once and the remaining terms cannot be of the order t^n since they do not permit to have the all the D_i probabilities time independent. Focusing now on the interested term, for which we can consider all the D_is probability distributions stationary, we have that

$$2^n n! \int_{nt^*}^{t} dt_n \int_{(n-1)t^*}^{t_n-t^*} dt_{n-1} \cdots \int_{t^*}^{t_2} dt_1 \langle D(t_n)D(t_{n-1})\dots D(t_1)\rangle^* = \big(2\langle D\rangle^*(t-t^*)\big)^n \tag{B.2.3}$$

where the $(t-t^*)$ factor can be obtain changing the variables:

$$t_i' = t_i - it^* \qquad \text{for} \quad 1 \le i \le n, \tag{B.2.4}$$

and the $n!$ present in the l.h.s. of (B.2.3) vanishes because of the n integrations. We can thus conclude that for the n-th moment in the long time regime (3.5.8) holds.

Bibliography

[1] W. Feller.*An Introduction to Probability Theory and Its Applications*. Jhon Wiley & Sons, Inc. 1957.

[2] O. Kallenberg.*Foundations of Modern Probability*. Springer. 1997.

[3] C. W .Gardiner. *Handbook of Stochastic Methods*. Springer. 1985.

[4] R. Durrett.*Essentials of Stochastic Processes*. 2021.

[5] P. Morters. Y. Peres. *Brownian Motion*. 2008.

[6] L. Peliti. *Statistical Mechanics in a Nutshell*. Princeton University Press. 2011.

[7] A. Einstein. "Investigations on the Theory of the Brownian Movement ". In: Dover Publication,Inc. 1905.

[8] M. Smoluchowski."Zur kinetischen Theorie der Brownschen Molekular Bewegung und der Suspensionen ". In: Annelen der Physik 1906.

[9] M. A. Islam. "Einstein–Smoluchowski Diffusion Equation: A Discussion ". In: Physica Scripta. 2004.

[10] R. Brown."A Brief Account of Microscopical Observations on the Particles Contained in the Pollen of Plants and the General Existence of Active Molecules in Organic and Inorganic Bodies". 1827.

[11] D. Gusak et al.*Theory of Stochastic Processes With Applications to Financial Mathematics and Risk Theory*.Springer. 2000.

[12] D. Levin and Y. Peres .*Markov Chains and Mixing Times*. American Mathematical Society.

[13] P. Doukhan. *Stochastic Models for Time*. Springer. 2018.

[14] E. A. C. and R. A. Fisher "Moments and Cumulants in the Specification of Distributions". In: International Statistical Institute (ISI). 1938.

[15] Kevin P. Balanda and H. L. MacGillivray ."Kurtosis: A Critical Review ". In: The American Statistician. 2012.

[16] A. Winkelbauer."Moments and Absolute Moments of the Normal Distribution". Institute of Telecommunications, Vienna University of Technology. 2013.

[17] L.F. Richardson ."Atmospheric Diffusion Shown on a Distance-Neighbour Graph".In: Royal Society. 1926.

[18] D. Nickelsen, H. Touchette ."Anomalous Scaling of Dynamical Large Deviations ".In: Physical Review Letters. 2018.

[19] G. I. Barenblatt.*Scaling, self-similarity, and intermediate asymptotics*. Cambridge University Press. 1996.

[20] B. Wang et al.“Anomalous yet Brownian ”. Edited by David Chandler, University of California, Berkeley. 2009.

[21] B. Wang et al .“When Brownian diffusion is not Gaussian”. In: Nature. 2012.

[22] J. C. Pardo.“A brief introduction to self-similar processes ”. https://www.researchgate.net/. 2007.

[23] P. Embrechts and M. Maejima.“An introduction to the Theory of Self-Similar Stochastic Processes”. In: International Journal of Modern Physics B. 2000.

[24] M.I. Freidlin, A.D. Wentzell,.“Random Perturbations of Dynamical Systems”. In:Grundlehren der Mathematischen Wissenschaften.1984.

[25] M.D. Donsker, S.R.S. Varadhan,.“Asymptotic evaluation of certain Markov process expectations for large time ”. In:Communications on Applied Mathematics and Computation.1983.

[26] H.Touchette .“Legendre-Fenchel transforms in a nutshell”. School of Mathematical Sciences, Queen Mary, University of London. 2005.

[27] H. Touchette.“The large deviation approach to statistical mechanics”.In: Physics Reports. 2009.

[28] S. R. S. Varadhan .“LARGE DEVIATIONS ”. In: Ann. Probab. 2010.

[29] .“On a new limit theorem in probability theory”. Translated by H.Touchette. In: Actualités scientiques et industrielles. 1938.

[30] R.S. Ellis .“An overview of the theory of large deviations and applications to statistical mechanics ”. In:Scandinavian Actuarial Journal. 1995.

[31] H. E. Daniels .“Saddlepoint Approximation in Statistics ”. In:The Annals of Mathematical Statistics. 1954.

[32] G. Ferré. *Large Deviations Theory in Statistical Physics: Some Theoretical and Numerical Aspects*. Université Marne La Vallée. 2019.

[33] J. Bertoin.*Subordinators: Examples and Applications*. Springer-Verlag Berlin Heidelberg. 1999.

[34] F. Garaj.*Methods of Superstatistics and Subordination in Theory of Option Pricing with Stochastic Volatility*. Czech Technical University in Prague. Faculty of Nuclear Science and Physical Engineering. 2018

[35] P. de Gennes *Scaling Concepts in Polymer Physics*. Cornell University.1979.

[36] R. Metzler .“Superstatistics and non-Gaussian diffusion”. In: The European Physical Journal. 2020.

[37] V. Sposini et al .“Being heterogeneous is disadvantageous: Brownian non-Gaussian searches”.In: Physical Review E. 2024.

[38] F. Seno et al.“Brownian yet Non-Gaussian Diffusion:From Superstatistics to Subordination of Diffusing Diffusivities”. In: Physical Review X. 2017.

[39] M. A. F. dos Santos and L. Menon Junior b.“Random diffusivity models for scaled Brownian motion”. In: Elsevier Ltd. 2020.

[40] A.L. Stella and F. Baldovin .“Anomalous scaling due to correlations:limit theorems and self-similar processes”. In: Journal of Statistical Mechanics: Theory and Experiment. 2010.

[41] J. Luytes thesis.“The Anomalous Behaviourof Drunken Walks and Broken Clocks,The Large Deviation Structure of Subordination ”. Katholieke Universiteit Leuven. 2021.

[42] C. Beck .“Generalized statistical mechanics for superstatistical systems”. In: Philosophical Transactions of The royal Society. 2010.

[43] A. L. Stella et al .“Anomalous Scaling Determines Universal Critical Singularities”. In: Physical Review Letters. 2023.

[44] C. Beck, E. G. D. Cohen .“Superstatistics”. In: Physica A: Statistical Mechanics and its Applications. 2003.

[45] J. M. Bernardo.*The Concept of Exchangeability and its Applications*. Far East J. Math. Sci. 1996.

[46] Wolfram Research, Inc..*Mathematica, Version 14.*. Wolfram Research, Inc..2024.