

Revealing the transport mechanisms from a single trajectory in living cells

Yann Lanoiselée

▶ **To cite this version:**

Yann Lanoiselée. Revealing the transport mechanisms from a single trajectory in living cells. Statistical Mechanics [cond-mat.stat-mech]. Université Paris Saclay (COmUE), 2018. English. NNT : 2018SACLX081 . tel-01997530

HAL Id: tel-01997530
https://pastel.hal.science/tel-01997530

Submitted on 29 Jan 2019

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

université PARIS-SACLAY

NNT : 2018SACLX081

Thèse de doctorat

Revealing the transport mechanisms from a single trajectory in living cells

Thèse de doctorat de l'Université Paris-Saclay
préparée à l'École Polytechnique

École doctorale n°573 Interfaces : approches interdisciplinaires, fondements, applications et innovation (INTERFACES)
Spécialité de doctorat : Physique

Soutenance de thèse présentée et soutenue à Palaiseau, le 1er Octobre 2018, par

YANN LANOISELÉE

Composition du Jury :

Antigoni Alexandrou Directrice de recherche, École Polytechnique (Laboratoire d'Optique et Biosciences)	Présidente
Olivier Bénichou Directeur de recherche, Université Pierre et Marie Curie (Laboratoire de Physique Théorique de la Matière condensée)	Rapporteur
Maxime Dahan Professeur, Institut Curie (Laboratoire de Physico Chimie Curie)	Invité
Denis Grebenkov Chargé de recherche, École Polytechnique (Laboratoire de Physique de la matiére Condensèe)	Directeur de thèse
David Lacoste Directeur de recherche, ESPCI (Laboratoire Gulliver)	Rapporteur
Ralf Metzler Professor, University of Potsdam (Institute of Physics and Astronomy)	Examinateur
Gleb Oshanin Directeur de recherche, Université Pierre et Marie Curie (Laboratoire de Physique Théorique de la Matière condensée)	Examinateur

Remerciements

Cette thèse a été pour moi une expérience passionnante. Le sujet correspondait exactement à ce que je voulais car il s'agissait de traiter des problèmes issus de la biologie au travers d'approches mathématiques et physiques élaborées. J'ai appris énormément sur les processus de diffusion dans les milieux hétérogènes, sur la description et l'analyse de ces milieux. J'ai appris à développer une démarche de physicien, j'ai ressenti la joie de comprendre un nouveau phénomène, mais aussi les efforts que représente l'aboutissement d'un travail scientifique rigoureux. Pour tout cela, je remercie en premier lieu mon directeur de thèse Denis S. Grebenkov. Le temps qu'il a accordé à nos discussions scientifiques et ses conseils avisés ont permis d'approfondir ma compréhension du sujet ainsi que la qualité de mes recherches. Au cours de ma thèse, j'ai beaucoup apprécié pouvoir aborder des sujets très variés, réalisant ainsi une marche aléatoire (corrélée) dans le milieu des marches aléatoires.

Je remercie aussi toutes les personnes avec qui j'ai pu collaborer durant cette thèse. Merci donc à Doris Heinrich, Maria Götz et Patrick Witzel pour le travail sur la diffusion non-Gaussienne de traceurs dans le cytoplasme qui a été une grande source d'inspiration. Je remercie aussi Olivier Dauchot et Guillaume Briand pour le travail sur les disques vibrés qui a permis d'observer a l'échelle macroscopique les mouvements habituellement invisible a l'œil nu. Aussi je remercie Agnieszka Wyłomańska, Grzegorz Sikora et Aleksandra Grzesiek pour la collaboration sur la régression non linéaire du TAMSD en présence de bruit blanc.

Cette thèse a été aussi l'occasion de rencontrer de nombreux scientifiques, et parmi eux je voudrais particulièrement remercier Aleksander Weron pour son invitation au centre Hugo Steinhaus à Wrocław, pour nos discussions intéressantes et pour son soutien précieux lors de mes candidatures en post-doctorat.

Durant ces trois années, j'ai eu aussi la chance de côtoyer des esprits formidables au sein de notre

groupe de recherche. Je remercie Liubov Tupikina pour toutes nos discussions passionnantes sur la diffusion anormale, sur les mathématiques en général, et pour le temps passé à jouer ensemble de la musique. Merci aussi à Nicolas Moutal pour son aide, sa pédagogie et ses critiques toujours pertinentes, avec qui je suis très fier d'avoir publié un article.
Merci aussi à Yves et Anne, mes collègues de bureau qui m'ont toujours soutenu pendant ces trois années. Je remercie tous mes amis des éclaireurs, mes amis de Tours et ceux du laboratoire, pour leur soutien indéfectible.

Un grand merci à Éric, Céline et mes chers neveux Marie-Anne, Ulysse et Apolline de m'avoir si gentiment fait une place dans leur foyer.
A tous mes frères Jean-Michel, Alain, Olivier et Éric, qui êtes depuis toujours mes références, mes tuteurs, merci.
A mes parents, qui m'ont laissé suivre les chemins qui se sont offerts à moi, merci.

À ma famille.

Contents

General Context and motivation **6**

Contribution of this thesis **14**

1 Theoretical Background **16**

1.1 Brownian motion 16

1.2 Anomalous subdiffusion 18

1.2.1 Theoretical models of anomalous diffusion 18

1.2.2 Recognizing the subdiffusive mechanism 29

1.3 Non-Gaussian diffusion 30

1.3.1 Kärger model 30

1.3.2 Diffusing diffusivity 31

1.3.3 The square of the n-dimensional Ornstein-Uhlenbeck process 31

2 Revealing weak ergodicity breaking from a single trajectory **33**

2.1 Introduction 33

2.1.1 The dynamical functional 34

2.1.2 Several issues 35

2.1.3 Improved estimators 36

2.2 Gaussian processes 38

2.2.1 Mean estimators 38

2.2.2 Variance of the estimators 39

2.2.3 Brownian motion with two diffusion coefficients 40

2.2.4 Main results 42

2.3 Continuous Time Random Walks . . . 43
2.3.1 Number of jumps . . . 44
2.3.2 Macroscopic limit . . . 46
2.3.3 Main results . . . 48
2.4 Several models of anomalous diffusion . . . 48
2.4.1 Diffusion on percolating clusters . . . 49
2.4.2 Scaled Brownian motion . . . 50
2.4.3 CTRW with exponential cut-off . . . 50
2.4.4 Geometric Brownian motion . . . 53
2.5 Application . . . 54
2.5.1 Measurement noise and practical issues . . . 54
2.5.2 Validation on anomalous diffusion models . . . 56
2.5.3 Application to experimental data . . . 58
2.6 Discussion . . . 62
2.6.1 Positions versus increments . . . 62
2.6.2 Role of the frequency . . . 64
2.6.3 Impact of outliers . . . 67
2.6.4 Discrete displacements . . . 67
2.7 Conclusions . . . 68

3 Unravelling intermittent features from a single trajectory 70
3.1 Introduction . . . 70
3.2 Local Convex Hull Method . . . 72
3.3 Numerical validation . . . 75
3.3.1 Two Brownian motions . . . 77
3.3.2 Brownian motion with a drift . . . 80
3.3.3 Two fractional Brownian motions . . . 82
3.3.4 Brownian motion and Ornstein-Uhlenbeck process . . . 83
3.3.5 Brownian motion and exponential flights . . . 86
3.3.6 Surface-mediated diffusion . . . 86
3.4 Discussion . . . 90

3.4.1 Infinite moments . . . 90
3.4.2 Parameters of the method . . . 91
3.4.3 Phase durations . . . 94
3.5 Conclusion . . . 97

4 Non-Gaussian diffusion in heterogeneous media 99
4.1 Introduction . . . 99
4.2 Model of non-Gaussian diffusion . . . 100
4.3 Solution of the Fokker-Planck equation . . . 106
4.3.1 Statement of the problem . . . 106
4.3.2 Reflecting boundary condition at $D = 0$ . . . 109
4.3.3 Subordination . . . 111
4.3.4 Absorbing boundary condition at $D = 0$ . . . 112
4.4 Asymptotic behavior . . . 115
4.4.1 Brownian limit . . . 115
4.4.2 Short-time behavior . . . 115
4.4.3 Large x behavior . . . 117
4.5 Statistical properties . . . 119
4.5.1 Moments and the non-Gaussian parameter . . . 119
4.5.2 Autocorrelation of squared increments . . . 120
4.5.3 Ergodicity and finite sample effects . . . 122
4.6 Discussion . . . 126
4.6.1 Fourth moment is not enough . . . 126
4.6.2 Anomalous diffusion . . . 126
4.7 Conclusion . . . 128

5 Application to experiments 129
5.1 Introduction . . . 129
5.2 Macroscopic realization of Brownian motion . . . 130
5.2.1 Experimental setup . . . 131
5.2.2 Statistical analysis . . . 132
5.2.3 Data description . . . 132

5.2.4 Distribution of increments . . . 133
5.2.5 Ergodicity hypothesis . . . 136
5.2.6 Velocity auto-correlations . . . 137
5.2.7 Estimation of diffusion coefficient . . . 138
5.2.8 Distribution of TAMSD . . . 139
5.2.9 Mean versus the most probable TAMSD . . . 141
5.3 Passive diffusion in the cytoplasm . . . 143
5.3.1 Presentation of the experiment . . . 144
5.3.2 Ergodicity . . . 146
5.3.3 Probability distribution . . . 147
5.3.4 Origin of non-Gaussian distribution . . . 151
5.3.5 Superstatistical description . . . 156
5.3.6 Cell-based variability . . . 159
5.4 Conclusion . . . 160
5.4.1 Conclusion on the vibrated disks . . . 160
5.4.2 Conclusion on passive intracellular transport . . . 161

6 Conclusion **170**

A Résumé en Français **177**

B Properties of random processes and how to test them **179**
B.1 Markov property . . . 179
B.1.1 Ljung-Box Test . . . 179
B.2 Stationarity . . . 180
B.2.1 Strong stationarity . . . 180
B.2.2 Weak stationarity . . . 180
B.2.3 Unit root . . . 181
B.2.4 Trend stationarity . . . 181
B.3 Ergodicity . . . 182
B.3.1 Statement of the problem . . . 182
B.3.2 Definition of ergodicity for dynamical systems . . . 183

B.3.3 Ergodicity Breaking Parameter 184
B.4 Time Averaged moments 184
B.5 Testing the distribution of increments 185
B.5.1 General goodness-of-fit tests 185
B.5.2 Shapiro-wilk test 186
B.6 Bayesian Inference 187
B.7 Identifying Intermittent Processes 187

C Matlab code for ergodicity testing **188**

D Some properties of the convex hull **189**

General Context and motivation

Living cells, the fondamental units of life, are among the most complex objects known at the micron scale. The chemical, functional, structural, and dynamical aspects are all intertwined in these objects. Their dimensions can vary significantly, from a few micrometers (*Escherichia Coli*) up to half a millimeter (*Amoeba Proteus*). They display a huge number of different components, whose sizes span on three orders of magnitude, from a nanometer for proteins to a micrometer for organelles. They sustain themselves by consuming nutrients from the surroundings and reproduce. Their inner structure adapts to the local mechanical stress from their environment, collaborate with cell partners and consume energy by elaborated cascades of chemical reactions. As an example, inside a mitochondrion, the Krebs cycle, which extracts from old proteins carbohydrates and lipids the ATP molecules, providing energy for other reactions, necessitates 10 steps to be completed through the help of 8 different enzymes. While being simple compared to other processes occurring in cells, the Kreb's cycle is a perfect example of how elaborated and challenging the fundamental understanding of cell functioning can be. The cell interior is composed of different sub-units, called organelles, which can be compared to small chemical factories.

The molecular biology of cell is already well developped, many processes undergoing the cell life [2] are known. Various functions of organelles are known as well as the breathtaking complexity of intertwined network of chemical reactions. Nowadays, one of the challenges is to understand the dynamic transport of different particles inside the cell. Indeed a complete description of the inner life of cell requires to understand not only the structural properties, as well as the chemical reactions but also how fast they are occurring, where and by which mechanisms.

At this scale the dynamics of particles is dominated by thermal fluctuations, i.e. never-ending

mutual collisions of particles with their respective surrounding, resulting in irregular trajectories of particles. This random nature of a particle's motion could be illustrated by that of a drunkard searching for it way home by randomly choosing the next tile on which to step on the infinite lattice of the city (see Fig. 1).



Figure 1: "The Drunkard", Salvador Dali, Gouache on paper (1922)

Under the condition that the particle is much larger than the surrounding objects, the motion is adequately described by Brownian motion. It is a stochastic process for which the probability

density $P(x,t)$ to find a particle at a distance x from its starting point is given by a Gaussian distribution. Due to isotropy, the Gaussian distribution is centered and symmetric.

The fluctuations of position as a function of time are solely described by the Mean Squared Displacement (MSD), denoted $\langle X^2(t)\rangle$, which is equal to

$$\langle x^2(t)\rangle = \int_{-\infty}^{\infty} x^2 P(x,t) = 2Dt \tag{1}$$

in one dimension. Here $\langle . \rangle$ denotes for the Ensemble Average, i.e. the average over all possible realizations of the squared position x^2 at time t weighted by the probability density $P(x,t)$. The MSD has a linear dependence on time, and the only parameter that matters in this problem is the diffusion coefficient D which characterizes the intensity of fluctuations of position at time t. For more details on Brownian motion see Sec. 1.1.

However the problem is much more complicated for the cell. The cell interior is extremely crowded, 30% to 40% [85,86,200] of its inner volume is occupied by large macromolecules. Such a crowding has an important effect on the dynamical properties of objects in the cytoplasm. Figure 2 illustrates well the crowdedness of the cell. Since biologists started to be able to look at the dynamics of individual trajectories of tracers through elaborated microscopy measurements, discrepancies from the celebrated linear MSD have been recorded. Typically one observes a power law dependence of the MSD on time called "anomalous" diffusion, the relation is

$$\langle X^2(t)\rangle = 2D_\alpha t^\alpha, \tag{2}$$

with D_α the generalized diffusion coefficient in units $m^2.s^{-\alpha}$ and α the scaling exponent.

The dynamics is subdiffusive (i.e. $\alpha < 1$) in many experiments testifying for slower space exploration. It is now consensual that anomalous subdiffusion can be attributed to crowding inside the cell. However the precise physical mechanism responsible for it remains debated. Several explanations have been given to explain this behavior, (see Sec. 1.2). In brief possible explanations of subdiffusion include: long periods on which the particle is stopped (see Sec.1.2.1.2), or due to negative correlation of the tracer's increments due to encounters with its environment (see Sec. 1.2.1.1) or finally because a large part of the volume is excluded so that the accessible space is fractal (see Sec. 1.2.1.3).

The observations of anomalous subdiffusion are numerous. In solution, the motion of streptavidin in the presence of dextrane obstacles is more and more subdiffusive as the concentration

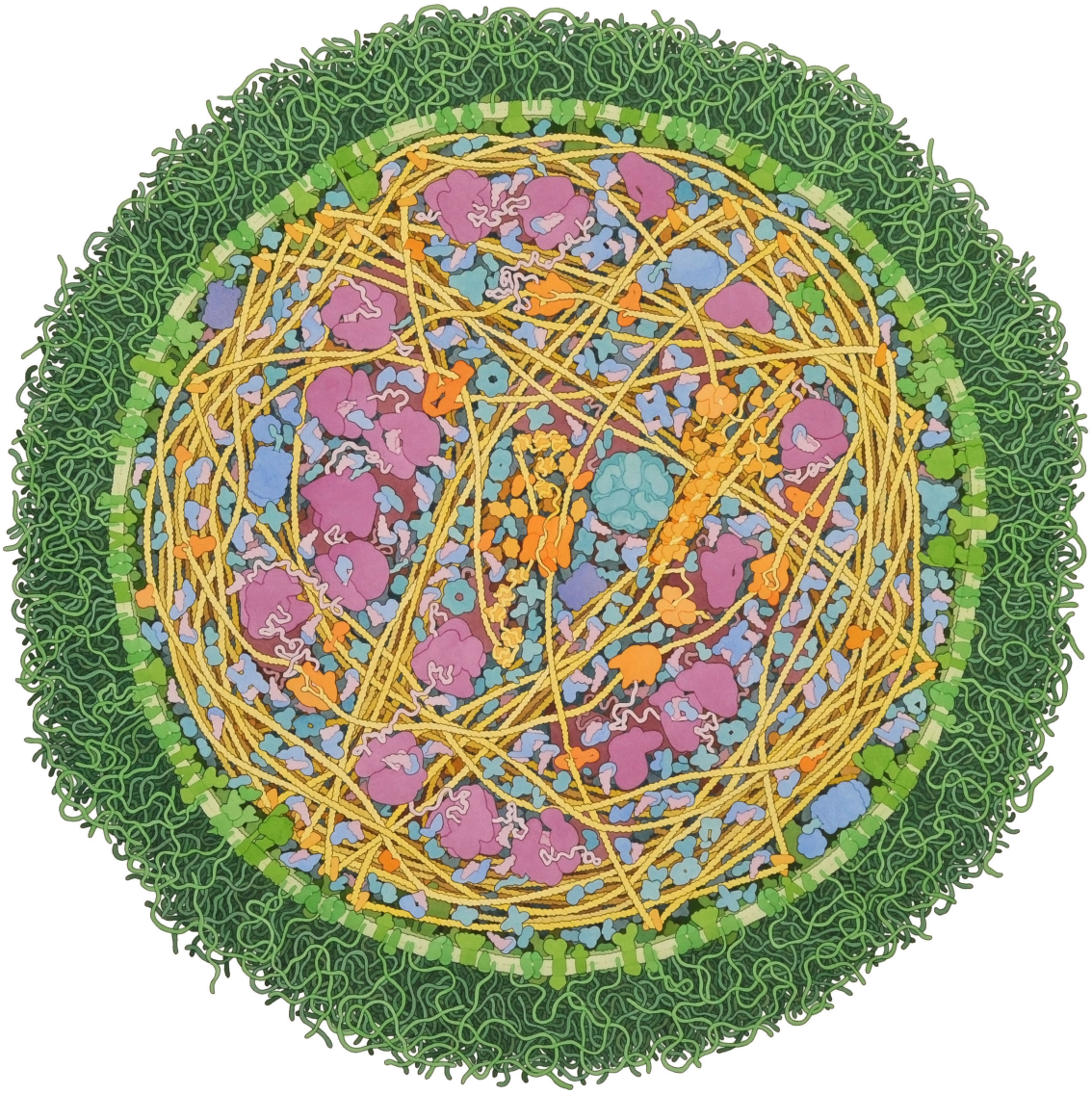


Figure 2: Artistic view of the Mycoplasma mycoides cell interior by David Goodsell (2011), Watercolor, the Scripps Research Institute

of dextrane decreases. The exponent α ranges from 1 at low concentration, to a plateau value $\alpha_{min} = 0.74 \pm 0.02$ at large concentration. From this constatation a phenomenological relation for the anomalous exponent has been provided

$$\alpha = \alpha_{min} + (1 - \alpha_{min}) e^{-\phi/\phi_0}, \tag{3}$$

where ϕ_0 is the typical decay depending on the type of molecule [9]. In Kv1.4 and Kv2.1 potassium ion channels, anomalous diffusion is attributed to obstructed diffusion (see Sec. 1.2.1.3) [233]. The effect of the actin network [288] when probes have a size a comparable to the actin's mesh size λ were reported to undergo successive trapping with power law waiting time modeled by Continuous Time Random Walk (see Sec.1.2.1.2), the exponent being related to the ratio a/λ.

But for smaller objects, the actin network does not act as a cage. However, labeled dextran in the cytoplasm of *HeLa* cells remains subdiffusive. The anomalous behavior is better described by anticorrelated fractional Brownian motion (see Sec. 1.2.1.1.2) [281]. Experiments on crowded fluids reveal the same anti-correlated motion [258, 282, 283]. Lipid granules inside the cytoplasm of yeast cells reveal anticorrelated motion which have been modeled by a generalized Langevin equation (see Sec. 1.2.1.1.1); the motion of the same granules in water, after lysis of the cell was reported to be Brownian, once again testifying for the role of crowding [265]. The motion of mRNA in cytoplasm of prokaryotic E. Coli bacterial cells presents the same negative correlations, but additionally, mRNA displays "periods of almost localized motion, separated by fast jumps to a new position" associated to successive interactions with objects in the cytoplasm [106]. In [141], the motion of lipid granules in *Schizosaccharomyces pombe* was shown to look like a Continuous Time Random walk (with a power law waiting time with a cut-off) inside a harmonic trap at short-time and a motion which resembles fractional Brownian motion.

When performing the inverse problem of identifying the motion origin from a trajectory, it is important to realize that the exponent alone does not allow to determine a microscopic mechanism. However the single trajectory statistics differ from one model of anomalous diffusion to another. So the analysis of single trajectories offers a possibility to identify the correct model.

In addition to the anomalous behavior of MSD, several recent experiment looked at the distribution of displacements and revealed unexpected features. According to recorded experimental trajectories in granular media, turbulent flow, intracellular transport [111, 256] and cell membrane motion [123, 143], the distribution of displacements is not Gaussian as one would expect, but displays exponential tails. This interesting feature is not exclusive to cells but has appeared in various complex media. Three typical shapes of the distribution of displacements have been observed: (i) flat distribution near zero with an exponential tail is found in granular

materials [209], turbulent flow [108], active gels [25, 202, 267], glassy material [48], (ii) exponential behavior in entangled F-actin networks [278, 279], log-return of stock prices [75], and (iii) stretched exponential form in granular gas [229], in crowded environments [102, 103] and from simulations of diffusion with interacting obstacles [235]. A common feature of these dynamics is that the displacement distribution becomes Gaussian in the long-time limit [195].
The exponential tails are reminiscent of heterogeneities in the medium. What is the physical mechanism at the origin of this observation remains an open question. However several phenomenological approaches have been able to catch the main features of these heterogeneities by considering a "diffusing diffusivity". The tracer experiences a medium in which its diffusion coefficient is either time or space dependent. At the experimental timescale, fluctuations of diffusivity make the displacements not identically distributed which induces departure from Gaussian distribution of displacements. Moreover, new questions arise on how to correctly measure and model this heterogeneity. What are the relevant statistical quantities to properly identify the mechanism responsible for this heterogeneity? How much the fluctuations of diffusivity affect the measurement of an effective averaged diffusion coefficient ?

The crowding makes the cell in general more difficult to explore and diffusivity heterogeneities seem to make some places even less accessible. However, the apparent difficulty of intracellular exploration does not prevent cells to be alive. This means that other mechanisms exist to shortcut this limitation. One of these mechanisms is well known. Vesicles are transported along microtubules (the cellular highways) by kinesin and dynein molecular motors [7, 33, 47, 107, 146, 237]. This transport consumes energy in the form of ATP molecules that allows the vesicles to be transported almost ballistically (i.e. the MSD exponent $\alpha = 2$) making much more efficient the trafficking of the strategic resources contained in these microscopic vehicles. Due to fluctuations, the vesicle can detach and resort to diffusive motion until the next active transport. The vesicle thus alternates between two phases of motion, it is an "intermittent process". Many biological processes are intermittent, i.e., they switch between two or several phases. The most common examples are the run-and-tumble motion of bacteria [21, 23, 259] and foraging strategies of animals, e.g., predators who adopt hunting strategies by alternating between slow careful search and fast displacement (see the review [18] and references therein). In microbiology, examples include alternating phases of three-dimensional bulk diffusion and

one-dimensional sliding along DNA chains of DNA-binding proteins [22,171,226,275], temporal trapping of tracers in polymer cages [95, 288], binding/unbinding of macromolecules to form temporal ligand-receptor pairs or to neutralize pathogens by antibodies [29,199,238,245,266,293], switching between distinct conformational states [46,98], alternating phases of bulk and surface diffusion [17,18,166,228,231], etc. The intermittent character of the motion makes the statistical analysis and biophysical interpretation of single particle trajectories even more challenging. If the intermittency is ignored, switching between several phases can be (mis)interpreted as a single effective phase with peculiar properties. For instance, switching between active (ballistic, $\alpha = 2$) and passive (diffusive, $\alpha = 1$) motion of a vesicle can be effectively understood as super-diffusion, with an intermediate scaling exponent $1 < \alpha < 2$. Moreover, intermittency can be also misleadingly interpreted as non-stationarity or non-ergodicity of the process according to statistical tests.

In order to observe the in vivo motion of an intracellular object, a common technique is to attach a fluorescent protein to it and then record its motion using dedicated microscopy technique. Due to the fact that some statistical properties are only accessible at the single trajectory level, the analysis is very challenging. In order to reduce the fluctuations of calculated relevant statistical quantities, many statistical tools need a large amount of trajectories in order to properly evaluate ensemble averages. In the scenarios we will consider during this thesis, this ensemble average is not achievable, thus even the simple MSD is not accessible. Therefore any estimator applied to a single random trajectory will itself be a random variable. Moreover, experimental noise always adds complexity to the problem. A possibility to reduce the variations is to time average the quantities over a trajectory.

From a trajectory of duration t_{exp} made of N increments of duration δ (such that $t_{exp} = N\delta$) recorded at discretized instants of time, the time average using a lag-time Δ (of duration $\Delta\delta$) is

$$\bar{f}(\Delta\delta) = \frac{1}{N-\Delta} \sum_{k=1}^{N-\Delta} f(x(k+\Delta) - x(k)), \tag{4}$$

where f is an arbitrary smooth function and ' $\bar{}$ ' denotes the time-average. However when performing this average, a question comes immediately in mind. Is the time-averaged quantity equal to the ensemble averaged one that I was aiming at calculating? In other words, we ask to which extent a single trajectory is representative of an ensemble of trajectories originated

from the same dynamics. In general the answer is negative. Two conditions must be verified for getting the positive answer: ergodicity and stationarity. The former deals with the equivalence of time and ensemble averages in the limit of infinitely long trajectories while the latter ensures the distribution of increments of the process to be independent of time. If both conditions are fulfilled, the time average with respect to any well-behaved function is equal to ensemble averaged counterpart. Since neither the mechanism underlying the motion nor the validity of both properties are known, they should be tested beforehand. Then, only a thorough exploration of the physical properties through various statistical tests can unravel the nature of the transport mechanism. From this mechanism one can formulate an appropriate physical model. Once an appropriate model is identified, one can estimate its parameters that are related to the intracellular transport, and then rely on the theoretical knowledge about the model to predict biological implications, e.g., biochemical reaction rates, translocation or transcription mechanisms, drug delivery, etc. [20, 35].

Contribution of this thesis

In this thesis, the overall question we ask is: What can one learn from a single realization of a random trajectory about the object that produced it and its environment? We put efforts in trying to be as close as possible to a non-biased analysis by suspending our preconceived ideas and taking a priori no assumption. Consequently, our approach consists in combining hypothesis testing to various statistics in order to understand the underlying dynamics.

For this purpose, in the rest of the first chapter, we describe the main models of anomalous diffusion and non-Gaussian diffusion, their properties and propagators. We present also the main statistical properties of continuous random processes in general in Appendix A. We also discuss the literature for testing various hypotheses and recognizing transport processes.

In the second chapter we propose improved ergodicity and mixing estimators in order to identify nonergodic dynamics from a single particle trajectory. The estimators are based on the time averaged characteristic function of the increments and can thus capture additional information on the process as compared to the conventional time averaged mean square displacement. The estimators are first investigated and validated for several models of anomalous diffusion such as ergodic fractional Brownian motion and diffusion on percolating clusters, and nonergodic continuous time random walks and scaled Brownian motion. The estimators are then applied to two sets of earlier published trajectories of mRNA molecules inside live *E. coli* cells and of Kv2.1 potassium channels in the plasma membrane. Since the estimators do not rely on ensemble averages, the nonergodic features can be revealed separately for each trajectory, providing a more flexible and reliable analysis of single-particle tracking experiments in microbiology.

In the third chapter, we propose a new model-free method to detect change points between distinct phases in a single random trajectory of an intermittent stochastic process. The local convex hull (LCH) is constructed for each trajectory point, while its geometric properties (e.g.,

the diameter or the volume) are used as discriminators between phases. The efficiency of the LCH method is validated for six models of intermittent motion, including Brownian motion with different diffusivities or drifts, fractional Brownian motion with different Hurst exponents, and surface-mediated diffusion. We discuss potential applications of the method for detection of active and passive phases in the intracellular transport, temporal trapping or binding of diffusing molecules, alternating bulk and surface diffusion, run and tumble (or search) phases in the motion of bacteria and foraging animals, and instantaneous firing rates in neurons.

In Chapter 4, we propose an analytically solvable model of non-Gaussian diffusion that generalizes the approaches [49,135,137] of diffusing diffusivity with easily physically interpretable parameters. Depending on the parameters, the distribution of displacements can be either flat or peaked at small displacements, with an exponential tail at large displacements. We show that the distribution converges slowly to a Gaussian one. We calculate statistical properties, derive the asymptotic behavior, and discuss some implications and extensions.

In Chapter 5 we apply the statistical machinery for analyzing experiments. First we study a macroscopic realization of planar Brownian motion by vertically vibrated disks. We perform a systematic statistical analysis of many random trajectories of individual disks. The distribution of increments is shown to be almost Gaussian, with slight deviations at large increments caused by inter-disk collisions. The velocity auto-correlation function takes both positive and negative values at short lag times but rapidly vanishes. We compare the empirical and theoretical distributions of time averaged mean square displacements and discuss distinctions between its mean and its most probable value. In the second experiment, we probe the cytoplasm by tracking passive cytoplasm-induced intracellular motion of tracers, which yield anomalous diffusion and non-Gaussian distributions of increments. The distributions collapse onto a single master curve even under modification of the cytoskeleton composition. This reveals a universal scaling, manifesting a generic feature of intracellular transport, which is independent of the cytoskeleton. Hence, we show that the crowded cytosol determines the dominant mechanism of intracellular transport, while the different cytoskeleton components control only the efficiency of intracellular transport. Finally, we attribute the non Gaussian transport features to spatio-temporal heterogeneities of the cytoplasm. In Chapter 6, we summarize the major results of this thesis and discuss perspectives for future research.

Chapter 1

Theoretical Background

1.1 Brownian motion

Robert Brown observed in a microscope a continuous jittery motion of minute particles ejected from the *Clarkia* pollen grains suspended in water [40]. Since a more systematic study by Jean Perrin [214, 215], the abundant experimental evidence of Brownian motion of microscopic particles has been established [34, 96, 112, 232]. The mathematical origin of this abundance lies in the central limit theorem which implies a universal probabilistic description of motion at mesoscopic time and length scales, regardless microscopic dynamics.

It is Einstein [83] who first described the problem by considering the motion of a colloid in suspension. The colloid experiences a symmetric force around its surface due to kicks by water molecules hence the random motion is symmetric. Thus the probability $P(x,t)$ of a colloid as a function of time is itself symmetrical and centered at zero. The mean displacement averaged over the concentration $P(x,t)$ is not informative as its value is zero at all time due to the mirror symmetry of the problem. The appropriate quantity to inquire into is the Mean Squared Displacement (MSD), which measures the averaged squared deviations of the concentration.

Einstein found that the Fick equation describes well the problem

$$\frac{\partial P(x,t)}{\partial t} = D\frac{\partial^2 P(x,t)}{\partial x^2} \tag{1.1}$$

where D is the diffusion coefficient in unit $m^2.s^{-1}$. In the case of Brownian motion, the MSD follows

$$\langle X^2(t)\rangle = 2dDt, \tag{1.2}$$

in d dimensions. The dependence on time is linear with time such that the only parameter involved in this complex process is the diffusion coefficient.

$$D = \frac{k_B T}{6\pi r \nu}, \tag{1.3}$$

with k_B the Boltzmann's constant, T is the absolute temperature, r the radius of the particle and ν the dynamic viscosity of the viscous medium. Through the relation $k_B = \frac{R}{N_A}$, where R is the gas constant and N_A the Avogadro number, the number of particle in a mole of water is connected to the macroscopic behavior of the colloid in suspension that allowed Jean Perrin [214] to calculate Avogadro's number and verify the atomic hypothesis.

Another derivation was proposed by Smoluchowski in 1906 [249] where he derived the Gaussian distribution as the continuum limit of a simple random walk.

Yet another approach was developed by Langevin [158] who described the motion of a colloid of mass m by Newton's second law of motion with friction constant γ in s^{-1} and an additional random force $\eta(t)$. Because the colloid is gigantic as compared to water molecules, the time for water to equilibrate is much faster than for the colloid. There is then a separation of scales making possible to consider the change of position dX in a time dt as a random change by virtue of the central limit theorem which states that the sum of random Independent Identically Distributed (IID) variable, with finite two first moments asymptotically converges in law to a Gaussian distribution. The Langevin equation

$$m\frac{d^2}{dt^2}x(t) = -\gamma m\frac{d}{dt}x(t) + \eta(t), \tag{1.4}$$

is the first stochastic equation. The two time scale argument allows one to consider the noise as Gaussian with mean zero due to the symmetry $\langle X(t)\rangle = 0$ and the covariance $\langle \eta(t)\eta(t^{'})\rangle = 2\gamma k_B T\delta(t - t^{'})$.

In the limit of large friction, inertial effect disappears leading to the simplified time-rescaled equation

$$\frac{d}{d\tau}x(\tau) = \eta(\tau). \tag{1.5}$$

For a complete review see [130].

1.2 Anomalous subdiffusion

In this section we discuss the three main models that rationalize of anomalous subdiffusion and how to recognize between them.

1.2.1 Theoretical models of anomalous diffusion

In this section are presented the three main models of anomalous diffusion, together with their physical interpretation and some of their properties. A good reference concerning this topic is the book from Klafter and Sokolov [153].

1.2.1.1 Memory induced anomalous diffusion

In this section, we present three main models in which anomalous diffusion arises from the memory of the increments. The negative - power law decaying - correlation of increments has been identified as a rationalizing transport in several experiments.

1.2.1.1.1 The generalized Langevin equation

The Generalized Langevin Equation models a particle experiencing a viscoelatic medium. The medium induces non Markovian friction forces of which decay is described by the friction kernel $\gamma(t)$. The equation reads

$$m\ddot{x}(t) - \int_{-\infty}^{t} m\gamma(t-s)\dot{x}(t)(s)ds = F(t), \tag{1.6}$$

where $F(t)$ is a random force with covariance $\langle F(t)F(t')\rangle = K(|t-t^{'}|)$ refered to as colored noise. The random force and the memory kernel are related through the fluctuation dissipation theorem:

$$K(t) = k_B T\gamma(t) \tag{1.7}$$

with k_B the Boltzmann constant and T the absolute temperature of the system. The particular form of the covariance depends on the microscopic details of the dynamics, for example hydrodynamic effects can be taken in account [94].

The Generalized Langevin equation is in general hard to solve. However in the particular case of a linear equation with Gaussian noise, the probability density of increments is Gaussian.

In this case the propagator is completely characterized by first and second moments. The first moment is zero the second moment can be obtained through a Laplace transformation with respect to time [264, 277].

1.2.1.1.2 fractional Brownian Motion

The fractional Brownian motion $W_H(t)$ [154,180] is a centered Gaussian process, which is defined by its covariance function:

$$\langle W_H(t)W_H(t^{'})\rangle = D_\alpha \left(|t|^{2H} + |t^{'}|^{2H} - |t - t^{'}|^{2H}\right), \tag{1.8}$$

where $0 < H < 1$ is the Hurst exponent, and D_α is the generalized diffusion coefficient.

From which the expression yields from a weighted average of the Wiener increments dW_t. Increments of a Wiener process are intergated on the interval $s \in (-\infty, t)$. The process is non-local in time and non-Markovian.

$$W_H(t) = W_H(0) + \frac{D_\alpha}{\Gamma(H+1/2)}\left[\int_{-\infty}^{0}\left(|t-s|^{H-1/2} - (-s)^{H-1/2}\right)dW_s + \int_0^t (t-s)^{H-1/2}dW_s\right] \tag{1.9}$$

The basic properties of the increments are

- Self-similarity: $\mathcal{P}(W_H(at)) = a^{2H}\mathcal{P}(W_H(t))$, the process looks the same at all scales;
- Stationarity: $\mathcal{P}(W_H(s+t) - W_H(s)) = \mathcal{P}(W_H(t))$, the distribution of its increments is time invariant.

The fBm is an important model in mathematics as it is the only self-similar Gaussian process with stationary increments.

The anomalous scaling of the MSD $\langle X^2(t)\rangle = 2D_\alpha t^{2H}$ originates from the long-time memory of its increments. Three types of behaviors are distinguishable: anti-persistence for $H < 0.5$ (i.e. rapidly changes direction), persistence for $H > 0.5$ (i.e. tendency to keep the same direction), and Markovian Brownian motion for $H = 0.5$.

This is the long-range memory process that is often used to model anti-persistent subdiffusive motion for $H < 1/2$ (e.g., the motion of a tracer in a visco-elastic medium with no characteristic timescale [26, 116, 258]) and persistent superdiffusive motion for $H > 1/2$ (e.g., active transport of cargos on microtubules by molecular motors [7, 33, 47, 69]).

The probability of finding an object at distance x at time t from x_0 at time t_0 is

$$P(x,t|x_0,t_0) = \frac{1}{\sqrt{4\pi D_\alpha (t-t_0)^{2H}}} \exp\left(-\frac{(x-x_0)^2}{4D_\alpha (t-t_0)^{2H}}\right) \tag{1.10}$$

In a discrete-time approximation, the increments of the fBm, $\Delta X_n = W_H(n+1) - W_H(n)$ called the fractional noise, can be simulated as a column vector $\mathbf{\Delta X}$ using the relation

$$\mathbf{\Delta X} = \mathbf{C}^{1/2}\mathbf{N}(0,1) \tag{1.11}$$

where C is the covariance matrix of the fractional noise constructed from the autocovariance function $K(k) = D_\alpha \left[|n-1|^{2H} - 2|n|^{2H} + |n+1|^{2H}\right]$ and $\mathbf{N}(0,1)$ is a column vector of independent standard Gaussian variables. However in pratice, computing the square root of the covariance matrix can be time consuming, especially for long trajectories. To overcome this difficulty, several strategies have been proposed, either exact or approximated [72].

A possible generalization of fBm is obtained by replacing the Gaussian noise by a stable noise yielding the "fractional Lévy stable motion" [41].

The fBm is a mathematical model which is not based on a physical ground. However it is equivalent to a force-free overdamped generalized Langevin equation (i.e. inertia is neglected) in the long-time limit. So the phenomenological approach offered by the fBm should be physically interpreted in term of the generalized Langevin equation.

1.2.1.1.3 Phantom polymers and Markovian embedding

Polymer dynamics can be modeled as a system of particles undergoing random motion within an interaction potential binding together the polymer. Let us consider the simplest model, when the interactions are approximated by harmonic potential and the beads are phantom so they can overlap. The problem is linear and can be solved analytically. The equations on each degree of freedom are Markovian; however the coupling between them induces some memory effects so the marginal dynamic of a single bead is no longer Markovian. The most classical Rouse model, where the monomers are arranged in series, displays a subdiffusive anomalous MSD exponent $\alpha = 1/2$ in the limit of infinitely many monomers [109].

The idea of Markovian embedding is to take advantage of this interaction induced anomaly to approximate an arbitrary correlation function by choosing an appropriate structure of the polymer. One can add imaginary particles interacting through harmonic potential with the

tracer of interest to model the effect of the complex surrounding viscoelastic media [109]. The structural properties of the system are encoded in the eigenvalues of the matrix describing the harmonic coupling that determine the decay of the VACF thus the MSD values. Choosing appropriately the coupling constants make possible to model any correlation, any impact of the viscoeclatic medium on the individual motion of the tracer. The description of a Gaussian process with complicated correlation function is mapped onto the problem of a multidimensional Markovian dynamics [212]. The Langevin equation for the system of $M+1$ particles, which describes the motion of the particle of interest x_0 with M imaginary ones $x_1, \dots x_M$ reads

$$\begin{cases} \eta_0 \dot{x} = f(x,t) - \sum_{i=1}^{M} k_i (x - x_i) + \sqrt{2\eta_0 k_B T}\xi_0(t) \\ \eta_i \dot{x_i} = k_i (x_i - x) + \sqrt{2\eta_i k_B T}\xi_i(t) \end{cases} \tag{1.12}$$

where for each particle $i \in [0, M]$, η_i is the viscous friction coefficient, k_i is the spring constant of the harmonic potential and ξ_i is a white noise with $\langle \xi_i(t) \rangle = 0$ and $\langle \xi_i(t_1)\xi_j(t_2) \rangle = \delta_{ij}\delta(t_1 - t_2)$. A force $f(x,t)$ applied on the system is applied only to the particle $i = 0$ and not the imaginary ones $i > 0$.

We can describe this system as a matrix equation

$$\dot{X} = -\mu X + F$$

where μ is the matrix describing the interactions between the monomers.

We search solutions of $\dot{X} = -\mu X$. They are of the form $X(t) = e^{-\mu t}B(t)$. Now we characterize the function $B(t)$ by constant variation method which gives, with the initial condition $U = X(0)$,

$$X(t) = e^{-\mu t}X(0) + e^{-\mu t}\int_0^t dt' e^{\mu t'} F(t')$$

If the matrix μ is symmetric, it is diagonalizable in the form $\mu = V\lambda V^{-1}$ where V is the eigenvector and λ the diagonal matrix of eigenvalues. The MSD reads

$$\begin{aligned} \langle (X_i(t) - X_i(0))^2 \rangle &= \sum_{n_1=1}^{M+1}\sum_{n_2=1}^{M+1}\sum_{m=1}^{M+1} V_{in_1}\left(e^{-\lambda_{n_1}t} - 1\right) V^{-1}_{n_1 m}\frac{k_B T}{k_m} V_{in_2}\left(e^{-\lambda_{n_2}t} - 1\right) V^{-1}_{n_2 m} \\ &+ \frac{1}{\lambda_{n_1} + \lambda_{n_2}}\left(1 - e^{-(\lambda_{n_1} + \lambda_{n_2})t}\right)\frac{k_B T}{\eta_m} \end{aligned} \tag{1.13}$$

In the long-time limit, when the velocity Autocorrelation Function (VACF) of every degree of freedom has vanished ($t \gg \max(\lambda_i)$) the MSD becomes linear (as Brownian motion). At this

timescale, interactions within the polymer are no longer affecting the marginal probability of each monomers.

Now it is instructive to write the Fokker-Planck equation for the system of Langevin equations. The expression for this $(M+1)$-dimensional dimensional Fokker-Planck equation reads

$$\frac{\partial P\left(\mathbf{x},t\right)}{\partial t}=\left(\sum_{i=1}^{M+1}\sum_{j=1}^{M+1}\frac{\partial}{\partial x_i}\mu_{ij}x_j(t)+\frac{k_BT}{\eta_i}\sum_{i=1}^{M+1}\frac{\partial^2}{\partial^2 x_i}\right)P\left(\mathbf{x},t\right). \tag{1.14}$$

Because the input noise is Gaussian, we expect to obtain a Gaussian distribution at the end. So the solution in the Fourier space is given by the characteristic function of a multidimensional Gaussian process

$$P\left(\mathbf{x},t,x_0,t_0\right)=\exp\left(i\sum_{k=0}^{M}\langle x_k\left(t-t_0\right)\rangle-\sum_{i=0}^{M}k_i^2\langle x_i^2\left(t-t_0\right)\rangle\right). \tag{1.15}$$

In order to start the system at equilibrium, the initial position of the i-th particle $x_i(0)$ is randomly choosen from a Gaussian distribution of mean $\langle x_i(0)\rangle=0$ and variance $\langle x_i^2(0)\rangle=k_BT/k_i$ [109] so the mean position over time $\langle x_i\left(t-t_0\right)\rangle=0$. The second term of the characteristic function corresponds to the MSD $\langle x_i^2\left(t-t_0\right)\rangle$ for each particle that has already been calculated for the Langevin equation in Eq. (1.13).

For the particle of interest x_0, the marginal probability density is then

$$P\left(x_0,t|x_0(t_0),t_0\right)=\frac{1}{\sqrt{2\pi\langle x_0^2\left(t-t_0\right)\rangle}}\exp\left(-\frac{\left(x_0-x_0(t_0)\right)^2}{2\langle x_0^2\left(t-t_0\right)\rangle}\right). \tag{1.16}$$

We obtain the analytical form of the distribution that takes in account correlations.

In order to be able to reproduce a power law for memory depicting anomalous diffusion, the following system have been proposed [109]

$$k_i=\frac{\eta_\alpha v_0^\alpha}{b^{\alpha(i-1)}\Gamma\left(1-\alpha\right)}$$

and

$$\eta_i=\frac{\eta_\alpha v_0^{\alpha-1}b^{(\alpha-1)(i-1)}}{\Gamma\left(1-\alpha\right)}$$

with $0<\alpha<1$ and $i\in[1,M]$. Here η_α is the viscous friction coefficient a single monomer in the medium and v_0 its initial speed. The parameter b defines the repartition of spring constants and friction cefficients among the scales. This approximation works for the time

period $t \in \left[\frac{1}{v_0}, \frac{b^{M-1}}{v_0}\right]$ [109].
The Markovian embedding approach is very interesting as it can model complicated correlation structures and reproduces the cross-over toward Brownian motion at long time.

1.2.1.2 Waiting time induced anomalous diffusion

Another possible explanation of anomalous diffusion due to molecular crowding is possible. A particle can exhibit long stalling periods (waiting times) between jumps. Under the assumption that the waiting times are power law distributed with infinite mean, the MSD is anomalous (subdiffusive).

Continuous Time Random Walk [32] is a general model introduced by Montroll and Weiss [201] to describe waiting situations. This approach raised a considerable interest in the physics community as it allowed to derive various diffusion equations and notably a class of fractional diffusion equations.

One can imagine that a molecule diffuses into a medium with a high concentration of macromolecules. Thus, the particle may be blocked by macromolecules and prevented from moving for a long time, which may be comparable to the time of the experiment. The continuous time random walk describes the discrete jumps happening in continuous time.

1.2.1.2.1 Renewal theory

The mathematical description of CTRW takes advantage of the renewal theory. For independent jumps, the probability $p_n(x)$ of being at x after n displacements is given by the renewal formula

$$p_n(x) = \int_{-\infty}^{\infty} p_{n-1}(x - x')p(x')dx' \tag{1.17}$$

which in the Fourier domain, by the convolution property, becomes $\hat{p}_n(q) = (\hat{p}(q))^n$. Given the waiting times density $\phi(t)$ between two consecutive jumps, one can characterize in a similar way the probability density $\phi_n(t)$ of the time at which the n^{th} step has occured:

$$\phi_n(t) = \int_0^t \phi_{n-1}(t - t')\phi(t')dt' \tag{1.18}$$

which in the Laplace domain reads $\tilde{\phi}_n(s) = (\phi(s))^n$. Then the probability for waiting a time t without moving, in the Laplace domain is $\tilde{\Phi}_0(s) = \frac{1-\tilde{\phi}(s)}{s}$ Finally, the complete propagator is

the probability of making n jumps in a time $t^{'}$ and then not moving for a time $t - t^{'}$ summed over all possible $n \in [0, \infty]$ which in the Laplace-Fourier domain is explicitly

$$\hat{\tilde{P}}(q,s) = \sum_{n=0}^{\infty} \hat{p}_n(q)\tilde{\phi}_n(s)\tilde{\Phi}_0(s). \tag{1.19}$$

Because $p(x)\phi(s) < 1$, one can use geometric series summation to get the Montroll-Weiss formula

$$\hat{\tilde{P}}(q,s) = \frac{1-\tilde{\phi}(s)}{s}\frac{1}{1-\hat{p}(q)\tilde{\phi}(s)}. \tag{1.20}$$

This expression is very general as it is independent of the distribution considered.

Assuming that $\phi(s)$ has a finite first moment (e.g. exponentially distributed) and $p(x)$ is centered and has a finite second moment, the Taylor expansion for long time and large displacement gives

$$\begin{aligned} \hat{p}(q) &\approx 1 - \frac{\sigma^2}{2}q^2 \\ \tilde{\phi}(s) &\approx 1 - \lambda s \end{aligned} \tag{1.21}$$

The solution in Fourier Laplace space is

$$\hat{\tilde{P}}(q,s) = \frac{1}{\frac{\sigma^2}{2\lambda}q^2 + s} \tag{1.22}$$

By using the properties of the Laplace and Fourier transformation, diffusion equation is retrieved by inversion of the transforms

$$\frac{\partial P}{\partial t} = D\frac{\partial^2 P}{\partial x^2} \tag{1.23}$$

where the diffusion coefficient $D = \frac{\sigma^2}{2\lambda}$. When the mean waiting time is finite, the process is asymptotically equivalent to a Brownian motion.

1.2.1.2.2 Time-fractional diffusion

If one chooses $\phi(s)$ to be a power law distribution (e.g. Pareto I), the mean waiting time becomes infinite. We then have to consider the following asymptotic expansion

$$\tilde{\phi}(s) \approx 1 - \tau^{\alpha}s^{\alpha}, \tag{1.24}$$

for which the solution is

$$\hat{\tilde{P}}(k,s) = \frac{\tau^{\alpha}s^{\alpha-1}}{\tau^{\alpha}s^{\alpha} + q^2\sigma^2/2}. \tag{1.25}$$

By inverse Laplace and Fourier transform, one can deduce the time-fractional diffusion equation

$$\frac{\partial P}{\partial t} = {}_0D_t^{1-\alpha} D_\alpha \frac{\partial^2 P}{\partial x^2}, \tag{1.26}$$

with the generalized diffusion coefficient$D_\alpha = \frac{\sigma^2}{2\tau^\alpha}$ in units $m^2.s^{-\alpha}$ and the ${}_0D_t^{1-\alpha}$ is the Riemann-Liouville fractional operator defined for an arbitrary function $f(t)$ as

$${}_0D_t^{1-\alpha} f(t) = \frac{1}{\Gamma(\alpha)} \frac{\partial}{\partial t} \int_0^t \frac{1}{(t-t')^{1-\alpha}} f(t')dt'. \tag{1.27}$$

The emergence of this operator is directly related to the fact that the discrete times at which jumps happen, form a fractal dust on the support of the time half-line [164]. This operator is non-local as the integral involves all the preceding times.

A more general derivation including a force deriving from a potential is given in [12], while the case of a time-dependent force is discussed in [124].

The exact solution to Eq. (1.26) is known to be [192, 290]

$$P(x,t) = \frac{1}{\sqrt{4\pi K_\alpha t^\alpha}} H_{1,2}^{2,0}\left[\frac{x^2}{4\pi K_\alpha t^\alpha} \middle| \begin{matrix} (1-\alpha/2, \alpha) \\ (0,1), (1/2,1) \end{matrix} \right], \tag{1.28}$$

where $H_{1,2}^{2,0}$ is the Fox H function [178].

One of the properties of the CTRW is its non-ergodicity which will be discussed in Annex B.3.

The literature on Continuous Time Random Walk is large as it makes possible a wide range of generalizations of the diffusion equations. It is possible to derive and solve space-time fractional diffusion equation [177] when the mean jump length is infinite.

Another intepretation is that the waiting times come from random interactions with the surrounding [181].

1.2.1.3 Excluded volume induced anomalous diffusion

Another mechanism which can be responsible for anomalous diffusion is that the space in which the motion occurs is obstructed by the crowded environment. Here we particularly look at the case of self-similar spaces where the accessible volume displays a scaling relation, in the sense that the mass $m(r)$ contained inside a sphere of radius r scales as

$$m(r) \propto r^{d_f}, \tag{1.29}$$

where d_f is the fractal dimension of the accessible space.

While only the dimensionality d is necessary to describe the properties of an Euclidean space, self-similar objects need 3 dimensions to be described: the dimension of the embedding space d, the fractal dimension d_f, and the spectral dimension d_s.

The fractal dimension describes how the self-similar object fills the Euclidean embedding space at different scales. For rigorous mathematical definition of the fractal dimension see [89, 91, 268].

The spectral dimensions can be defined by considering that each site of the supporting object is a particle interacting with the other through an harmonic potential. For the whole system, the spectral density distribution $p(\omega)$ of the system follows the relation

$$p(\omega) \propto \omega^{d_s - 1}, \tag{1.30}$$

at low frequencies ω [222].

Another point of view is that the probability for a random walker to return to its starting site at time t is

$$P(0, t) \propto t^{-2d_s}, \tag{1.31}$$

which is directly related to the anomalous exponent $\alpha = 2d_s$.

When considering diffusion on such structures the walk dimension d_w needs to be introduced. It corresponds to the fractal dimension of the random walk. The mean squared displacement reads

$$\langle X^2(t) \rangle \propto t^{\alpha} = t^{2/d_w}, \tag{1.32}$$

where $d_w = 2$ for normal diffusion, leading to the well known linear dependence on time for Brownian motion. Another interesting property is that the probability for the walker to return at its initial position is one (i.e. recurrence property) only if $d_w > d_f$. For some fractal structures (see examples in Fig. 1.1), there is a relationship between three of these dimensions, known as the Alexander-Orbach relation [3],

$$d_s/2 = d_f/d_w. \tag{1.33}$$

Interestingly, the spectral dimension is independent of the dimension of the embedding space d and the relation is valid for any size of the system.

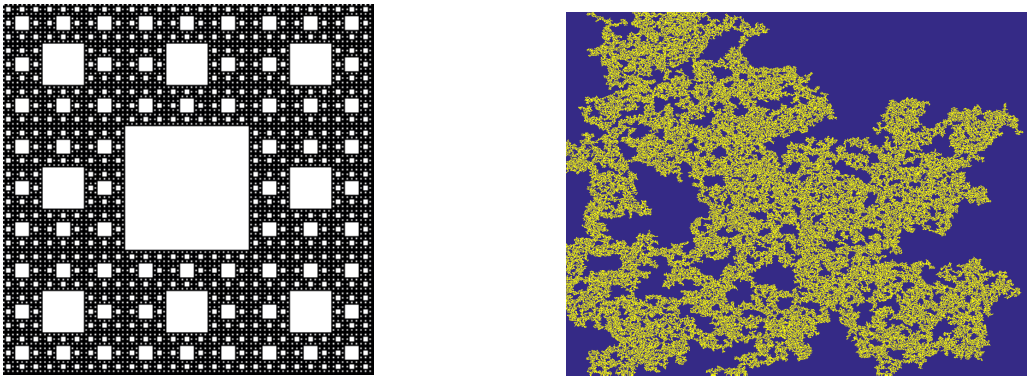


Figure 1.1: *Left.* Sierpinski gasket of fractal dimension $d_f = \frac{\ln 8}{\ln 3}$ *Right.* Percolation cluster on the square lattice at $p = 0.5928$ just above the critical value p_c.

However this relation breaks in fractals in which the diffusive properties asymptotically depend on the system size such as diffusion in a cluster of Diffusion Limited Aggregation [133] or on a fractal trees [70] (see examples in Fig. 1.2). For a finite system, the generalized diffusion coefficient D_α depends directly on the number M of sites in the system [252]. When the time is less than the time needed for the MSD to saturate (i.e. exploration of the whole system), the MSD reads

$$\langle X^2(t)\rangle = M^\beta t^\alpha \tag{1.34}$$

where $\beta = \frac{2}{d_s}(d_s/d_f - \alpha)$. For a fractal verifying the Alexander-Orbach relation, $d_s/d_f = \alpha$ such that the dependence vanishes.

1.2.1.3.1 Diffusion in percolation clusters

A percolation cluster is an object obtained by randomly filling a lattice. Consider a lattice in $\mathbb{R}^d$ with $d > 1$, each site is filled with a probability p independently of the others. This structure can display a so-called percolation transition, meaning that there exists a critical value p_c such that if $p > p_c$ there exists with probability one a simply connected cluster of infinite size crossing the whole system. Near the critical value $p = p_c + \epsilon$, the infinite cluster is known to display a self similar pattern. The critical p_c as well as the fractal dimension d_f, depend on the lattice type. Alexander and Orbach conjectured that $d_s = 4/3$ is any dimension [3]. Diffusion in percolation

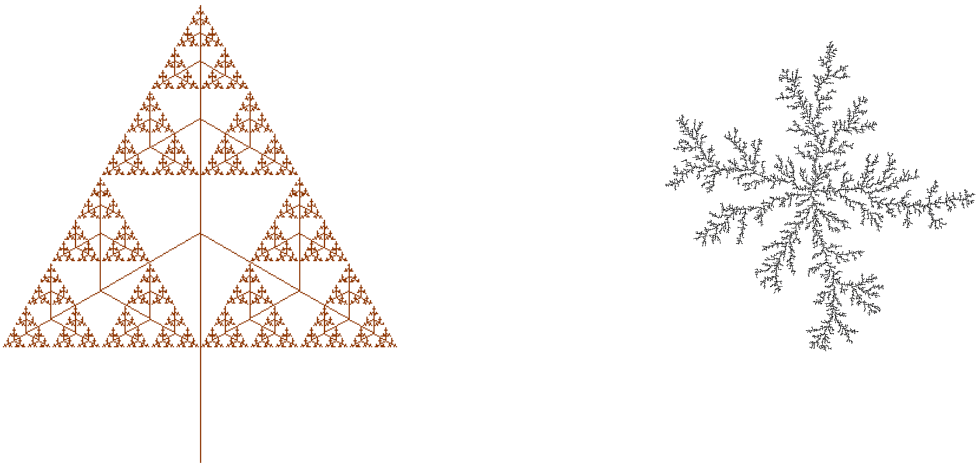


Figure 1.2: Two examples where the Alexander-Orbach relation does not hold. *Left.* A fractal tree at the sixth generation with three new branches at each generation. *Right.* A diffusion limited aggregation cluster simulated on a square lattice with 10000 aggregated particles [133, 287].

clusters have been reviewed in [120].

1.2.1.3.2 A scaling equation

An equation for diffusion in fractals has been proposed through scaling arguments and analogy with a conductivity problem [244]. Starting from a spherical Fokker-Planck equation in d_f dimensions in which is added a radius-dependent diffusivity $D(r) = r^{1-d_f}\left(\frac{\partial R(r)}{\partial(r)}\right)^{-1} = r^{-\theta}$ with $\theta = d_f + \alpha - 2$. The diffusion coefficient is deduced from the resistance in a fractal which scales like $R(r) = r^{-\alpha}$. O'Shaughnessy and Procaccia proposed the Fokker-Planck equation

$$\frac{\partial}{\partial t}P(r,t) = \frac{1}{r^{d_f-1}}\frac{\partial}{\partial r}\left(D(r)r^{d_f-1}\frac{\partial P}{\partial r}\right), \tag{1.35}$$

where $D(r) = D_0 r^{-\theta}$ is a space dependent diffusion coefficient, which can be solved analytically:

$$P(r,t) = \frac{2+\theta}{d_f\Gamma(d_f/(2+\theta))}\left(\frac{1}{D_0(2+\theta)^2 t}\right)^{d_f/(2+\theta)}\exp\left(-\frac{r^{2+\theta}}{D_0(2+\theta)^2 t}\right) \tag{1.36}$$

for which the MSD is anomalous. The probability of return $p(0,t) \propto t^{-d_f/(d_f+\alpha)}$ gives the spectral dimension $d_s = 2d_f/(d_f+\alpha)$ and the walk dimension $d_w = d_f + \alpha$.

This equation reproduces the anomalous subdiffusion but fails to reproduce the propagator obtained for a random walk on a self avoiding walk (which is fractal) as well as in the infinite cluster of percolation at criticality [120]. To overcome this problem Giona et al. [105] proposed a fractional equation which reproduces the asymptotic probability density function. This equation is a "halved equation", obtained by taking the couple of solutions of normal diffusion in Laplace domain, removing the diverging solution and manually replacing the order of the time derivative (from 1/2 to $1/d_w$). The equation in spherical coordinates, under radial symmetry assumption reads

$$ {}_0D_t^{1/d_w} P(r,t) = -D \frac{1}{r^{d_f/d_w - 1/2}} \frac{\partial}{\partial r} \left(r^{d_f/d_w - 1/2} P(r,t) \right) \tag{1.37} $$

with ${}_0D_t^{1/d_w}$ the Riemann-Liouville operator defined in Eq.(1.27). For which the solution is not Gaussian and scales as

$$ P(r,t) = A t^{-d_f/d_w} \exp\left(-c\left(r/R(t)\right)^{\gamma}\right) \tag{1.38} $$

with $\gamma = \frac{d_w}{d_w - 1}$ and $R(t) = \sqrt{\langle X^2(t) \rangle} \propto t^{2/d_w}$.

However this equation holds only asymptotically and does not reduce to normal diffusion when $d_w = 2$.

To solve the problem, Metzler *et al.* [191] proposed and solved an equation generalizing both approaches, valid at any time in any embedding dimension

$$ {}_0D_t^{2/d_w} P(r,t) = \frac{1}{r^{d_f - 1}} \frac{\partial}{\partial r} \left(D(r) r^{d_f - 1} \frac{\partial P}{\partial r} \right), \tag{1.39} $$

from which the solution reproduces the expected propagator [120].

1.2.2 Recognizing the subdiffusive mechanism

Recognizing nature of the anomalous motion from a single trajectory can be difficult. It is a priori clear that the observed motion has only one source of subdiffusion. One can imagine situations where the mechanisms appear together, e.g. a viscoelastic fractal medium with energetic disorder [188].

However, several observables based on a single-trajectory statistics were proposed. The distributions of first-passage times [59, 147] and of the maximum excursion [36, 261] are different for the three main subdiffusive models discussed above offering thus an option to distinguish them. Also, fundamental moments [262] allow one to distinguish between the structural and the

energetic part of the disorder. A possibility to detect diffusion in fractal media is to evaluate if the explored space is fractal with a dimension $d_f < d$ [189] (see also [190] and references therein). At the level of a single trajectory, elaborate statistical tools have been developed to recognize fractional Brownian motion (fBm) [42], to distinguish between fBm and Continuous Time Random Walk (CTRW) [87, 174, 283], and to reveal ergodicity breaking [175].

1.3 Non-Gaussian diffusion

In this section we review some former theoretical contributions to non-Gaussian diffusion.

1.3.1 Kärger model

One fo the first contributions on the theoretical side is the Kärger model [93, 145].

The Kärger model [145] has been developed to study diffusion in a medium in which a particle can randomly switch between two domains with distinct diffusion coefficients D_1 and D_2, with the exchange rates K_{12} and K_{21}. By solving two coupled diffusion-reaction equations, the Fourier transform of the propagator can be derived [145]

$$\tilde{P}_{KM}(q,t) = (1-p')\exp(-q^2 D_1'(q)t) + p'\exp(-q^2 D_2'(q)t), \tag{1.40}$$

with

$$\begin{aligned} D_1'(q) &= \frac{1}{2}\left(D_1 + D_2 + \frac{1}{q^2}(K_{12}+K_{21}) - \left(\left(D_2 - D_1 + \frac{1}{q^2}(K_{21}-K_{12})\right)^2 + \frac{4K_{12}K_{21}}{q^4}\right)^{1/2}\right), \\ D_2'(q) &= \frac{1}{2}\left(D_1 + D_2 + \frac{1}{q^2}(K_{12}+K_{21}) \left(\left(D_2 - D_1 + \frac{1}{q^2}(K_{21}-K_{12})\right)^2 + \frac{4K_{12}K_{21}}{q^4}\right)^{1/2}\right), \\ p' &= \frac{1}{D_2'(q) - D_1'(q)}(p_1 D_1 + p_2 D_2 - D_1'(q)), \end{aligned} \tag{1.41}$$

where p_1 and p_2 are relative volume fractions of two domains.

The probability density distribution is not Gaussian at short-time but converges to Gaussian at long-times. The non Gaussian parameter $\gamma(t)$ measures the departure from Gaussianity:

$$\gamma(t) = \frac{1}{3}\frac{\langle X^4(t)\rangle}{\langle X^2(t)\rangle^2} - 1, \tag{1.42}$$

which is equal to the excess kurtosis divided by 3 (the kurtosis of the Gaussian distribution). By definition, the non-Gaussian parameter is zero for the Gaussian distribution. The analytical

expression of the non-Gaussian parameter from Kärger model, which was derived in [139], and also studied in [93], is

$$\gamma_{KM}(t) = \eta \frac{2}{t/\tau} \left(1 - \frac{1}{t/\tau} \left(1 - e^{-t/\tau} \right) \right), \tag{1.43}$$

with the coefficient $\eta = \frac{p_1 p_2 (D_1 - D_2)^2}{(p_1 D_1 + p_2 D_2)^2}$.

1.3.2 Diffusing diffusivity

Chubynsky and Slater [55] proposed the concept of "diffusing diffusivity" where the diffusion coefficient evolves as a continuous random process. Therefore, the probability $\pi(D, t)$ to have a diffusivity D at time t obeys the diffusion-advection equation

$$\frac{\partial}{\partial t} \pi(D, t) = \frac{\partial}{\partial D} \left[s(D) \pi(D, t) + \frac{\partial}{\partial D} \left[d(D) \pi(D, t) \right] \right], \tag{1.44}$$

where $s(D)$ is the drift term and $d(D)$ is the diffusive term. Due to physical constraints (positivity of kinetic energy), the equation is solved on the real half-line $\mathbb{R}^+$ with reflecting boundary at $D = 0$. The problem was partly solved by considering the diffusivity in its stationary regime so the distribution of displacements is a superposition of Gaussian distributions weighted by the distribution of diffusivity known as superstatistics [14, 15]. From the chosen stationary distribution of diffusivity

$$\pi(D) = \frac{1}{D_0} \exp\left(-D/D_0 \right), \tag{1.45}$$

Chubynsky and Slater deduced the purely exponential propagator

$$P(x, t) = \frac{1}{2\sqrt{2 D_0}} \exp\left(-\frac{|x|}{2 D_0 t} \right). \tag{1.46}$$

This propagator was in agreement with the short-time behavior experimentally found in [278] but could not reproduce the convergence to Gaussian distribution at long time.

1.3.3 The square of the n-dimensional Ornstein-Uhlenbeck process

Jain and Sebastian solved the time-dependent problem in the case when diffusivity is the distance to origin of an n-dimensional Ornstein-Uhlenbeck (OU) process. The OU process is a Gaussian process modeling the diffusive motion of a particle trapped by a harmonic potential. This is a widely used model of particle interactions (e.g., a Rouse or bead-spring model in polymer

physics [67, 74]) and of the trapping effect of optical tweezers [26, 116, 118, 156, 286]. The one-dimensional OU process can be defined as a solution of the Langevin equation

$$dX_t = -\frac{1}{\tau}X_t dt + \sqrt{2D}dW_t, \tag{1.47}$$

where τ is the correlation time (which is related to the inverse of spring constant), D is the diffusion coefficient, and W_t is the standard Brownian motion.

For the solution they used a phase space integration technique [135, 137]. The probability distribution is given by

$$\tilde{P}(q,t) = \int_{-\infty}^{\infty} dq e^{iqx} \left(\frac{4\alpha e^{-(\alpha-1)t/\tau}}{(\alpha+1)^2 + (\alpha-1)^2 e^{-2\alpha t/\tau}} \right)^n, \tag{1.48}$$

with $\alpha = \sqrt{1+4q^2}$. This solution generalizes Chubinsky and Slater's one which is retrieved for $n = 2$. Additionally they were able to reproduce the convergence to Gaussian distribution at long time, as illustrated by the non-Gaussian parameter

$$\gamma(t) = \frac{2\tau}{nt} - \frac{2\tau^2}{nt^2}(1 - e^{-2t/\tau}). \tag{1.49}$$

Chechkin *et al.* [49] solved the same problem using the subordination technique and pointed out that superstatistical description matches diffusing diffusivity only at short times as it cannot reproduce convergence to a Gaussian distribution at long times. They also showed, in the superstatistical approach that taking a stretched exponential form of the distribution of diffusion coefficients they also obtained a stretched exponential propagator $P(x,t)$ with a power law tail in agreement with [143].
In all of these cases, the diffusivity fluctuations do not affect the MSD behavior which remains linear with respect to time.

Jain and Sebastian also generalized the approach to the case where diffusivity is an Ornstein-Uhlenbeck process driven by a one-sided (positive) Lévy noise with infinite first moment [136]. In this case the distribution was expressed as a combination of hypergeometric and exponential functions and the MSD is superdiffusive with exponent $\alpha > 1$. However this behavior has not been observed experimentally, yet.

Chapter 2

Revealing weak ergodicity breaking from a single trajectory

2.1 Introduction

Statistical analysis of a single random realization of an unknown stochastic process has become indispensable in various fields, from geosciences to microbiology and finances. In these fields, multiple realizations of a process are either impossible due to the non-repeatable unique character of observations (e.g., earth vibrations or stock prices), or undesirable due to spatial heterogeneity or time evolution of the medium (e.g., motion inside living cells). One therefore needs to resort to single observations to construct a mathematical or physical model of the unknown process. In the perspective of identifying the transport mechanism without prior knowledge, the identification of weak ergodicity breaking is a central point. The assessment of ergodicity property has a two-fold interest. First, it is a necessary point for a correct assessment of an appropriate stochastic process as weak ergodicity can be broken in several ways [16, 181]. Second, the calibration of model's parameters has to rely on the (implicit) ergodicity assumption to inter-change ensemble and time averages. However, the ergodicity can fail in active or aging systems such as living cells [13, 141, 280], viscoelastic media [288] or blinking nanocrystals [38, 182, 183]. A finite length of acquired trajectories and randomness of estimators make challenging verifications of ergodicity in single-particle tracking (SPT) experimental data. At the same time, this is a necessary step towards reliable biophysical interpretations: if the ergodicity breaking remains undetected,

any conclusion based on time averages along a single trajectory may be strongly misleading. We investigate in this chapter how to reveal weak ergodicity breaking from a single trajectory realization by generalizing the approach by Marcin Magdziarz and Aleksander Weron [175].

2.1.1 The dynamical functional

Magdziarz and Weron proposed to use the dynamical functional for testing ergodicity for a stationary infinitely divisible (SID) process, that is, a process $X(n)$ whose distribution can be written as the sum of independent identically distributed (IID) random variables.. Consider the process $Y(n)$ constructed from the increments of $X(n)$ such that $Y(n) = X(n+1) - X(n)$ with $n \in [0, N]$ where $N+1$ is the number of increments of the trajectory. The dynamical function $D(n)$ is the characteristic function of the variable $Y(n) - Y(0)$, for a frequency $\omega = 1$:

$$D(n) = \langle \exp[i(Y(n) - Y(0))\rangle \tag{2.1}$$

where $\langle \ldots \rangle$ denotes the ensemble average [175]. For a SID mixing process the dynamical functional verifies

$$\lim_{n\to\infty} D(n) = |\langle \exp[iY(0)]\rangle|^2. \tag{2.2}$$

From this result, they introduced the functional

$$E(n) \equiv \langle \exp(i[Y(n) - Y(0)])\rangle - |\langle \exp(iY(0))\rangle|^2. \tag{2.3}$$

This functional fully characterizes the mixing and ergodic properties of SID processes: a SID process is mixing (resp., ergodic) if and only if $E(n) \to 0$ (resp., of $n^{-1}\sum_{k=0}^{n-1} E(k) \to 0$) as $n \to \infty$. However in experimental conditions, one cannot acquire an infinitely long trajectory and often an ensemble of trajectories is hardly accessible.

In order to apply the functional 2.3 to a single trajectory, Magdiarz and Weron made the following argument: if $Y(n)$ is ergodic, the ensemble and time average can be interchanged, so the ensemble averaged dynamical functional will vanish as well as the time averaged one for an infinitely long trajectory. The estimator is obtained just by replacing the ensemble by the time average [175]:

$$\hat{E}(n) \equiv \frac{1}{N-n+1}\sum_{k=0}^{N-n} e^{i[Y(k+n)-Y(k)]} - \left|\sum_{k=0}^{N}\frac{e^{iY(k)}}{N+1}\right|^2. \tag{2.4}$$

The smallness of $\hat{E}(n)$ (resp., of $n^{-1}\sum_{k=0}^{n-1}\hat{E}(k)$) for large n is the necessary condition for mixing (resp., ergodicity) whereas violation of this condition reveals the mixing/ergodicity breaking. We emphasize that the estimators based on a single trajectory allow one to reject, with some degree of certainty, the mixing or ergodicity hypothesis but they cannot affirm it. For instance, the smallness of the estimator $\hat{E}(n)$ does not imply mixing.

2.1.2 Several issues

Applying a functional to a single random trajectory of finite length results in a random time series that need to be characterized through its statistical properties. The smallness of the estimators needs to be quantified (e.g., by comparing the mean estimator to its standard deviation or by determining the confidence intervals, see [138]). The mean mixing estimator from [175] for a Brownian motion is

$$\langle\hat{E}(n)\rangle = c^n - \frac{1}{N+1} - \frac{2c}{(N+1)(1-c)}\left(1 - \frac{1-c^{N+1}}{(N+1)(1-c)}\right), \tag{2.5}$$

with $c = e^{-\sigma^2/2}$ where σ is the standard deviation of the increments. Several remarks can be made.

• First, regardless of the process considered, there is an intrinsic $-\frac{1}{N+1}$ bias due to the finite length of the trajectory, which has to be removed.

• Second, the estimator depends directly on the scale of fluctuations σ: a trajectory with a smaller diffusion coefficient will have its estimator vanishing slower so that a process with a small diffusivity can be wrongly classified as non ergodic. The impact of the dependence of the estimator on σ and on a finite trajectory length, due to which the estimators can be relatively large even for mixing/ergodic processes, should be reduced and controlled.

• Finally, the estimator fails to identify the nonergodic continuous time random walk (CTRW) as nonergodic. Figure 2.1 shows that the estimator averaged over an ensemble of CTRW trajectories remains very close to zero, not detecting the nonergodic behavior. This last point was the major motivation for our work because macromolecular crowding strongly affects the intracellular and membrane transport [73, 129, 251] and can yield nonergodic features which are often modeled by nonergodic CTRW [141, 288]. As a consequence, a reliable analysis of single-particle tracking in living cells urges for developing statistical tests to identify single trajectories for which time averages are not representative, either due to ergodicity breaking, or nonstationarity of the

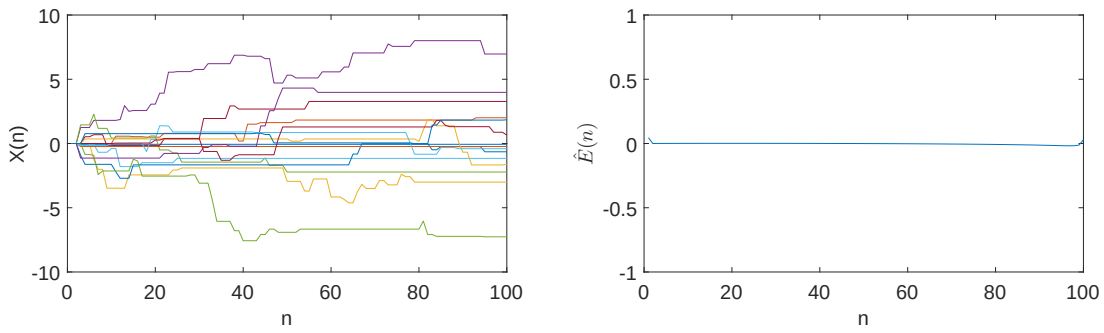


Figure 2.1: *Left.* Illustration of 15 one-dimensional trajectories of non ergodic CTRW with Gaussian jumps and power law waiting times $\phi(t) \propto 1/t^{1-\alpha}$ where the exponent is $\alpha = 0.7$. *Right.* The corresponding mixing estimator averaged over 10^5 realizations of the CTRW with the same parameters.

process.

2.1.3 Improved estimators

We resolve the above issues by modifying the mixing estimator as

$$\begin{aligned}\hat{E}_\omega(n,N) \equiv & \frac{1}{N-n+1}\sum_{k=0}^{N-n} e^{i\omega[X(k+n)-X(k)]} \\ & - \frac{1}{N(N+1)}\left|\sum_{k=0}^{N} e^{i\omega[X(k)-X(0)]}\right|^2 + \frac{1}{N}.\end{aligned} \tag{2.6}$$

The first term can be interpreted as the time averaged characteristic function of the increment $X(k+n)-X(k)$ at lag time n, while the second term ensures that the estimator is strictly 0 for a constant process $X(n) = X_0$ (in addition, the mean estimator is strictly 0 for a process with independent $X(n)$). The ergodicity estimator generalizes to

$$\hat{F}_\omega(n,N) \equiv \frac{1}{n}\sum_{k=1}^{n} \hat{E}_\omega(k,N), \tag{2.7}$$

where the summation over k was shifted from the original range $0,\ldots,n-1$ for convenience.

There are three modifications with respect to the original estimators: (i) we consider all Fourier modes, not only $\omega = 1$, (ii) we partly remove the bias by subtracting the constant term from the second sum and changing accordingly the normalization, and (iii) most importantly, we

apply the estimators to the *long-time increments* (or *positions*) of a trajectory, not to the *short-time increments* (or *velocities*), see further discussion in Sec. 2.6.1. Note that each coordinate of a two- or three-dimensional trajectory is considered separately, i.e., $X(n)$ is restricted to be a one-dimensional process. Formally, the inclusion of the frequency ω can be seen as rescaling of the process [1]. While such a rescaling does not change the mixing or ergodic property of the process and is thus redundant in the limit $n \to \infty$, high-frequency modes with large ω become important for a finite length trajectory. Even though the changes between Eq. (2.4) and Eq. (2.6) are minor, the application of the improved estimators to the positions of a tracer is the key feature. Rederiving the properties of the new estimators from scratch, we manage to reveal the nonergodic character of both model and experimental diffusive processes from a single trajectory.

To investigate the improved mixing and ergodicity estimators (2.6, 2.7), we consider several models of anomalous diffusion: fractional Brownian motion (fBm) [154, 180], diffusion on percolating clusters [100], CTRW [32, 153]. The first two processes are ergodic while the latter is not. We also consider two intermediate cases which exhibit nonergodic features at short times (non-equality between time and ensemble averages) but are actually ergodic as in the long time limit both averages coincide. These processes, formerly qualified as "mildly nonergodic" (which is mileading), are the scaled Brownian motion (sBm) [142,263] and the CTRW with exponential cut-off. These in between cases are important to analyze because in experimental conditions, a short trajectory may be classified as nonergodic due to finite size effect while at longer time the classification would change. Finally, we will test geometric Brownian motion (gBm) which was reported to be nonergodic [217].

In Sec. 2.2 we present the analytical results on the mean and variance of the estimator for Gaussian processes, while corresponding results for CTRW are developped in Sec. 2.3. In Sec. 2.4 are discussed the ergodic properties of other models of anomalous diffusion. The application of the estimator on simulated and experimental data is presented in Sec. 2.5. Section 2.6 presents some technicalities. The results of this chapter have been published in [159].

[1] Historically, first papers on mixing/ergodic properties of SID processes required a dynamical functional in a form similar to Eq. (2.3) to vanish as $n \to \infty$ for any frequency ω (for details, see [234], Chapter 14.4). In other words, to validate the mixing property, one needed to check the behavior of the functional for any ω that was impractical. Magdziarz and Weron have made a significant improvement by reducing the analysis to a single value of ω (they set $\omega = 1$). Here, we re-introduce the frequency ω as a mean to normalize the process but we still consider a single value of ω (which may different from 1).

2.2 Gaussian processes

In this section the main analytical results on mean and variance of the estimator concerning Gaussian processes are presented. The reader not interested in the mathematical derivation can skip directly to the conclusion in Sec. 2.2.4.

2.2.1 Mean estimators

The fBm at discrete time steps is a centered Gaussian process with $\langle [X(k+n)-X(k)]^2 \rangle = \sigma^2 n^{2H}$, where σ^2 is the variance of one-step displacement, and $0 < H < 1$ is the Hurst exponent. Using the identity

$$\langle e^{i\omega[X(k+n)-X(k)]} \rangle = e^{-\frac{1}{2}\omega^2 \langle [X(k+n)-X(k)]^2 \rangle}, \tag{2.8}$$

which is valid for any discrete centered Gaussian process, we compute

$$\langle \hat{E}_\omega(n,N) \rangle = c^{n^{2H}} - 2\sum_{k=1}^{N} \frac{N+1-k}{N(N+1)} c^{k^{2H}}, \tag{2.9}$$

$$\langle \hat{F}_\omega(n,N) \rangle = \frac{1}{n}\sum_{k=1}^{n} c^{k^{2H}} - 2\sum_{k=1}^{N} \frac{N+1-k}{N(N+1)} c^{k^{2H}}, \tag{2.10}$$

where $c = e^{-\omega^2\sigma^2/2}$. For Brownian motion ($H = 1/2$), one gets explicitly

$$\langle \hat{E}_\omega(n,N) \rangle = c^n - \frac{2c\big(1 - \frac{1-c^{N+1}}{(N+1)(1-c)}\big)}{N(1-c)}, \tag{2.11}$$

$$\langle \hat{F}_\omega(n,N) \rangle = c\frac{1-c^n}{n(1-c)} - \frac{2c\big(1 - \frac{1-c^{N+1}}{(N+1)(1-c)}\big)}{N(1-c)}. \tag{2.12}$$

Both $\langle \hat{E}_\omega(n,N) \rangle$ and $\langle \hat{F}_\omega(n,N) \rangle$ monotonously decrease with n and approach to the limit given by the second term, but the decrease of the mean ergodicity estimator is much slower (as $1/n$). For both estimators, the second term presents the bias which vanishes as either $N \to \infty$ or $\omega \to \infty$. While the trajectory length N is fixed by experimental setup, the frequency ω of the estimator can be increased at will. Note that the mean of the original estimator (2.4) contains the term $-1/(N+1)$ that could not be removed by varying ω.

For fixed n and N, the mean mixing estimator $\langle \hat{E}_\omega(n,N) \rangle$ from Eq. (2.11) as a function of ω exhibits nonmonotonous behavior. When $n \ll N$ and $N \gg 1$, the estimator can be approximated as $c^n - \frac{2}{N(1-c)}$ that reaches the maximum at $\sigma\omega_c \simeq (8/(nN))^{1/4}$. As a consequence, the estimator

with $\omega = 1$ would classify Brownian motion with $\sigma \lesssim (8/N)^{1/4}$ as a nonergodic process. This finite length effect can be eliminated by varying the frequency ω (see Sec. 2.6.2).

2.2.2 Variance of the estimators

The variance of the mixing estimator,

$$\mathrm{var}\{\hat{E}_\omega(n,N)\} = \langle \left|\hat{E}_\omega(n,N)\right|^2\rangle - \left|\langle \hat{E}_\omega(n,N)\rangle\right|^2, \tag{2.13}$$

can be formally expressed in terms of the covariance matrix, with

$$\langle \hat{E}_\omega(n,N)\rangle = \frac{1}{N-n+1}\sum_{k=0}^{N-n} C_{k+n,k,0,0} - \frac{1}{N(N+1)}\sum_{k_1\neq k_2=0}^{N} C_{k_1,k_2,0,0}, \tag{2.14}$$

and

$$\begin{aligned}\langle |\hat{E}_\omega(n,N)|^2\rangle =& \frac{1}{(N-n+1)^2}\sum_{k,k'=0}^{N-n} C_{k+n,k,k'+n,k'} \\ &- \frac{2}{(N-n+1)N(N+1)}\sum_{k=0}^{N-n}\sum_{k_1\neq k_2=0}^{N} C_{k+n,k,k_1,k_2} \\ &+ \frac{1}{N^2(N+1)^2}\sum_{k_1\neq k_2,k_3\neq k_4}^{N} C_{k_1,k_2,k_3,k_4},\end{aligned} \tag{2.15}$$

where

$$C_{k_1,k_2,k_3,k_4} = e^{-\frac{1}{2}\omega^2\langle (X(k_1)-X(k_2)-X(k_3)+X(k_4))^2\rangle} \tag{2.16}$$

includes the elements of the covariance matrix. Even for Brownian motion, the combinatorial computation of all terms in Eq. (2.15) is tedious while the formulas are cumbersome. It is more instructive to investigate the variance in the limit $\omega \to \infty$, in which only the terms $C_{k,k,k',k'} = C_{k,k',k,k'} = 1$ do not vanish exponentially. Keeping only these terms, one easily finds

$$\mathrm{var}\{\hat{E}_\infty(n,N)\} = \frac{1}{N-n+1} - \frac{1}{N(N+1)}. \tag{2.17}$$

The variance is of the order of $1/N$ for small n but progressively grows up to 1 at $n = N$. For long trajectories ($N \gg 1$), the second term can be neglected.

The same technique yields the large ω asymptotic behavior of the variance of the ergodicity estimator:

$$\mathrm{var}\{\hat{F}_\infty(n,N)\} = \frac{1}{n^2}\sum_{k=1}^{n}\frac{1}{N-k+1} - \frac{1}{N(N+1)}, \tag{2.18}$$

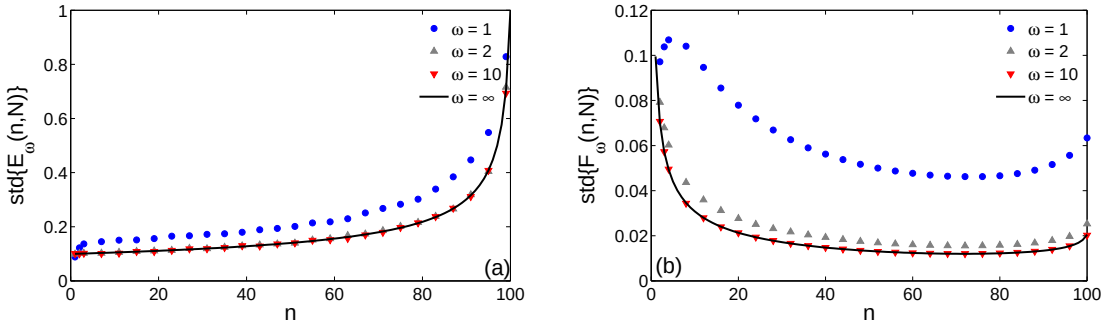


Figure 2.2: The standard deviation of the mixing **(a)** and ergodicity **(b)** estimators at several values of ω for Brownian motion with $N = 100$ and $\sigma = 1$. The large ω asymptotic limits given by square roots of Eqs. (2.17, 2.18) are shown by solid lines.

in which the second term can be neglected for $N \gg 1$. When $1 \ll n \ll N$, the sum can be approximated as $\ln(N/(N-n+1)) \simeq n/N$, i.e., the variance decreases with n as $1/(nN)$, in contrast to the increasing variance of the mixing estimator in Eq. (2.17). In other words, summing contributions from different lag times greatly reduces fluctuations so that the ergodicity estimator applied to a single trajectory yields less noisy results.

For the case of Brownian motion, Fig. 2.2 illustrates how the standard deviation of the mixing and ergodicity estimators approaches their asymptotic limits (given by square roots of Eqs. (2.17, 2.18)) as ω increases. One can see that the asymptotic formulas become accurate approximations for $\omega\sigma \gtrsim 2$.

2.2.3 Brownian motion with two diffusion coefficients

To illustrate the potential impact of time-dependent diffusion coefficient onto the mixing estimator, we consider the trajectory concatenating two Brownian trajectories with distinct diffusion coefficients D_1 and D_2. In the discrete case, one can generate such process by adding independent Gaussian variables with variance $\sigma_1^2 = 2D_1\delta$ up to the step $m-1$, and then completing the second part by independent Gaussian variables with $\sigma_2^2 = 2D_2\delta$, $m - 1/2$ being the "border" between two parts. The computation of the mean mixing estimator is cumbersome but straightforward:

$$\langle \hat{E}_\omega(n,N) \rangle = \frac{S_1}{N-n+1} - \frac{2S_2}{N(N+1)}, \tag{2.19}$$

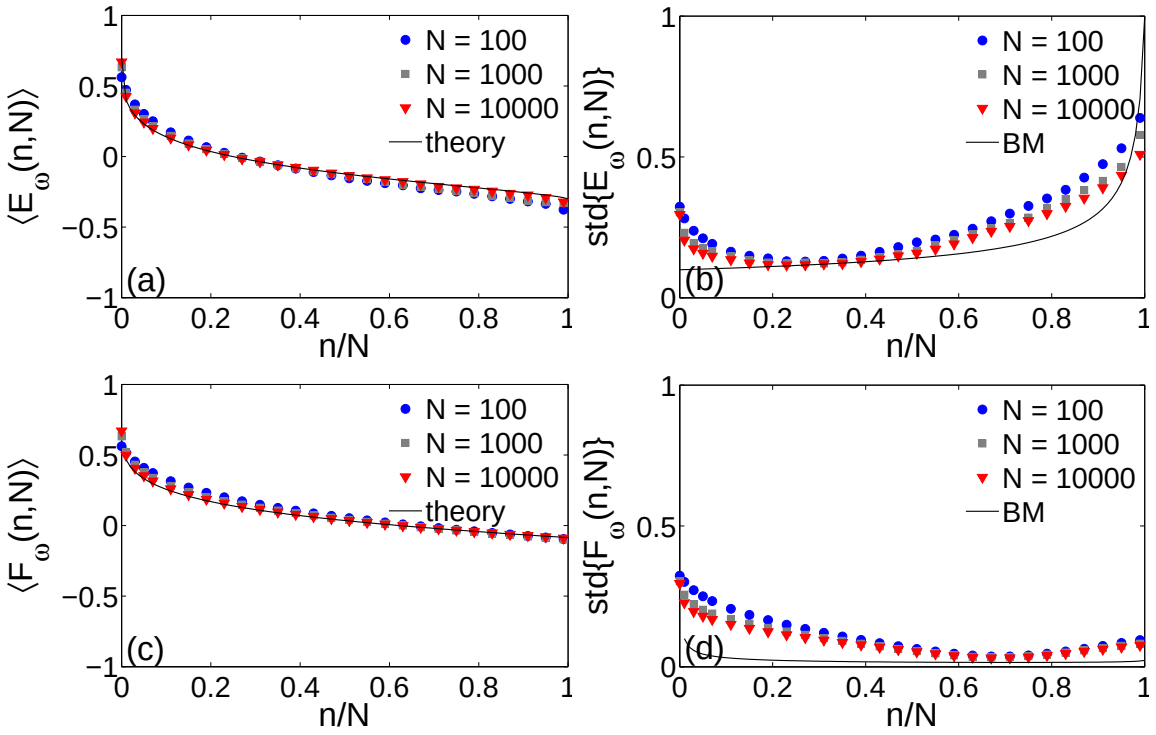


Figure 2.3: The mean (**a,c**) and standard deviation (**b,d**) of the mixing (**a,b**) and ergodicity (**c,d**) estimators at $\omega = 10$ as a function of n/N for CTRW with $\alpha = 0.7$, $\sigma = 1$, and three trajectory lengths $N = 100, 1000, 10000$ (symbols). For comparison, solid line shows the theoretical limits (2.51, 2.52) of the mean estimators for CTRW and the standard deviation for Brownian motion (given by square roots of Eqs. (2.17, 2.18) with $N = 100$).

where

$$S_1 = \begin{cases} m \le N/2 & \begin{cases} (m-n)c_1^n + (N+1-m-n)c_2^n + \frac{c_2(c_1^n - c_2^n)}{c_2 - c_1} & (n < m), \\ c_2^n (N+1-n-m + \frac{c_2^{1-m}(c_1^m - c_2^m)}{c_1 - c_2}) & (m \le n \le N-m), \\ c_1^m c_2^{n-m+1} \frac{1-(c_2/c_1)^{N+1-n}}{c_1 - c_2} & (n > N-m), \end{cases} \\ m > N/2 & \begin{cases} (m-n)c_1^n + (N+1-m-n)c_2^n + \frac{c_2(c_1^n - c_2^n)}{c_1 - c_2} & (n \le N-m), \\ (m-n)c_1^n + c_2 c_1^n \frac{1-(c_2/c_1)^{N+1-m}}{c_1 - c_2} & (N-m < n < m), \\ c_1^m c_2^{n-m+1} \frac{1-(c_2/c_1)^{N+1-n}}{c_1 - c_2} & (n \ge m), \end{cases} \end{cases} \tag{2.20}$$

$$S_2 = -(N+1) + \frac{m-(m+1)c_1+c_1^{m+1}}{(1-c_1)^2} + \frac{N-m+1-(N-m+2)c_2+c_2^{N-m+2}}{(1-c_2)^2} + \frac{c_2(1-c_1^m)(1-c_2^{N-m+1})}{(1-c_1)(1-c_2)}, \tag{2.21}$$

$$\tag{2.22}$$

with $c_1 = e^{-\omega^2\sigma_1^2/2}$ and $c_2 = e^{-\omega^2\sigma_2^2/2}$. The estimator vanishes as $\omega \to \infty$, as expected. Setting $m = 1$ (or $m = N+1$), one retrieves the mean mixing estimator in Eq. (2.11) for Brownian motion.

However, normalizing the process by the standard deviation of its increments does not resolve the problem of false nonmixing classifications. In fact, setting ω^2 to be the inverse of the variance of increments of the whole trajectory, $\omega^{-2} = \mu\sigma_1^2 + (1-\mu)\sigma_2^2$ (with $\mu = (m-1)/N$), yields

$$c_1 = e^{-\frac{1}{2}[\mu + (1-\mu)\nu]^{-1}}, \qquad c_2 = e^{-\frac{1}{2}\nu[\mu + (1-\mu)\nu]^{-1}}, \tag{2.23}$$

where $\nu = \sigma_2^2/\sigma_1^2$. As a consequence, varying μ and ν, one can make c_1 or c_2 small enough so that the mixing estimator would not appear as small, wrongly suggesting nonmixing behavior (similar conclusion holds for nonergodicity).

2.2.4 Main results

The estimator for fractional Brownian motion (including Brownian motion) vanishes in the long time limit verifying the expected ergodicity property. The mean mixing estimator in Eg. 2.9 decays exponentially with n with increasing variance while the ergodicity one from Eq. 2.2.1 decays as $1/n$ with diminishing variance. When the process alternates between two Brownian motions with different diffusion coefficients, both mean estimators vanish as $N \to \infty$ and $w \to \infty$ but the normalization by standard deviation may not be sufficient for correct assessment from a short trajectory.

2.3 Continuous Time Random Walks

We now turn to CTRW for which long stalling periods between successive jumps lead to the nonergodic behavior [16] (see also [194] and references therein). The reader only interested in the results can skip directly to the conclusion in Sec. 2.3.3. Using the renewal technique, we derive the exact expressions for the mean mixing and ergodicity estimators. We consider the CTRW with independent centered Gaussian jumps of variance σ^2, separated by independent waiting times with a prescribed probability density $\psi(t)$. Throughout this Section, we use the continuous-time version of the estimators, in which k, n, and N are placed by t, Δ, and T, respectively:

$$\begin{aligned}\hat{E}_\omega(\Delta,T) = \frac{1}{T-\Delta}\int_0^{T-\Delta} dt\ e^{i\omega[X(t+\Delta)-X(t)]} \\ - \frac{1}{T^2}\int_0^T dt_1 \int_0^T dt_2\ e^{i\omega[X(t_1)-X(t_2)]}.\end{aligned} \tag{2.24}$$

In order to obtain the mean $\langle\hat{E}_\omega(\Delta,T)\rangle$, one needs to compute the expectation

$$h_\omega(t,\Delta) \equiv \langle e^{i\omega[X(t+\Delta)-X(t)]}\rangle. \tag{2.25}$$

The mean mixing estimator can then be expressed as

$$\langle\hat{E}_\omega(\Delta,T)\rangle = \frac{H_\omega(T-\Delta,\Delta)}{T-\Delta} - \frac{2}{T^2}\int_0^T d\Delta\ H_\omega(T-\Delta,\Delta), \tag{2.26}$$

where

$$H_\omega(t,\Delta) = \int_0^t dt'\ h_\omega(t',\Delta). \tag{2.27}$$

Note that $H_\omega(T-\Delta,\Delta)/(T-\Delta)$ is the mean time averaged characteristic function of increments of a CTRW.

For CTRWs, the expectation in Eq. (2.25) includes the average over the jump distribution and the average over the waiting time distribution. The first average is elementary for a Gaussian jump distribution:

$$h_\omega(t,\Delta) = \langle e^{-\frac{1}{2}\omega^2\sigma^2\mathcal{N}(t,t+\Delta)}\rangle_\psi, \tag{2.28}$$

where $\mathcal{N}(t, t+\Delta)$ is the random number of jumps between times t and $t+\Delta$, and $\langle \ldots \rangle_\psi$ denotes the average over waiting times. In other words, $h_\omega(t,\Delta)$ is the Laplace transform of the probability density of $\mathcal{N}(t, t+\Delta)$ with respect to $\frac{1}{2}\omega^2\sigma^2$.

2.3.1 Number of jumps

The further computation of the function $h_\omega(t,\Delta)$ relies on renewal techniques [153]. First, the joint probability $P_{k,n}(t,\Delta)$ for getting k jumps in the interval $(0,t)$ and $n>0$ jumps in the interval $(t, t+\Delta)$ can be written as

$$\begin{aligned} P_{k,n}(t,\Delta) &= \int_0^t dt' \psi_k(t') \int_t^{t+\Delta} dt_1 \psi(t_1 - t') \int_{t_1}^{t+\Delta} dt_2 \psi(t_2 - t_1) \ldots \\ &\quad \ldots \times \int_{t_{n-1}}^{t+\Delta} dt_n \psi(t_n - t_{n-1})\ \Psi_0(t+\Delta - t_n), \\ &= \int_0^t dt' \psi_k(t') \int_0^{\Delta} dt_1 \psi(t_1 + t - t') \int_{t_1}^{\Delta} dt_2 \psi(t_2 - t_1) \ldots \\ &\quad \ldots \times \int_{t_{n-1}}^{\Delta} dt_n \psi(t_n - t_{n-1})\ \Psi_0(\Delta - t_n), \end{aligned} \tag{2.29}$$

where $\Psi_0(t)$ is the probability of no jump until time t, and $\psi_k(t)$ is the probability density for the k-th jump at time t. The Laplace transform with respect to Δ yields

$$\begin{aligned} \tilde{P}_{k,n}(t,s) &\equiv \int_0^\infty d\Delta\ e^{-s\Delta} P_{k,n}(t,\Delta) \\ &= \int_0^t dt' \psi_k(t')\ \tilde{\psi}_{t-t'}(s)\ \left[\tilde{\psi}(s)\right]^{n-1} \frac{1-\tilde{\psi}(s)}{s}, \end{aligned} \tag{2.30}$$

where

$$\tilde{\psi}_t(s) \equiv \int_0^\infty dt'\ e^{-st'} \psi(t'+t). \tag{2.31}$$

The second Laplace transform with respect to t yields

$$\begin{aligned} \tilde{\tilde{P}}_{k,n}(s',s) &\equiv \int_0^\infty dt\ e^{-s't} \tilde{P}_{k,n}(t,s) \\ &= \left[\tilde{\psi}(s')\right]^k\ \tilde{\tilde{\psi}}(s',s)\ \left[\tilde{\psi}(s)\right]^{n-1} \frac{1-\tilde{\psi}(s)}{s}, \end{aligned} \tag{2.32}$$

where

$$\tilde{\tilde{\psi}}(s',s) \equiv \int_0^\infty dt e^{-st} \int_0^\infty dt' e^{-t's'} \psi(t+t') = \frac{\tilde{\psi}(s) - \tilde{\psi}(s')}{s'-s}. \tag{2.33}$$

For the special case $n = 0$, one gets

$$P_{k,0}(t,\Delta) = \int_0^t dt' \ \psi_k(t') \ \Psi_0(t+\Delta-t'), \tag{2.34}$$

from which

$$\begin{aligned} \tilde{\tilde{P}}_{k,0}(s',s) &= \left[\tilde{\psi}(s')\right]^k \ \tilde{\tilde{\Psi}}_0(s',s) \\ &= \left[\tilde{\psi}(s')\right]^k \ \frac{1}{ss'}\left[1 - \frac{s\tilde{\psi}(s') - s'\tilde{\psi}(s)}{s-s'}\right]. \end{aligned} \tag{2.35}$$

The average of the joint distribution $P_{k,n}(t,\Delta)$ over k yields the marginal distribution of n steps in the interval $(t, t+\Delta)$:

$$P_n(t,\Delta) \equiv \sum_{k=0}^\infty P_{k,n}(t,\Delta), \tag{2.36}$$

from which

$$\begin{aligned} \tilde{\tilde{P}}_0(s',s) &= \frac{1}{1-\tilde{\psi}(s')} \ \frac{1}{ss'}\left[1 - \frac{s\tilde{\psi}(s') - s'\tilde{\psi}(s)}{s-s'}\right], \\ \tilde{\tilde{P}}_n(s',s) &= \frac{1}{1-\tilde{\psi}(s')} \ \frac{\tilde{\psi}(s) - \tilde{\psi}(s')}{s'-s} \ \left[\tilde{\psi}(s)\right]^{n-1} \frac{1-\tilde{\psi}(s)}{s}. \end{aligned} \tag{2.37}$$

As a consequence, one gets

$$\begin{aligned} \tilde{\tilde{h}}_\omega(s',s) &= \sum_{n=0}^\infty e^{-\frac{1}{2}\omega^2\sigma^2 n} \tilde{\tilde{P}}_n(s',s) \\ &= \frac{1}{ss'(e^{\frac{1}{2}\omega^2\sigma^2} - \tilde{\psi}(s))}\left[1 - \tilde{\psi}(s) + \frac{e^{\frac{1}{2}\omega^2\sigma^2} - 1}{1-\tilde{\psi}(s')}\left(1 - \frac{s\tilde{\psi}(s') - s'\tilde{\psi}(s)}{s-s'}\right)\right]. \end{aligned} \tag{2.38}$$

Setting $\omega = 0$, one retrieves the probability normalization: $\tilde{\tilde{h}}_0(s',s) = 1/(ss')$ from which $h_0(t,\Delta) = 1$, as expected. The generating function $\tilde{\tilde{h}}_\omega(s',s)$ can be used to compute the moments of $\mathcal{N}(t,t+\Delta)$.

To retrieve $h_\omega(t,\Delta)$, one needs to perform the double inverse Laplace transform. In general, the Laplace inversion has to be performed numerically. For the special case of the exponential

waiting time distribution, the inversion becomes simple. Setting $\psi(t) = e^{-t/\delta}/\delta$ so that $\tilde{\psi}(s) = 1/(1+s\delta)$, one finds exactly

$$\tilde{\tilde{h}}_\omega(s', s) = \frac{1}{s'(s + (1 - e^{-\frac{1}{2}\omega^2\sigma^2})/\delta)}, \tag{2.39}$$

from which

$$h_\omega(t, \Delta) = \exp\bigl(-(1 - e^{-\frac{1}{2}\omega^2\sigma^2})\Delta/\delta\bigr). \tag{2.40}$$

This function corresponds to the Poisson probability distribution of the number of jumps in the interval $(t, t+\Delta)$ that does not depend on t as expected:

$$P_n(t, \Delta) = e^{-\Delta/\delta}\frac{(\Delta/\delta)^n}{n!}. \tag{2.41}$$

2.3.2 Macroscopic limit

In the macroscopic limit of large t and Δ, the above expressions can be simplified by considering small s and s'. In general, the Laplace-transformed probability density $\tilde{\psi}(s)$ behaves as

$$\tilde{\psi}(s) \simeq 1 - (\delta s)^\alpha + \ldots \quad (s \to 0), \tag{2.42}$$

where $0 < \alpha \leq 1$ is the scaling exponent, and δ is a time scale of one jump. This behavior incorporates both normal diffusion ($\alpha = 1$) with a finite mean waiting time δ and anomalous diffusion ($\alpha < 1$) with a heavy-tailed waiting time density: $\psi(t) \simeq \frac{\delta^\alpha}{|\Gamma(-\alpha)|}\, t^{-1-\alpha}$ as $t \to \infty$. In the macroscopic limit, the variance σ^2 of one jump scales as $2D_\alpha\delta^\alpha$, where D_α is the generalized diffusion coefficient. As a consequence, $e^{\frac{1}{2}\omega^2\sigma^2} \simeq 1 + D_\alpha\omega^2\delta^\alpha + \ldots$ In the lowest order in δ, Eq. (2.38) reads

$$\tilde{\tilde{h}}_\omega(s', s) \simeq \frac{1}{s'}\,\frac{s^{\alpha-1}}{D_\alpha\omega^2 + s^\alpha} + \frac{D_\alpha\omega^2(s'^{\alpha-1} - s^{\alpha-1})}{s'^\alpha(s - s')(D_\alpha\omega^2 + s^\alpha)}. \tag{2.43}$$

In particular, the derivative with respect to $\frac{1}{2}\omega^2\sigma^2 = D_\alpha\delta^\alpha\omega^2$ yields the mean number of jumps

$$\langle\tilde{\tilde{\mathcal{N}}}(s', s)\rangle \simeq \frac{s^{-\alpha} - s'^{-\alpha}}{s\delta^\alpha(s - s')}, \tag{2.44}$$

from which

$$\langle\mathcal{N}(t, t+\Delta)\rangle \simeq \frac{1}{\delta^\alpha}\int_0^\Delta dt'\frac{(t+t')^{\alpha-1}}{\Gamma(\alpha)} = \frac{(t+\Delta)^\alpha - t^\alpha}{\delta^\alpha\Gamma(\alpha+1)}, \tag{2.45}$$

as expected.

For normal diffusion ($\alpha = 1$), the second term in Eq. (2.43) vanishes while the double inverse Laplace transform of the first term yields

$$h_\omega(t,\Delta) \simeq \exp\bigl(-D_1\omega^2\Delta\bigr) \tag{2.46}$$

that approximates Eq. (2.40) in the macroscopic limit.

In turn, when $\alpha < 1$, the dominant contribution comes from the second term of Eq. (2.43), especially in the limit of large ω:

$$\tilde{\tilde{h}}_\infty(s',s) \simeq \frac{s'^{\alpha-1} - s^{\alpha-1}}{s'^{\alpha}(s-s')}. \tag{2.47}$$

Using the identity for the double Laplace transform,

$$\mathcal{L}_{s_1}\mathcal{L}_{s_2}\left\{\int\limits_0^{t_1} dt' f(t')g(t_1-t'+t_2)\right\} = \tilde{f}(s_1)\frac{\tilde{g}(s_2)-\tilde{g}(s_1)}{s_1-s_2}, \tag{2.48}$$

one can invert the above relation by setting $\tilde{f}(s) = s^{-\alpha}$ and $\tilde{g}(s) = s^{\alpha-1}$:

$$h_\infty(t,\Delta) \simeq \frac{\sin(\pi\alpha)}{\pi}\ \frac{{}_2F_1(\alpha,\alpha;\alpha+1;(1+\Delta/t)^{-1})}{\alpha(1+\Delta/t)^\alpha}, \tag{2.49}$$

where the integral in Eq. (2.48) was expressed in terms of the hypergeometric function ${}_2F_1(a,b;c;z)$:

$$\int\limits_0^1 dx\ \frac{x^{b-1}(1-x)^{c-b-1}}{(1-zx)^a} = \frac{\Gamma(b)\Gamma(c-b)}{\Gamma(c)}\ {}_2F_1(a,b;c;z).$$

To compute $H_\infty(t,\Delta)$ (i.e., the integral of $h_\infty(t,\Delta)$ over t in Eq. (2.27)), one can add the extra factor s'^{-1} to Eq. (2.47) and then again perform the Laplace inversion:

$$H_\infty(t,\Delta) \simeq t\frac{\sin(\pi\alpha)}{\pi\alpha}\ \frac{{}_2F_1(\alpha,\alpha+1;\alpha+2;(1+\Delta/t)^{-1})}{(1+\alpha)(1+\Delta/t)^\alpha} \tag{2.50}$$

that gives the first term in Eq. (2.26).

Using the formula 7.512.3 from [110] to integrate $H_\infty(T-\Delta,\Delta)$ over Δ, one shows that the second term in Eq. (2.26) is equal to $1-\alpha$. We obtain therefore the asymptotic limit of the mean mixing estimator as $\omega \to \infty$

$$\langle\hat{E}_\infty(\Delta,T)\rangle \simeq \alpha - 1 + \frac{\sin(\pi\alpha)}{\pi\alpha(1+\alpha)}\ (1-\Delta/T)^\alpha\ {}_2F_1\bigl(\alpha,\alpha+1;\alpha+2;1-\Delta/T\bigr). \tag{2.51}$$

As expected for a nonergodic CTRW, the estimator does not vanish even for infinitely long trajectories. In fact, when $T \to \infty$, the last term approaches 1 so that $\langle\hat{E}_\infty(\Delta,\infty)\rangle = \alpha$,

independently of Δ. The same limit is formally obtained as $\Delta \to 0$: $\langle \hat{E}_\infty(0,T)\rangle = \alpha$. In the opposite limit $\Delta \to T$, the last term vanishes, yielding the negative value $\alpha - 1$. In other words, when the lag time Δ varies from 0 to T, the mean mixing estimator decreases from α to $\alpha - 1$, in sharp contrast to the case of ergodic processes.

From Eq. (2.51), one also deduces the asymptotic behavior of the mean ergodicity estimator

$$\langle \hat{F}_\infty(\Delta,T)\rangle \simeq \frac{1}{\Delta}\int_0^\Delta d\Delta'\, \langle \hat{E}_\infty(\Delta',T)\rangle = \alpha - 1 + \frac{\sin(\pi\alpha)}{\pi\alpha(1+\alpha)}\,\frac{T}{\Delta}\int_{1-\Delta/T}^{1} dx\, x^\alpha\, {}_2F_1\big(\alpha,\alpha+1;\alpha+2;x\big). \tag{2.52}$$

We recall that Eqs. (2.51, 2.52) are derived in the macroscopic limit when Δ and T greatly exceed the time step scale δ. These expressions present one of the main analytical results of the chapter. The estimators do not vanish even for infinitely long trajectories: $\langle \hat{E}_\infty(n,\infty)\rangle = \langle \hat{F}_\infty(n,\infty)\rangle = \alpha$, independently of n. In sharp contrast to ergodic processes, the mean mixing estimator decreases from α to $\alpha - 1$ when n varies from 0 to N (Fig. 2.3a). Longer the trajectory (larger N), closer the numerical curves to the limiting relations (2.51, 2.52). Figure 2.3b,d shows the standard deviations of both estimators that weakly depend on the trajectory length N, in contrast to Eq. (2.13) for discrete Gaussian processes. As a consequence, increasing N does not improve the estimation quality in the case of nonergodic CTRW.

2.3.3 Main results

The mean mixing and ergodicity estimators have been calculated, their expressions are given in Eq. 2.51, and Eq. 2.52. Both mean estimators never vanish even in the limit $N \to \infty$ and $w \to \infty$ verifying the expected behavior for non ergodic CTRW.

2.4 Several models of anomalous diffusion

In this Section, we describe the results of nonergodicity testing on several models of anomalous diffusion and on geometric Brownian motion.

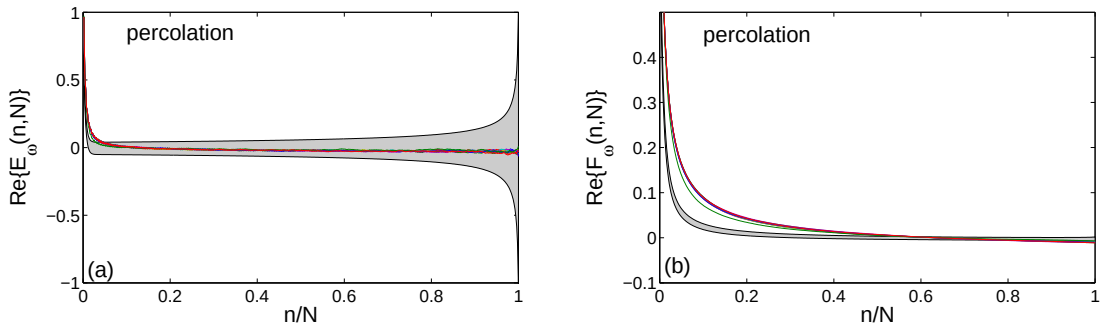


Figure 2.4: The mean mixing (**a**) and ergodicity (**b**) estimators at $\omega = 1$ as a function of n for one coordinate $X(n)$ of two-dimensional random walk of length $N = 500$ on a critical percolating cluster on a 1000×1000 square lattice, with the lattice step $\sigma = 1$. Each curve presents the mean computed by averaging over 1000 trajectories generated on one cluster. The curves obtained for ten random clusters are almost identical. For comparison, light gray shadowed region delimits the mean plus and minus the standard deviation of each estimator computed for Brownian motion according to Eqs. (2.11, 2.12, 2.17, 2.18).

2.4.1 Diffusion on percolating clusters

Anomalous diffusion on fractals which can mimic a multiscale hierarchical structure of the medium, is a common model for interpreting single particle tracking experiments. In particular, diffusion on percolating clusters has numerous applications for modeling transport phenomena in porous media [255]. For illustrative purposes, we only consider two-dimensional percolating clusters on a square lattice at the critical threshold probability $p_c \approx 0.59\ldots$ These clusters are known to have the fractal dimension $d_f = 91/48 \approx 1.896\ldots$ and to yield the anomalous diffusion with $\alpha \approx 0.7$ [100,205]. To test the mixing and ergodicity estimators, we first generate 10 random clusters on a 1000×1000 square lattice, and then simulate 1000 random walk trajectories of length $N = 500$ for each cluster. Figure 2.4 shows the mean mixing and ergodicity estimators obtained by averaging over 1000 trajectories, each curve representing the average over one cluster. These curves are almost indistinguishable, suggesting a very weak dependence on the particular random realization of the cluster. Both estimators do not reveal nonmixing/nonergodic behavior as expected. A small bias (deviation from 0) at large n can be attributed to the relatively short trajectory length: the bias is reduced for longer trajectories (not shown).

2.4.2 Scaled Brownian motion

We also consider the scaled Brownian motion (sBm), a simple model of anomalous diffusion, in which the diffusion coefficient varies with time as $D(t) = \alpha D_\alpha t^{\alpha-1}$, where $0 < \alpha < 2$ is the scaling exponent, and D_α is the generalized diffusion coefficient [142,263]. This is a nonstationary Gaussian process obtained by rescaling Brownian motion $W(t)$: $X(t) = \sqrt{2D_\alpha}\, W(t^\alpha)$, for which $\langle [X(t_1) - X(t_2)]^2 \rangle = 2D_\alpha |t_1^\alpha - t_2^\alpha|$. The mean mixing estimator is then

$$\langle \hat{E}_\omega(n, N) \rangle = \frac{1}{N-n+1} \sum_{k=0}^{N-n} c^{(k+n)^\alpha - k^\alpha} - \frac{2}{N(N+1)} \sum_{k_1=0}^{N-1} \sum_{k_2=k_1+1}^{N} c^{k_2^\alpha - k_1^\alpha}, \tag{2.53}$$

where $c = e^{-\omega^2\sigma^2/2}$ and $\sigma^2 = 2D_\alpha \delta^\alpha$, δ being the time step (the expression for the mean ergodicity estimator follows from its definition). In the limit of large $\omega\sigma$, both mean estimators vanish, showing no evidence for nonmixing/nonergodic behavior. At the same time, one can easily check that the ensemble averaged MSD, $\langle X^2(t) \rangle = 2D_\alpha t^\alpha$, differs from the time averaged MSD along the trajectory of length T

$$\frac{1}{T-t} \int_0^{T-t} dt_0 \, \langle [X(t_0+t) - X(t_0)]^2 \rangle \simeq 2D_\alpha T^{\alpha-1} t \tag{2.54}$$

(for $t \ll T$). In contrast to CTRW, this weak ergodicity breaking progressively vanishes as the trajectory length T (or N) goes to infinity [142, 263]. Figuratively speaking, the scaled Brownian motion falls in between ergodic and nonergodic processes, with nonergodic features appearing only for finite length trajectories. The mixing and ergodicity estimators do not capture this peculiar behavior that illustrates the difference between weak ergodicity breaking and non equality of ensemble and time average at short time.

2.4.3 CTRW with exponential cut-off

It is instructive to consider CTRW with exponential cut-off to "switch" between ergodic and nonergodic behavior. In order to simulate such CTRW trajectories, we generate random waiting times with the Pareto type III distribution defined by the cumulative function

$$F(t) = 1 - \left(1 + \frac{t}{\delta_s}\right)^{-\alpha} \exp\left(-\frac{t}{T_c}\right), \tag{2.55}$$

where δ_s is the time scale (fixed to be 1), α the scaling exponent, and T_c is the cut-off time. As discussed in [31], the cumulative function (2.55) can be explicitly inverted to generate the

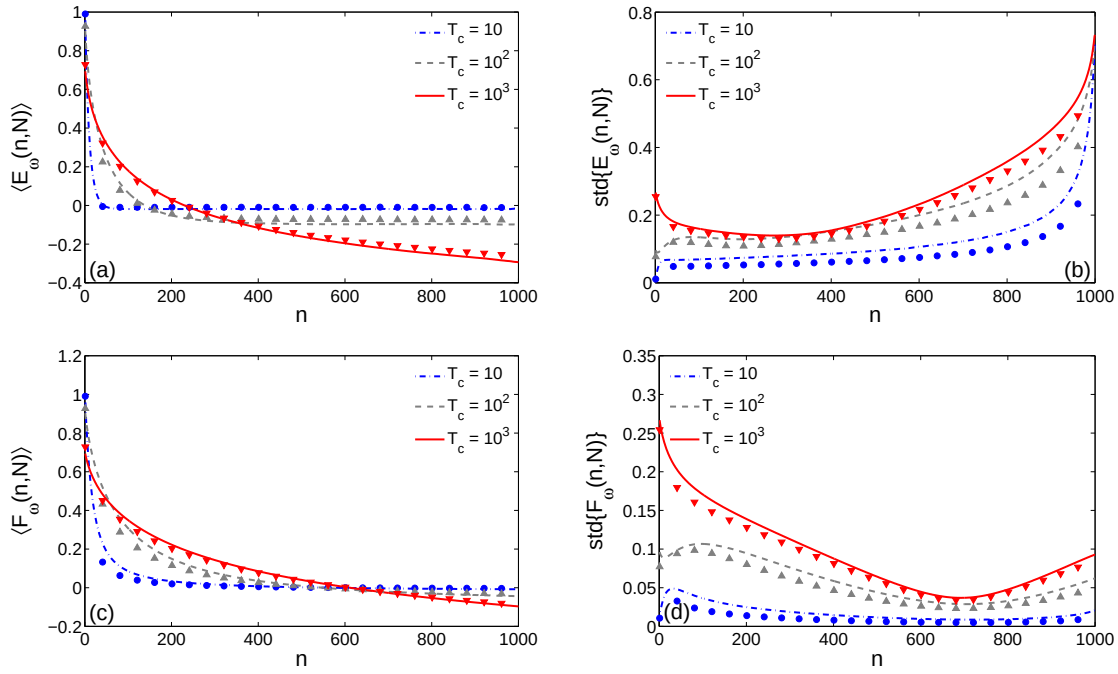


Figure 2.5: The mean (**a,c**) and standard deviation (**b,d**) of the mixing (**a,b**) and ergodicity (**c,d**) estimators at $\omega = 1$ (lines) and $\omega = 10$ (symbols) for CTRW with $N = 1000$, $\alpha = 0.7$, $\sigma = 1$, $\delta_s = 1$, and exponential cut-off at $T_c = 10, 10^2, 10^3$.

waiting times τ from the uniformly distributed variable η:

$$\tau = \alpha T_c \, W\left(\frac{\beta}{\alpha} \, e^{\beta/\alpha}(1-\eta)^{-1/\alpha}\right) - \delta_s, \tag{2.56}$$

where $\beta = \delta_s/T_c$, and $W(x)$ is the real branch of the Lambert function satisfying $W(x)e^{W(x)} = x$. In the limit $T_c \to \infty$ (no cut-off), one retrieves the standard Pareto waiting times generated as $\tau = \delta_s\big[(1-\eta)^{-1/\alpha} - 1\big]$.

The characteristic function and the moments of τ can be written as [31]

$$\begin{aligned} \langle e^{ik\tau} \rangle &= 1 + ik\delta_s \, e^{\beta - ik\delta}(\beta - ik\delta)^{\alpha-1}\Gamma(1-\alpha, \beta - ik\delta), \\ \langle \tau^n \rangle &= \delta_s^n e^{\beta} \sum_{j=1}^{n} \binom{n}{j} \frac{j(-1)^{n-j}}{\beta^{j-\alpha}} \Gamma(j-\alpha, \beta), \end{aligned}$$

where $\Gamma(k, z)$ is the incomplete Gamma function. In particular, the mean waiting time,

$$\langle \tau \rangle = \delta_s \, e^{\delta_s/T_c}(\delta_s/T_c)^{\alpha-1}\Gamma(1-\alpha, \delta_s/T_c), \tag{2.57}$$

asymptotically behaves as

$$\langle \tau \rangle \simeq \begin{cases} T_c(\delta_s/T_c)^{\alpha}\Gamma(1-\alpha), & (T_c \gg \delta_s), \\ T_c, & (T_c \ll \delta_s). \end{cases} \tag{2.58}$$

The generated sequence of waiting times, $\{\tau_k\}$, is then applied to produce positions of CTRW at equal time steps δ by assigning the same random position x_k to X_n over a time interval between $\tau_1 + \ldots + \tau_k$ and $\tau_1 + \ldots + \tau_k + \tau_{k+1}$:

$$X_n = \begin{cases} 0, & 0 \le n\delta < \tau_1, \\ x_k, & \sum_{j=1}^{k} \tau_j \le n\delta < \sum_{j=1}^{k+1} \tau_j, \end{cases} \tag{2.59}$$

where $x_k = x_{k-1} + \chi_k$, with χ_k being independent Gaussian displacements with mean zero and variance σ^2.

Figure 2.5 shows the mean and standard deviation of the mixing and ergodicity estimators as a function of n computed numerically for CTRW with exponential cut-off. When the cut-off time T_c is significantly smaller than the trajectory length (here, $N = 1000$), the mean mixing estimator $\langle \hat{E}_\omega(n, N) \rangle$ vanishes very rapidly with n (dash-dotted line or circles), as expected for Brownian motion. In turn, for larger T_c, long waiting times break the ergodicity and mixing (dashed and solid lines or triangles). Note also that $\langle \hat{E}_\omega(n, N) \rangle$ at $\omega = 1$ and $\omega = 10$ almost

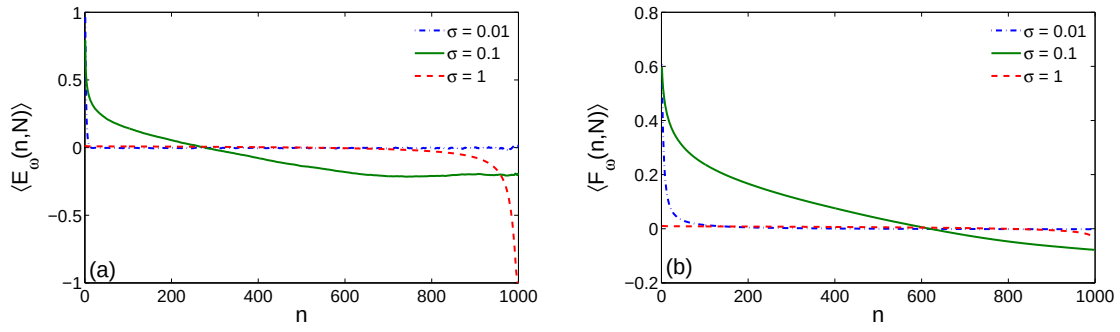


Figure 2.6: The mean mixing **(a)** and ergodicity **(b)** estimators at $\omega = 1$ as a function of n for geometric Brownian motion with $\mu = 0$ and $\sigma = 0.01$, $\sigma = 0.1$, and $\sigma = 1$. Each generated trajectory was normalized by the empirical standard deviation of its increments in order to get comparable results for different σ. The mean was computed numerically by averaging over 1000 trajectories.

coincide (similar for $\langle \hat{F}_\omega(n, N) \rangle$). For the ergodic case ($T_c = 10$), this is a visual artifact because the mixing estimator at $\omega = 10$ decreases much faster than that at $\omega = 1$. In turn, the weak dependence of the estimator on ω for large ω is expected for nonergodic CTRW due to the nontrivial limiting relations (2.51, 2.52).

The standard deviation of the mixing estimator is close to that given by square root of Eq. (2.17) for Brownian motion for $T_c = 10$ while it is larger for the nonergodic cases $T_c = 10^2$ and $T_c = 10^3$ (similar for the ergodicity estimator). As for the mean value, the standard deviation does not much depend on ω (once ω is large enough). The minimum of the standard deviation at an intermediate n for the nonergodic cases can be related to vanishing of the mean value when it crosses the horizontal axis.

2.4.4 Geometric Brownian motion

We consider geometric (or exponential) Brownian motion $X(t)$ which is the basic model in finance, in particular, in the Black-Scholes model for option pricing [131]. This stochastic process can be expressed as

$$X(t) = X(0) \exp\left((\mu - \sigma^2/2)t + \sigma W(t)\right), \tag{2.60}$$

where μ and σ are drift and standard deviation, $W(t)$ is a Wiener process (a standard Brownian motion), and the term $(\mu - \sigma^2/2)t$ is explicitly added in order to prevent exponential growth of the variance of $X(t)$. This nonstationary process was recently reported to be nonergodic [217] that may eventually affect current views on trading strategies [216]. After a random exploration time, the trajectory of geometric Brownian motion tends to remain close 0 for a very long time that explains the nonergodic behavior.

Figure 2.6 shows the mean mixing and ergodicity estimators at $\omega = 1$ as a function of n for geometric Brownian motion with $\mu = 0$. During the simulation time ($N = 1000$), the gBm with $\sigma = 0.01$ does not have enough time to be stuck near 0, yielding a rapid decay of both estimators (dash-dotted line). This is a finite length effect: much longer trajectories would attend the "trapped" state near 0 and thus exhibit nonergodic behavior (not shown). Instead of increasing the trajectory length, we take larger σ for which the trapped state is reached earlier on average. The nonmixing/nonergodic behavior is clearly seen for $\sigma = 0.1$ and $\sigma = 1$. The latter case is also instructive to illustrate that the estimator can take very small values for moderate n for a nonergodic process. This is due to the fact that a large part of the trajectory is almost 0. Inspecting the whole dependence on n can thus be informative.

2.5 Application

2.5.1 Measurement noise and practical issues

Since both estimators vanish in the limit $\omega \to \infty$ for mixing and ergodic processes but remain nonzero for nonergodic CTRW, the estimation at very large ω might be thought as optimal. However, this strategy is not convenient in practice because of measurement artifacts such as localization errors, blurring, or electronic noises [8, 24]. To account for some of these effects, the intrinsic trajectory $X(t)$ can be superimposed with a measurement noise $\epsilon(t)$: $\tilde{X}(t) = X(t) + \epsilon(t)$. If $\epsilon(t)$ is a white Gaussian noise independent of the tracer's dynamics $X(t)$, two contributions are factored out:

$$\langle e^{i\omega[\tilde{X}(n+k)-\tilde{X}(k)]}\rangle = \langle e^{i\omega[X(n+k)-X(k)]}\rangle \; e^{-\sigma_\varepsilon^2\omega^2}, \tag{2.61}$$

where the second factor is the average over the white noise of variance σ_ε^2 (when the measurement noise is not Gaussian, its effect onto the estimators can be different and needs further analysis).

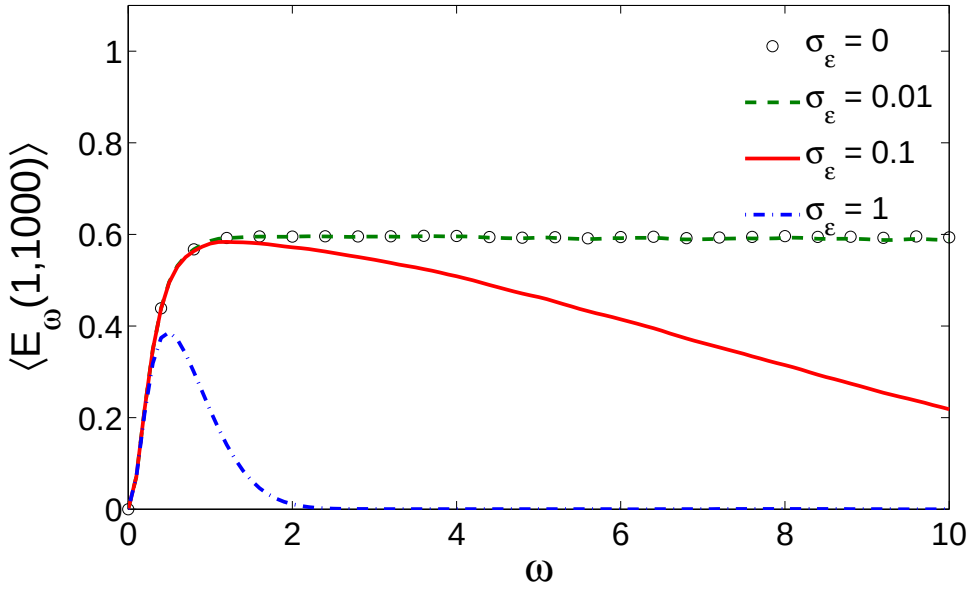


Figure 2.7: The mean mixing estimator $\langle \hat{E}_\omega(1,1000)\rangle$ as a function of ω for CTRW with $\alpha = 0.7$ and $\sigma = 1$, and four levels of white Gaussian noise: $\sigma_\varepsilon = 0$ (circles) and $\sigma_\varepsilon = 0.01, 0.1, 1$ (lines). Note that the plateau at 0.6 is smaller than the value $\alpha = 0.7$ expected from Eq. (2.51) because of the finite length effect ($N = 1000$).

Even if the intrinsic dynamics is not mixing/ergodic, both estimators will be strongly attenuated in the limit of large ω by the second factor coming from the ergodic white noise. This is illustrated on Fig. 2.7 which shows the mean mixing estimator as a function of frequency ω for a CTRW corrupted by white Gaussian noise with different σ_ε. When there is no noise ($\sigma_\varepsilon = 0$), the mean estimator rapidly saturates on a plateau, in agreement with the above theoretical analysis. In turn, the presence of noise attenuates the estimator. This effect is not yet seen at $\sigma_\varepsilon = 0.01$ because the factor $e^{-\omega^2\sigma_\varepsilon^2}$ remains close to 1 for the considered range of ω, but it is clearly seen for $\sigma_\varepsilon = 0.1$ and $\sigma_\varepsilon = 1$.

To limit this noise-induced attenuation, the rule of thumb consists in keeping $\omega\sigma$ of the order of 1, σ being the empirical standard deviation of increments. In practice, one can normalize the trajectory by σ and then consider ω between 1 and 3. On one hand, this normalization helps to eliminate false nonergodicity identifications due to too small σ. On the other hand, if the noise level σ_ε is much smaller than σ, the condition $\omega\sigma \sim 1$ ensures that the estimator is not

much attenuated due to noise (i.e., the factor $e^{-\sigma_\varepsilon^2\omega^2}$ remains close to 1). At the same time, one can construct counterexamples for which this normalization is not enough. For instance, the trajectory concatenating two Brownian motions with different diffusion coefficients can be made appearing as nonergodic (Section 2.2.3). Extending this construction to diffusion in a heterogeneous medium consisting of regions of random sizes and random diffusivities, one can produce a truly nonergodic process under simple assumptions on the distribution of the sizes and diffusivities [172, 184].

2.5.2 Validation on anomalous diffusion models

In order to validate the proposed statistical tool, we apply the mixing and ergodicity estimators to four single trajectories generated according to four anomalous diffusion models with the same exponent $\alpha = 2H = 0.7$: fBm, CTRW, sBm, and diffusion on percolating cluster (in the last case, the exponent is random and distributed around 0.7 due to the random shape of percolating clusters, see Appendix 2.4 for details). Each trajectory is generated with the same one-step variance $\sigma^2 = 1$. To render the comparison closer to the experimental situation, all trajectories were corrupted by white Gaussian noise with standard deviation $\sigma_\varepsilon = 0.2$. Figure 2.8 shows four simulated trajectories and the corresponding curves of the mixing and ergodicity estimators. One can clearly distinguish the nonmixing feature of a CTRW trajectory even from noisy curves of the mixing estimator. These curves allow one to reject the mixing hypothesis with a high degree of certainty. Similarly, the nonergodic behavior is seen from the ergodicity estimator. In turn, the estimator curves for three other trajectories vanish as n increases so that the mixing/ergodicity hypothesis cannot be rejected. However, it does not imply mixing or ergodicity for these models. Moreover, the "mild nonergodicity" of the scaled Brownian motion is not detected by both estimators due to a relatively short trajectory length. This is not surprising because the one-step standard deviation varies from 1 at $n = 1$ to $500^{(\alpha-1)/2} \approx 0.4$ at $n = 500$. To detect the nonergodic-like behavior of this model, one needs higher variations and thus much longer trajectories.

To illustrate the statistical variability of estimators, we repeat the same analysis for ten simulated trajectories for each model. Note that Janczura and Weron have managed to reveal the ergodic property of fBm from empirical ensemble averages over many trajectories (the smallest analyzed sample containing ten trajectories) [138]. Here, we aim at probing nonergodicity *indi-*

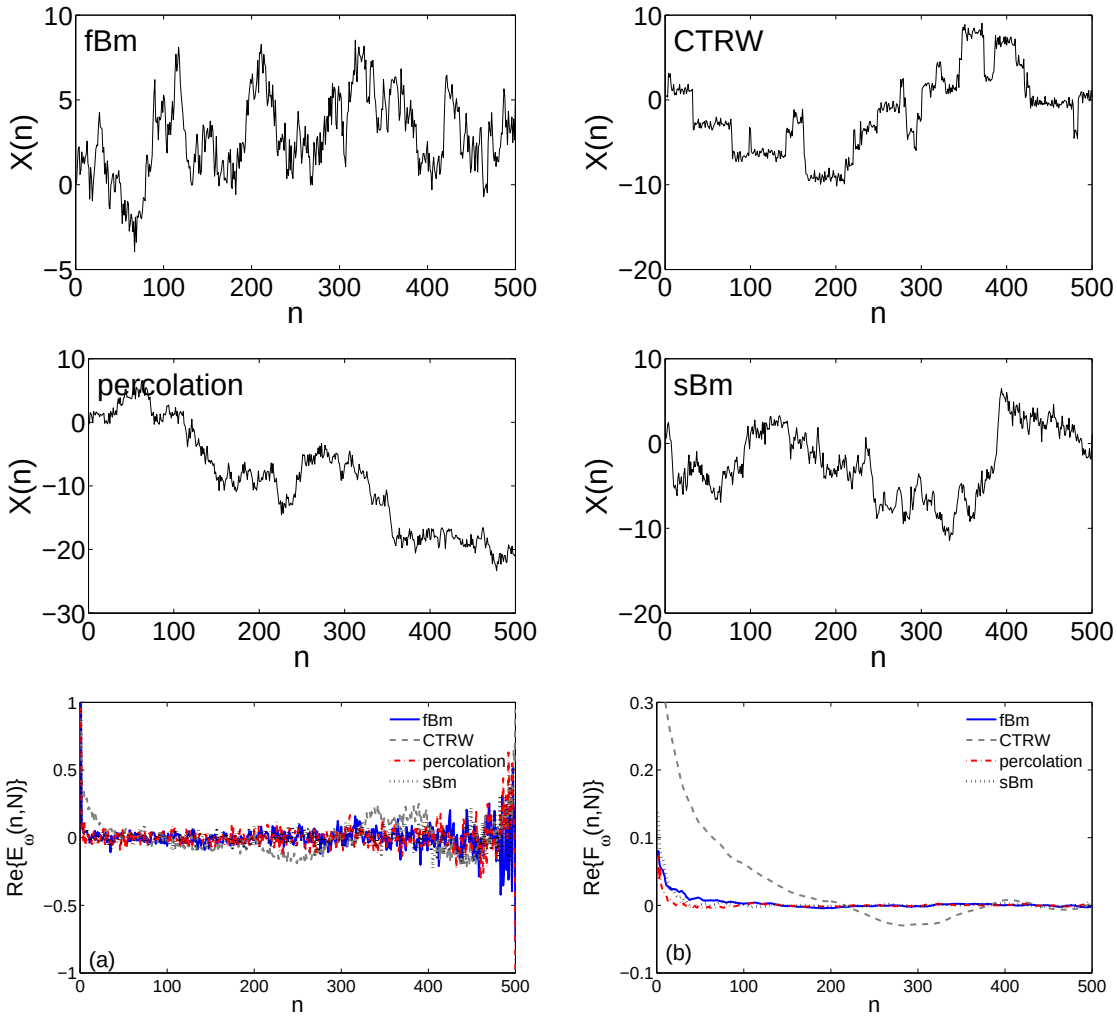


Figure 2.8: **(Top)**. Examples of trajectories of four anomalous diffusion models: fBm, CTRW, diffusion on percolating clusters, and sBm, with the same scaling exponent $\alpha = 0.7$ and one-step size $\sigma = 1$. Each trajectory was corrupted by the white Gaussian noise of level $\sigma_\varepsilon = 0.2$ and then normalized by the empirical standard deviation of its increments. **(Bottom)**. The real part of the mixing and ergodicity estimators at $\omega = 2$ applied to these four trajectories.

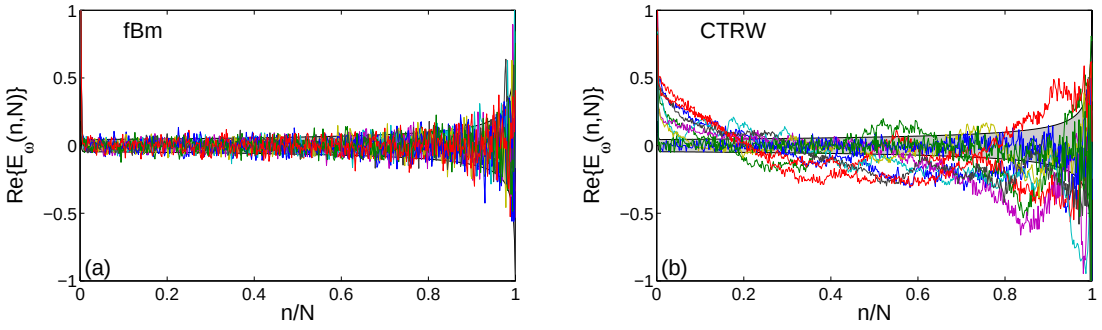


Figure 2.9: The real part of the mixing estimator at $\omega = 2$ for ten simulated trajectories from two anomalous diffusion models: ergodic fBm with $2H = 0.7$ (**a**) and nonergodic CTRW with the same exponent $\alpha = 0.7$ (**b**) (in both cases, we set $\sigma = 1$ and $N = 500$). All trajectories were corrupted by white Gaussian noise with standard deviation $\sigma_\varepsilon = 0.1$. Each trajectory is normalized by the empirical standard deviation of its increments. Light gray shadowed region delimits the typical range of fluctuations for Brownian motion, i.e., the mean plus and minus the standard deviation.

vidually for each single trajectory. Figure 2.9 shows the results for the mixing estimator (only for fBm and CTRW), while the ergodicity estimator is illustrated in Fig. 2.10. Both estimators allow one to clearly identify the nonergodic character of CTRW from a single trajectory, even for the trajectory length as small as $N = 500$. Since the ergodicity estimator yields much smoother curves, it is more appropriate for the analysis of single-particle tracking experiments (note that although mixing and ergodicity are not equivalent, they are satisfied or violated simultaneously in many diffusive processes).

In addition, both estimators were also successfully tested on CTRW with exponential cut-off which allows one to "switch" between ergodic and nonergodic behavior (Section 2.4.3), and on the nonergodic geometric Brownian motion (Section 2.4.4) that plays a major role in finances [216, 217].

2.5.3 Application to experimental data

Now, we apply both estimators to two samples of experimental trajectories: (i) mRNA molecules inside live *E. coli* cells [106], and (ii) Kv2.1 potassium channels in the plasma membrane [280].

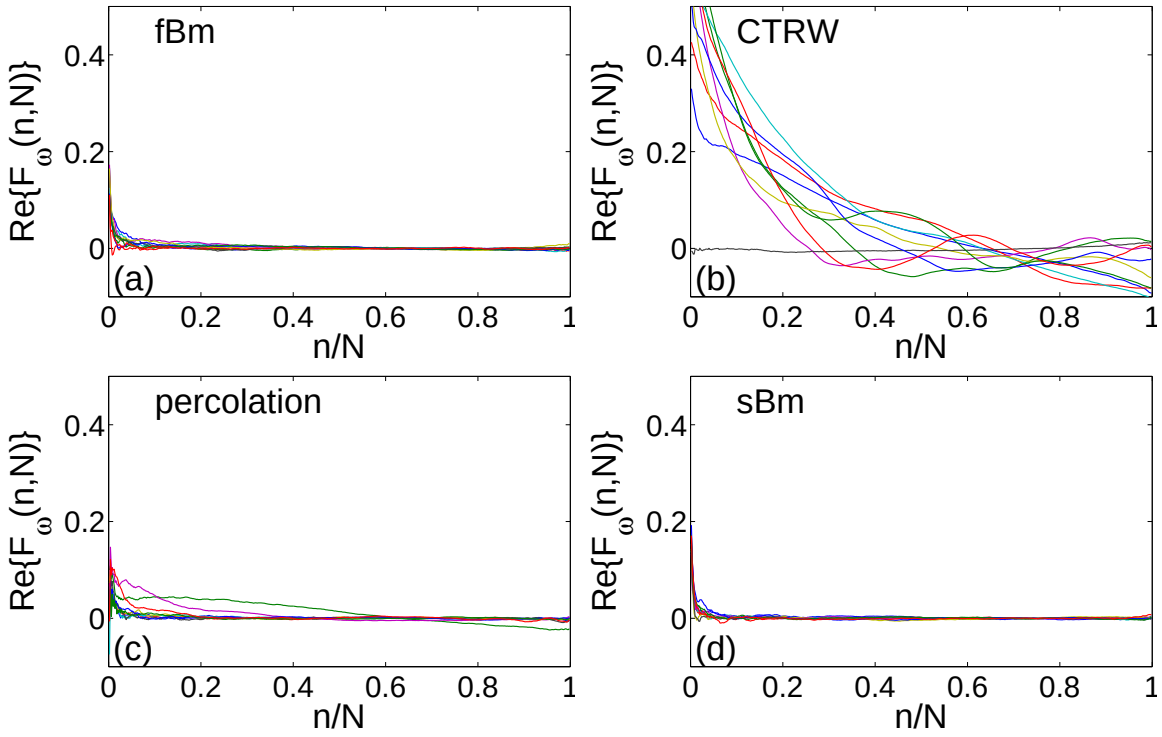


Figure 2.10: The real part of the ergodicity estimator at $\omega = 2$ for ten simulated trajectories from four anomalous diffusion models with the same exponent $\alpha = 2H = 0.7$: fBm (**a**), CTRW (**b**), diffusion on percolating cluster (**c**), and sBm (**d**) (in all cases, we set $\sigma = 1$ and $N = 500$). All trajectories were corrupted by white Gaussian noise with standard deviation $\sigma_\varepsilon = 0.1$. Each trajectory was normalized by the empirical standard deviation of its increments.

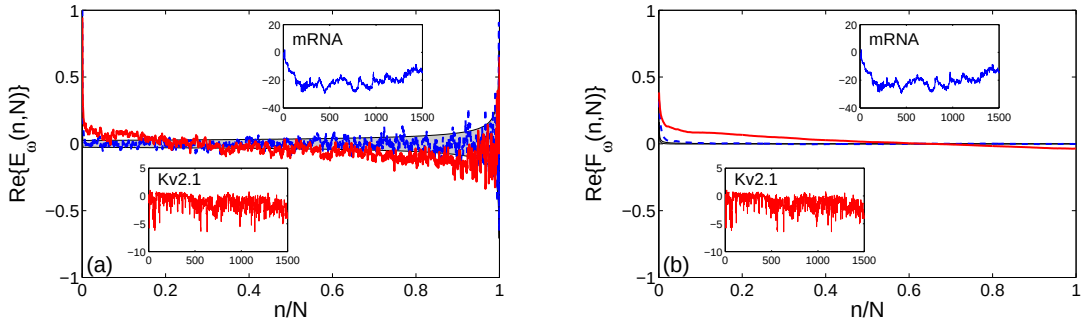


Figure 2.11: The real part of the mixing **(a)** and ergodicity **(b)** estimators at $\omega = 2$ as a function of n for two experimental trajectories of around $N = 1500$ points: the motion of mRNA molecule inside live *E. coli* cell [106] and Kv2.1 potassium channel anomalous dynamics in the plasma membrane [280]. Both trajectories (shown in insets) were normalized by the empirical standard deviation of their increments. Light gray shadowed region delimits the typical range of fluctuations of estimators for Brownian motion, i.e., the mean plus and minus the standard deviation.

Figure 2.11a shows that the mixing estimator applied to a trajectory of mRNA molecule (dashed line) rapidly decreases with n and then fluctuates around 0 within the typical range of fluctuations for Brownian motion. In other words, this test does not reveal nonmixing behavior, in agreement with conclusion of Ref. [175]. Similarly, the ergodicity estimator in Fig. 2.11b rapildy vanishes at large n, as expected for ergodic dynamics. However, many trajectories are needed to confirm that the dynamics is indeed mixing/ergodic (see further discussion in Sec. 2.7). In turn, both mixing and ergodicity estimators for a trajectory of the Kv2.1 potassium channel (solid line) lie outside the typical range and do not vanish as n increases. This is a signature of the nonmixing/nonergodic behavior, in agreement with the conclusion of Weigel *et al.* based on the analysis of ensemble averages over multiple trajectories [280].

We also applied both estimators to the trajectories of optically trapped (sub)micron-sized beads in living cells and actin solutions [26, 116]. These tests (not shown) did not reveal nonergodic behavior, as expected for a harmonically trapped particle.

The important advantage of these single-particle estimators is the possibility to probe nonergodicity for each trajectory. Figure 2.12 shows the real part of the ergodicity estimator at

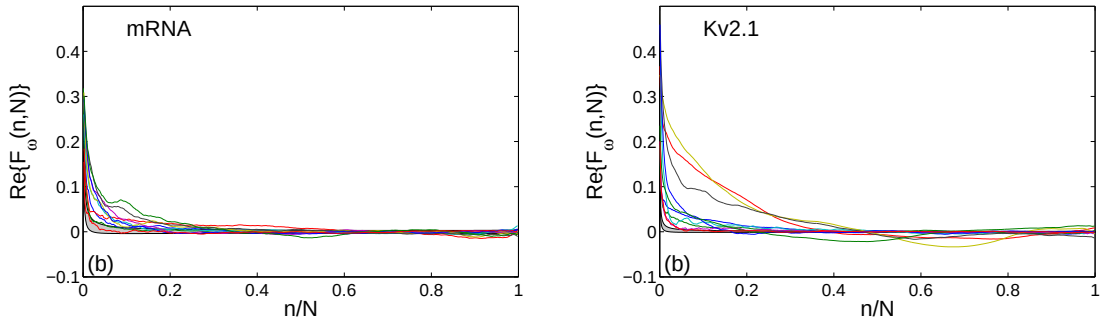


Figure 2.12: The real part of the ergodicity estimator at $\omega = 2$ for ten experimental trajectories of two samples: the motion of mRNA molecules inside live *E. coli* cells [106] and Kv2.1 potassium channel anomalous dynamics in the plasma membrane [280]. Each trajectory is normalized by the empirical standard deviation of its increments. The trajectory length ranges between 400-500 points for the first set and around 1500 points for the second set. Light gray shadowed region delimits the typical range of fluctuations for Brownian motion, i.e., the mean plus and minus the standard deviation.

$\omega = 2$ for ten experimental trajectories from two samples: the motion of mRNA molecules inside live *E. coli* cells [106] and Kv2.1 potassium channel anomalous dynamics in the plasma membrane [280]. One can see that the dispersion of the estimator curves is higher for the Kv2.1 potassium channel, in particular, several trajectories can be classified as nonergodic. As a consequence, the ergodicity estimator allows one to reject, with some degree of certainty, the ergodicity hypothesis. Weigel *et al.* identified transient binding to the actin cytoskeleton as a plausible biophysical origin of nonergodicity in this dynamics. In turn, the estimator curves for mRNA molecule trajectories are less dispersed and vanish as n grows, showing no evidence for nonergodic behavior.

It is also instructive to compare Fig. 2.12 to Fig. 2.10 for simulated trajectories. One can see that the ergodicity estimator curves for both experimental samples behave somewhat in between those for fBm and CTRW. In fact, their dispersion is larger than that for fBm but smaller than that for CTRW.

2.6 Discussion

2.6.1 Positions versus increments

The form of both mixing and ergodicity estimators originate from the analysis of stationary infinitely divisible (SID) processes by Magdziarz and Weron [175]. For instance, the mixing estimator aims at testing whether two increments $Y(k+n)$ and $Y(k)$ become asymptotically independent as n increases. However, the use of these short-time increments (or velocities) may be inconvenient for the analysis of anomalous intracellular transport because the most common nonergodic model of anomalous diffusion, the CTRW, cannot be classified as nonergodic since most of its increments are zero. Figure 2.13 illustrates this problem by comparing the increments-based method (as originally proposed by Magdziarz and Weron) and our positions-based method. For this purpose, we first apply the mixing and ergodicity estimators to the increments of a single CTRW trajectory, $\{X(1)-X(0), X(2)-X(1), \ldots, X(N)-X(N-1)\}$, and then to the positions of the same trajectory, $\{X(0), X(1), \ldots, X(N)\}$ (note that the improved estimators operate with long-time increments, namely, $X(k+n)-X(k)$ and $X(k)-X(0)$ stand in Eq. (2.6)). As expected, both original estimators applied to the short-time increments are very close to 0 that does not allow one to reveal its nonergodic character. In turn, the estimators applied to the positions of the same CTRW reveal its nonergodicity/nonmixing behavior. It is worth noting that the two methods can both work well (correctly identifying CTRW trajectories as nonergodic) or both fail (missing such identifications). This is related to the high sensitivity of the estimators to the presence of long stalling periods in a finite length trajectory. However, the use of the positions-based estimators results in statistically more reliable identifications of CTRW as nonergodic.

While we kept essentially the same form of both mixing and ergodicity estimators, their application to the positions of a trajectory has changed the theoretical paradigm. In particular, the improved mixing estimator probes now the asymptotic independence of the positions $X(k+n)$ and $X(k)$ or, equivalently, of the long-time increments $X(k+n)-X(0)$ and $X(k)-X(0)$ at different lag times. Alternatively, this estimator can be seen as the time averaged characteristic function of the increment $X(k+n)-X(k)$ at lag time n. In other words, the key modification consists in looking at the long-time increments instead of short-time ones (such as $X(k+1)-$

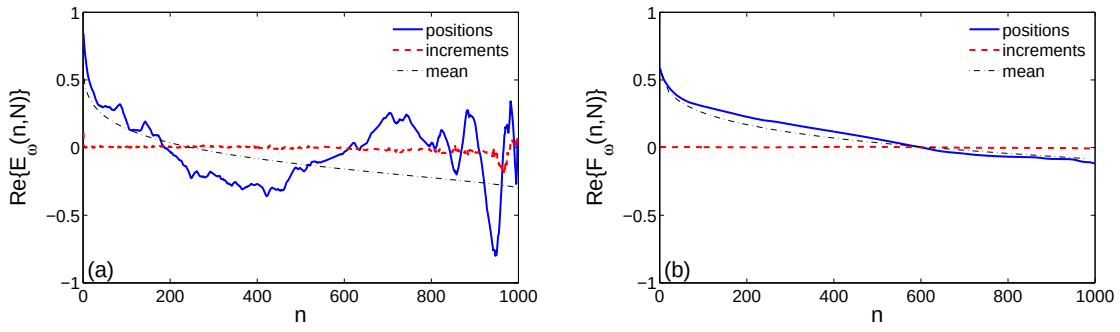


Figure 2.13: The real parts of the mixing (**a**) and ergodicity (**b**) estimators at $\omega = 1$ applied to the increments of a single CTRW trajectory (dashed line) and to the positions of the same trajectory (solid line), with $N = 1000$, $\alpha = 0.7$, and $\sigma = 1$. In both cases, $X(n)$ are normalized by the empirical standard deviation of the increments. Thin dash-dotted line presents the mean estimators from Eqs. (2.51, 2.52).

$X(k)$). To some extent, the difference between the increments-based and the positions-based estimators resembles the difference in estimation of the velocity auto-correlation function and the mean square displacement, the latter being less noisy and thus more robust and easier to estimate.

To further illustrate the difference between two methods, we plot in Fig. 2.14 the real part of the original ergodicity estimator [175] applied to the increments of ten experimental trajectories of mRNA molecules and of Kv2.1 potassium channels [2]. Given that the mRNA and Kv2.1 trajectories were recorded in different length units, the increments of each trajectory were normalized by their standard deviation. Following [175], only the first 200 lag times are plotted. One can see that the original estimators look very similar for two sets of experimental data, suggesting the ergodic behavior in both cases. In contrast, our improved estimator reveals more distinct behavior in two samples (Fig. 2.12).

[2] Note that Magdziarz and Weron have analyzed the same set of experimental trajectories of mRNA molecules reported in [106]. As a consequence, Fig. 2.14a resembles Fig. 6a from [175], except for the opposite sign which may be related to different normalizations of the increments and the shifted summation range ($k = 1, \ldots, n$ instead of $k = 0, \ldots, n-1$) in the ergodicity estimator that was actually used in [175] (private communication by M. Magdziarz).

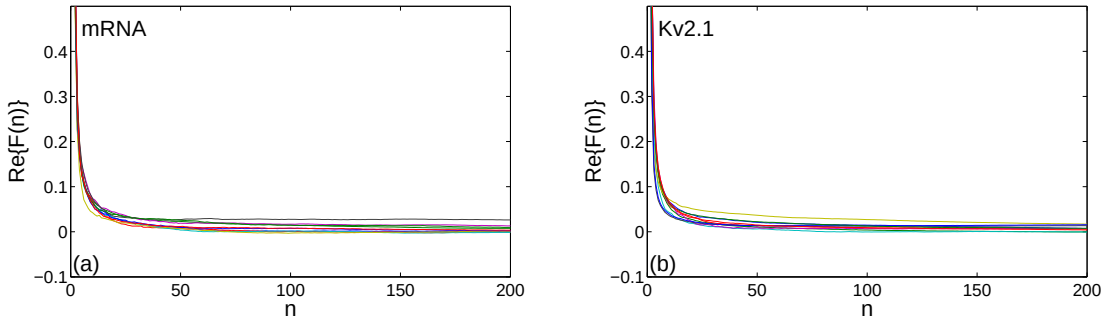


Figure 2.14: The real part of the original ergodicity estimator, $\mathrm{Re}\{n^{-1}\sum_{k=0}^{n-1}\hat{E}(k)\}$ from [175], applied to ten experimental trajectories of two samples: **(a)** the motion of mRNA molecules inside live *E. coli* cells [106] and **(b)** Kv2.1 potassium channel anomalous dynamics in the plasma membrane [280]. The estimator is applied to the increments of each trajectory which are normalized by their empirical standard deviation. The trajectory length ranges between 400-500 points for the first set and around 1500 points for the second set.

2.6.2 Role of the frequency

The introduction of variable frequency ω is important to avoid wrong identifications of nonmixing/nonergodic behavior due to deviations of the mixing or ergodicity estimators from 0. This problem can be relevant for both the original increments-based method and our positions-based method. In order to illustrate this point, we consider two examples: fractional Gaussian noise (fGn) and Brownian motion.

If $X(n)$ are centered stationary Gaussian increments defined by a covariance function $\langle X(k_1)X(k_2)\rangle = \gamma(|k_1-k_2|)$, the mean mixing and ergodicity estimators read

$$\begin{aligned}\langle\hat{E}_\omega(n,N)\rangle &= e^{-\omega^2\eta(n)} - 2\sum_{k=1}^{N}\frac{N-k+1}{N(N+1)}e^{-\omega^2\eta(k)},\\ \langle\hat{F}_\omega(n,N)\rangle &= \frac{1}{n}\sum_{k=1}^{n}e^{-\omega^2\eta(k)} - 2\sum_{k=1}^{N}\frac{N-k+1}{N(N+1)}e^{-\omega^2\eta(k)},\end{aligned}$$

where $\eta(n) = \gamma(0) - \gamma(n)$. For instance, if $X(n)$ are the increments of fractional Brownian motion (the so-called fractional Gaussian noise), one has

$$\gamma(n) = \frac{\sigma^2}{2}\left[|n-1|^{2H} + (n+1)^{2H} - 2n^{2H}\right]. \tag{2.62}$$

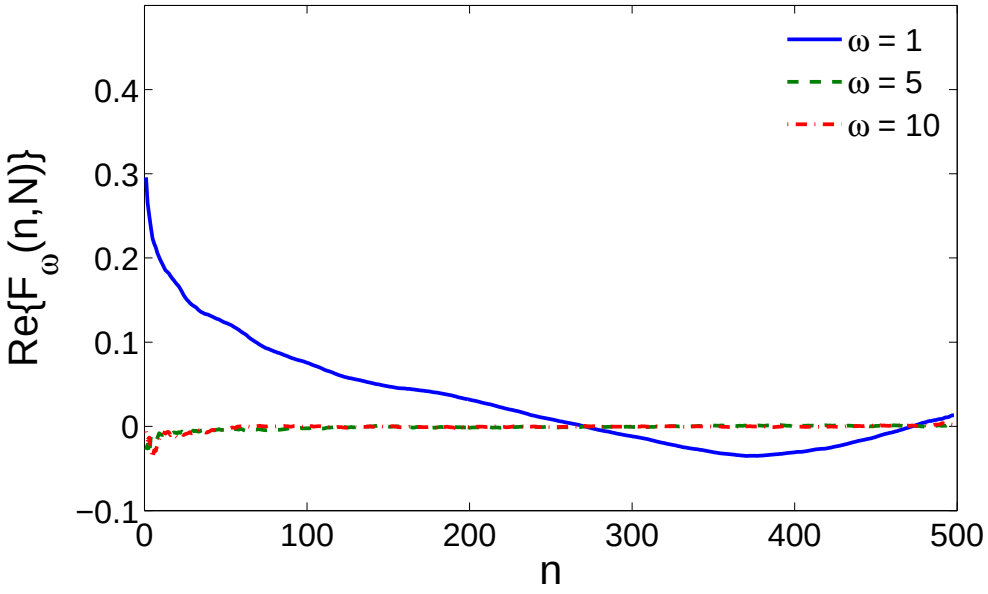


Figure 2.15: The real part of the ergodicity estimator as a function of n for a single fGn trajectory with $N = 500$, $H = 0.9$, and $\sigma = 1$. At $\omega = 1$ (solid line), the estimator deviates from zero that may wrongly suggest nonergodic behavior. Increasing ω, one can eliminate this false conclusion.

For $H = 1/2$, one retrieves the discrete white noise with $\gamma(n) = \sigma^2\delta_{n,0}$ that implies $\langle \hat{E}_\omega(n,N)\rangle = \langle \hat{F}_\omega(n,N)\rangle = 0$. More generally, both mean estimators are strictly zero for independent $X(n)$.

In turn, for correlated increments, the original estimators at $\omega = 1$ can deviate from 0 even for ergodic processes. Figure 2.15 shows the real part of the ergodicity estimator for a single realization of ergodic fGn. At $\omega = 1$, the estimator deviates from zero that may wrongly be interpreted as nonergodic behavior. Increasing the frequency ω allows one to re-establish the vanishing behavior of the estimator and thus to avoid such false conclusions. While the deviations are relatively small for the considered case of fGn (and are even smaller for other choices of H and σ), they could be stronger for some other ergodic processes. In general, the frequency ω is needed to control the amplitude of increments.

The effect of amplitude can be much more important when the estimators are applied to the positions of a trajectory instead of its increments. In particular, even Brownian motion can be falsely identified as nonergodic if the diffusion coefficient is small. This statement is

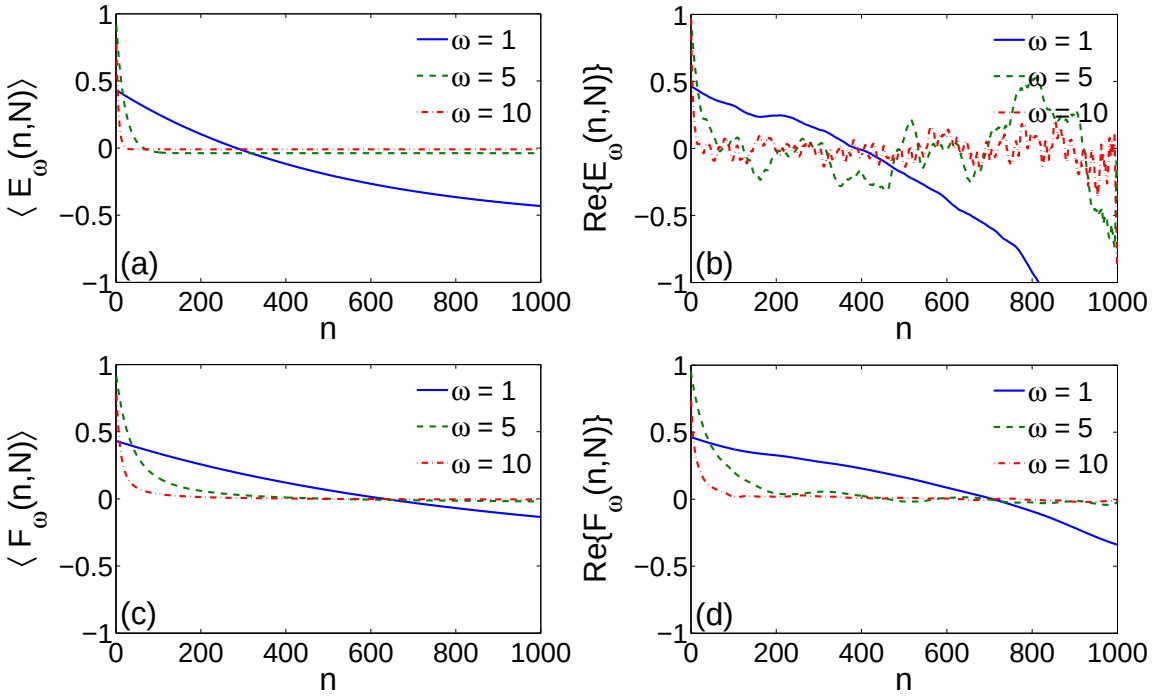


Figure 2.16: The mean mixing **(a)** and ergodicity **(c)** estimators as a function of n for a Brownian motion trajectory with $N = 1000$ and $\sigma^2 = 0.04$. At $\omega = 1$ (solid line), both estimators are not small that may wrongly suggest nonmixing/nonergodic behavior. Increasing ω, one can eliminate this false conclusion. The same conclusions can be made from the real part of the mixing **(b)** and ergodicity **(d)** estimators for a single Brownian motion trajectory with the same parameters.

illustrated on Fig. 2.16 that shows the mean mixing and ergodicity estimators as well as the real part of both estimators for a single Brownian motion trajectory with $N = 1000$ and $\sigma^2 = 0.04$. Distinct deviations from 0 are clearly seen at $\omega = 1$. This is a finite length effect that would disappear for a much longer trajectory. For a fixed N, this false conclusion can be removed by increasing ω that is equivalent to rescaling the variance σ^2. In practice, we suggest to normalize the trajectory by the empirical standard deviation of its increments, in which case the frequency ω can be kept in the order of 1.

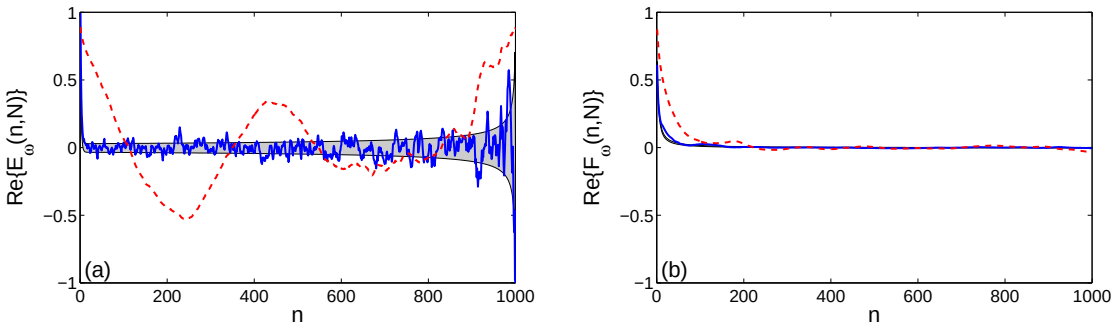


Figure 2.17: The real part of the mixing **(a)** and ergodicity **(b)** estimators at $\omega = 1$ as a function of n for an original Brownian motion trajectory with $N = 1000$ and $\sigma = 1$ (solid line) and for the same trajectory corrupted by resetting ten randomly chosen points to 0 (dashed line). In both cases, the trajectory was normalized by the empirical standard deviation of its increments. Light gray shadowed region delimits the mean plus and minus the standard deviation of each estimator computed for Brownian motion according to Eqs. (2.11, 2.12, 2.17, 2.18).

2.6.3 Impact of outliers

In addition to noise, experimental data can contain "outliers", i.e., singular erroneous points coming from instrumental or software failures, data recording or transmission problems, or human factors. The problem of outliers is particularly relevant for financial data [131]. These erroneous points whose statistical properties strongly deviate from the remaining "normal" points, may appear as nonergodic features in the analysis of a single finite length trajectory by means of the mixing/ergodicity estimators. In order to illustrate the potential impact of outliers, we generate a Brownian trajectory with $N = 1000$ steps and then reset 10 randomly chosen points of this trajectory to 0. This procedure results in 20 increments of anomalously large amplitude (as compared to the other increments). While only 1% of data points are assigned as outliers, both mixing and ergodicity estimators suggest nonmixing/nonergodic behavior, as illustrated on Fig. 2.17.

2.6.4 Discrete displacements

Digital acquisition systems or lattice models can produce discretely spaced positions of experimental or simulated trajectories, resulting in a periodic (nonvanishing) behavior of both

estimators as functions of the frequency ω. To illustrate this "artifact", we consider the one-dimensional projection of a d-dimensional random walk in which $X(n)$ is the sum of independent random variables χ_k taking the values $\pm\sigma$ with probability $1/(2d)$, or 0 with probability $1-1/d$. For this process, $\langle e^{i\omega(X(n+k)-X(k))}\rangle = [\phi(\omega)]^n$, where $\phi(\omega)$ is the characteristic function of χ_k: $\phi(\omega) = 1-(1-\cos(\omega\sigma))/d$. The mean mixing and ergodicity estimators are still given by Eqs. (2.11, 2.12), with $c = \phi(\omega)$. One can see that both mean estimators are 2π-periodic functions of $\omega\sigma$. As a consequence, too large values of $\omega\sigma$ do not improve the quality of estimation. Moreover, the estimators would lead to false nonmixing or nonergodicity identifications at the values $\omega\sigma = 2\pi m$ (with an integer m), at which $c = 1$. In turn, there is an optimal value minimizing c which corresponds to $\omega\sigma = \pi/2 \simeq 1.57$ for this model. Note that if the trajectory was normalized by the standard deviation of its increments, $\sigma/\sqrt{d}$, the characteristic function would be $\phi(\omega) = 1-(1-\cos(\omega/\sqrt{d}))/d$, and the optimal value would be $\omega = \sqrt{d}\,\pi/2$. Since the underlying model is not known for experimental data, the rule of thumb consists in analyzing the estimators for several values of ω in the range between 1 and 3.

2.7 Conclusions

In summary, we proposed, investigated and validated the improved mixing and ergodicity estimators based on long-time increments (or positions) of a single trajectory. Aiming applications to experimental trajectories which *a priori* are neither stationary, nor infinitely divisible, we extended the range of applicability of the estimators beyond SID processes. The new estimators rely on the time averaged characteristic function of the increments that can bring complementary information as compared to the time averaged MSD or VAF, especially for non-Gaussian processes. We showed that the ergodicity estimator vanishes for basic ergodic models of anomalous diffusion (fBm, diffusion on fractals) and remains nonzero for nonergodic models (CTRW, geometric Brownian motion). The significant advantage of the present method is that nonmixing or nonergodic behavior can be revealed from a single trajectory, with no need in ensemble averages over many trajectories that may be difficult or even impossible to collect.

As discussed by Magdziarz and Weron [175], the smallness of the estimators is the necessary but not sufficient condition for mixing and ergodicity. In other words, these statistical tests can reveal nonmixing or nonergodic properties from a single trajectory but many trajectories

are needed to confirm mixing or ergodicity. For instance, a constant process $X(n) = x_0$ with a random initial position x_0 is not ergodic, and this property cannot in principle be revealed from a single trajectory. In particular, the mixing and ergodicity estimators are strictly zero for this process but their smallness is not a signature of ergodicity but rather a failure in detecting nonergodicity. In general, the finite trajectory length and the randomness of the estimators based on a single trajectory remain the major challenges for inference problems that aim at characterizing an unknown stochastic process from its single random realization. Deviations of the estimators from 0 can be either a signature of nonmixing/nonergodic or nonstationary dynamics, or a consequence of too short trajectory, or a measurement artifact (e.g., the presence of outliers, see Sec. 2.6.3), or a specific feature of the process (e.g., discrete displacements, see Sec. 2.6.4).

While probing nonergodicity of an unknown stochastic process from a single finite length trajectory remains debatable from the mathematical point of view, the ergodicity and mixing estimators partially answer the important practical question whether the time average along a single trajectory is representative or not. However the scaled Brownian motion is a counter example where, only at short time, the time and ensemble average are different, such a departure is not recognized by our estimators because the process is ergodic. If the estimators do not vanish with the lag time n, conclusions based on time averages can be strongly misleading. In particular, large values of the estimators indicate peculiar properties of the stochastic process that should warn scientists against a blind use of time averages. Along with other single-particle methods [7, 77, 148, 149, 173, 188, 189, 194, 240, 261, 269, 273, 283], the mixing and ergodicity estimators provide a powerful statistical tool for a more reliable interpretation of SPT experiments.

An important perspective of this work consists in formulating nonergodicity and nonmixing test statistics (that can be converted to an estimated yes/no answer) and in quantifying their properties (p-values, confidence intervals, etc.) [168–170]. For this purpose, the probability distribution of the estimators needs to be investigated for various stochastic processes. In addition, the role of small sample statistics can be analyzed for the improved estimators (see [138] for details).

Chapter 3

Unravelling intermittent features from a single trajectory

3.1 Introduction

In the previous chapter, we discussed a test for detecting weak ergodicity breaking in order to verify the interpretability of time averaged quantities obtained from a single trajectory. In some cases, the ergodicity hypothesis can fail due to the intermittent nature of the process observed (switching between two or several phases of motion). Even in conditions where each phase of the process is stationary and ergodic, the length of the trajectory can be too short for an effective regime to establish resulting in large fluctuations of a quantity averaged along a trajectory. For some applications, given that the intermittent nature of the process is known, a statistical description of the effective dynamic can be sufficient. A contrario, for identifying the underlying transport mechanism, it is no longer sufficient. Many procedures have been proposed to recognize the different phases of motion (see Annex A), but most of them are aimed to test a specific statistical change along the property.

In this chapter, we go a step further by addressing the challenging question of detection of change points between distinct phases in a single random trajectory without prior knowledge of the underlying stochastic model. The phases are distinguished by their dynamical properties such as different (i) diffusivities, (ii) drifts, (iii) auto-correlations of increments, (iv) distributions of increments, (v) dimensionalities (e.g., bulk/surface), (vi) isotropic/anisotropic character, or

(vii) space accessibility (e.g., restricted character due to reflecting obstacles or confinement). The basic idea of such model-free methods consists in considering a *local* functional of the trajectory, $Q(n)$, which depends on a relatively small number of points around the point $\boldsymbol{x}_n$. When applied to successive points along the trajectory, this local functional transforms the trajectory into a new time series, which can then be used either to characterize the dynamics (e.g., to get local drift or diffusivity), or to discriminate between different phases of the motion. For instance, the points $\boldsymbol{x}_n$ with $Q(n)$ below some threshold can be assigned to one phase while the remaining points are assigned to the other phase. In contrast to Bayesian methods, the functional $Q(n)$, which is used to process the data, does not depend on the model, though the efficiency of this binary classification clearly depends on the choice of the functional $Q(n)$. The simplest choices,

$$Q_\mu(n) = \frac{1}{2\Delta}(\boldsymbol{x}_{n+\Delta} - \boldsymbol{x}_{n-\Delta}), \tag{3.1}$$

$$Q_\sigma(n) = \frac{1}{2\Delta+1}\sum_{k=-\Delta}^{\Delta} \|\boldsymbol{x}_{n+k+1} - \boldsymbol{x}_{n+k}\|^2, \tag{3.2}$$

are respectively the drift estimator and the variance estimator (over a window of size $2\Delta+1$) of isotropic Brownian motion. Another standard functional is the angle between successive increments of the trajectory, which is often used to detect ballistic parts of the motion [146] or even a combination of them [7]. Having a differential form (i.e., involving differences between points), these functionals are rather sensitive to noise, exhibit large fluctuations and thus yield large statistical uncertainties. This drawback can be reduced by increasing the window size Δ. However, too large Δ would make the functional $Q(n)$ less sensitive to phase alternations, especially for short phases. Several improvements have been proposed, e.g., the root mean square estimator for the variance [114] or the scaling exponent extracted from the local TA MSD [7]. Nevertheless, these improved estimators still have the differential form and thus remain sensitive to noise.

To overcome this fundamental limitation, we propose the geometric properties of a *local convex hull* (LCH) as robust discriminators of different phases. In contrast to earlier used differential-like estimators, the convex hull is intrinsically an integral-like characteristic that is thus less sensitive to noise, as shown below. This new method is applicable for the analysis of single trajectories of any dimensionality. The method is model-free because it relies exclusively on geometric properties of the trajectory. We demonstrate its efficiency in recognizing different

intermittent dynamical scenarios inspired from biology. The results of this chapter have been published in [160].

In this chapter we first describe the LCH method in Sec. 3.2, how it works and the different quantities associated to it. In Sec. 3.3 we test the efficiency of the method on six scenarios of intermittent motion inspired from biology. Finally we discuss some limitations and concerns for the future developments of the method in Sec. 3.4.

3.2 Local Convex Hull Method

The convex hull of a finite set of points, $\{\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n\} \subset \mathbb{R}^d$, is the set of all convex combinations based on these points:

$$\mathrm{Conv}(\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n) = \left\{ \sum_{k=1}^{n} \alpha_k \boldsymbol{x}_k \mid \alpha_k \geq 0, \ \sum_{k=1}^{n} \alpha_k = 1 \right\}. \tag{3.3}$$

In simpler terms, it is the minimal convex shape that encloses all the points $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n$. In the planar case, the convex hull of a finite set of points $\{\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n\}$ is a convex polygon whose vertices are some of the points in this set. The construction of a convex hull is thus reduced to identifying these points in a clockwise or counter-clockwise order. In higher dimensions, the convex hull is a convex polytope (e.g., a convex polyhedron in three dimensions) whose vertices are some of the points in the input set. Among several efficient algorithms [66, 208], we chose the quickhull algorithm [10] because of its available implementation in Matlab as functions "convhull" (in two dimensions) and "convhulln" (in higher dimensions). The convex hull has been used for home range estimations of animal territories [101, 223, 289] and for computation of fractal dimensions [206]. Some basic properties of the convex hull applied to stochastic processes are summarized in Appendix C.

We propose to consider the diameter and the volume of the local convex hull, computed over $2\Delta + 1$ trajectory points $\boldsymbol{x}_{n-\Delta}, \boldsymbol{x}_{n-\Delta+1}, \ldots, \boldsymbol{x}_{n+\Delta}$, as two functionals:

$$Q_d(n) = \mathrm{diam}(\mathrm{Conv}(\boldsymbol{x}_{n-\Delta}, \ldots, \boldsymbol{x}_{n+\Delta})), \tag{3.4}$$

$$Q_v(n) = \mathrm{vol}_d(\mathrm{Conv}(\boldsymbol{x}_{n-\Delta}, \ldots, \boldsymbol{x}_{n+\Delta})) \tag{3.5}$$

(the volume is replaced by the area in two dimensions). Note that the diameter of the LCH is simply the largest distance between any two points in the sequence $\{\boldsymbol{x}_{n-\Delta}, \ldots, \boldsymbol{x}_{n+\Delta}\}$, i.e.,

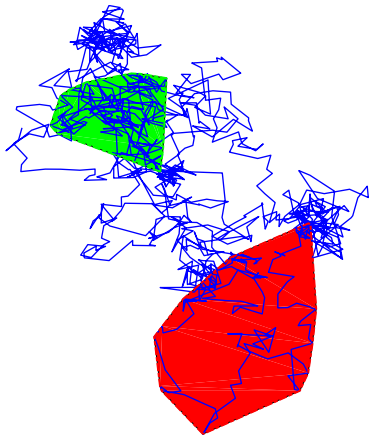


Figure 3.1: Illustration of the local convex hull method applied to an intermittent planar trajectory, alternating a "fast" phase of Brownian motion (with $D = 1/4$) and a "slow" phase of an Ornstein-Uhlenbeck process (with $D = 1/4$ and $k = 0.1$, see Sec. 3.3.4), of equal durations $T_1 = T_2 = 40$. Only two LCHs (based on $2\Delta + 1 = 41$ points) are shown by shadowed regions that correspond to the fast (dark red) and slow (light green) phases.

it is particularly easy to compute. Any significant change in the dynamics (i.e., switching between phases) is expected to be reflected in a notable change in the geometric form of the trajectory that is captured here by the local convex hull (Fig. 3.1). For instance, an increase in the diffusion coefficient or in the drift leads to larger increments and thus a larger LCH, with larger $Q_d(n)$ and $Q_v(n)$. In addition, anisotropic features of the motion, such as dimensionality reduction or presence of a reflecting wall, would be reflected in an anisotropic shape and in a reduced volume of the LCH. Depending on the observable chosen, the identification will be more sensitive to a particular aspect of the motion. The diameter is more sensitive to correlations and diffusion coefficients, whereas the volume is more sensitive to changes in the dimensionality and anisotropy. The window size Δ is a parameter of the method that controls the compromise between reactivity and robustness in change points detection: smaller Δ facilitates identification of short phases (the method is more reactive) but makes it less robust, as the estimators become more sensitive to fluctuations, noises, and outliers.

According to Eqs. (3.4, 3.5), the point $\boldsymbol{x}_n$ contributes to $2\Delta+1$ estimators $Q_d(k)$ (or $Q_v(k)$) with $k \in [n-\Delta, n+\Delta]$. It is therefore natural to classify the point $\boldsymbol{x}_n$ by using all the estimators

$Q_d(k)$ (or $Q_v(k)$), to which it contributed. We define thus the discriminator $S_d(n)$ as a weighted sum of $Q_d(k)$ over Δ left and Δ right neighbors of $\boldsymbol{x}_n$ (similar for $S_v(n)$). The simplest choice consists in setting equal weights to each contribution:

$$S_d(n) = \frac{1}{2\Delta+1} \sum_{k=n-\Delta}^{n+\Delta} Q_d(k), \tag{3.6}$$

$$S_v(n) = \frac{1}{2\Delta+1} \sum_{k=n-\Delta}^{n+\Delta} Q_v(k). \tag{3.7}$$

Other weighting schemes (e.g., exponential) are as well possible. A priori, it is not clear why the point n should be at the center of the window when calculating $Q_d(n)$ (resp. $Q_v(n)$). In fact, choosing equal weight for each contribution makes $S_d(n)$ (resp. $S_v(n)$) independent of the choice of the position of n in the functional $Q_d(n)$ (resp. $Q_v(n)$), which solves the ambiguity.

The LCH method for change points detection consists in two steps. In the first step, for a given trajectory with N points, discriminators $S_d(n)$ and $S_v(n)$ are computed for all points $\boldsymbol{x}_n$ with n from $2\Delta+1$ to $N-2\Delta$ (the first and last 2Δ points of the trajectory are discarded and remain unclassified). In other words, these discriminators transform the trajectory in $\mathbb{R}^d$ into two new one-dimensional time series. Due to the integral-like nature of the LCH, these times series are expected to be less erratic than trajectories and more sensitive to changes in the dynamics. In the second step, these time series are used to detect change points between two phases. In the basic setting, the points $\boldsymbol{x}_n$ with $S_d(n) > S_d$ (or $S_v(n) > S_v$) are classified as belonging to the "fast" phase whereas the points $\boldsymbol{x}_n$ with $S_d(n) \leq S_d$ (or $S_v(n) \leq S_v$) as belonging to the "slow" phase. The choice of the threshold value S_d (resp. S_v) is crucial for an efficient detection of change points. Without *a priori* knowledge of the underlying stochastic process, we set S_d (resp. S_v) to be the empirical mean of $S_d(n)$ (resp., of $S_v(n)$) over the trajectory, i.e.,

$$\begin{aligned} S_d &= \frac{1}{N-4\Delta} \sum_{n=2\Delta+1}^{N-2\Delta} S_d(n), \\ S_v &= \frac{1}{N-4\Delta} \sum_{n=2\Delta+1}^{N-2\Delta} S_v(n). \end{aligned} \tag{3.8}$$

In Sec. 3.4.2, we discuss this point and possible improvements.

We emphasize that the LCH method can be applied to a *single* trajectory of essentially *any* intermittent stochastic process. In turn, the quality of change points detection depends

on how distinct the geometric properties of two phases are, as well as on durations of these phases. We also stress that the method does not rely on specific properties of the underlying stochastic model, nor does it identify the properties of each phase. This analysis can be done by conventional methods after the change points detection. Finally, the same two phases do not need to be repeated along the trajectory. Although many distinct phases could in principle be present in a single trajectory, we focus on the particular case of two alternating phases (that we qualitatively called "fast" and "slow").

Note that the convex hull diameter can be considered as an alternative to the maximal excursion, i.e., the greatest distance from the origin that a particle reaches until time t [27]. The latter has been shown to have a narrower distribution than the conventional TA MSD, and thus proposed as an estimator of the scaling exponent that can be applied to single particle trajectories [261].

3.3 Numerical validation

In this section we investigate the efficiency of the LCH method by simulating six models of intermittent processes with the following phases:

1. two planar Brownian motions with distinct diffusion coefficients D_1 and D_2 (change in diffusivity);
2. two planar Brownian motions with the same diffusion coefficient D, with and without drift (change in directionality);
3. two planar fractional Brownian motions with distinct Hurst exponents H_1 and H_2 (change in auto-correlations);
4. planar Brownian motion and Ornstein-Uhlenbeck process (change in auto-correlations);
5. planar Brownian motion and exponential flights (change in the distribution of increments);
6. surface-mediated diffusion, with alternating phases of three-dimensional bulk diffusion and two-dimensional surface diffusion (change in dimensionality).

Although many other intermittent processes could be considered, we focus on the above cases as representative examples.

The simulations are performed as follows.

1. First, we generate thousand random trajectories, each with thousand points (i.e., $N = 1000$). The intermittency is implemented by partitioning these points into two alternating phases by assigning random, exponentially distributed durations, with prescribed mean durations T_1 and T_2 of two phases. Throughout this section, we consider two phases of equal mean duration, $T_1 = T_2 = T$, while the situation of unequal phases is discussed in Sec. 3.4.3. A white Gaussian noise of standard deviation σ_n is added to each trajectory point in order to check the robustness of the LCH method to measurement noises. The noise level σ_n is set to be proportional to the empirical standard deviation σ of increments, computed for each trajectory. Note that the noise has a stronger impact onto the "slow" phase than onto the "fast" phase.

2. Second, we fix the window size Δ and compute the diameter, $Q_d(n)$, and the volume, $Q_v(n)$, of the LCH at all time steps n from $\Delta+1$ to $N-\Delta$ for each trajectory. These estimators are then transformed into the weighted discriminators $S_d(n)$ and $S_v(n)$ according to Eqs. (3.6, 3.7) for n from $2\Delta+1$ to $N-2\Delta$. The default value of the window size Δ is 10, other sizes being considered in Sec. 3.4.2.

3. Third, for each trajectory, we assign the points with $S_d(n) > S_d$ (resp. $S_v(n) > S_v$) to the "fast" phase and the points with $S_d(n) \leq S_d$ (resp. $S_v(n) \leq S_v$) to the "slow" phase, where S_d (resp. S_v) is the empirical mean diameter (resp., volume) of the LCH given by Eq. (3.8). We stress that the threshold value is computed separately for each trajectory. In this way, we obtain two *unrelated* binary classifications, the one based on the diameter and the other based on the volume. *A priori*, it is unclear which classification results in a better detection of change points. In practice, one would use one of these classifications, depending on the anticipated properties of the phases.

4. Finally, to quantify the efficiency of the LCH method, we introduce the recognition score, R, as the fraction of points that have been correctly classified. This score is computed for each trajectory and then averaged over all simulated trajectories. The score 0.5 would be obtained by a completely random classification, whereas the score 1 corresponds to the perfect classification. We will analyze how the efficiency of the LCH method is affected by

the phase duration T, the window size Δ, and noise level σ_n. We emphasize that the phase classification is performed *individually* for each trajectory, whereas the ensemble average appears only at the last step to assess the quality of the method through the recognition score.

Throughout the chapter, we employ dimensionless units. In particular, time is identified with the point index n.

3.3.1 Two Brownian motions

We start with heterogeneous diffusion, in which a particle diffuses in a composite medium with high and low diffusivities. A simplified model of this dynamics is a Brownian motion switching between two diffusion coefficients D_1 and D_2. This model can also represent the motion of a polymer that switches between two conformational states having distinct hydrodynamic radii (e.g., a compact globular structure versus an extended fibrous one).

Figure 3.2 a shows a single planar trajectory with two alternating phases of slow ($D_1 = 1/2$) and fast ($D_2 = 2$) diffusion, each phase of the mean duration $T = 100$. The weighted LCH diameter $S_d(n)$ with the window size $\Delta = 10$ is shown in Fig. 3.2b. The change points are identified as crossings of $S_d(n)$ by the dashed horizontal line which indicates the empirical mean S_d over that trajectory. As expected, the detected change points are slightly delayed with respect to the actual change points (shown by changes of color/brightness in the curve). This delay is caused by the fact that the LCH at the actual change point, n_c, includes half points of one phase and half points of the other phase. Only when $n = n_c + 2\Delta + 1$, the LCH gets rid off the points of the previous phase. As a consequence, the delay is expected to be of the order of 2Δ, as qualitatively confirmed by this figure. The delay results in a false classification of some points, as illustrated by pink shadowed regions in Fig. 3.2b. The false classification can also result from spontaneous crossings by $S_d(n)$ of the dashed horizontal line (e.g., see the shadowed region at the time step around 640). This is just a random nature of the motion: the particle in the "fast" (resp., "slow") phase starts to explore the space slower (resp., faster) than usual due to stochastic fluctuations.

Figure 3.2c shows the recognition score R as a function of the mean phase duration T for the diameter-based discriminator $S_d(n)$. When the phase duration is too short (say, $T = 10$),

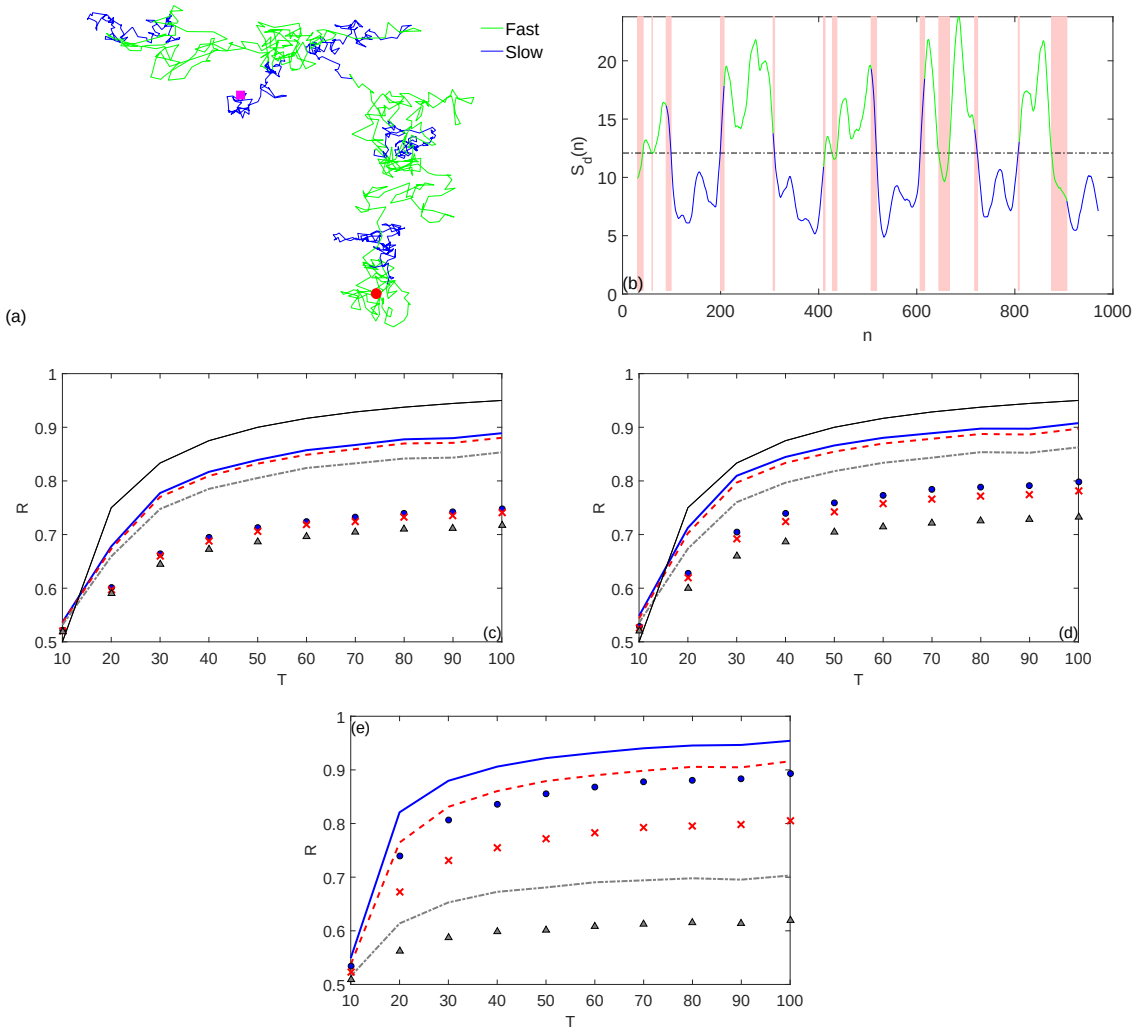


Figure 3.2: **Model 1 (a)** A single trajectory of planar Brownian motion alternating a "slow" phase ($D_1 = 1/2$, dark blue) and a "fast" phase ($D_2 = 2$, light green), of the mean phase duration $T = 100$. Circle and square indicate the starting and ending points. **(b)** The weighted LCH diameter $S_d(n)$ with the window size $\Delta = 10$, applied to this trajectory. Pink shadow highlights the false classification zones. Dashed horizontal line shows the empirical mean S_d over that trajectory. **(c,d,e)** Recognition score R of the diameter-based discriminator $S_d(n)$ **(c)**, the volume-based discriminator $S_v(n)$ **(d)**, and the TA MSD-based discriminator **(e)** as a function of the mean phase duration T. Lines show the results for the case $D_1 = 1/2$, $D_2 = 2$ with three noise levels σ_n: 0 (blue solid), 0.5σ (red dashed), and σ (gray dash-dotted) (σ being the empirical standard deviation of increment calculated for each trajectory). Symbols correspond to the case $D_1 = 1/2$, $D_2 = 1$ with the same levels of noise σ_n: 0 (circles), 0.5σ (crosses), and σ (triangles). Thin black solid line shows the hypothetical curve $1 - \Delta/(2T)$ for qualitative comparison.

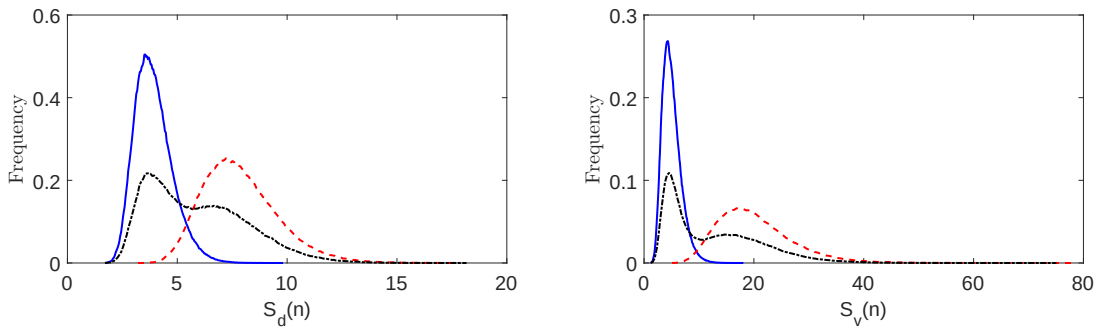


Figure 3.3: Distribution of $S_d(n)$ (**Left**) $S_v(n)$ (**Right**) for one trajectory of $N = 10^6$ steps of planar Brownian motion with diffusion coefficient $D = 1/2$ (blue curve) and $D = 2$ (dashed red curve) as well as an intermittent Brownian motion alternating the two previous diffusion coefficients (dashed black curve) with mean phase duration $\lambda = 100$ calculated for $\Delta = 10$.

the LCH of window size $\Delta = 10$ includes 21 consecutive points and thus almost always contains points from both phases. In this extreme case, the method is clearly unable to detect change points, in agreement with the obtained recognition score close to 0.5. When the phase duration T is comparable to $2\Delta+1 = 21$ (the number of points used to construct the LCH) and two phases are quite distinct (the example $D_1 = 1/2$, $D_2 = 2$), the recognition score is around 0.67, i.e., two thirds of points are correctly classified. The recognition score further increases up to 88% as the mean phase duration grows up to $T = 100$. If the delay in detection was equal on average to $c\Delta$, the curve would be $(T - c\Delta)/T = 1 - c\Delta/T$, as illustrated by thin black line for $c = 1/2$. One can see that this hypothetical curve over-estimates the recognition score probably because of additional spontaneous false classifications due to the overlapping of the distribution of the LCH for Brownian motion alone with $D_1 = 1/2$ and $D_2 = 2$, as shown in Fig. 3.3. According to Fig. 3.3, both distributions of $S_d(n)$ overlap more than that of $S_v(n)$ explaining why the hypothetical curve fits better the case of the volume. We expect that when the distributions do not overlap at all, the hypothetical curve would fully reproduce the simulated results.

Figure 3.2d presents the recognition score for the volume-based classification. While this classification slightly outperforms the diameter-based one, the behavior is very similar to that shown in Fig. 3.2c. This is not surprising because both phases and thus the shapes of the LCHs are isotropic so that the volume and the diameter of the LCH bear essentially the same

information.

Figures 3.2c,d also show that the LCH estimators are robust again measurement noises. In fact, adding the white Gaussian noise of amplitude σ_n which is comparable to the amplitude of one-step increments, has almost no effect on the recognition score. This is an important advantage of the LCH method, which is based on integral-like characteristics of the trajectory, as compared to conventional techniques based on differential-like estimators such as local TA MSD, which are more sensitive to noise. This point is illustrated in Fig. 3.2e, which shows the recognition score of the classification scheme, in which the LCH-based estimator $Q_d(n)$ is replaced by the local TA MSD estimator $Q_\sigma(n)$ from Eq. (3.2). We use the same window size $\Delta = 10$ as for the LCH estimators. When there is no noise, the local TA MSD estimator outperforms the LCH-based estimators $Q_d(n)$ and $Q_v(n)$. This is not surprising as the diffusivity estimator $Q_\sigma(n)$ is known to be optimal for Brownian motion [24, 114, 274]. However, the presence of noise drastically deteriorates the quality of change points detection by TA MSD. In turn, the effect of the same level of noise onto the LCH-based estimator is much weaker.

When the diffusion coefficients of two phases become closer, the recognition score is reduced. This is illustrated in Fig. 3.2c,d,e by symbols that show the recognition score for the case $D_1 = 1/2$ and $D_2 = 1$. In the ultimate case $D_1 = D_2$, the two phases become identical, and any phase detection would meaningless, yielding the recognition score close to 0.5 (not shown).

3.3.2 Brownian motion with a drift

In the second example, we consider a planar Brownian motion alternating two phases (of the same diffusivity), without and with a drift μ in a fixed direction. This is a very basic model for active intracellular transport, in which cargos can attach to motor proteins, be transported ballistically along microtubules, spontaneously detach, and resume diffusion [33, 35]. If the intermittent character of this process was ignored, switching between two phases would effectively look like a superdiffusive motion with a scaling exponent α between 1 and 2.

As the drift tends to elongate the trajectory in one direction, one should be able to identify the drifted phase via anisotropic and larger LCH. To test the efficiency of the LCH method, we generate a planar Brownian motion alternating two phases, one of which has a small drift μ in a fixed direction. As previously, durations of both phases are random and exponentially distributed variables with the mean phase duration T. Figure 3.4 shows an example of a single

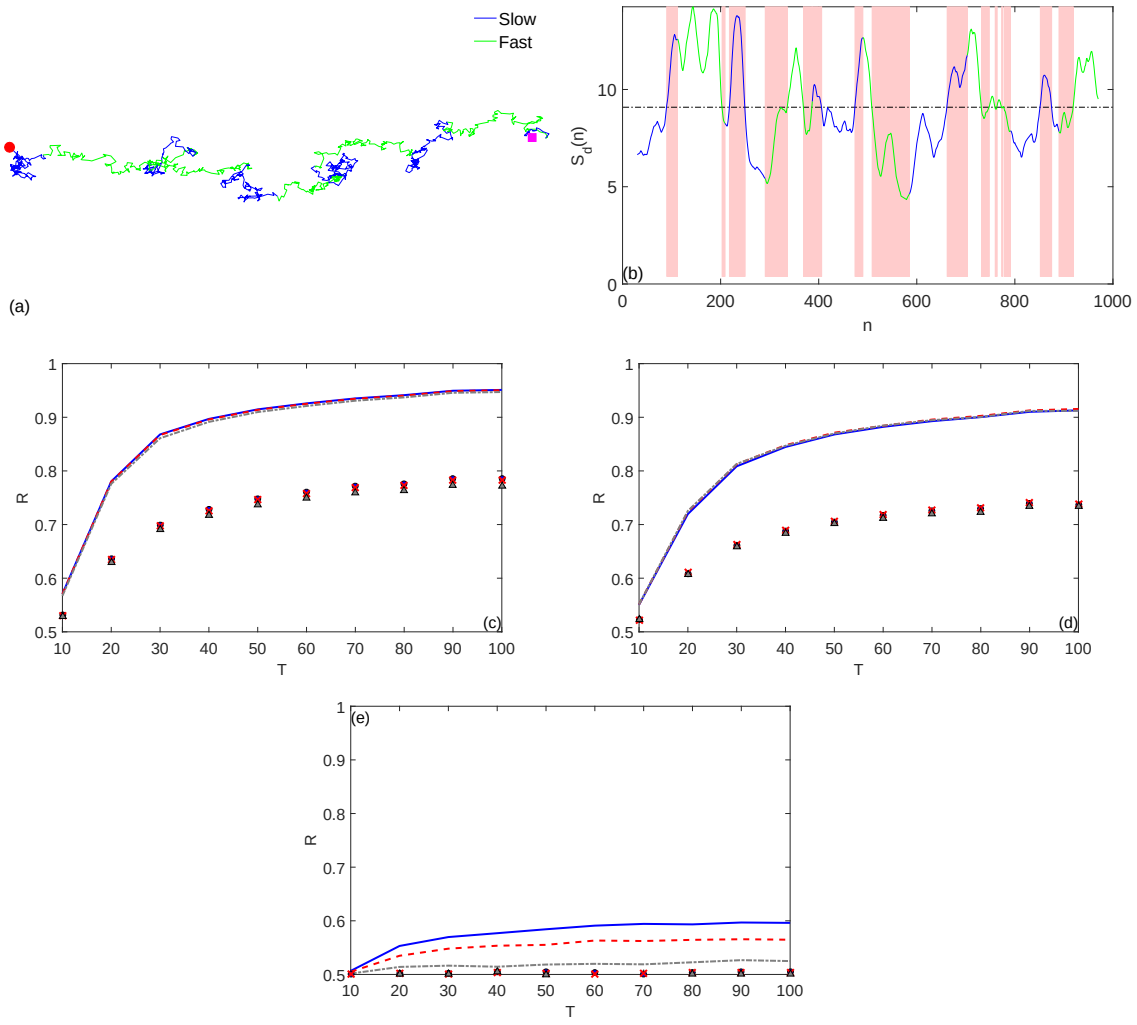


Figure 3.4: **Model 2 (a)** A single trajectory of an intermittent process alternating a "slow" phase of planar Brownian motion without drift (dark blue) and a "fast" phase with drift ($\mu = 0.5$, light green), of the mean phase duration $T = 100$ (both phases with $D = 1/2$). Circle and square indicate the starting and ending points. **(b)** The weighted LCH diameter $S_d(n)$ with the window size $\Delta = 10$ applied to this trajectory. Pink shadow highlights the false classification zones. Dashed horizontal line shows the empirical mean S_d over that trajectory. **(c,d,e)** Recognition score R of the diameter-based discriminator $S_d(n)$ **(c)**, the volume-based discriminator $S_v(n)$ **(d)**, and the TA MSD-based discriminator **(e)** as a function of the mean phase duration T. Lines show the results for the case $\mu = 0.5$ with three noise levels σ_n: 0 (blue solid), 0.5σ (red dashed), and σ (gray dash-dotted) (σ being the empirical standard deviation of increment calculated for each trajectory). Symbols correspond to the case $\mu = 0.1$ with the same levels of noise σ_n: 0 (circles), 0.5σ (crosses), and σ (triangles).

trajectory of this process, the weighted LCH diameter $S_d(n)$, and the recognition scores for the diameter-based and volume-based discriminators. For a relatively strong drift, $\mu = 0.5$ (with the one-step standard deviation $\sigma = 1$), both estimators efficiently detect the change points. For a much smaller drift $\mu = 0.1$, the recognition scores are decreased but remain good enough. For instance, one attains the recognition score of 80% at the mean phase duration $T = 100$. In both cases, the diameter-based discriminator outperforms the volume-based discriminator, and both discriminators are very robust against noise. As expected, the local TA MSD estimator, which is *a priori* not adapted to detect drift, shows a poor performance (Fig. 3.4e).

3.3.3 Two fractional Brownian motions

For each planar trajectory, X and Y coordinates were generated by concatenating independent "pieces" (phases) of one-dimensional fBm with alternating phases of subdiffusive ($H_1 = 0.35$) and superdiffusive ($H_2 = 0.7$) motion, with $D_H = 1/2$ in both cases. As previously, durations of phases are independent exponentially distributed random variables, with a prescribed mean phase duration T. In other words, the numerical algorithm consists in (i) generating a sequence of independent exponentially distributed durations τ_1, τ_2, ..., from which the integer change points are defined as $t_k = t_{k-1} + \lfloor \tau_k \rfloor$ (with $t_0 = 0$); and (ii) generating successive phases $\{x_{t_{k-1}+1}, \ldots, x_{t_k}\}$ and $\{y_{t_{k-1}+1}, \ldots, y_{t_k}\}$ (with $k = 1, 2, \ldots$) as one-dimensional fBms, with the covariance defined in Chapter 1, where H is equal to H_1 for even k and to H_2 for odd k. The starting point of the "piece" k is the ending point of the "piece" $k-1$. The whole trajectory is thus composed of points

$$\Big\{\underbrace{(x_1, y_1), \ldots, (x_{t_1}, y_{t_1})}_{\text{"piece" 1: fast phase}}, \underbrace{(x_{t_1+1}, y_{t_1+1}), \ldots, (x_{t_2}, y_{t_2})}_{\text{"piece" 2: slow phase}},$$

$$\underbrace{(x_{t_2+1}, y_{t_2+1}), \ldots, (x_{t_3}, y_{t_3})}_{\text{"piece" 3: fast phase}}, \ldots, \underbrace{(x_{t_K+1}, y_{t_K+1}), \ldots, (x_N, y_N)}_{\text{"piece" K}}\Big\}.$$

Note that the last "piece" (with index K) is truncated to get the whole trajectory with N points. We emphasize that this algorithm yields the successive phases that are independent from each other, and there is no cross-correlation in increments along X and Y coordinates. In this model, switching between two phases mimics changes in auto-correlations of increments.

Figure 3.5 shows an example of such intermittent trajectory, the weighted LCH diameter $S_d(n)$ with the window size $\Delta = 10$ applied to this trajectory, and the recognition score for both

the diameter-based and the volume-based classifications. The results are similar to that shown in Fig. 3.2 for intermittent Brownian motion. When the distinction between two phases is lower (the example with $H_1 = 0.35$ and $H_2 = 0.5$), the recognition scores are reduced (symbols) but remain satisfactory. These scores are particularly favorably compared to that of the TA MSD discriminator which cannot detect phases at all (Fig. 3.5e). Although the conventional TA MSD estimator may be modified to show a better performance, such a modification would implicate additional knowledge on the model. The results for the case $H_1 = 0.5$ and $H_2 = 0.7$ (normal versus superdiffusive motion) are very similar to the latter case and thus not shown.

3.3.4 Brownian motion and Ornstein-Uhlenbeck process

For each planar trajectory, X and Y coordinates were generated as two independent intermittent processes, alternating a "fast" phase of Brownian motion ($k = 0$) and a "slow" phase of Ornstein-Uhlenbeck process ($k > 0$), with $D = 1/2$ for both phases. Phase durations were independent exponentially distributed random variables, with a prescribed mean phase duration T. This intermittent process can model the motion of a particle evolving in a medium where it freely moves until it interacts with an attracting trap, fluctuates near this trap via the OU process, liberates itself until the next interaction, etc. In neurosciences, the OU process is traditionally used as a model of the instantaneous firing rate of neurons in the neocortex [4, 5]. Recently, a change-point detection procedure to detect changes in the spiking activity of neurons has been proposed [186]. In this section, we check the ability of the LCH method to detect the phases of motion governed by the OU process.

Figure 3.6 shows an example of such intermittent trajectory, the weighted LCH diameter $S_d(n)$ with the window size $\Delta = 10$ applied to this trajectory, and the recognition score for both the diameter-based and the volume-based classifications. The results are similar to that shown in Fig. 3.2 for intermittent Brownian motion. When k is getting smaller, the OU process becomes close to Brownian motion, and the recognition scores are reduced (symbols). The TA MSD discriminator shows poorer performance (Fig. 3.6e).

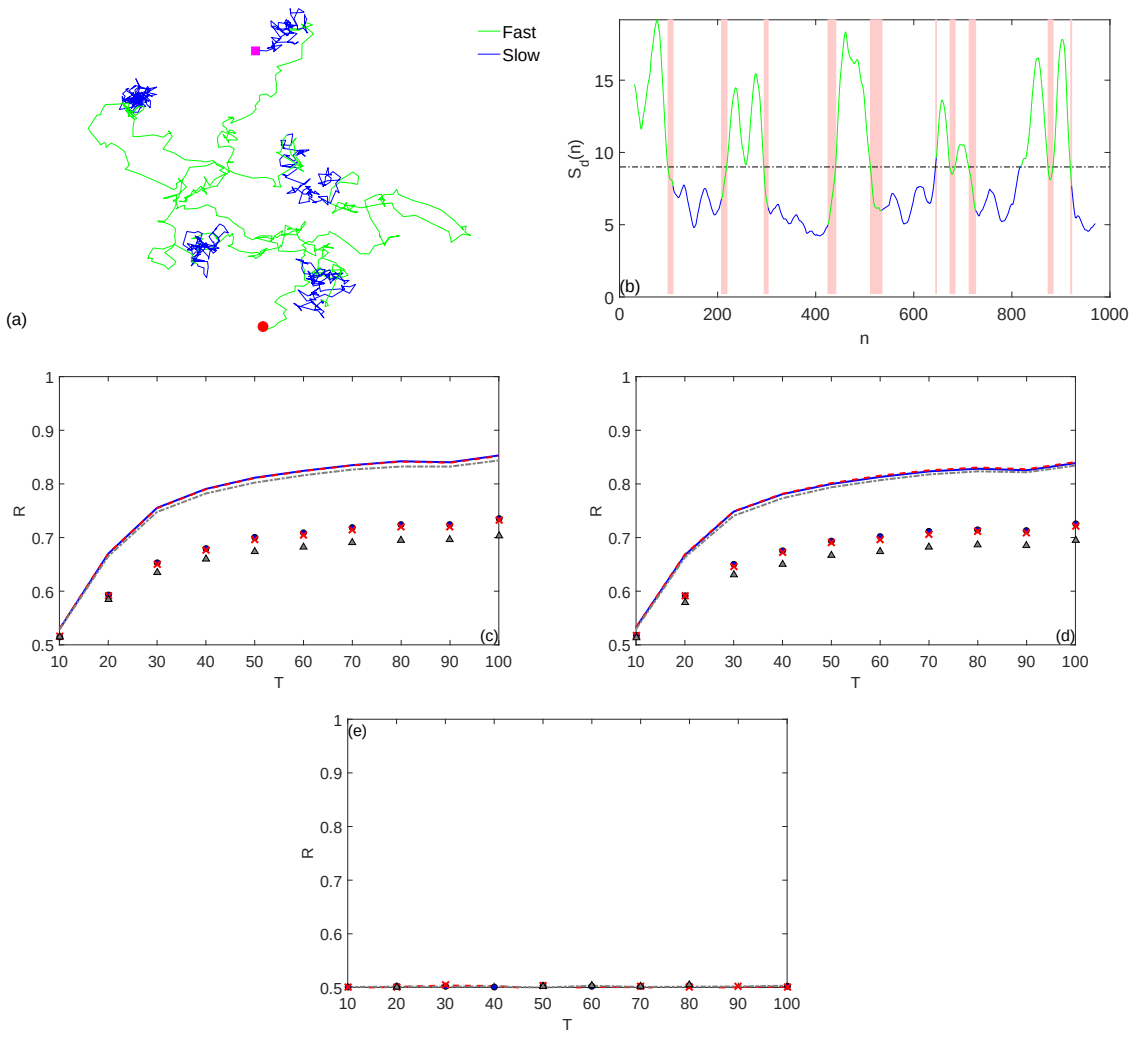


Figure 3.5: **Model 3 (a)** A single trajectory of planar fBm, alternating a "slow" phase ($H_1 = 0.35$, dark blue) and a "fast" phase ($H_2 = 0.7$, light green), of the mean phase duration $T = 100$ and $D_H = 1/2$. Circle and square indicate the starting and ending points. **(b)** The weighted LCH diameter $S_d(n)$ with the window size $\Delta = 10$ applied to this trajectory. Pink shadow highlights the false classification zones. Dashed horizontal line shows the empirical mean S_d over that trajectory. **(c,d,e)** Recognition score R of the diameter-based discriminator $S_d(n)$ **(c)**, the volume-based discriminator $S_v(n)$ **(d)**, and the TA MSD-based discriminator **(e)** as a function of the mean phase duration T. Lines show the results for the case $H_1 = 0.35$, $H_2 = 0.7$ with three noise levels σ_n: 0 (blue solid), 0.5σ (red dashed), and σ (gray dash-dotted) (σ being the empirical standard deviation of increment calculated for each trajectory). Symbols correspond to the case $H_1 = 0.35$, $H_2 = 0.5$ with the same levels of noise σ_n: 0 (circles), 0.5σ (crosses), and σ (triangles).

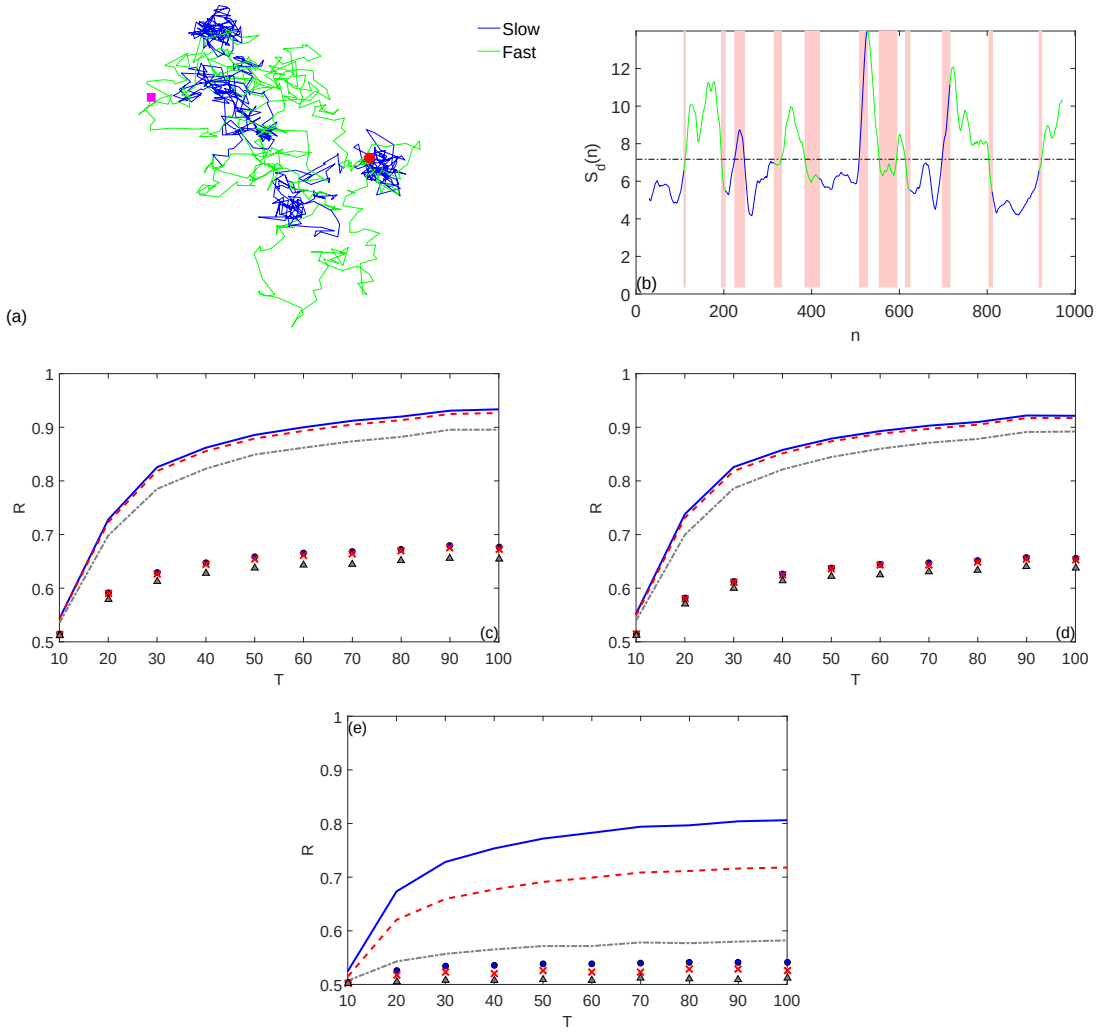


Figure 3.6: **Model 4 (a)** A single trajectory of planar motion, alternating a "fast" phase ($k = 0$, dark blue) and a "slow" phase ($k = 0.1$, light green), with $D = 1/2$ and the mean phase duration $T = 100$. Circle and square indicate the starting and ending points. **(b)** The weighted LCH diameter $S_d(n)$ with the window size $\Delta = 10$ applied to this trajectory. Pink shadow highlights the false classification zones. Dashed horizontal line shows the empirical mean S_d over that trajectory. **(c,d,e)** Recognition score R of the diameter-based discriminator $S_d(n)$ **(c)**, the volume-based discriminator $S_v(n)$ **(d)**, and the TA MSD-based discriminator **(e)** as a function of the mean phase duration T. Lines show the results for the case $k = 1$ with three noise levels σ_n: 0 (blue solid), 0.5σ (red dashed), and σ (gray dash-dotted) (σ being the empirical standard deviation of increment calculated for each trajectory). Symbols correspond to the case $k = 0.1$ with the same levels of noise σ_n: 0 (circles), 0.5σ (crosses), and σ (triangles).

3.3.5 Brownian motion and exponential flights

In order to test the efficiency of the LCH method in detection of change points between phases with distinct distributions of increments, we consider an intermittent process, alternating a "slow" phase of planar Brownian motion (with $D = 1/2$) and a "fast" phase of two-dimensional exponential flights. In the "fast" phase, an increment at each time step was generated independently from the others as $(r\cos\theta, r\sin\theta)$, with the exponential distribution of flight length r (with the mean length ℓ) and the uniform distribution of angle θ. In other words, Gaussian increments (in the "slow" phase) are just replaced by such exponentially distributed increments (in the "fast" phase).

Figure 3.7 shows an example of such intermittent trajectory, the weighted LCH diameter $S_d(n)$ with the window size $\Delta = 10$ applied to this trajectory, and the recognition score for both the diameter-based and the volume-based classifications. The results are similar to that shown in Fig. 3.2 for intermittent Brownian motion. The second choice of the mean flight length, $\ell = \sqrt{2}$, ensures that the variances of exponential flights and of Gaussian jumps are equal. In this case, two phases differ only in the distribution of increments (Gaussian versus exponential). Although these two distributions are relatively close (e.g., both distributions prohibit very large increments), the achieved classification is reasonably good. The performance of the TA MSD discriminator is comparable to that of the LCH for small noises but is reduced significantly for a larger noise.

We emphasize that our realization of exponential flights is different from the run-and-tumble model of bacteria motion, in which case a walker performs a ballistic motion for a random time and then undergoes random rotations [21, 23]. We expect that the recognition score would be even higher for the run-and-tumble motion (as compared to our model) because two phases are geometrically more distinct.

3.3.6 Surface-mediated diffusion

In many biological and chemical applications, particles can adsorb to and desorb from the surface and thus alternate between bulk and surface diffusions. For instance, this so-called surface-mediated diffusion has been suggested as an efficient search mechanism for DNA-binding proteins [22, 171, 226, 275]. The statistics of durations of bulk and surface phases plays an

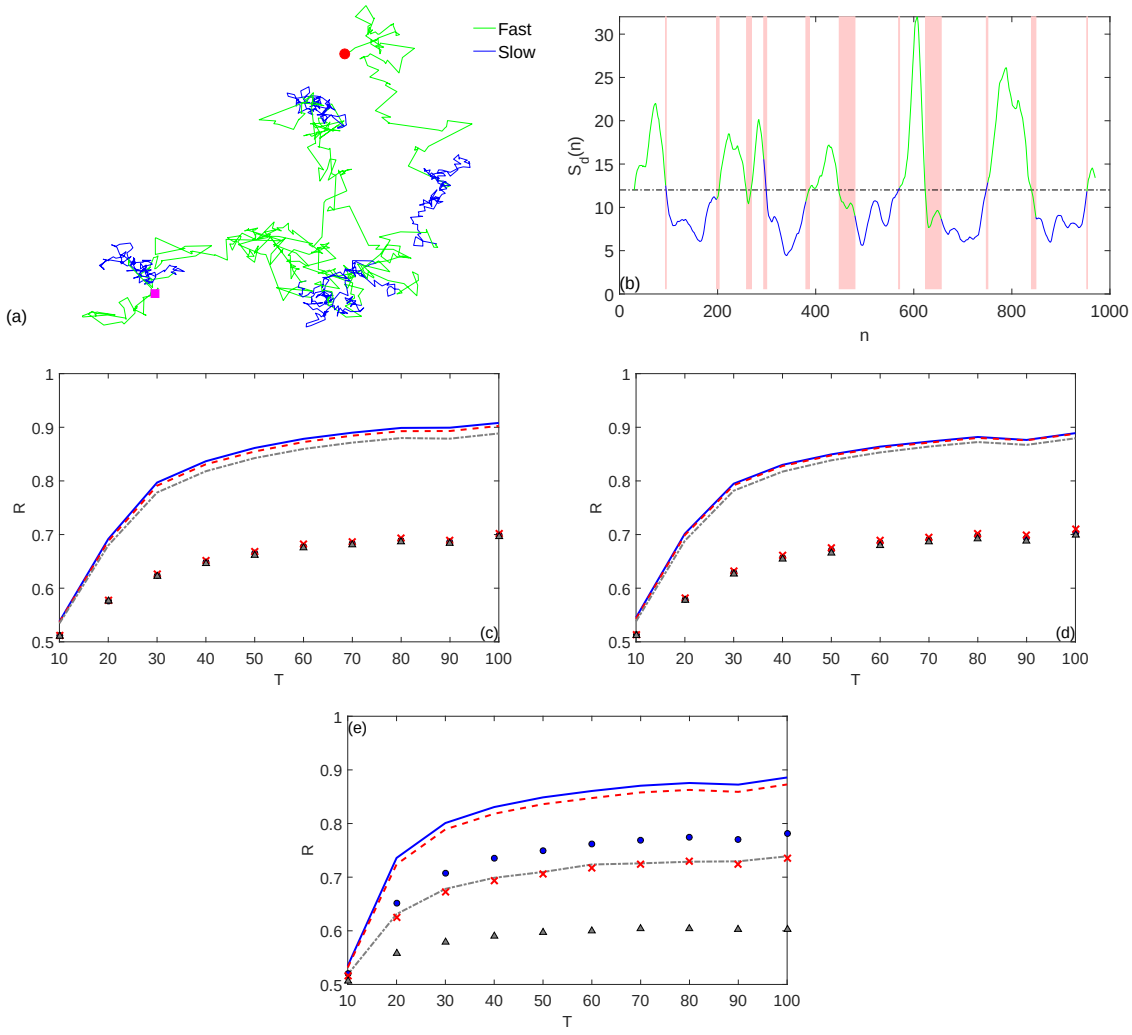


Figure 3.7: **Model 5 (a)** A single trajectory of an intermittent process, alternating a "slow" phase of planar Brownian motion (with $D = 1/2$, dark blue) and a "fast" phase of two-dimensional exponential flights (with $\ell = 3$, light green), of the mean phase duration $T = 100$. Circle and square indicate the starting and ending points. **(b)** The weighted LCH diameter $S_d(n)$ with the window size $\Delta = 10$ applied to this trajectory. Pink shadow highlights the false classification zones. Dashed horizontal line shows the empirical mean S_d over that trajectory. **(c,d,e)** Recognition score R of the diameter-based discriminator $S_d(n)$ **(c)**, the volume-based discriminator $S_v(n)$ **(d)**, and the TA MSD-based discriminator **(e)** as a function of the mean phase duration T. Lines show the results for the case $\ell = 3$ with three noise levels σ_n: 0 (blue solid), 0.5σ (red dashed), and σ (gray dash-dotted) (σ being the empirical standard deviation of increment calculated for each trajectory). Symbols correspond to the case $\ell = \sqrt{2}$ with the same levels of noise σ_n: 0 (circles), 0.5σ (crosses), and σ (triangles).

important role, in particular, in predicting the mean first passage times [17, 18, 228, 231]. For this reason, we test the LCH method on the surface-mediated diffusion inside a three-dimensional sphere of radius R. The particle starts from the origin of the sphere and undergoes Brownian motion in the bulk with diffusion coefficient D_{3d}, until the first arrival onto the surface. From this moment, the particle adsorbs to the surface and diffuses on the surface with the diffusion coefficient D_{2d}. The surface diffusion occurs during a random exponentially distributed waiting time (with the rate λ). After desorption, the particle is ejected into the bulk to the distance $a = 0.05R$ from the boundary and then resumes its bulk diffusion. Here we consider the equal diffusion coefficients $D_{2d} = D_{3d} = 10^{-3}$.

Figure 3.8 shows an example of such intermittent trajectory, the weighted LCH diameter $S_d(n)$ with the window size $\Delta = 10$ applied to this trajectory, and the recognition score for both the diameter-based and the volume-based classifications. The significant difference of the surface-mediated diffusion as compared to the earlier considered models of intermittent processes is that the bulk phase duration is not an exponentially distributed random variable, it is determined by the statistics of first arrivals onto the surface. In turn, the duration of surface diffusion can be controlled by the desorption rate λ. Given that the mean surface duration is $1/\lambda$, we formally set $T = 1/\lambda$. This distinction explains new features in the recognition scores shown in Fig. 3.8c,d. First, one can see that the volume-based discriminator greatly outperforms the diameter-based one. This is not surprising because the volume of the LCH is much more sensitive to the dimensionality reduction than the diameter. Second, in both cases, the recognition score does not grow monotonously with the mean phase duration T. In fact, the mean duration of the bulk phase, $(R^2 - (R-a)^2)/(6D)$, is fixed (and equal to 16.25 in our example), whereas the mean duration of the surface phase is progressively increased. As a consequence, as T grows, it becomes more difficult to detect the short bulk phases that results in the decrease of the recognition score. A similar behavior can be seen in Fig. 3.12 for intermittent Brownian motion (Model 1) when the durations of two phases differ significantly. Note also that the impact of noise is stronger than in other cases. The performance of the TA MSD discriminator is good but poorer than that of the volume-based estimator. This good performance is explained by the fact that the change in dimensionality from three dimensions to two dimensions also reduces the MSD from $6Dt$ and $4Dt$ that is captured by TA MSD estimator.

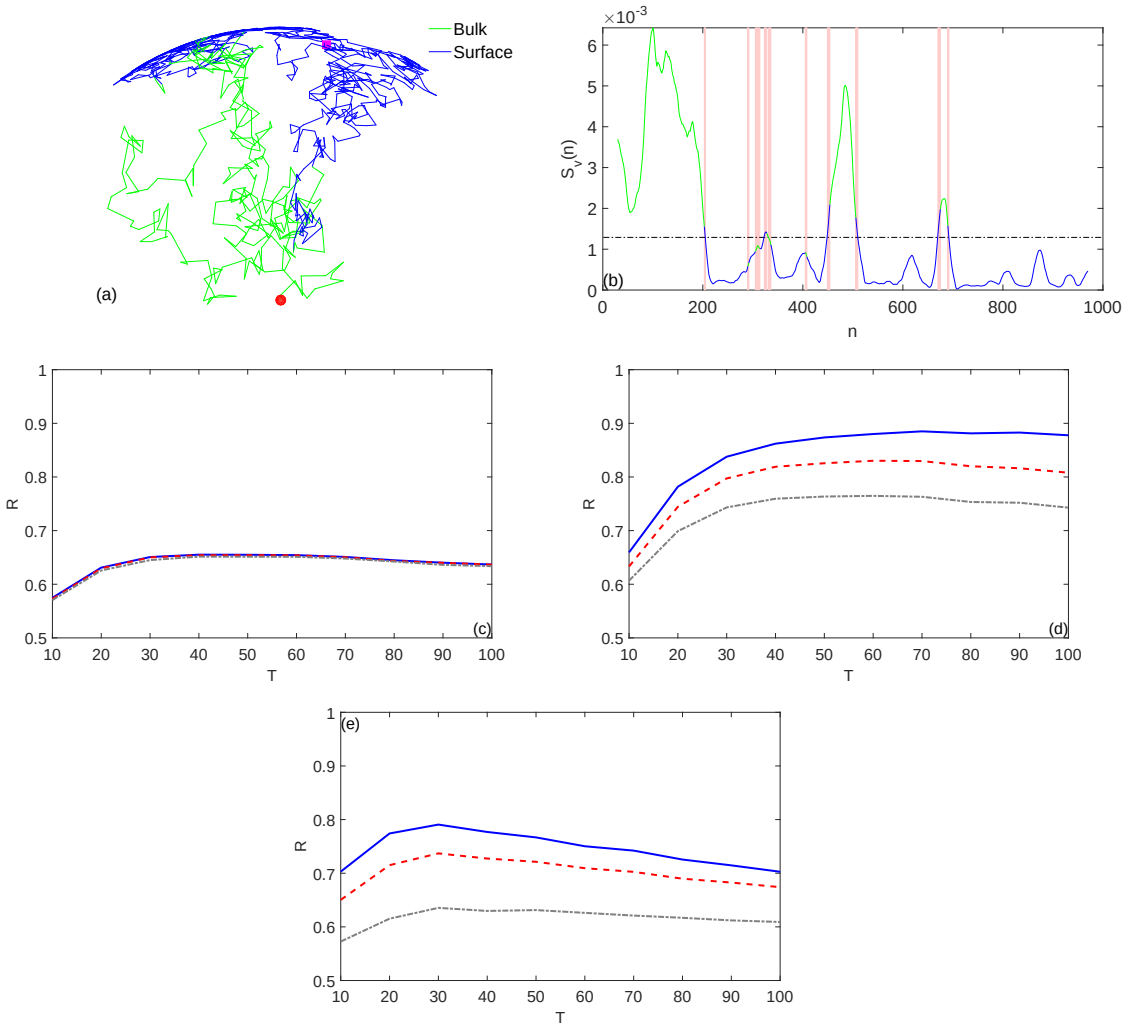


Figure 3.8: **Model 6 (a)** A single trajectory of surface-mediated diffusion in a three-dimensional sphere of radius $R = 1$, alternating bulk diffusion (light green) and surface diffusion (dark blue), with $D_{2d} = D_{3d} = 10^{-3}$ and $\lambda = 10^{-2}$. Circle and square indicate the starting and ending points. **(b)** The weighted LCH diameter $S_d(n)$ with the window size $\Delta = 10$ applied to this trajectory. Pink shadow highlights the false classification zones. Dashed horizontal line shows the empirical mean S_d over that trajectory. **(c,d,e)** Recognition score R of the diameter-based discriminator $S_d(n)$ **(c)**, the volume-based discriminator $S_v(n)$ **(d)**, and the TA MSD-based discriminator **(e)** as a function of the mean duration of the surface phase $T = 1/\lambda$. Three curves correspond to three noise levels σ_n: 0, 0.5σ, and σ (σ being the empirical standard deviation of increment calculated for each trajectory).

3.4 Discussion

The numerical validation has shown that the LCH method can detect change points between the phases that differ either by amplitudes of increments, or by the presence of drift, or by auto-correlations between increments, or by distribution of increments, or by dimensionality of the explored space. Moreover, the method is robust against noise due to the integral-like nature of the LCH-based estimators. In this section, we discuss the choice of the parameters, several limitations and future improvements of the method.

3.4.1 Infinite moments

Even in the case where one of the phases has infinite mean displacement length (e.g. random walk with heavy tailed distribution of increments), the calculation of the discriminator $S_d(n)$ (resp. $S_v(n)$) is still possible. We illustrate this point in Fig. 3.9 by considering an intermittent process alternating Brownian motion with standard deviation of increments $\sigma = 1$ and Lévy walk that is, random walk with random uniformly distributed direction and random jump lengths picked from a stable distribution. The stable distribution is defined by its characteristic function

$$P(k) = \exp\left(-i\mu k + |\sigma k|^{\alpha}(1 - i\,\beta \mathrm{sign}(k)\tan(\pi\alpha/2))\right), \tag{3.9}$$

for $\alpha \neq 1$, where $\mu \in [0, \infty]$ is the position of the maximum of the distribution, the exponent $\alpha \in [0, 2]$ describes the decay of distribution's tails, $\sigma \in [0, \infty]$ is the scale of fluctuations and β is the asymmetry parameter such that $\beta = 0$ is symmetric while $\beta = -1, 1$ is totally skewed. In the special case $\alpha = 2$, $\beta = 0$, one retrieves the characteristic function of a Gaussian distribution. In our example we choose the parameters $\alpha = 0.7$, $\beta = 1$, $\sigma = 1$, and $\mu = 0$, such that the displacement length is always positive with infinite mean. We also corrupt the trajectories with white Gaussian noise, but this time its standard deviation σ_n is fixed to be proportional to the standard deviation of increment of the Brownian motion. We choose three levels of noise σ_n: $0, \sigma, 10\sigma$, the latter case is extreme because the noise level is ten fold higher than the mean displacement length of Brownian motion.

In this configuration, a small amount of extreme events from the Lévy walk dominates the threshold S_d (resp. S_v) that determines the discrimination. In this case using the empirical

average as a threshold leads to distinguishing the extreme events from the rest but not exactly the two states of motion. As an example, Fig. 3.9 (a) shows that $S_d(n)$ is dominated by two contributions from Lévy walk. This is a standard situation when dealing with heavy-tailed processes. As a consequence the recognition with $S_d(n)$ in Fig. 3.9 (c) (resp. $S_v(n)$ in Fig. 3.9 (e)) is not satisfactory.

A possible solution is to consider the logarithm of the discriminator, $\log S_d(n)$ (resp. $\log S_v(n)$), and to use as a criterion the empirical average of the transformed time series. Figure 3.9 (b) shows that this non-linear transformation reveals the alternation of the phases of motion for which the new threshold is more adapted. Results for $\log S_d(n)$ in Fig.3.9 (d) (resp. $\log S_v(n)$ in Fig.3.9 (f)) are particularly good considering the difficulty of the task. The reason for this transformation to be efficient is that it evaluates how the discriminator is localized on the different scales. The discriminator for the Brownian motion alone remains roughly on a single scale due to the finite standard deviation of the increments while the Lévy walk is scale invariant therefore still displays significant fluctuations in log-scale.

3.4.2 Parameters of the method

Although the LCH method is model-free, there are two parameters to be chosen: the window size Δ and the threshold S_d (or S_v) used for the binary classification. Let us discuss the choice of these parameters in more detail.

Throughout Sec. 3.3, we set $\Delta = 10$. Figure 3.10 shows the effect of the window size Δ on the recognition score R for intermittent Brownian motion (Model 1) with $D_1 = 1/2$, $D_2 = 2$, and the mean phase durations $T_1 = T_2 = 40$. First, one can see that the recognition score as a function of Δ is not monotonous, i.e., there is an optimal window size Δ_c that maximizes the recognition score. This optimality results from a compromise between the reactivity and the robustness of the method. When Δ is too large, the LCH contains too many points that leads to larger delays between actual and detected change points and thus increases the fraction of false classifications. In turn, when Δ is too small, the method is reactive (delays are short) but also too sensitive to stochastic fluctuations within one phase; as a consequence, the fraction of false classifications is also higher due to spontaneous crossings of the discriminator $S_d(n)$ (or $S_v(n)$) of the mean level S_d (or S_v). This latter effect is drastically enhanced in the presence of noise (see how the curves with larger noise levels σ_n are diminished at short Δ). As a consequence,

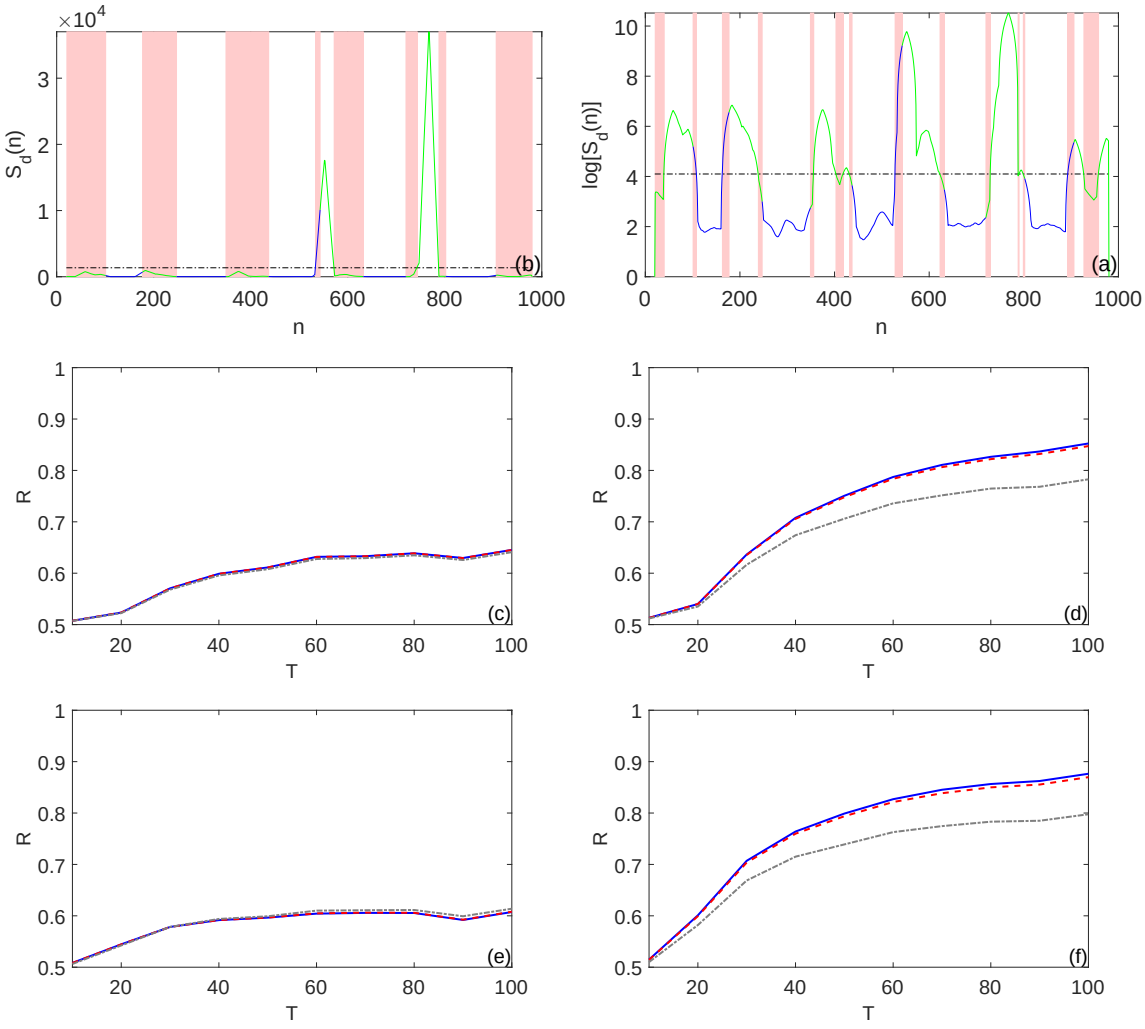


Figure 3.9: **(a,b)** Results for a two-dimensional intermittent process, alternating a "slow" phase of Brownian motion with standard deviation of increments in each dimension $\sigma = 1$ (blue) and a "fast" phase of Lévy walk with stable increments of parameters $\alpha = 0.7$, $\beta = 1$, $\sigma = 1$, and $\mu = 0$ (light green). **(a)** The weighted LCH diameter $S_d(n)$ with the window size $\Delta = 10$ applied to a single trajectory of length $N = 1000$. Dotted line shows the empirical average of S_d. **(b)** The logarithm of $S_d(n)$ from (a). Dotted line shows the empirical average of $\log S_d$. **(a,b)** Pink shadow highlights the false classification zones. **(c,e)** Recognition score R of the diameter-based discriminator $S_d(n)$, and **(c)** and the volume-based discriminator $S_v(n)$ **(e)**. **(d,f)** Recognition score R of the diameter-based discriminator $\log S_d(n)$ **(d)** and the volume-based discriminator $\log S_v(n)$ **(f)**. **(c,d,e,f)** Plot as a function of the mean phase duration T. Lines show the results with three noise levels σ_n: 0 (blue solid), σ (red dashed), and 10σ (gray dash-dotted) (σ being the being the standard deviation of increments of the Brownian phase).

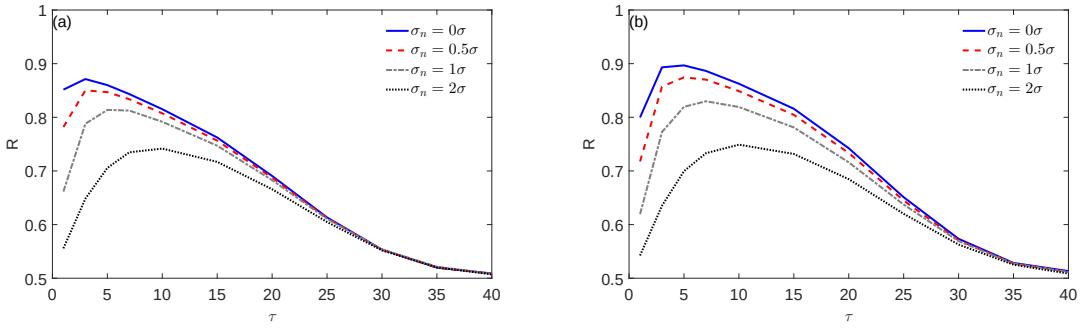


Figure 3.10: Effect of the window size Δ on the recognition score R for intermittent Brownian motion (Model 1) with $D_1 = 1/2$, $D_2 = 2$, and the mean phase durations $T_1 = T_2 = 40$. **(a)** The diameter-based discriminator $S_d(n)$; **(b)** the volume-based discriminator $S_v(n)$. Four curves correspond to four noise levels σ_n: 0, 0.5σ, σ, and 2σ (σ being the empirical standard deviation of increment calculated for each trajectory).

the optimal window size Δ_c is increased for noisier data. Clearly, the Δ_c should also depend on the phase duration. One cannot therefore choose the optimal window size without *a priori* knowledge about the noise and phase durations. In practice, the range $5 \leq \Delta \leq 10$ seems to be the reasonable choice of the window size. Note that the recognition score versus Δ for the intermittent fBm (Model 3) exhibits very similar behavior (not shown).

The second parameter of the method is the threshold S_d (or S_v) that is used to distinguish "slow" and "fast" phases. Without prior knowledge about the process, we chose the arithmetic mean of the discriminator $S_d(n)$ (or $S_v(n)$) over the trajectory as the threshold. However, this choice is not necessarily optimal. For instance, one could use another mean (e.g., quadratic, harmonic or geometric, as we used in Sec. 3.4.1), or set the threshold to be proportional to the mean, or choose another function or constant. In order to justify this empirical choice, we consider a receiver operating characteristic (ROC) curve for both diameter-based and volume-based discriminators. For this purpose, we compute the true positive rate (the fraction of "fast" phase points that were identified as "fast") and the false positive rate (the fraction of "slow" phase points that were identified as "fast") by varying the threshold from the minimal to the maximal value of the discriminator $S_d(n)$ (or $S_v(n)$). Figure 3.11 shows the ROC curves for six considered models and for two discriminators at the mean duration time $T = 40$. An ideal discriminator

would yield the true positive rate at 1 and the false positive rate at 0 (the left upper corner), whereas a random discriminator would fill the diagonal. The threshold at the minimal value of the discriminator classifies all points as belonging to the "fast" phase (as $S_d(n) \geq \min_n\{S_d(n)\}$) that corresponds to the right upper corner. Similarly, the threshold at the maximal value of the discriminator classifies all points as belonging to the "slow" phase (as $S_d(n) \leq \max_n\{S_d(n)\}$) that corresponds to the left lower corner. The intermediate thresholds yield a concave ROC curve lying above the diagonal. The diamond symbols indicate the threshold at the mean value S_d (or S_v) that we suggest and use in this chapter. One can see that this value is the closest to the left upper corner and thus the optimal choice, at least for the considered models and sets of parameters.

The choice of the threshold is related to another important question about the choice of the LCH-based geometric property as the discriminator. The local convex hull captures changes in the mutual arrangement of points, in a somewhat similar way as our eyes do. The diameter and the volume are the basic geometric characteristics of the LCH that reflect, respectively, the overall size and anisotropy of points. Although these characteristics appear to be natural, one can use any function of these (or other) characteristics as well. It is still unclear what is the optimal function of the local convex hull to distinguish between two phases. In other words, among all possible functions of the LCH, which one would yield the highest recognition score. We expect that the optimal choice that maximizes the recognition score, depends on the stochastic model of both phases as illustrated in Sec. 3.4.1 where one of the phases is heavy a tailed process.

3.4.3 Phase durations

In all examples studied in Sec. 3.3 (except for surface-mediated diffusion), phase durations were chosen as independent exponentially distributed random variables with equal mean phase durations: $T_1 = T_2 = T$. This means that all phases have distinct but similar durations, whose average and standard deviation are equal to T.

When two phases have significantly different durations, the detection of the shorter phase can be problematic. To illustrate this point, we consider again the intermittent Brownian motion (Model 1), in which one phase duration is kept fixed, whereas the other phase duration is variable (in contrast to earlier figures, here the phase durations are fixed, not exponentially distributed).

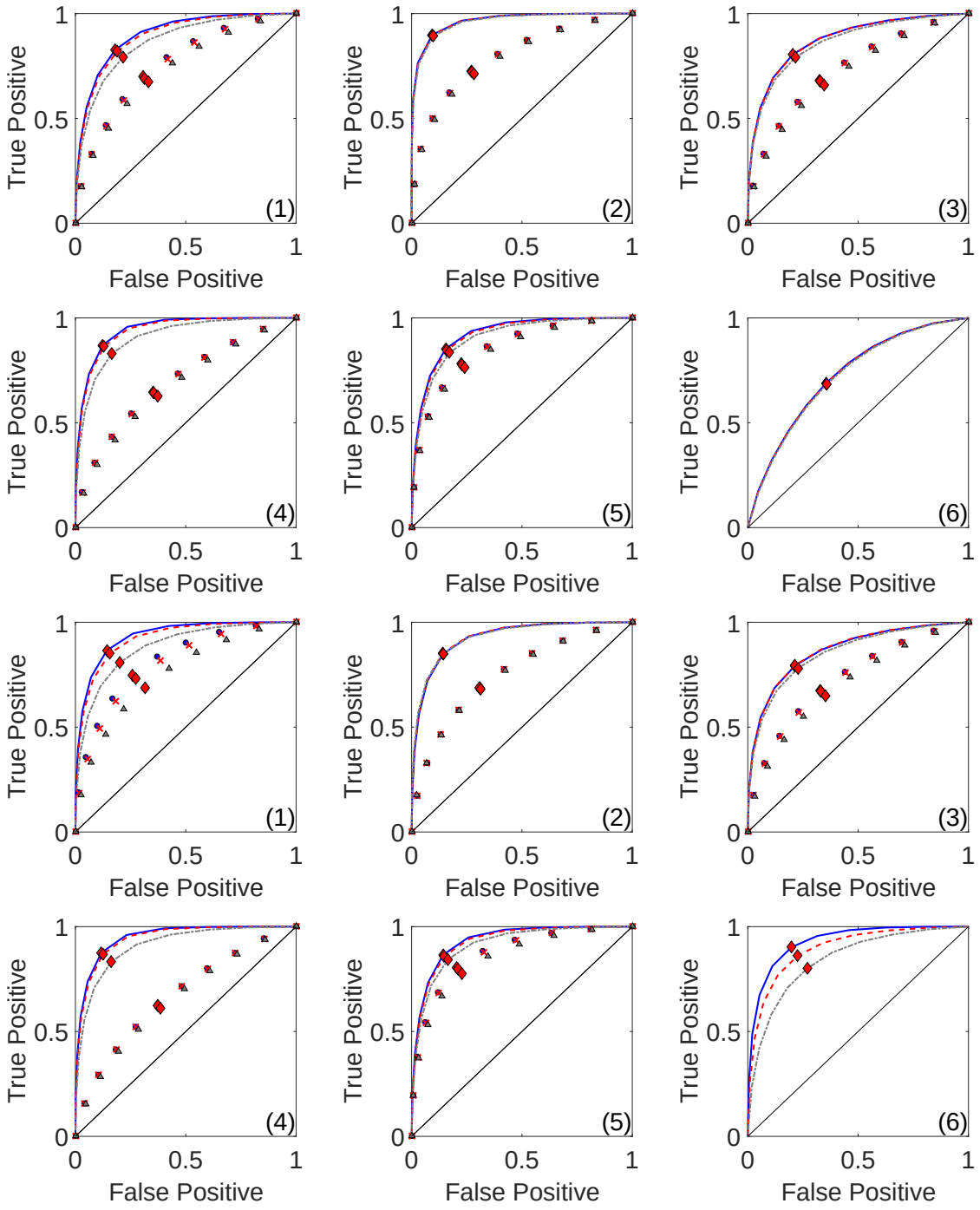


Figure 3.11: Effect of the threshold on the recognition quality for six models (following the enumeration at the beginning of Sec. 3.3), for the diameter-based discriminator (top 6 plots) and the volume-based discriminator (bottom 6 plots). The left upper corner and the right lower corner correspond to the threshold at the minimal and the maximal values, whereas diamonds indicate the mean value S_d (or S_v) used as the threshold throughout the chapter. Legend of six curves on each plot is the same as in Figs. 3.2, 3.4, 3.5, 3.6, 3.7, 3.8 for the recognition score (three levels of noise and two sets of model parameters, except for the Model 6).

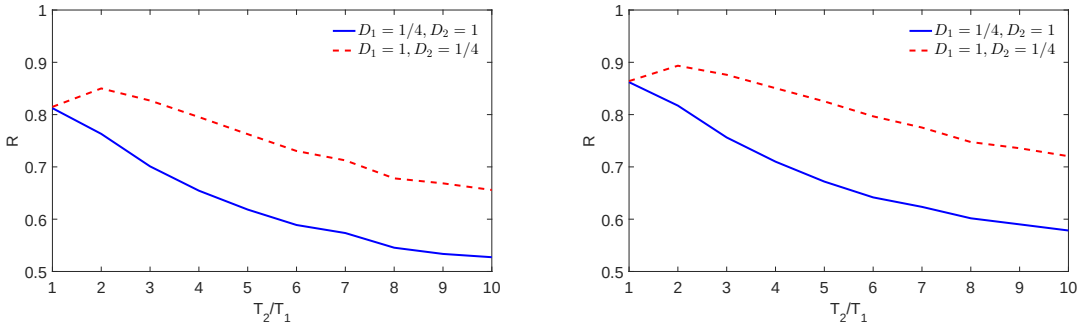


Figure 3.12: Effect of unequal phase durations for planar Brownian motion alternating "slow" and "fast" phases (Model 1). Recognition score R of the diameter-based discriminator $S_d(n)$ **(a)** and the volume-based discriminator $S_v(n)$ **(b)** as a function of the second phase duration T_2 while the first phase duration T_1 is kept fixed: $T_1 = 40$. Solid line: the first phase is "slow" ($D_1 = 1/2$), the second phase is "fast" ($D_2 = 2$); dashes line: the first phase is "fast" ($D_1 = 2$), the second phase is "slow" ($D_2 = 1/2$).

Figure 3.12 shows the recognition score as a function of T_2/T_1, with fixed $T_1 = 40$. Solid line corresponds to the case, in which the first phase is "slow" and the second phase is "fast". As the duration of the fast phase increases, its points start to dominate in the geometric properties of the LCH, and thus to shift the threshold S_d (or S_v) to higher values. As a consequence, the detection of shorter slow phases becomes more difficult, and the recognition score decreases. In the opposite (dashed line) case when the first phase is "fast" and the second phase is "slow", the situation is slightly different. One can see that the recognition score first increases and then decreases. We conclude that short "slow" phases are on average more difficult to detect than short "fast" phases. In both cases, the method is incapable of detecting the phases shorter than or comparable to the window size Δ, which in turn should not be smaller than $5 - 10$ steps.

If we assume that there are no phases shorter than some T_0 ($\approx 10 - 20$), any shorter phase detected by the algorithm can be attributed to a spontaneous crossing of $S_d(n)$ (or $S_v(n)$) of the mean level. The recognition quality can thus be improved by re-classifying such inappropriate phases. For instance, if a too short slow phase is detected between two fast phases, the slow phase can be re-classified into the fast one, i.e., these three consecutive phases are merged and classified as a single fast phase. This post-processing correction is particularly important when one aims

at getting the statistics of slow and fast phase durations. In fact, a spurious short slow phase of the above example cut a long fast phase into two shorter pieces and thus significantly affected the statistics of phase durations. For a better performance, such post-processing corrections or more elaborate statistical techniques (such as an estimation of likelihood of short phases) need to be elaborated and tested on an application-specific basis. Note also that even if the corrected short phase was not spurious (and thus the correction was erroneous), the correction may still improve the overall recognition score, as it eliminates two false classifications related to the delays. We emphasize that this correction procedure has not been used in this chapter.

3.5 Conclusion

We introduced a new model-free local convex hull method for detecting change points in a single-particle trajectory. The LCH is constructed on a small number of consecutive points of the trajectory and thus reflects the dynamics *locally*. In this way, the trajectory is transformed into two time series, the diameter and the volume of the LCH along the trajectory. These time series are then used for a binary classification into "slow" and "fast" phases.

The LCH method was validated on six common models of intermittent processes: Brownian motion with two diffusivities, Brownian motion with and without drift, fractional Brownian motion with different Hurst exponents, Brownian motion with and without harmonic potential, Brownian motion and exponential flights, and surface-mediated diffusion. For all these models, we computed the recognition score R as the fraction of successfully classified points. We showed that R grows with the mean phase duration and reaches the values as high as 90% at $T = 100$ when two phases were quite distinct for a trajectory of $N = 1000$ steps. We analyzed how distinct two phases should be for a robust classification. We showed that recognition errors are mostly localized at the transition between two phases, but can also occur due to the overlapping of the distributions of LCH from the pure processes. We also showed that, due to the integral-like character of the LCH, the method is much more robust against noise than conventional methods such as, e.g., time averaged mean square displacement. In particular, the TA MSD discriminator has outperformed the LCH discriminators only in the case of two alternating Brownian motions and only without noise. In all other cases, the performance of LCH estimators was much higher.

In the case where one of the phases displays infinite moments, the discrimination procedure

should be based on the logarithm of $S_d(n)$ (resp. $S_v(n)$) be compared with the average of $\log S_d(n)$ (resp. $\log S_v(n)$). In general choosing a window $\Delta = 10$ is a good compromise between reactivity and robustness of the detection. The choice of the threshold to be the arithmetic mean of $S_d(n)$ (resp. $S_v(n)$) is the best in the cases considered (when the moments of displacements are finite). However this choice is not universal, it can depend on the process itself and on the proportion of each phase. Nevertheless this choice seems to be robust when the proportion of one phase is greater than 10% of the trajectory.

The LCH method, being based on purely geometric features of the trajectory, can be applied to a wide range of relevant intermittent processes such as animal foraging, active/passive motion inside the cell, run and tumble motion of bacteria or motion exhibiting a change in dimensionality. The LCH method also has the potential to be successfully applied to detect changes in instantaneous firing rates in neurons. Further development will be directed to a more automatized version of this test. First it would be interesting to develop a strategy for verifying that the process is indeed intermittent before applying the LCH method. A second development would be to find an appropriate window size Δ directly from the data, or to consider all possible Δ in the spirit of persistent homology [104]. Finally an LCH-based functional sensitive to anisotropy should be relevant for many situations.

Chapter 4

Non-Gaussian diffusion in heterogeneous media

4.1 Introduction

There is an abundance of empirical observations of non-Gaussian diffusion [25, 48, 75, 102, 103, 108, 111, 123, 202, 209, 229, 235, 256, 267, 278, 279] (see Chapter 1). The observation of the distribution of displacements with exponential tails in multiple systems suggests that it is reminiscent of heterogeneous complex media despite differences in experimental and microscopic setups. Aiming at modeling these exponential tails, we proceed in this chapter by a mesoscopic approach, which describes a system with time-dependent macroscopic quantities. More particularly we focus on the case when non-Gaussian diffusion originates from local changes in diffusive properties of the medium. We propose an analytically solvable generalization of former models [49, 55, 135, 137, 278], in which the parameters have a clear physical interpretation, then we discuss the relevant statistical properties and investigate them analytically and numerically. Within this chapter, analytical results are corroborated by Monte Carlo simulations using Milstein scheme [127].

We present in Sec. 4.2 a three parameter model of non-Gaussian diffusion in which diffusivity is fluctuating around an average value $\bar{D}$ (m^2/s) (which constitutes the effective diffusion coefficient at long time), with the correlation time τ (s) and the amplitude of fluctuations σ (m/s). The description is formulated in terms of coupled Langevin equations from which the

characteristic function of displacements is derived in an exact explicit form. The shape of the distribution is tuned by one dimensionless parameter

$$\nu = \frac{\bar{D}}{\sigma^2 \tau}, \tag{4.1}$$

which compares the diffusivity fluctuation time $\bar{D}/\sigma^2$ and the diffusivity correlation time τ. Depending on ν, the distribution of displacements can be close to exponential ($\nu = 1$), parabolic ($\nu > 1$) or peaked ($\nu < 1$) at the origin. In all cases, the distribution of displacements exhibits an exponential tail with eventual power law corrections. We show that this description leads to a linear dependence of the MSD on time, while fluctuations of time-averaged MSD span up at long times depicting the effect of heterogeneous diffusivity. In Sec. 4.4 we derive the asymptotic behavior of the probability distribution as a function of time and space in several cases. Then in Sec. 4.5 we derive and analyze statistical properties of the model: (i) the autocorrelation of squared increments which describes memory loss of diffusivity, (ii) the convergence to a Gaussian distribution through the non-Gaussian parameter, (iii) the ergodic and self-averaging properties. In Sec. 4.6 we show that two models with the same non-Gaussian parameter can be different and briefly discuss a generalization to anomalous diffusion. Finally, we conclude the investigations in Sec. 4.7.

4.2 Model of non-Gaussian diffusion

We propose a model of a tracer motion in a heterogeneous medium, in which the diffusivity is a stochastic process instead of being a constant. In order to justify this description, let us consider a single particle tracking measurement of duration t_{exp} with a timestep Δt between two position recordings. If the motion occurs in a homogeneous environment, the distribution of displacements becomes Gaussian very fast, in a time t_{loc} of equilibration of the tracer with its local environment. For a heterogeneous medium, in which the diffusivity can vary spatio-temporally (noted $D_{x,t}$), we introduce the time t_{sys} for a particle to explore the whole medium and to average diffusivities experienced in the medium. On one hand, if $t_{loc} \ll t_{sys} < \Delta t$, increments of the motion are already coarse-grained at a measurement timestep Δt and therefore are Gaussian. On the other hand, if $t_{loc} < \Delta t \ll t_{sys}$, the motion is not necessarily Gaussian because diffusivity evolves in time, and the tracer continuously moves from one equilibrium state

to another. This can be interpreted as the effect of spatio-temporal heterogeneities in the medium seen from the point of view of a single particle. To simplify the analysis we describe diffusivity as a stochastic process in time, D_t, with the idea that the stochasticity is an annealed simplification of the spatio-temporal disorder. The particle experiences a fluctuating diffusivity around an average value $\bar{D}$ toward which the time-averaged effective diffusion coefficient converges at long times (i.e. $t \gg t_{sys}$). Two physical constraints for a fluctuating diffusivity are (i) the distribution of displacements converges to a Gaussian one at long times, so diffusivity should have a stationary distribution in the long-time limit, with the average value $\bar{D}$; (ii) diffusivity as a measure of local kinetic energy of the tracer should be non-negative.

We propose to model time-dependent diffusivity D_t as a Feller process [92, 99] or square root process, also known in financial literature as the Cox-Ingersoll-Ross process (CIR) [61]. This process has been developed in order to rationalize fluctuations of volatility in price asset returns. In the CIR model, the diffusivity fluctuates in a harmonic potential centered on $\bar{D}$ and remains non-negative thanks to the balance between the pulling of harmonic potential and the noise reduction of diffusivity-dependent fluctuations at small D_t. Moreover, the stationary distribution of diffusivity is known to be a Gamma distribution. For the sake of clarity, we first formulate the model for one-dimensional motion and then show its straightforward extension to the multi-dimensional isotropic case. For a tracer starting at x_0 with diffusivity D_0, the corresponding coupled Langevin equations read:

$$\begin{cases} dx_t = \sqrt{2D_t}dW_t^{(1)}, \\ dD_t = \frac{1}{\tau}(\bar{D} - D_t)dt + \sigma\sqrt{2D_t}dW_t^{(2)}, \end{cases} \tag{4.2}$$

where x_t and D_t are stochastic time-dependent position and diffusivity of the tracer, $dW_t^{(1)}$ and $dW_t^{(2)}$ are increments of independent Wiener processes (white noises). The model includes three parameters: the average diffusivity $\bar{D}$ (in m^2/s), the correlation time τ (in s) and the amplitude of fluctuations σ (in m/s).

The approach by Chubynsky and Slater [55] is retrieved by setting a diffusivity bias $s(D) = -\frac{1}{\tau}(D - \bar{D})$ and a diffusivity of diffusivity $d(D) = \sigma^2\sqrt{2D}$, although in our model, a reflecting boundary is necessary only at $D = 0$ (see below). Jain and Sebastian [135] and Chechkin *et al.* [49] considered the diffusivity as the distance from the origin of an n-dimensional Ornstein-Uhlenbeck process, which is a particular case of our model. To emphasize this point we

present the derivation of the Cox-Ingersoll-Ross model starting from the n-dimensional Ornstein-Uhlenbeck process in the next paragraph.

Let us consider a collection of independent Ornstein-Uhlenbeck processes indexed by $i \in [1, n]$, each of them obeying the Langevin equation:

$$dY_t^{(i)} = -\frac{1}{2}\beta Y_t^{(i)} dt + \sigma dW_t^{(i)}, \tag{4.3}$$

where β is the inverse correlation time, σ is the level of fluctuations, and $W_t^{(i)}$ are independent Wiener processes. Following [49, 135] the diffusing diffusivity is modeled as

$$D_t = \sum_{i=1}^{n} \left(Y_t^{(i)}\right)^2. \tag{4.4}$$

Let $f(y_1, y_2..., y_n) = \sum_{i=1}^{n} y_i^2$ so that $\frac{\partial}{\partial y_i} f = 2y_i$ and $\frac{\partial^2}{\partial y_i \partial y_j} f = 2\delta_{ij}$. According to Itô formula, we get

$$\begin{aligned}
dD_t &= \sum_{i=1}^{n} \frac{\partial}{\partial y_i} f dY_t^{(i)} + \frac{1}{2}\sum_{i,j=1}^{n} \frac{\partial^2}{\partial y_i \partial y_j} f dY_t^{(i)} dY_t^{(j)} \\
&= \sum_{i=1}^{n} 2Y_t^{(i)} \left(-\frac{1}{2}\beta Y_t^{(i)} dt + \sigma dW_t^{(i)}\right) + \sum_{i=1}^{n} \left(-\frac{1}{2}\beta Y_t^{(i)} dt + \sigma dW_t^{(i)}\right)^2 \\
&= \sum_{i=1}^{n} 2Y_t^{(i)} \left(-\frac{1}{2}\beta Y_t^{(i)} dt + \sigma dW_t^{(i)}\right) + n\sigma^2 dt \\
&= \left(n\sigma^2 - \beta \sum_{i=1}^{n} \left(Y_t^{(i)}\right)^2\right) dt + 2\sigma \sum_{i=1}^{n} Y_t^{(i)} dW_t^{(i)} \\
&= \left(n\sigma^2 - \beta D_t\right) dt + 2\sigma \sum_{i=1}^{n} Y_t^{(i)} dW_t^{(i)} \\
&= \left(n\sigma^2 - \beta D_t\right) dt + 2\sigma \sqrt{D_t} \sum_{i=1}^{n} \frac{Y_t^{(i)}}{\sqrt{D_t}} dW_t^{(i)}.
\end{aligned}$$

The stochastic process W_t defined as $W_t = \sum_{i=1}^{n} \int_0^t \frac{Y_s^{(i)}}{\sqrt{D_s}} dW_s^{(i)}$ is a martingale because it has no drift [187]. For its increments, $dW_t = \sum_{i=1}^{n} \frac{Y_t^{(i)}}{\sqrt{D_t}} dW_t^{(i)}$, we verify that $dW_t dW_t = \sum_{i=1}^{n} \frac{\left(Y_t^{(i)}\right)^2}{D_t} \left(dW_t^{(i)}\right)^2 = dt$, so the increments are properly normalized. We conclude that W_t is a Wiener process. We now can rewrite the above equation as:

$$dD_t = \left(n\sigma^2 - \beta D_t\right) dt + \sqrt{2}\sigma \sqrt{2D_t} dW_t. \tag{4.5}$$

Setting $\tilde{\sigma} = \sqrt{2}\sigma$, $\beta = 1/\tau$ and $n = \frac{\bar{D}}{\tilde{\sigma}^2\tau} = \nu$ one finally retrieves the Cox-Ingersoll-Ross equation

$$dD_t = \frac{1}{\tau}\left(\bar{D} - D_t\right)dt + \tilde{\sigma}\sqrt{2D_t}dW_t. \tag{4.6}$$

It is now then evident that the results of former studies can be reproduced for integer values $n = \frac{\bar{D}}{\sigma^2\tau}$ and the range of applicability is thus widened because parameters in our model are continuous: $\{\tau, \bar{D}, \sigma\} \in (0, \infty)$. In particular, the case $\nu < 1$, which yields a peaked distribution of displacements and the most peculiar properties of heterogeneous diffusion (see below), was not accessible so far.

We introduce the propagator $P(x, D, t|x_0, D_0)$, i.e. the probability density for a tracer to be at x with diffusivity D at time t, when started from x_0, D_0 at $t = 0$. The corresponding forward Fokker-Planck equation in the Itô interpretation reads

$$\frac{\partial}{\partial t}P(x, D, t|x_0, D_0) = \frac{1}{\tau}\frac{\partial}{\partial D}\left[(D - \bar{D})P\right] + \frac{\partial^2}{\partial x^2}(DP) + \sigma^2\frac{\partial^2}{\partial D^2}(DP), \tag{4.7}$$

with the initial condition $P(x, D, t = 0|x_0, D_0) = \delta(x - x_0)\delta(D - D_0)$.
Following Dragulescu and Yakovenko [75], this equation is solved by performing the Fourier transform with respect to position x, and the Laplace transform with respect to diffusivity $D \geq 0$:

$$\tilde{P}(q, s, t|x_0, D_0) = \int_{-\infty}^{\infty} dx \int_0^{\infty} dD e^{-iqx - Ds} P(x, D, t|x_0, D_0), \tag{4.8}$$

where q and s are the dual variables to position and diffusivity, respectively. Inserting Eq. (4.8) into Eq. (4.7) leads to the first order partial differential equation:

$$\frac{\partial}{\partial t}\tilde{P} + \left(\sigma^2 s^2 + \frac{1}{\tau}s - q^2\right)\frac{\partial}{\partial s}\tilde{P} = -\frac{1}{\tau}\bar{D}s\tilde{P} + J_D(D = 0, t), \tag{4.9}$$

subject to the initial condition $\tilde{P}(q, s, t = 0|x_0, D_0) = e^{-iqx_0}e^{-sD_0}$. The last term

$$J_D(D = 0, t) = \left(\frac{\bar{D}}{\tau} - \sigma^2\right)P(q, D = 0, t|X_0, D_0) \tag{4.10}$$

can be interpreted as a probability density flux across the boundary $D = 0$ in the phase space (x, D). In the case of an *absorbing* diffusivity boundary at $D = 0$, there is an atom of probability measure at $D = 0$ which "accumulates" absorbed trajectories. The probability of having $D = 0$ grows with time and is related to the function $J_D(D = 0, t)$ which can be deduced from initial conditions by solving an integral equation (see Sec. 4.3.4). In our model, the diffusivity cannot

be physically zero and the distribution must have a nontrivial stationary solution to match convergence to a Gaussian distribution at long times, so we choose *reflecting* boundary condition at $D = 0$ and thus $J_D(D = 0, t) = 0$. In this case, the solution for any $\nu \geq 0$ is (see Sec. 4.3.2 for detailed derivation)

$$\tilde{P}(q, s, t|x_0, D_0) = F(x_0, D_0, s) \left(\frac{\sigma^2 \tau}{\omega} \left[\left(s + \frac{1+\omega}{2\sigma^2\tau} \right) - \left(s + \frac{1-\omega}{2\sigma^2\tau} \right) e^{-\omega t/\tau} \right] e^{\left(\frac{\omega-1}{2}\right) t/\tau} \right)^{-\nu}, \tag{4.11}$$

with

$$F(x_0, D_0, s) = \exp\left[-iqx_0 - \frac{D_0}{2\sigma^2\tau} \left(-1 - \omega + \frac{2\omega}{1 - \xi e^{-\omega t/\tau}} \right) \right], \tag{4.12}$$

$$\xi = 1 - \frac{2\omega}{1 + \omega + 2\sigma^2\tau s}, \tag{4.13}$$

and

$$\omega = \sqrt{1 + 4\sigma^2\tau^2 q^2}. \tag{4.14}$$

The inverse Fourier and Laplace transforms yield $P(x, D, t|x_0, D_0)$.

However, this solution provides too detailed information which can hardly be confronted to single particle tracking data with no direct access to diffusivities D and D_0. We thus integrate the solution over D (which is equivalent to set $s = 0$) to get the marginal distribution of positions. We also assume that the tracer's initial diffusivity D_0 is taken from its stationary Gamma distribution $\Pi(D_0)$ (see Sec. 4.3.2):

$$\Pi(D_0) = \frac{\nu^\nu D_0^{\nu-1}}{\Gamma(\nu)\, \bar{D}^\nu} \exp\left(-\frac{\nu}{\bar{D}} D_0 \right), \tag{4.15}$$

where the shape parameter ν is defined in Eq. (4.1). The average over D_0 yields the marginal distribution

$$P(x, t|x_0) = \frac{1}{2\pi} \int_{-\infty}^{\infty} dq\, e^{-iqx_0} \tilde{P}(q, t), \tag{4.16}$$

with

$$\tilde{P}(q, t) = \left(e^{-\frac{1}{2}(\omega-1)t/\tau} \frac{4\omega}{(\omega+1)^2} \left(1 - \left(\frac{\omega-1}{\omega+1} \right)^2 e^{-\omega t/\tau} \right)^{-1} \right)^\nu. \tag{4.17}$$

An alternative solution using the subordination concept, inspired from [49], is given in Sec. 4.3.3.

When particles undergo isotropic motion in $\mathbb{R}^d$, the formula for the distribution of displacements

remains almost unchanged, except that one has to perform multi-dimensional inverse Fourier transform in $\mathbb{R}^d$, with $\mathbf{q}, \mathbf{x}$ and $\mathbf{x_0}$ being vectors:

$$P(\mathbf{x}, t|\mathbf{x_0}) = \int\limits_{\mathbb{R}^d} \frac{d^d\mathbf{q}}{(2\pi)^d}\, e^{i\mathbf{q}(\mathbf{x}-\mathbf{x_0})} \tilde{P}(|\mathbf{q}|, t), \tag{4.18}$$

with $w = \sqrt{1 + 4\sigma^2\tau^2|\mathbf{q}|^2}$. Since the characteristic function $\tilde{P}(|\mathbf{q}|, t)$ depends only on $|\mathbf{q}|$, one can use spherical coordinates and integrate out the angular variables, yielding

$$P(r, t) = \frac{r^{1-d/2}}{(2\pi)^{d/2}} \int\limits_0^\infty dq\, q^{d/2} J_{\frac{d-2}{2}}(qr)\, \tilde{P}(q, t), \tag{4.19}$$

where $J_\alpha(x)$ is the Bessel function of the first kind, $r = |\mathbf{x} - \mathbf{x_0}|$, and $\tilde{P}(q, t)$ is given by Eq. (4.17). In what follows, we focus on the one-dimensional case, bearing in mind straightforward extensions to the multi-dimensional case.

Figure 4.1 illustrates the convergence of the distribution of displacements to a Gaussian one as t increases. Note also a perfect agreement between the theoretical formula (4.16) and Monte Carlo simulations.

Figure 4.2 shows the effect of the shape parameter ν on the distribution of displacements at time $t = 1$. The parameter ν changes the shape of the distribution. When $\nu \leq 1$, fluctuations are strong compared to both the average diffusivity $\bar{D}$ and the correlation time τ. In this case, the probability of diffusivity close to zero is large that makes the distribution of displacements peaked near $x = 0$. In turn, the distribution gets closer and closer to Gaussian as $\nu \to \infty$ (see Sec. 4.4.1).

On the length scale $\sigma\tau$, the diffusivity remains roughly the same. Intuitively, if $\sqrt{\bar{D}t} \ll \sigma\tau$, a particle has not enough time to explore the system. The distribution $P(x, t|x_0)$ could be considered as a superstatistical description of independent particles with constant but randomly chosen diffusion coefficients (see Sec. 4.4.2). Inversely, when $\sqrt{\bar{D}t} \gg \sigma\tau$, the particle has enough time to explore the medium and the distribution progressively becomes Gaussian. We introduce thus the time-dependent dimensionless diffusion length:

$$\mu(t) = \frac{\sqrt{\bar{D}t}}{\sigma\tau}. \tag{4.20}$$

As $\mu(t) \to \infty$, the particle explores the space beyond the correlation length, and the distribution gets closer to a Gaussian one. We show in Fig. 4.3 how $\mu(t)$ impacts the shape of the distribution.

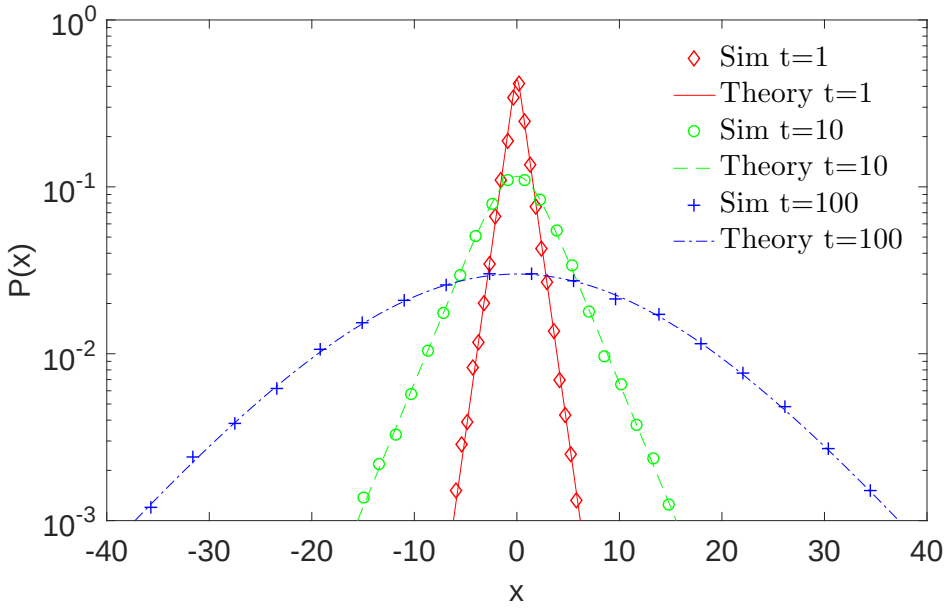


Figure 4.1: Distribution of displacements at times $t = \{1, 10, 100\}$. Here $\tau = 10$, $\bar{D} = 1$, $\sigma = 1/\sqrt{\tau}$ and thus $\nu = 1$. Theoretical results (lines) are compared to Monte Carlo simulations (symbols) with $M = 10^6$ particles.

For instance at $\mu(t) = 1$, the distribution is already almost Gaussian. When $\mu(t)$ decreases, the distribution becomes more peaked. Note that the quantity $\mu(t)$ is directly related to the non-Gaussian parameter (see Eq. (4.85) below).

Figure 4.4 illustrates four random trajectories and corresponding displacements. The envelop of time series of displacements shows patterns of fluctuations correlated on timescale τ. For small τ, the envelop becomes constant as in the Brownian motion case.

4.3 Solution of the Fokker-Planck equation

4.3.1 Statement of the problem

In this section we present the detailed solution to the two-dimensional forward Fokker-Planck equation (Eq. (4.7)). This section concerns the necessary mathematical steps for obtaining the solution presented and commented in Sec. 4.2.

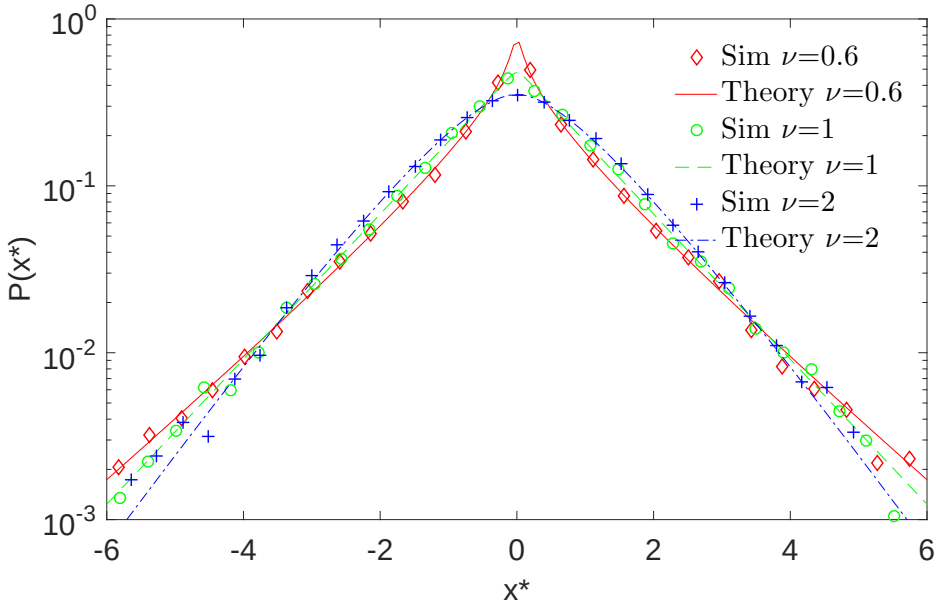


Figure 4.2: Distribution of normalized displacements, with $x^* = x/\sqrt{\bar{D}t}$, at fixed time $t = 1$ for different parameters $\nu = \{0.6, 1, 2\}$. For each case, we kept $\bar{D} = 1$ and $\tau = 100$ and varied σ. Theoretical results (lines) are compared to Monte Carlo simulations (symbols) with $M = 10^6$ particles.

The Fokker-Planck equation on position and diffusivity with the initial condition $P(x, D, t = 0|x_0, D_0) = \delta(x - x_0)\delta(D - D_0)$. The equation on $P(x, D, t|x_0, D_0)$ can be formally expressed in term of a two-dimensional probability density flux $\vec{J} = \{J_x, J_D\}$

$$\frac{\partial}{\partial t}P(x, D, t|x_0, D_0) = -\text{div}\left(\vec{J}\right), \tag{4.21}$$

with components

$$J_x = -D\frac{\partial}{\partial x}P, \tag{4.22}$$

$$J_D = -\frac{1}{\tau}\left(D - \bar{D}\right)P - \sigma^2\frac{\partial}{\partial D}\left[DP\right]. \tag{4.23}$$

This equation can be solved by transforming the position x into the Fourier space, and the diffusivity, defined on the real half line $D \in [0, \infty)$, into the Laplace space:

$$\tilde{P}(q, s, t|x_0, D_0) = \int_{-\infty}^{\infty} dx e^{-iqx} \int_0^{\infty} dD e^{-Ds} P(x, D, t|x_0, D_0). \tag{4.24}$$

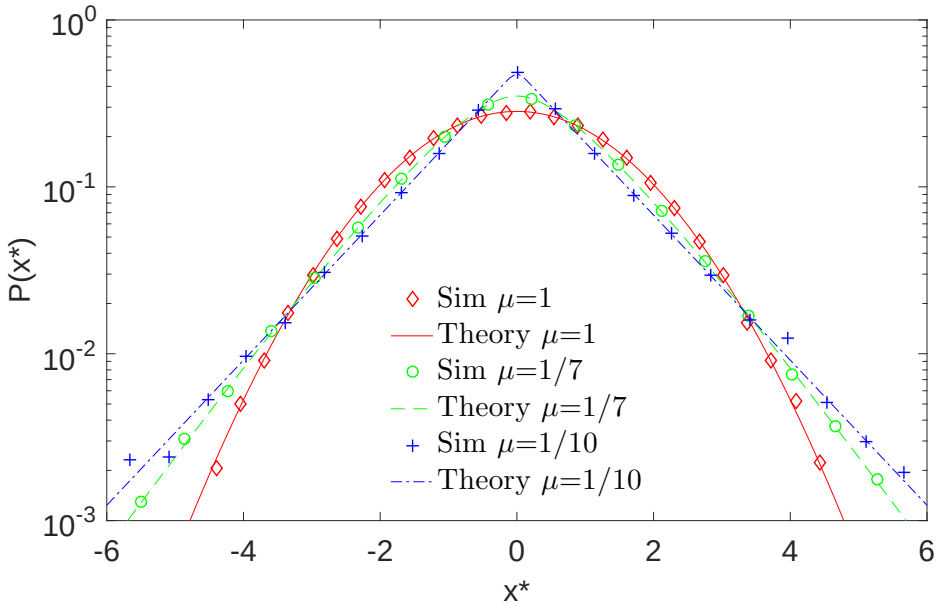


Figure 4.3: Distribution of normalized displacements, with $x^* = x/\sqrt{\bar{D}t}$, at fixed time $t = 1$, with $\mu(t)$ from Eq. (4.20) being varied in the range $\{1/10, 1/7, 1\}$ corresponding to $\nu = \{1, 2, 100\}$, by changing σ and keeping $\bar{D} = 1$ and $\tau = 100$. Theoretical results (lines) are compared to Monte Carlo simulations (symbols) with $M = 10^6$ particles.

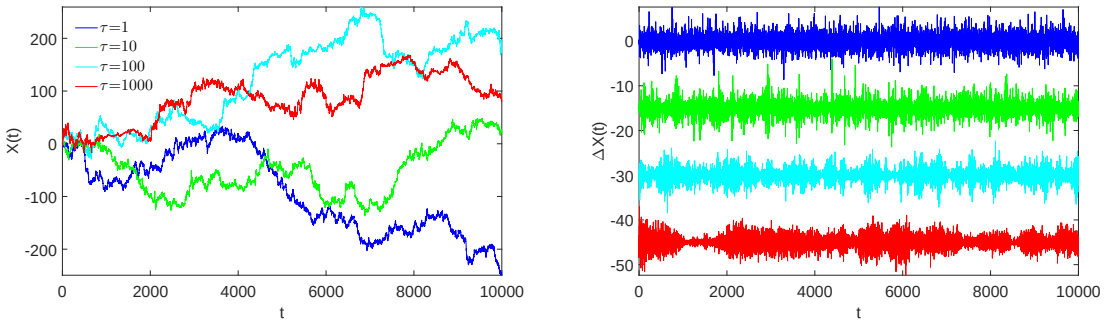


Figure 4.4: *Left.* Trajectories simulated for several values of $\tau = \{1, 10, 100, 1000\}$. Here ν is kept equal to one, with $\bar{D} = 1$ and $\sigma = 1/\sqrt{\tau}$. *Right.* Corresponding time series of position increments with lag-time $\delta t = 1$. For clarity, the time series are artificially shifted (with increasing τ values from top to bottom), but remain with zero mean.

so that Eq. (4.21) becomes:

$$\frac{\partial}{\partial t}\tilde{P} = \int_{-\infty}^{\infty} dx e^{-iqx} \int_{0}^{\infty} dD e^{-Ds} \left(D\frac{\partial^2}{\partial x^2}P - \frac{\partial}{\partial D}J_D \right) \tag{4.25}$$

This leads to the first-order partial differential equation:

$$\frac{\partial}{\partial t}\tilde{P} + G(s)\frac{\partial}{\partial s}\tilde{P} = -\frac{\bar{D}}{\tau}s\tilde{P} + J_D(D=0,t), \tag{4.26}$$

with

$$G(s) = \sigma^2 \left(s - s_1\right)\left(s - s_2\right) \tag{4.27}$$

where

$$s_1 = \frac{-1+\omega}{2\sigma^2\tau}, \quad s_2 = \frac{-1-\omega}{2\sigma^2\tau}, \tag{4.28}$$

and

$$\omega = \sqrt{1 + 4\sigma^2\tau^2 q^2}. \tag{4.29}$$

The initial condition is now $\tilde{P}(q, s, t = 0|x_0, D_0) = e^{-iqx_0}e^{-sD_0}$. The last term in Eq. (4.26) is the probability density flux at $D = 0$ which can be equivalently written $J_D(D = 0, t) = \sigma^2(\nu - 1)P(q, D = 0, t|x_0, D_0)$. First we solve the problem for the model of reflecting boundary in Sec. 4.3.2 and then demonstrate the effect with absorbing boundary condition at $D = 0$ in Sec. 4.3.4.

4.3.2 Reflecting boundary condition at $D = 0$

First we solve the problem in the case of reflecting boundary condition, i.e. without flux at $D = 0$: $J_D(D = 0, t) = 0$. For $\nu \geq 0$, we search for the solution of the equation

$$\frac{\partial}{\partial t}\tilde{P} + G(s)\frac{\partial}{\partial s}\tilde{P} = -\frac{\bar{D}}{\tau}s\tilde{P}, \tag{4.30}$$

in the form:

$$\tilde{P}(q, s, t|x_0, D_0) = f\left(t - g\left(s\right)\right)h\left(s\right), \tag{4.31}$$

with three unknown functions f, g, h. Nontrivial solutions are found by solving

$$\begin{cases} 1 - g'G = 0, \\ \\ Gh' + \frac{\bar{D}}{\tau}sh = 0, \end{cases} \tag{4.32}$$

which gives

$$\begin{cases} g(s) = \frac{\tau}{\omega} \ln\left(\frac{s-s_1}{s-s_2}\right), \\ \\ h(s) = (s-s_1)^{\frac{-\bar{D}}{\omega}s_1}(s-s_2)^{\frac{\bar{D}}{\omega}s_2}. \end{cases} \tag{4.33}$$

Now we use the initial condition to deduce the function f:

$$\tilde{P}(q,s,t=0|x_0,D_0) = e^{-iqx_0}e^{-sD_0} = f\left(-g\left(s\right)\right)h\left(s\right), \tag{4.34}$$

from which we get

$$f(z) = \frac{e^{-iqx_0}e^{-D_0 g^{-1}(-z)}}{h\left(g^{-1}\left(-z\right)\right)}, \tag{4.35}$$

or equivalently,

$$f(z) = \left(\frac{\omega}{\sigma^2\tau}\right)^{\nu} e^{-iqx_0} \exp\left(-D_0\frac{s_1 - s_2e^{-\omega z/\tau}}{1-e^{-\omega z/\tau}}\right)(1-e^{-\omega z/\tau})^{-\nu}e^{-\bar{D}s_1 z/\tau}. \tag{4.36}$$

The solution is finally

$$\tilde{P}(q,s,t|x_0,D_0) = F(x_0,D_0,s)\left(\frac{\omega}{\sigma^2\tau}\right)^{\nu}\left(s-s_2-(s-s_1)e^{-\omega t/\tau}\right)^{-\nu}\exp\left(-\bar{D}\left(\frac{\omega-1}{2\sigma^2\tau}\right)t/\tau\right), \tag{4.37}$$

with ν from Eq. (4.1) and

$$F(x_0,D_0,s) = \exp\left(-iqx_0 - D_0\frac{s_1 - s_2\frac{s-s_1}{s-s_2}e^{-\omega t/\tau}}{1-\frac{s-s_1}{s-s_2}e^{-\omega t/\tau}}\right). \tag{4.38}$$

Substituting s_1 and s_2 in Eq. (4.37) and Eq. (4.38), we get Eq. (4.11).

In practice, it is hard to access directly the time-dependent diffusivity. It is therefore convenient to integrate the joint probability density over diffusivity to get the marginal distribution $\tilde{P}(q,t|x_0,D_0)$, which can be obtained in the Laplace domain by simply setting $s=0$:

$$\tilde{P}(q,t|x_0,D_0) = F(x_0,D_0,s=0|x_0)\left(\frac{\left(\frac{2\omega}{1+\omega}\right)}{1-\left(1-\frac{2\omega}{1+\omega}\right)e^{-\omega t/\tau}}\right)^{\nu}\exp\left(-\bar{D}\left(\frac{\omega-1}{2\sigma^2\tau}\right)t/\tau\right). \tag{4.39}$$

Another issue is the dependence on the initial diffusivity D_0. If the system is in a stationary regime for the diffusivity, one can average over the stationary distribution $\Pi(D_0)$. This distribution can be obtained from Eq. (4.11) by averaging over position (i.e. by setting $q=0$), then

taking the limit $t \to \infty$ and using the inverse Laplace transform relation

$$\Pi(D_0) = \mathcal{L}^{-1}\left[\left(s + \frac{1}{\sigma^2\tau}\right)^{-\nu}\right] = \frac{\nu^\nu}{\Gamma(\nu)\bar{D}^\nu} D_0^{\nu-1} \exp\left(-\frac{D_0}{\sigma^2\tau}\right), \tag{4.40}$$

also known from [92]. Then, the average over initial diffusivity reads

$$\tilde{P}(q,t|x_0) = \int_0^\infty \Pi(D_0)\tilde{P}(q,t|x_0,D_0)dD_0. \tag{4.41}$$

Taking the integral, we deduce Eq. (4.17).

4.3.3 Subordination

In this section we present an alternative derivation of the propagator $P(q,t|x_0)$ in Eq. (4.41) using Subordination technique. Subordination is an elegant mathematical tool to describe complex processes, in particular anomalous diffusion [262,284]. Chechkin *et al.* [49] applied it in the diffusing diffusivity context by observing that the Fokker-Planck equation

$$\frac{\partial}{\partial t}P(x,t) = D(t)\frac{\partial^2}{\partial x^2}P(x,t), \tag{4.42}$$

can be written in the subordinated form:

$$\begin{cases} \dfrac{\partial p(x,u)}{\partial u} = \dfrac{\partial^2}{\partial x^2}p(x,u), \\ \dfrac{\partial u}{\partial t} = D(t), \end{cases} \tag{4.43}$$

where $p(x,u) = \frac{1}{\sqrt{4\pi u}} \exp\left(-\frac{x^2}{4u}\right)$ is the Gaussian propagator. Let $T(u,t)$ be the probability density of $u(t) = \int_0^t D(s)ds$. The solution of Eq. (4.42) can be expressed as

$$P(x,t) = \int_0^\infty p(x,u)T(u,t)du = \int_0^\infty \frac{e^{-\frac{x^2}{4u}}}{\sqrt{4\pi u}}T(u,t)du. \tag{4.44}$$

Now the Fourier transform with respect to x yields:

$$\tilde{P}(q,t) = \int_0^\infty T(u,t)e^{-q^2u}du = \tilde{T}(q^2,t), \tag{4.45}$$

where $\tilde{T}(q^2,t)$ denotes the Laplace transform of T with respect to $s = q^2$. In our model, the description of diffusivity is made in term of the Cox-Ingersoll-Ross equation which reads

$$\frac{\partial \Pi(D,t|D_0)}{\partial t} = \frac{1}{\tau}\frac{\partial}{\partial D}\left[(D-\bar{D})\Pi\right] + \sigma^2\frac{\partial^2}{\partial D^2}\left(D\Pi\right). \tag{4.46}$$

In the Laplace domain, one has

$$\frac{\partial}{\partial t}\tilde{\Pi}(s,t) + G(s)\frac{\partial}{\partial s}\tilde{\Pi}(s,t) = -\frac{1}{\tau}\bar{D}s\tilde{\Pi}(s,t), \tag{4.47}$$

with $G(s) = s(\sigma^2 s + \frac{1}{\tau})$. The initial condition is now $\tilde{\Pi}(s, t = 0|D_0) = e^{-sD_0}$. The integral $\tilde{T}(s,t) = \int_0^t \tilde{\Pi}(s,t')dt'$ is known from [76]

$$T(s,t|D_0) = \left[\frac{e^{t^*/2}}{\cosh(\omega_s t^*/2) + \frac{1}{\omega_s}\sinh(\omega_s t^*/2)}\right]^\nu \exp\left[-\frac{sD_0\tau}{\omega_s}\frac{2\sinh(\omega_s t^*/2)}{\cosh(\omega_s t^*/2) + \frac{1}{\omega_s}\sinh(\omega_s t^*/2)}\right], \tag{4.48}$$

with $t^* = t/\tau$ and $\omega_s = \sqrt{1 + 4s\sigma^2\tau^2}$. According to Eq. (4.45), one deduces thus the characteristic function as a function of initial diffusivity D_0:

$$\tilde{P}(q,t|D_0) = \left[\frac{e^{t^*/2}}{\cosh(\omega t^*/2) + \frac{1}{\omega}\sinh(\omega t^*/2)}\right]^\nu \exp\left[-\frac{D_0 q^2\tau}{\omega}\frac{2\sinh(\omega t^*/2)}{\cosh(\omega t^*/2) + \frac{1}{\omega}\sinh(\omega t^*/2)}\right], \tag{4.49}$$

with $\omega = \sqrt{1 + 4q^2\sigma^2\tau^2}$. After integration over initial diffusivity the characteristic function yields

$$\tilde{P}(q,t) = \left[\frac{e^{t^*/2}}{\cosh(\omega t^*/2) + \frac{1}{\omega}\sinh(\omega t^*/2)}\right]^\nu \left(1 + \frac{2\sigma^2 q^2\tau^2\sinh(\omega t^*/2)}{\omega\cosh(\omega t^*/2) + (1 - 2\omega\sigma^2 q^2\tau^2)\sinh(\omega t^*/2)}\right)^\nu. \tag{4.50}$$

This is an alternative representation of the characteristic function $\tilde{P}(q,t)$ from Eq. (4.16).

4.3.4 Absorbing boundary condition at $D = 0$

In the case with absorbing boundary condition at $D = 0$, the Fokker-Planck equation represents the evolution with time of the probability density of being at position x with diffusivity D starting at x_0, D_0, without ever having diffusivity $D = 0$. In other words, we look at the propagator of particles which have never stopped. For simplicity we adopt the notation $J_D(D = 0, t) = \sigma^2(\nu - 1)\phi(t)$. This problem is solved using the method of characteristics. The solution of the equation

$$dt = \frac{ds}{G(s)} = \frac{d\tilde{P}}{\sigma^2(\nu - 1)\phi(t) - \frac{s\bar{D}}{\tau}\tilde{P}} \tag{4.51}$$

gives

$$s = \frac{s_1 - s_2 C_1 e^{\omega t/\tau}}{1 - C_1 e^{\omega t/\tau}}, \tag{4.52}$$

with a constant C_1. This expression is used to deduce the homogeneous solution of the second equation in Eq. (4.51):

$$\tilde{P}_h(q,s,t) = \left[C_1 e^{\omega t/\tau}\right]^{-\frac{\nu}{2}\left(1-\frac{1}{\omega}\right)} \left[1 - C_1 e^{\omega t/\tau}\right]^{\nu}. \tag{4.53}$$

Together with the particular solution we find

$$\tilde{P}(q,s,t|x_0,D_0) = \tilde{P}_h(q,s,t) \left[C_2 + \sigma^2(\nu-1) \int_0^t dt' \frac{\phi(t')}{\left[C_1 e^{\omega t'/\tau}\right]^{-\frac{\nu}{2}\left(1-\frac{1}{\omega}\right)} \left[1 - C_1 e^{\omega t'/\tau}\right]^{\nu}} \right]. \tag{4.54}$$

We pose that C_2 is an arbitrary function H of C_1: $C_2 = H(C_1)$. The initial condition determines the function H:

$$H(u) = e^{-iqx_0} \exp\left(-D_0 \frac{s_1 - s_2 u}{1-u}\right) [u]^{\frac{\nu}{2}\left(1-\frac{1}{\omega}\right)} [1-u]^{-\nu}. \tag{4.55}$$

For brevity we make the substitution $\rho = \frac{s-s_1}{s-s_2}$. Inserting Eq. (4.52) and Eq. (4.55) to Eq. (4.54) leads to the solution

$$\begin{aligned} \tilde{P}(q,s,t|x_0,D_0) &= e^{-iqx_0} \exp\left(-D_0 \frac{s_1 - s_2\rho e^{-\omega t/\tau}}{1-\rho e^{-\omega t/\tau}}\right) \left(e^{-\omega t/\tau}\right)^{\frac{\nu}{2}\left(1-\frac{1}{\omega}\right)} \left(\frac{1-\rho e^{-\omega t/\tau}}{1-\rho}\right)^{-\nu} \\ &\quad + \sigma^2(\nu-1) \int_0^t dt' \frac{\phi(t')}{\left(e^{-\omega(t-t')/\tau}\right)^{-\frac{\nu}{2}\left(1-\frac{1}{\omega}\right)}} \left(\frac{1-\rho e^{-\omega(t-t')/\tau}}{1-\rho}\right)^{-\nu}, \end{aligned} \tag{4.56}$$

with the function $\phi(t)$ to be determined. Following Feller [92], we require that the inverse Laplace transform of the propagator exists. At large s, $\tilde{P}$ becomes

$$\begin{aligned} \tilde{P}(q,s\to\infty,t|x_0,D_0) &= s^{-\nu} \left(\frac{1-e^{-\omega t/\tau}}{s_1-s_2}\right)^{-\nu} \left(e^{-\omega t/\tau}\right)^{\frac{\nu}{2}\left(1-\frac{1}{\omega}\right)} \\ &\quad \times \left[e^{-iqx_0} \exp\left(-D_0 \frac{s_1 - s_2\rho e^{-\omega t/\tau}}{1-\rho e^{-\omega t/\tau}}\right) \right. \\ &\quad \left. + \sigma^2(\nu-1) \int_0^t dt' \frac{\phi(t')}{\left(e^{-\omega t'/\tau}\right)^{\frac{\nu}{2}\left(1-\frac{1}{\omega}\right)}} \left(\frac{1-e^{-\omega(t-t')/\tau}}{1-e^{-\omega t/\tau}}\right)^{-\nu} \right]. \end{aligned} \tag{4.57}$$

At this stage one sees that for $\nu \geq 1$, the existence of the inverse Laplace transform is ensured by the $s^{-\nu}$ factor so taking $\phi(t) = 0$ gives the exact solution, and the particle never gets $D = 0$.

When $\nu < 1$, the finite solution at $D = 0$ exists only if the sum cancels in the right hand side, leading to the integral equation (cf. [92]):

$$\int_0^t dt' \frac{\phi(t')}{\left(e^{\omega t'/\tau}\right)^{\frac{\nu}{2}\left(1-\frac{1}{\omega}\right)}} \left(\frac{1-e^{-\omega(t-t')/\tau}}{1-e^{-\omega t/\tau}}\right)^{-\nu} = -\frac{e^{-iqx_0}}{\sigma^2(\nu-1)} \exp\left(-D_0 \frac{s_1 - s_2 e^{-\omega t/\tau}}{1-e^{-\omega t/\tau}}\right). \quad (4.58)$$

Using the variable change $1 - e^{-\Omega t} = z^{-1}$ and $1 - e^{-\Omega t'} = \xi^{-1}$ and defining

$$g(u) = \frac{\tau}{\omega}\phi(u) e^{iqx_0} e^{D_0 s_2} \sigma^2(\nu-1) \frac{1}{\left(1-u^{-1}\right)^{\frac{\nu}{2}\left(1-\frac{1}{\omega}\right)} u(u-1)^{1-\nu}}, \quad (4.59)$$

we get

$$\int_0^\infty d\xi g(\xi)(\xi - z)^{-\nu} = e^{-D_0(s_1-s_2)z}. \quad (4.60)$$

The solution g is an exponential function, from which

$$\phi(z) = \omega \frac{e^{-iqx_0} e^{-D_0 s_2}}{\sigma^2 \tau(\nu-1)} \frac{[D_0(s_1-s_2)]^{1-\nu}}{\Gamma(1-\nu)} e^{-D_0(s_1-s_2)z} \left(1-z^{-1}\right)^{\frac{\nu}{2}\left(1-\frac{1}{\omega}\right)} z(z-1)^{1-\nu}. \quad (4.61)$$

The complete solution for the case $\nu < 1$ now reads

$$\begin{aligned} \tilde{P}(q,s,t,x_0,D_0) &= e^{-iqx_0} \exp\left(-D_0 \frac{s_1 - s_2\rho e^{-\omega t/\tau}}{1-\rho e^{-\omega t/\tau}}\right) \left(e^{-\omega t/\tau}\right)^{\frac{\nu}{2}\left(1-\frac{1}{\omega}\right)} \left(\frac{1-\rho e^{-\omega t/\tau}}{1-\rho}\right)^{-\nu} \quad (4.62) \\ &\times \frac{1}{\Gamma(1-\nu)} \gamma\left(1-\nu, \frac{D_0\omega}{\sigma^2\tau} \frac{(1-\rho)}{\left(1-\rho e^{-\omega t/\tau}\right)} \frac{e^{-\omega t/\tau}}{\left(1-e^{-\omega t/\tau}\right)}\right), \end{aligned}$$

where $\gamma(a,x)$ is the lower incomplete gamma function: $\gamma(a,x) = \int_0^x du\, e^{-u} u^{a-1}$. The probability of $D > 0$ can be obtained by integrating over x and D (setting respectively $q = 0$ and $s = 0$). As a consequence, we retrieve the Feller's formula for the probability [92] of getting $D = 0$ up to time t

$$\pi(D=0,t|D_0) = 1 - \frac{1}{\Gamma(1-\nu)} \gamma\left(1-\nu, \frac{D_0}{\sigma^2\tau} \frac{e^{-t/\tau}}{\left(1-e^{-t/\tau}\right)}\right). \quad (4.63)$$

Once the process reaches the absorbing boundary at $D = 0$, it remains trapped there, so that the probability $\pi(D = 0, t|D_0)$ is a norm decreasing function of time. Figure 4.5 shows the behavior of the probability $D = 0$ as a function of ν. At all times, this probability is zero at $\nu = 1$ and is equal to 1 at $\nu = 0$. As time increases, the diffusivity distribution is getting localized at $D = 0$.

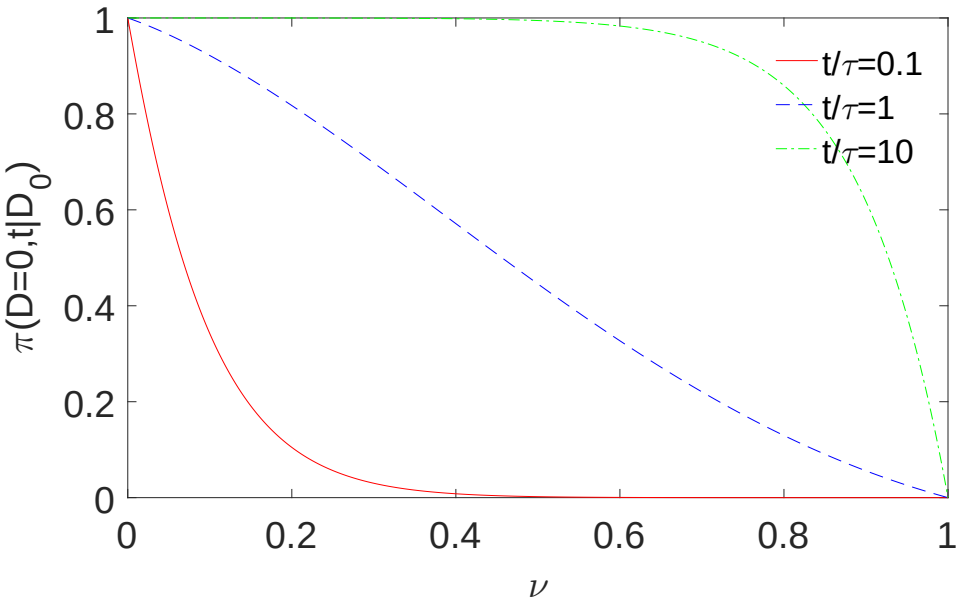


Figure 4.5: Probability of $D = 0$ as a function of ν for three different dimensionless times $t/\tau = \{0.1, 1, 10\}$. ν is varied from 0 to 1 by keeping $\sigma = 1$, $\tau = 10$ and adjusting $\bar{D} = \nu\sigma^2\tau$ with initial diffusivity $D_0 = \bar{D}$.

4.4 Asymptotic behavior

4.4.1 Brownian limit

We first consider the limiting case $\sigma\tau \to 0$, which can either be interpreted as diffusivity behaving deterministically ($\sigma \to 0$) or the mean reversion significantly stronger than fluctuations of diffusivity ($\tau \to 0$). In this limit one recovers $\tilde{P}(q,t) = e^{-q^2\bar{D}t}$, from which the Gaussian propagator for Brownian motion is retrieved:

$$P(x,t|x_0) = \frac{1}{\sqrt{4\pi\bar{D}t}} \exp\left(-\frac{(x-x_0)^2}{4\bar{D}t}\right). \tag{4.64}$$

This distribution also corresponds to the limit $\nu \to \infty$.

4.4.2 Short-time behavior

The superstatistical approach [14, 15] consists in writing the distribution of displacements as a superposition of Gaussian distributions weighted by a stationary distribution of diffusivity. In a

recent work, Chechkin *et al.* [49] showed that non-Gaussian diffusion can be described at short times by superstatistics. Since diffusivity does not evolve much during the correlation time τ, one can imagine an ensemble of particles with independent diffusivities. In our model, we use this relation to establish the short-time behavior.

One can relate our model to the superstatistical approach in the following terms. At short times we have

$$x_t = \int_0^t \sqrt{2D_s} dW_s^{(1)} \approx \sqrt{2D_0} W_t^{(1)}, \tag{4.65}$$

and consider D_0 in the stationary regime. This short-time description looses track of the dynamics. We calculate $P_0(r,t)$, the probability density to be at distance r from the starting point in d dimensions, where the subscript 0 highlights that it is a short-time description:

$$P_0(r,t) = \int_0^\infty dD_0 \, \Pi(D_0) \, \frac{1}{(4\pi D_0 t)^{d/2}} \exp\left(-\frac{r^2}{4D_0 t}\right), \tag{4.66}$$

with the stationary distribution $\Pi(D_0)$ of the CIR model from Eq. (4.15), which gives

$$P_0(r,t) = \frac{2^{1-\nu-d/2}\nu^{d/2}}{\Gamma(\nu)(\pi\bar{D}t)^{d/2}} \left(r\sqrt{\frac{\nu}{\bar{D}t}}\right)^{\nu-d/2} K_{\nu-d/2}\left(r\sqrt{\frac{\nu}{\bar{D}t}}\right), \tag{4.67}$$

where $K_\alpha(x)$ is the modified Bessel function of the second kind. Using the small x expansion of $K_\alpha(x)$, one gets for $\nu > d/2$

$$P_0(r=0,t) = \frac{\Gamma(\nu - d/2)}{\Gamma(\nu)(4\pi t)^{d/2}} \left(\frac{\nu}{\bar{D}}\right)^{d/2}. \tag{4.68}$$

In the case $\nu = 1$ and $d = 1$, the distribution is purely exponential

$$P_0(r,t) = \frac{1}{2\sqrt{\bar{D}t}} \exp\left(-\frac{|r|}{\sqrt{\bar{D}t}}\right), \tag{4.69}$$

(note that in this case the displacement r is distributed over $(-\infty, \infty)$ that explains the extra factor 1/2).

This approach is applicable at short times ($\mu(t) < 1$) but fails at long times because the underlying processes are fundamentally different. One can compare our model to this approach by calculating the non-Gaussian parameter

$$\gamma(t) = \frac{1}{3}\frac{\langle X^4(t)\rangle}{\langle X^2(t)\rangle^2} - 1, \tag{4.70}$$

which is equal to the excess kurtosis divided by 3 (the kurtosis of the Gaussian distribution). By definition, the non-Gaussian parameter is zero for the Gaussian distribution. For superstatistics with $d = 1$ and $x_0 = 0$, the MSD is $\langle x^2(t)\rangle_0 = 2\bar{D}t$ and the fourth moment $\langle x^4(t)\rangle_0 = 12t^2\bar{D}^2\frac{\nu+1}{\nu}$, which leads to the non-Gaussian parameter:

$$\gamma_0(t) = \frac{1}{\nu}. \tag{4.71}$$

In contrast to our model (see Sec. 4.5.1), the distribution of displacements $P_0(r,t)$ spreads at all times but does not change its shape: changing time just rescales space coordinates of the distribution. From this argument it is clear that the only way to reproduce convergence to a Gaussian distribution at long times is to make the stationary distribution $\Pi(D_0)$ of diffusivity time-dependent, which does not make sense. This is a branching point among non-Gaussian models, as constant or vanishing non-Gaussian parameter implies different miscroscopic mechanisms. Note that the distinction between interpretations can also be made using the autocorrelation of diffusivity: it is a Dirac delta distribution $\delta(\tau)$ in a superstatistical approach and an exponentially vanishing function in our model (see Sec. 4.5.2).

4.4.3 Large x behavior

We investigate the asymptotic behavior of the propagator at large x. Let us first consider the particular case $\nu = 1$. As the propagator $P(x,t|x_0)$ is obtained as the inverse Fourier transform of $\tilde{P}(q,t)$, it is instructive to search for the poles of $\tilde{P}(q,t)$ in the complex plane of q in order to compute the inverse Fourier transform by the residue theorem. We write

$$\tilde{P}(q,t) = \frac{\omega\, e^{t^*/2}}{f_+(\omega)\, f_-(\omega)}, \tag{4.72}$$

where

$$f_+(\omega) = \omega\cosh(t^*\omega/4) + \sinh(t^*\omega/4), \tag{4.73}$$

$$f_-(\omega) = \omega\sinh(t^*\omega/4) + \cosh(t^*\omega/4). \tag{4.74}$$

Setting $\omega = i4\alpha/t^*$, we search for α at which these functions vanish, i.e.,

$$f_+(\omega) = i(4\alpha/t^*)\cos(\alpha) + \sin(\alpha) = 0, \tag{4.75}$$

$$f_-(\omega) = -(4\alpha/t^*)\sin(\alpha) + \cos(\alpha) = 0. \tag{4.76}$$

Both equations have infinitely many solutions. The solutions of the first equation lie in the intervals $\bigcup_{k=-\infty}^{\infty}(\pi/2+k\pi,\pi+k\pi)$ (including the trivial solution $\alpha=0$), whereas the solutions of the second equation lie in the intervals $\bigcup_{k=-\infty}^{\infty}(k\pi,\pi/2+k\pi)$. Since $\omega=0$ is not a pole of $\tilde{P}(q,t)$ (as it is compensated by the numerator), we exclude this point. The pole with the smallest absolute value is thus given as the smallest positive solution of the second equation that we rewrite as

$$\alpha_{t^*}\sin\alpha_{t^*}=\frac{t^*}{4}\cos\alpha_{t^*}. \tag{4.77}$$

The smallest positive solution of this equation, α_{t^*}, is a monotonously increasing function of t^*, ranging from 0 at $t^*=0$ to $\pi/2$ at $t^*=\infty$. The corresponding value of ω will determine the asymptotic exponential decay of the propagator.

Since $i4\alpha_{t^*}/t^*=\omega=\sqrt{1+4q^2\sigma^2\tau^2}$, we identify the pole in the q plane:

$$q_0=\pm i\beta_{t^*}\frac{1}{2\sigma\tau}\,,\qquad \beta_{t^*}=\sqrt{1+(4\alpha_{t^*}/t^*)^2}\,. \tag{4.78}$$

Applying the residue theorem, we get

$$P(x,t|x_0)=\int_{-\infty}^{\infty}\frac{dq}{2\pi}e^{iq(x-x_0)}\tilde{P}(q,t)=2\pi i\sum_n\frac{e^{iq_n(x-x_0)}}{2\pi}\mathrm{res}_{q_n}\{\tilde{P}(q,t)\}, \tag{4.79}$$

where the sum runs over the poles. The asymptotic behavior at large $|x-x_0|$ is determined by the pole with the smallest $|q_0|$. We get thus

$$P(x,t|x_0)\propto\exp\left(-\frac{|x-x_0|\beta_{t^*}}{2\sigma\tau}\right)\qquad(|x-x_0|\to\infty), \tag{4.80}$$

This agrees with experimental observations of a distribution of displacements with exponential tails [75, 278, 279].

One can also compute the prefactor by evaluating the residue of $\tilde{P}(q,t)$ at $q=q_0$. Note that for large t^*, one has $\alpha_{t^*}\approx\pi/2$, and thus the dependence on t^* is eliminated, yielding $\beta_{t^*}\simeq 1$ as $t^*\to\infty$. In turn, when t^* is small, one has $\alpha_t^*\simeq\sqrt{t^*}/2$, and thus $\beta_{t^*}\simeq\sqrt{1+4/t^*}\to\infty$. As a consequence, the distribution becomes more and more narrowed, as expected. We emphasize that this analysis is not rigorous enough, as the relation between q and ω involves the square root and thus requires some cuts in the complex plane to avoid multiple branches.

When ν is a strictly positive integer, the above analysis remains applicable. However, the pole is not simple (as for $\nu=1$) but has a degree ν. The degree $\nu>1$ results in a more

complicated computation of the residue and, more importantly, in power law corrections to the exponential decay

$$P(x,t|x_0) \propto |x-x_0|^{\nu-1} \exp\left(-\frac{|x-x_0|\beta_{t^*}}{2\sigma\tau}\right) \qquad (|x-x_0| \to \infty). \tag{4.81}$$

We also emphasize that the current analysis only focuses on the dependence on $|x-x_0|$ and does not capture the complete dependence on t^* which enters through different coefficients.

We expect that the same asymptotic behavior remains valid for any $\nu > 0$ (even non-integer), although its rigorous demonstration requires much finer analysis and is beyond the scope of this chapter. We conclude that the propagator exhibits a universal exponential decay at large increments, whereas the value of ν determines the power law corrections.

4.5 Statistical properties

In this section we describe the statistical properties of our model.

4.5.1 Moments and the non-Gaussian parameter

First we calculate the second and fourth moments using the relation

$$\langle X^k(t)\rangle = (-i)^k \left(\frac{\partial^k}{\partial q^k}\tilde{P}(q,t)\right)\Big|_{q=0}, \tag{4.82}$$

where $\langle . \rangle$ denotes the expectation. The second moment reads

$$\langle X^2(t)\rangle = 2\bar{D}t. \tag{4.83}$$

We observe thus the mean squared displacement grows linearly with time, as in the Brownian case. In Sec. 4.6, an extension to anomalous diffusion through scaling arguments is proposed.

The process described in this chapter possesses many characteristics which are not deducible from the MSD. So we go further and calculate the fourth moment:

$$\langle X^4(t)\rangle = 12\bar{D}^2t^2 + 24\sigma^2\bar{D}\tau^2 t + 24\sigma^2\bar{D}\tau^3\left(e^{-t/\tau}-1\right). \tag{4.84}$$

From the second and fourth moments in Eqs.(4.83,4.84), we calculate the non-Gaussian parameter in Eq. (4.70)

$$\gamma(t) = \frac{2\sigma^2\tau^2}{\bar{D}t}\left(1-\frac{1}{t/\tau}\left(1-e^{-t/\tau}\right)\right). \tag{4.85}$$

As $t \to \infty$, the distribution slowly converges to a Gaussian distribution, as $1/t$. The theoretical formula is verified by simulations (Fig. 4.6). The leading term can be expressed in terms of $\mu(t)$ as $\frac{2\sigma^2\tau^2}{\bar{D}t} = \frac{1}{2}\mu(t)^{-2}$ (see Eq. (4.20)), which shows that non-Gaussianity is related to space exploration, but the complete description also requires to take into account the correction terms from memory effects. Interestingly, we obtained the same form of $\gamma(t)$ as in the Kärger model [93,145] with a finite number of equilibrium states (i.e. diffusivities), due to the averaging over diffusivity disorder (see also Sec. 4.6.1). The same results are evidently valid for the diffusivity modeled as the distance from the origin of an n-dimensional Ornstein-Uhlenbeck process [49,137].

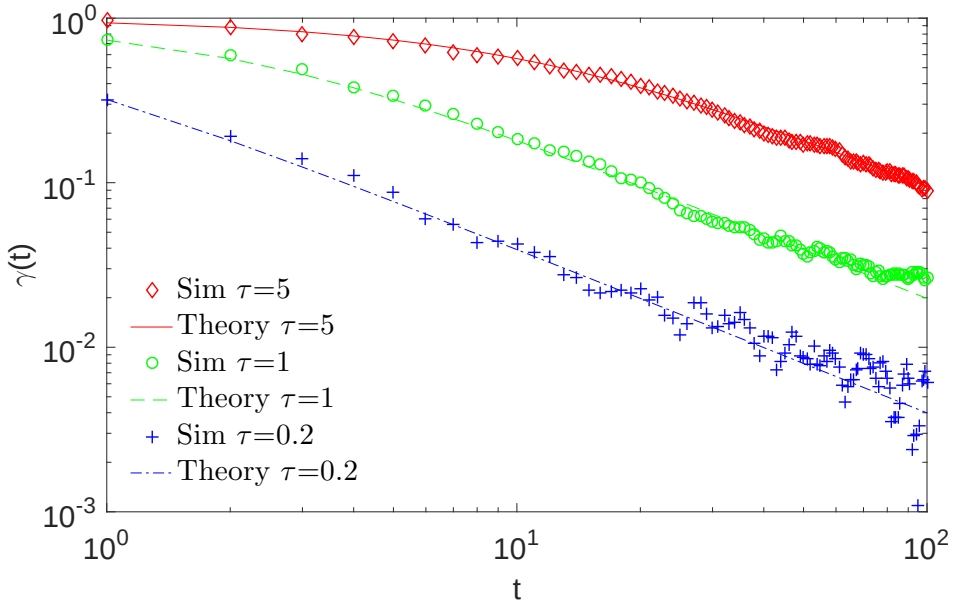


Figure 4.6: The non-Gaussian parameter calculated from Eq. (4.85) (lines) and from Monte Carlo simulations (symbols) with $M = 10^5$ particles, for different values of $\tau = \{0.2, 1, 5\}$, while keeping $\nu = 1$, $\bar{D} = 1$ and $\sigma = 1/\sqrt{\tau}$.

4.5.2 Autocorrelation of squared increments

Diffusing diffusivity models introduce a new level of complexity, far beyond the reach of the mean squared displacement analysis, and new tools are needed to describe such processes. A wide range of models with fluctuating volatility (or diffusivity in physical language) have already

been studied in finance [50, 61, 126]. Since the square of an increment is a local measure of diffusivity, its autocorrelations can reveal information on memory effects of diffusivity. On one hand, it is possible to evaluate the autocorrelation of diffusivity directly from a given trajectory by calculating the autocorrelation of its squared increments. On the other hand, this quantity is accessible theoretically. Let us go through this calculation.

We have $dx_t = \sqrt{2D_t}dW_t^{(1)}$ and the diffusivity in the integral form reads

$$D_t = D_0 e^{-t/\tau} + \bar{D}(1 - e^{-t/\tau}) + e^{-t/\tau} \int_0^t e^{s/\tau} \sqrt{2D_s} dW_s^{(2)}. \tag{4.86}$$

We define the centered squared increments $dx_t^{2*} = dx_t^2 - \langle dx_t^2 \rangle$.

Their autocorrelation is then

$$\begin{aligned} \langle dx_t^{2*} dx_{t+\Delta}^{2*} \rangle &= \langle dx_t^2 dx_{t+\Delta}^2 \rangle - \langle dx_t^2 \rangle \langle dx_{t\Delta}^2 \rangle \qquad (4.87) \\ &= 4\langle D_t D_{t+\Delta} \rangle \left\langle \left(dW_t^{(1)} dW_{t+\Delta}^{(1)} \right)^2 \right\rangle - 4\langle D_t \rangle \langle D_{t+\Delta} \rangle \left\langle \left(dW_t^{(1)} \right)^2 \right\rangle \left\langle \left(dW_{t+\Delta}^{(1)} \right)^2 \right\rangle. \end{aligned}$$

For $\Delta = 0$, we calculate

$$\langle (dx_t^{2*})^2 \rangle = 12\langle D_t^2 \rangle - 4\langle D_t \rangle^2, \tag{4.88}$$

which is obtained directly from Eq. (4.86):

$$\langle (dx_t^{2*})^2 \rangle = 12 \left[\sigma^2 \bar{D}\tau \left(1 - e^{-t/\tau} \right)^2 + 2\sigma^2 \tau D_0 \left(e^{-t/\tau} - e^{-2t/\tau} \right) \right] + 8 \left(D_0 e^{-t/\tau} + \bar{D} \left(1 - e^{-t/\tau} \right) \right)^2. \tag{4.89}$$

In the case $\Delta > 0$, as $dW_t^{(1)}$ is independent from $dW_{t+\Delta}^{(1)}$, one has

$$\left\langle \left(dW_t^{(1)} dW_{t+\Delta}^{(1)} \right)^2 \right\rangle = \left\langle \left(dW_t^{(1)} \right)^2 \right\rangle \left\langle \left(dW_{t+\Delta}^{(1)} \right)^2 \right\rangle \tag{4.90}$$

which leads to

$$\langle dx_t^{2*} dx_{t+\Delta}^{2*} \rangle = 4\langle D_t D_{t+\Delta} \rangle - 4\langle D_t \rangle \langle D_{t+\Delta} \rangle. \tag{4.91}$$

The autocorrelation of squared increments is explicitly related to the autocorrelation of diffusivity as

$$\langle dx_t^{2*} dx_{t+\Delta}^{2*} \rangle = 4e^{-(2t+\Delta)/\tau} \int_0^t \int_0^{t+\Delta} e^{(s_1+s_2)/\tau} \langle \sqrt{D_{s_1} D_{s_2}} \rangle \langle dW_{s_1}^{(2)} dW_{s_2}^{(2)} \rangle, \tag{4.92}$$

from which

$$\langle dx_t^{2*} dx_{t+\Delta}^{2*} \rangle = 4e^{-\Delta/\tau} \left[\sigma^2 \bar{D}\tau \left(1 - e^{-t/\tau} \right)^2 + 2\sigma^2 \tau D_0 \left(e^{-t/\tau} - e^{-2t/\tau} \right) \right]. \tag{4.93}$$

To summarize, in our model we find

$$\langle dx_t^{2*} dx_{t+\Delta}^{2*} \rangle = \begin{cases} 12\sigma^2\tau \left(1 - e^{-t/\tau}\right)^2 \left[\bar{D} + 2D_0 \frac{e^{-t/\tau}}{1-e^{-t/\tau}}\right] + 8\left((D_0 - \bar{D})e^{-t/\tau} + \bar{D}\right)^2 & (\Delta = 0), \\ 4e^{-\Delta/\tau}\left[\sigma^2\bar{D}\tau\left(1 - e^{-t/\tau}\right)^2 + 2\sigma^2\tau D_0\left(e^{-t/\tau} - e^{-2t/\tau}\right)\right] & (\Delta > 0). \end{cases} \tag{4.94}$$

One notes the exponentially vanishing dependence on initial conditions. In the long-time limit $t \to \infty$, one simply gets

$$\lim_{t\to\infty} \langle dx_t^{2*} dx_{t+\Delta}^{2*} \rangle = \begin{cases} 12\sigma^2\bar{D}\tau + 8\bar{D}^2 & (\Delta = 0), \\ 4\sigma^2\bar{D}\tau e^{-\Delta/\tau} & (\Delta > 0). \end{cases} \tag{4.95}$$

The mean-reverting property of the Feller process results in the exponential autocorrelation of diffusivity. If an experimentally measured autocorrelation of squared increments is not exponentially vanishing, the mean reverting property cannot be described by a simple harmonic potential centered on $\bar{D}$, and thus another model (or an extension of the present model) should be considered.

4.5.3 Ergodicity and finite sample effects

Data analysis is usually performed with time-averaged quantities because of small data samples. Then a natural question of equivalence between time and ensemble averages arises: "Is a time-averaged quantity from one particle representative of other particles from the same system?". For a system at thermodynamical equilibrium, the time average over an infinitely long trajectory matches the ensemble average over an infinite number of particles, this statement is known as the ergodicity hypothesis. This hypothesis is not satisfied in aging random media [32].

From the Langevin equation (4.2), one can directly see that our model is ergodic: as the diffusivity is fluctuating around its average, fluctuations will be averaged out in the limit of infinitely long trajectories. But for a finite duration of experiment, what can be said about ergodicity of the system?

If the experiment duration t_{exp} is shorter than the time to explore heterogeneities of the system, $t_{exp} < t_{sys}$, different tracers probe regions with different diffusivities. As a consequence, on such a timescale, tracers would appear as experiencing different dynamics, so that one could wrongly conclude that the dynamics of the system is nonergodic. Inversely, if the experiment is sufficiently long (i.e. $t_{exp} \gg t_{sys}$), tracers have enough time to visit every region of the system,

and one concludes correctly that the ergodicity hypothesis is fulfilled. The experiment duration plays therefore an important role and should be chosen accurately.

To illustrate this point we study two quantities characterizing ergodicity by different strategies. We show that depending on the parameters of the model, the results of the tests can sound contradictory. First we use the Ergodicity Breaking parameter $EB(\Delta, t_{exp})$ [43, 122, 196] which quantifies the dispersion of the time-averaged MSD $\bar{\delta^2}(\Delta, t_{exp})$ [114, 115, 246] with

$$\bar{\delta^2}(\Delta, t_{exp}) = \frac{1}{t_{exp} - \Delta} \sum_{n=1}^{t_{exp}-\Delta} (x_{n+\Delta} - x_n)^2 \tag{4.96}$$

as a function of the experiment duration t_{exp} (i.e. the trajectory length) evaluated with a time-lag Δ:

$$EB(\Delta, t_{exp}) = \frac{\langle (\bar{\delta^2}(\Delta, t_{exp}))^2 \rangle}{\langle \bar{\delta^2}(\Delta, t_{exp}) \rangle^2} - 1. \tag{4.97}$$

For an ergodic process, $\lim_{t_{exp}\to\infty} EB(\Delta, t_{exp}) = 0$ for any Δ, meaning that for a fixed Δ, the distribution of TAMSD converges to a Dirac delta distribution with $\bar{\delta^2}(\Delta, t_{exp} \to \infty) = \langle X^2(\Delta) \rangle$.

Figure 4.7 shows that fluctuations of TAMSD are impacted by two characteristics: the shape parameter ν and the correlation time τ. The smaller the parameter ν, the longer it takes for the EB parameter to vanish. Indeed, for $\nu \leq 1$, diffusivity can be small with high probability that would slow down the dynamics. The correlation time τ also influences the convergence of $EB(\Delta, t_{exp})$: larger τ implies longer time to recover from small diffusivities and thus slower dynamics. Setting $\nu = 1$ and varying τ, the EB parameter has a transient behavior until $\approx 2\tau$ and decays as a power law $1/t_{exp}$ as in the Brownian case for which the exact formula, in the discrete case, is $EB(\Delta, t_{exp}) = \frac{(2\Delta+1/\Delta)}{3(t_{exp}-\Delta+1)}$ [220]. Note that a slow decrease of the EB parameter due to disorder was also discussed for fluctuating diffusivity [53] and diffusion in a periodic potential [152].

We also discuss the ergodicity test based on the dynamical functional developed in Chapter 2. In Fig. 4.8, the estimator $\tilde{F}_\omega(\Delta, t_{exp})$ decays fast so that the temporal disorder due to diffusivity does not affect much this quantity, in contrast to the EB parameter. If this estimator vanishes for a single particle trajectory, one can expect asymptotic independence and ergodicity. This implies that getting longer data indeed increases the accuracy of time averaged quantities (smaller $EB(\Delta, t_{exp})$).

The ergodicity breaking parameter shows that the distribution of TAMSD slowly converges to a delta distribution. In turn, the ergodicity estimator $\tilde{F}_\omega(\Delta, t_{exp})$ indicates that the process

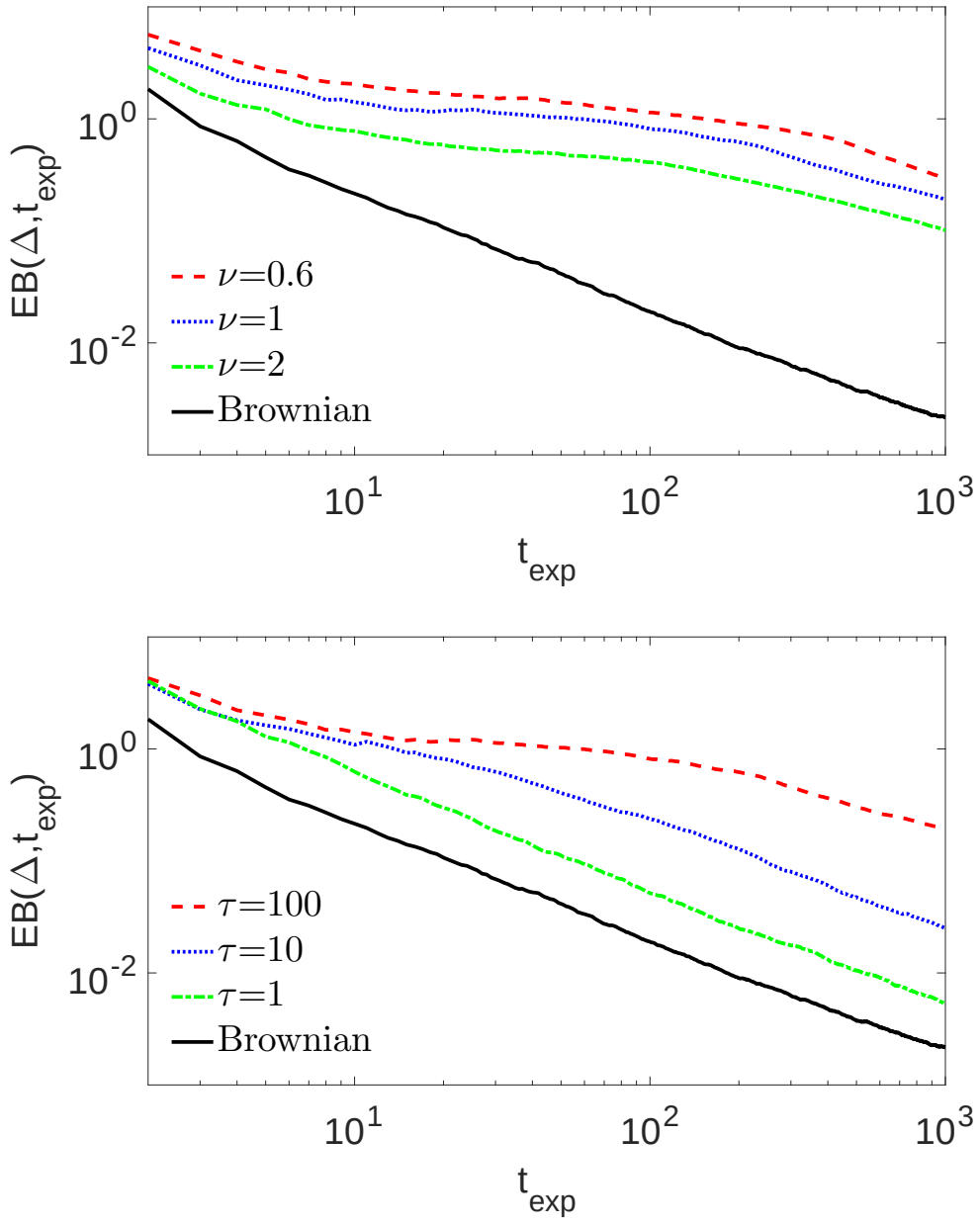


Figure 4.7: Ergodicity breaking parameter calculated by averaging over $M = 10^3$ simulated trajectory of length $t_{exp} = 10^3$. The result for Brownian motion (solid line) is also plotted for comparison. *Top.* Results for variable shape parameter $\nu = \{0.6, 1, 2\}$ by varying τ, with $\bar{D} = 1$ and $\sigma = 1$ being constant.*Bottom.* Results for variable correlation time $\tau = \{1, 10, 100\}$ while keeping $\nu = 1$, $\bar{D} = 1$ and $\sigma = 1/\sqrt{\tau}$.

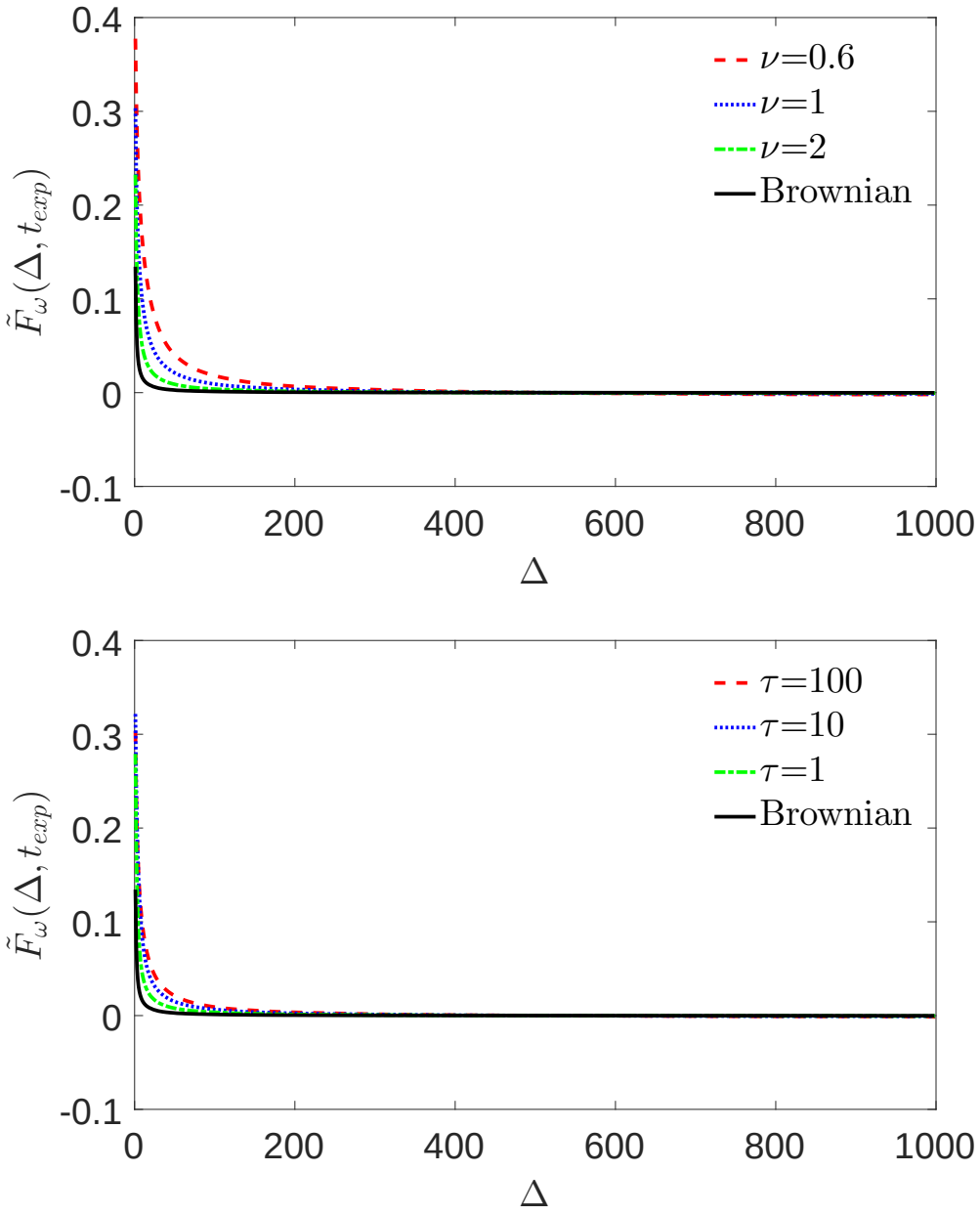


Figure 4.8: Mean ergodicity estimator $\tilde{F}_\omega(\Delta, t_{exp})$, calculated with $M = 10^3$ simulated trajectories of length $t_{exp} = 10^3$. The mean estimator for Brownian motion (solid line) is also plotted for comparison. *Top.* Different values of the shape parameter $\nu = \{0.6, 1, 2\}$ by varying τ, with $\bar{D} = 1$ and $\sigma = 1/\sqrt{\tau}$. *Bottom.* Different values of $\tau = \{1, 10, 100\}$ while keeping $\nu = 1$, $\bar{D} = 1$ and $\sigma = 1/\sqrt{\tau}$.

looses its memory and implies that the TAMSD distribution narrows with increasing t_{exp} (without specifying how). These two quantities do not answer the ergodicity question in the same way, they are complementary. If one needs to know the degree of dispersion of TAMSD, the EB parameter has to be used. The estimator $\tilde{F}_\omega(\Delta, t_{exp})$, which can be applied to a single trajectory, does not quantifies fluctuations, but allows to verify ergodicity, even in the presence of dynamic disorder because it relies on the estimation of the characteristic function of the process.

4.6 Discussion

4.6.1 Fourth moment is not enough

The analytical expression of the non-Gaussian parameter from Kärger model (see Chapter 1) has the same functional form as $\gamma(t)$ from Eq. (4.85):

$$\gamma_{KM}(t) = \frac{2\eta}{t/\tau}\left(1 - \frac{1}{t/\tau}\left(1 - e^{-t/\tau}\right)\right), \tag{4.98}$$

with the coefficient $\eta = \frac{p_1 p_2 (D_1 - D_2)^2}{(p_1 D_1 + p_2 D_2)^2}$, which corresponds in our case to $\frac{\sigma^2 \tau}{D}$, and τ is the exchange time: $\tau = 1/K_{12} = 1/K_{21}$.

Figure 4.9 compares distributions for the Kärger model and our approach. In the case $\nu > 1$, both distributions are very close at all times. In the case $\nu \leq 1$, obtained here by setting different relative volumes p_1 and p_2, the Kärger model does not reproduce the peak at 0. In other words, the Kärger model as a superposition of only two Gaussian distributions does not match our model with infinitely many Gaussian distributions (see Sec. 4.4.2).

We conclude that these two distributions having identical second and fourth moments, are still different. While it was known that mean squared displacement is not sufficient to characterize a model, here we illustrate that even the fourth moment (and the non-Gaussian parameter) is not enough.

4.6.2 Anomalous diffusion

In biology there are many experimental evidences of anomalous diffusion [13, 129, 181, 233, 281] when the mean squared displacement scales as a power law with time $\langle X^2(t)\rangle = 2D_\alpha t^\alpha$, where D_α is the generalized diffusion coefficient and α is the anomalous exponent. We propose an

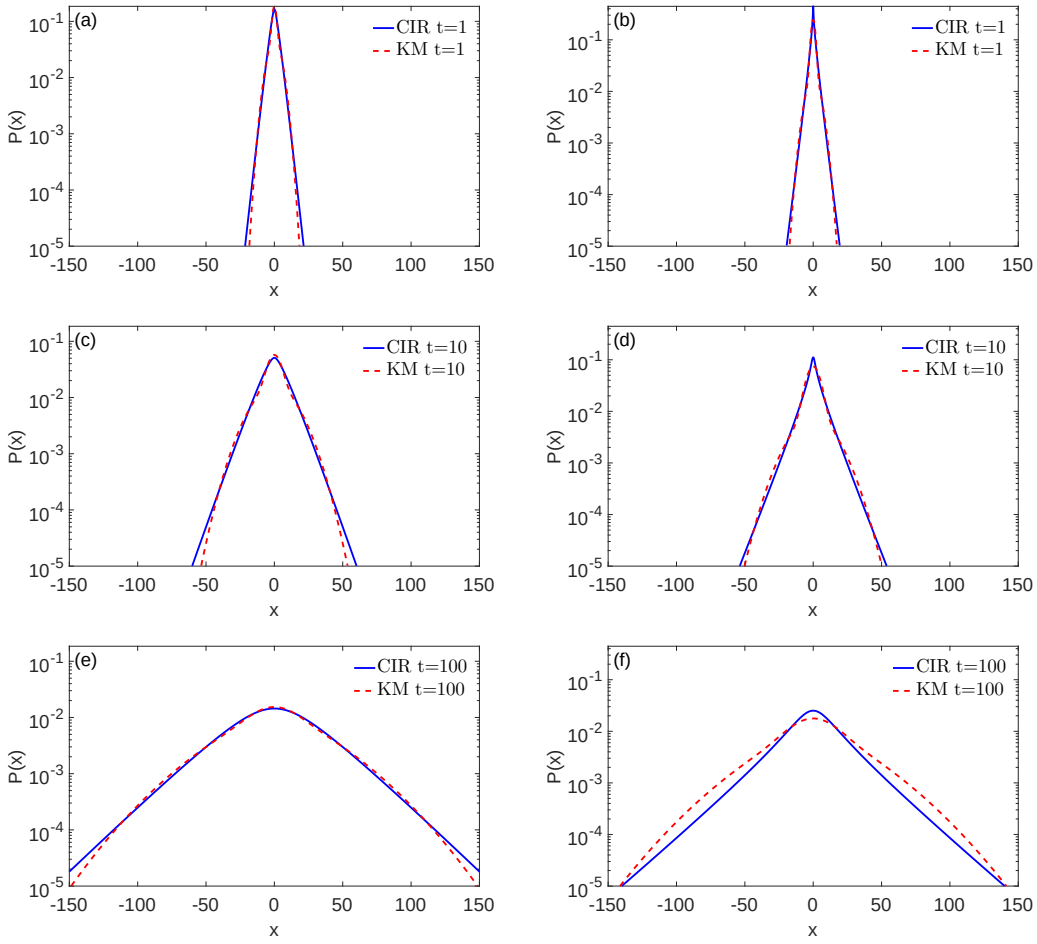


Figure 4.9: Distribution of displacements for the Kärger model (dashed lines) and our model (solid lines) at time $t = 1$ (top), $t = 10$ (middle), $t = 100$ (bottom). We choose the parameters for the Kärger model and deduce $\bar{D} = p_1 D_1 + p_2 D_2$ and $\sigma = \sqrt{\bar{D}\eta/\tau}$. *Left column.* Parameters of the Kärger model are $p_1 = 1/2$, $p_2 = 1/2$, $D_1 = 1$ and $D_2 = 10$ and $\tau = 10$, leading to $\nu \approx 1.5$. *Right column.* Parameters of the Kärger model are $p_1 = 4/5$, $p_2 = 1/5$, $D_1 = 1$ and $D_2 = 10$ and $\tau = 10$, leading to $\nu \approx 0.6$.

extension by a simple scaling of the time $t/\tau \Rightarrow (t/\tau)^\alpha$, so that $\tilde{P}(q,t)$ from Eq. (4.17) is replaced by

$$\tilde{P}(q,t) = \left(e^{-\frac{1}{2}(\omega-1)(t/\tau)^\alpha} \frac{4\omega}{(\omega+1)^2} \left(1 - \left(\frac{\omega-1}{\omega+1} \right)^2 e^{-\omega(t/\tau)^\alpha} \right)^{-1} \right)^\nu . \tag{4.99}$$

The non-Gaussian parameter would now depend on α:

$$\gamma(t) = \frac{2\sigma^2\tau^{1+\alpha}}{\bar{D}t^{\alpha}}\left(1 - \frac{1}{(t/\tau)^{\alpha}}\left(1 - e^{-(t/\tau)^{\alpha}}\right)\right). \tag{4.100}$$

As expected, in the subdiffusive case $\alpha < 1$, the convergence to a Gaussian distribution is slower as compared to the superdiffusive case $\alpha > 1$, because larger α means faster exploration of space.

4.7 Conclusion

We presented a model of non-Gaussian diffusion, based on coupled Langevin equations. We derived the explicit exact formula for the distribution of displacements in the Fourier-Laplace domain and studied different asymptotic regimes. We showed that this distribution exhibits exponential tails and converges slowly, as $1/t$, to a Gaussian one. Depending on the shape parameter ν, the distribution can be flat ($\nu > 1$) or peaked ($\nu \leq 1$) at zero. The MSD evolves linearly with time in spite of non-Brownian character of the motion. We pointed that the ergodicity estimator $\tilde{F}_{\omega}(\Delta, t_{exp})$ catches ergodic nature of the process while the random nature of diffusivity makes fluctuations of TAMSD to span up at long times as demonstrated by the ergodicity breaking parameter $EB(\Delta, t_{exp})$ making the TAMSD a bad estimator of the average diffusion coeffficient $\bar{D}$. We used the autocorrelation of squared increments to determine the autocovariance structure of diffusivity. Given that small diffusivities are made much more probable in the case $0 < \nu < 1$ (which was not accessible in former models), the underlying process exhibits a richer phenomenology. We expect that this model will help to understand more deeply dynamical heterogeneities observed in experiments. An important perspective is to relate the correlation structure of the stochastic diffusivity D_t with spatial correlations structure of the medium [272]. One can also analyze the first passage time (FPT) statistics in our model of heterogeneous diffusion to reveal the impact of the diffusing diffusivity. Although the mean squared displacement grows linearly with time, the distribution of FPT can be sensitive to the related annealed disorder (e.g., see [125] for models of quenched disorder).

Chapter 5

Application to experiments

5.1 Introduction

In this chapter we present the analysis of two experiments realized in collaboration with two experimental groups. In both experiments, individual random trajectories of particles are recorded. Using single trajectory analysis, developed in this thesis, we propose a physical interpretation of the mechanism underlying the tracer's motion.

The first experiment (Sec. 5.2) concerns a system of vertically vibrated disks undergoing random motion. It has been designed by O. Dauchot and carried out by G. Briand at ESPCI in Paris, France. The analysis has been performed by Y. Lanoiselée and D. S. Grebenkov. The results of this experiment are decribed in [161] (under revision).

The second experiment (Sec. 5.3) explores the effect of cytoskeleton components on the random motion of an inert tracer inside the cytoplasm. The experiment has been designed by D. Heinrich (Leiden University, Netherlands) and T. Franosch (Innsbruck University, Austria) and carried out in D. Heinrich's group by M. Götz (Leiden University, Netherlands). The statistical analysis has been jointly performed by Y. Lanoiselée, P. Witzel (Leiden University, Netherlands) and D. S. Grebenkov. The text of Sec. 5.3 was based on and adapted from the joint publication (submitted).

5.2 Macroscopic realization of Brownian motion

Experimental observations of Brownian motion in the macroscopic world are rarer. In fact, it is quite difficult to design an experiment with macroscopic objects that would result in Brownian trajectories. On one hand, the motion is strongly influenced by inertial effects, resulting in ballistic segments of the trajectory at the macroscopic scale (e.g., the motion of balls in a billiard). On the other hand, the number of interacting objects in a macroscopic system is much smaller than the number of water molecules involved in the motion of a microscopic particle, whereas the separation between the time scale of an elementary displacement and the duration of the measurement is not large enough. As a consequence, the motion of macroscopic objects is not enough randomized by their collisions. In particular, the dynamics of granular matter is typically far from Brownian motion [51, 54, 113, 134, 241]. For instance, there is a rather narrow range of packing fractions, for which the motion of spherical beads is fluid-like: in the low density regime, collisions between beads are rare while the mean free path is long so that too large experimental setups would be needed to observe a Brownian trajectory; in the high density regime, inter-bead collisions are often but collective modes of motion (e.g., crystallization or jamming) become dominant.

From the practical point of view, a well-controlled experimental realization of a macroscopic diffusive motion with an excellent statistics of long trajectories can serve as a benchmark for testing various statistical tools developed for the analysis of single particle trajectories (see [7, 24, 97, 150, 159, 160, 175, 193, 197, 198, 220, 273] and references therein). In fact, it is essential to disentangle finite time average and finite sampling effects when performing single probe experiments in biology (e.g., the intracellular transport or the motion of proteins on cell membranes). While statistical tools are commonly tested on simulated trajectories, a macroscopic realization of diffusive motions can present a rare opportunity to confront simulations and theoretical results to an experimental situation with true experimental noise, uncertainties, resolution issues, etc.

In this first part of the chapter, we report an experimental observation of the diffusive motion realized by macroscopic disks of 4 mm diameter on a vertically vibrating plate (see Sec. 5.2.1). Vibrations pump in the system the kinetic energy that substitutes thermal energy that drives the motion in a microscopic system. We undertake a systematic statistical analysis of the acquired

trajectories of individual disks (Sec. 5.2.2). In particular, we analyze the distribution of one-step displacements, the ergodicity, the velocity auto-correlation function, and the distribution of time averaged mean square displacements (TAMSD). This analysis shows that the macroscopic motion of disks exhibits small deviations from Brownian motion at short times but approaches it at longer times.

5.2.1 Experimental setup

The experimental system, made of vibrated disks, has been described in details previously [68]. We recall here the key ingredients of the set-up. Experiments with shaken granular particles are notoriously susceptible to systematic deviations from pure vertical vibration and special care must be taken to avoid them. First, to ensure the rigidity of the tray supporting the particles, we use a 110 mm thick truncated cone of expanded polystyrene sandwiched between two nylon disks. The top disk (diameter 425 mm) is covered by a glass plate on which lay the particles. The bottom one (diameter 100 mm) is mounted on the slider of a stiff square air-bearing (C40-03100-100254, IBSPE), which provides virtually friction-free vertical motion and sub-micron amplitude residual horizontal motion. The vertical alignment is controlled by set screws. The vibration is produced with an electromagnetic servo-controlled shaker (V455/6-PA1000L,LDS), the accelerometer for the control being fixed at the bottom of the top vibrating disk, embedded in the expanded polystyrene. A 400 mm long brass rod couples the air-bearing slider and the shaker. It is flexible enough to compensate for the alignment mismatch, but stiff enough to ensure mechanical coupling. The shaker rests on a thick wooden plate ballasted with 460 kg of lead bricks and isolated from the ground by rubber mats (MUSTshock 100x100xEP5, Musthane). We have measured the mechanical response of the whole setup and found no resonances in the window $70 - 130$ Hz. We use a sinusoidal vibration of frequency $f = 95$ Hz and set the relative acceleration to gravity $\Gamma = a(2\pi f)^2/g = 2.4$, where the vibration amplitude a at a peak acceleration is 100 μm. Using a triaxial accelerometer (356B18, PCB Electronics), we checked that the horizontal to vertical ratio is lower than 10^{-2} and that the spatial homogeneity of the vibration is better than 1%.

The particles are micro-machined copper-beryllium disks (diameter $d = 4 \pm 0.03$ mm). The contact with the vibrating plate is that of an extruded cylinder, resulting in a total height $h = 2.0$ mm. They are sandwiched between two thick glass plates separated by a gap $H = 2.4$ mm and

laterally confined in an arena of diameter 320 mm. A CCD camera with a spatial resolution of 1728 x 1728 pixels and standard tracking software is used to capture the motion of the particles at a frame rate of 25 Hz. In a typical experiment, the motion of the disks is recorded during 600 seconds, producing 15 000 images. The resolution on the position $\vec{r}$ of the particles is better than 0.05 particle diameter (i.e., 0.2 mm).

In the following, particle trajectories are tracked within a circular region of interest (ROI) of diameter $50d = 200$ mm, far from the border of the arena, where the long-time averaged density field is homogeneous. The average packing fractions ϕ measured inside the ROI ranges from 0.3 to 0.64, and the total number of particles ranges from 1000 to 2500. As the onset of spatial order typically takes place at $\phi_\dagger \simeq 0.71$, we always deal with a liquid state.

5.2.2 Statistical analysis

We performed a systematic statistical analysis of the acquired random trajectories. Examples of such trajectories are shown in Fig. 5.1.

5.2.3 Data description

We analyzed 14 samples with different surface packing fractions ϕ, ranging between 0.298 and 0.637 (Table 5.1). The time step (i.e., the duration of one displacement) is fixed by the acquisition frequency: $\delta = 1/25$ Hz $= 0.04$ s. The positions are measured in units of the disk diameter, $d = 4$ mm. To avoid boundary effects, only the disks within the ROI were used for the analysis. In particular, a trajectory is terminated when the disk leaves the ROI, and a new trajectory is initiated when a disk enters the ROI. As a consequence, the acquired trajectories have very different lengths varying from 1 to 15 000. To improve the statistical accuracy of our results, we discarded all the trajectories whose length was shorter than 1000. The disks exhibited multiple mutual collisions during the experiments. Although the collective motion of these disks might be studied as the dynamics of interacting particles in a large phase space, we look at this problem from the single-particle point of view and treat each disk as a single particle interacting with its complex dynamic environment. This view is typical for single-particle tracking experiments in microbiology when one can record only the motion of a labeled (e.g., fluorescent) particle, whereas the dynamics of all other constitutes of the cytoplasm remains inaccessible.

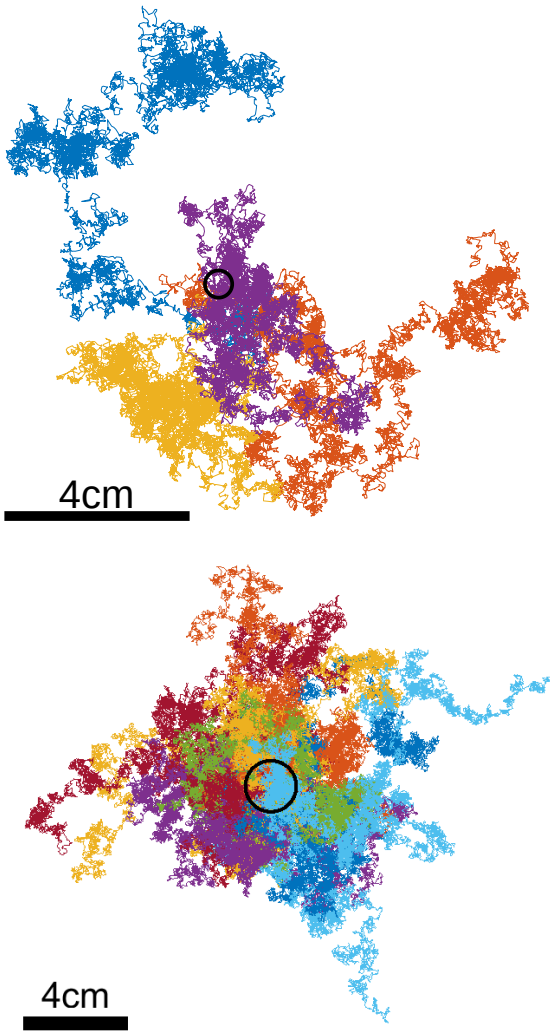


Figure 5.1: **(Top)** Example of random trajectories of 4 disks that were originally located close to each other (inside the black circle) and then diffused during 600 seconds (i.e., 15 000 points in each trajectory). **(Bottom)** Example of random trajectories of 34 disks that were originally located close to each other and then diffused during 464.6 seconds (i.e., 11 615 points in each trajectory).

5.2.4 Distribution of increments

We start by verifying whether the one-step increments obey a Gaussian distribution. For each sample, we collected the one-step increments along X and Y axes for each trajectory in the sample and constructed their histogram. Having checked for the isotropy of the statistics, we

sample	ϕ	std/d	D (in mm^2/s)
1	0.298	0.0956	1.83
2	0.324	0.0938	1.76
3	0.350	0.0932	1.74
4	0.376	0.1048	2.20
5	0.402	0.1025	2.10
6	0.428	0.1107	2.45
7	0.454	0.1052	2.21
8	0.480	0.1005	2.02
9	0.507	0.0978	1.91
10	0.533	0.0950	1.81
11	0.559	0.0926	1.71
12	0.585	0.0942	1.78
13	0.611	0.0936	1.75
14	0.637	0.0895	1.60

Table 5.1: Summary of experimental data: the sample index, the surface fraction ϕ, the standard deviation of one-step one-dimensional increments (in units of the disk diameter $d =$ 4 mm), and the corresponding diffusion coefficient: $D = \text{std}^2/(2\delta)$, with $\delta = 0.04$ s. For comparison the maximal disk packing fraction, corresponding to the close-packed hexagonal lattice, is $\pi/(2\sqrt{3}) \simeq 0.9069$; and the crystallization transition for equilibrium hard disks takes place at $\phi_\dagger \simeq 0.71$.

focus on one-dimensional increments and merge increments along X and Y coordinates in order to get a representative statistics even for large increments. Figure 5.2(a) shows these histograms (presented in the form of probability densities at the semilogarithmic scale) for 14 samples. These densities are close to each other and exhibit a parabolic shape reminiscent of a Gaussian distribution. The standard deviations of one-step increments are summarized in Table 5.1. These values are also close to each other and show no systematic dependence on the packing fraction. At first sight, there is no systematic variation of probability densities with the packing fraction. This suggests that the randomness of motion essentially comes from the rotational

symmetry of the disk, which undergoes a displacement in a random direction after each kick by the vibrating plate. Note that the frequency of plate vibrations is 4 times higher than the acquisition frequency meaning that each displacement results from 4 random kicks.

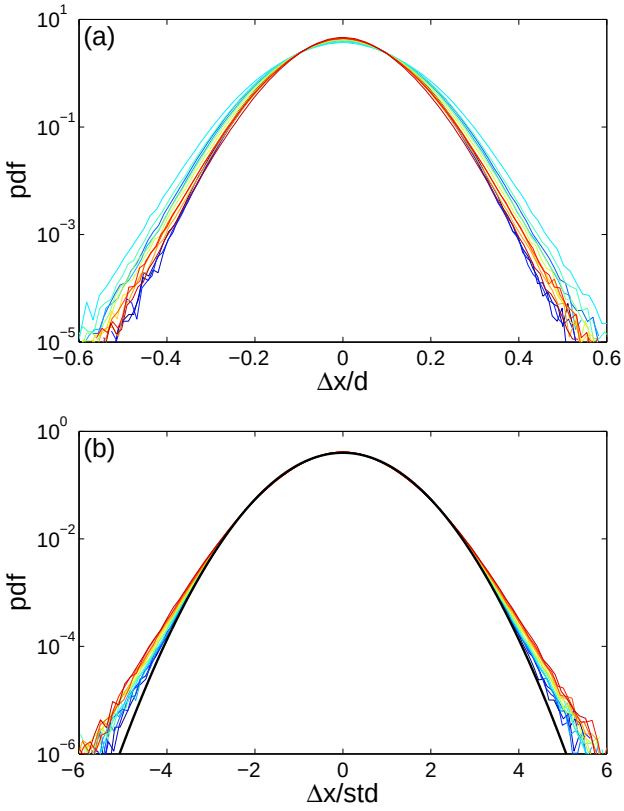


Figure 5.2: **(a)** The empirical distributions of one-step one-dimensional increments for 14 samples (in units of the disk diameter, $d = 4$ mm). **(b)** The empirical distributions of rescaled one-step increments for 14 samples. Thick black curve shows the standard Gaussian density $e^{-x^2/2}/\sqrt{2\pi}$. Color of thin curves changes from dark blue for the lowest packing fraction ϕ to dark red for the highest one.

Despite their delicate machining, the precise contact of the disks with the vibrating plate is influenced by minor asperities, which differ from disk to disk but also depend on the location of the disks on the vibrating plate. In order to reduce these factors of diversity, we rescale the one-step increments from one trajectory by the empirical standard deviation of these increments. Such a rescaling partly levels off eventual heterogeneities between trajectories. Once calculated,

the rescaled increments along X and Y coordinates are merged from different trajectories in each sample. The obtained distributions are presented in Fig. 5.2(b). One can see that the distributions for all 14 samples almost collapse and remain close to the standard Gaussian density $\exp(-x^2/2)/\sqrt{2\pi}$. However, now that heterogeneities between trajectories have been levelled off by the rescaling, one distinguishes small but statistically significant deviations for large increments. These deviations progressively increase with the packing fraction, and can therefore be attributed to disk-disk collisions.

5.2.5 Ergodicity hypothesis

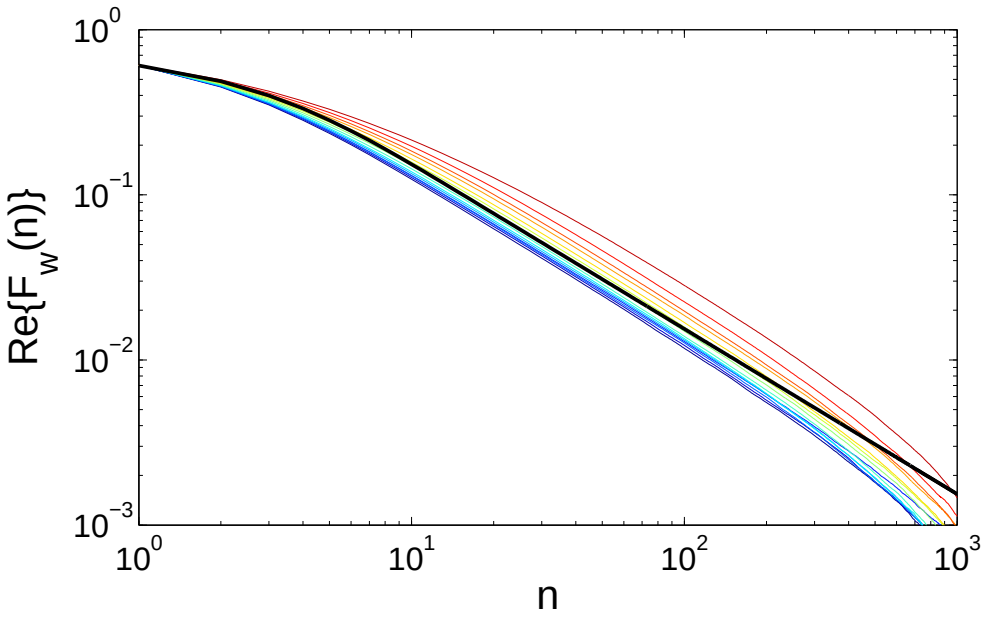


Figure 5.3: The real part of the ergodicity estimator, $\mathrm{Re}\{\hat{F}_\omega(n)\}$ (with $\omega = 1/\sigma$), averaged over trajectories in each of 14 samples (thin lines). Color of thin curves changes from dark blue for the lowest packing fraction ϕ to dark red for the highest one. Thick black line shows the mean value of this estimator for Brownian motion.

We analyze whether the system of vibrated disks can be considered as being at equilibrium. In practice, we test the ergodicity hypothesis which is a necessary but not sufficient condition for equilibrium. The ergodicity hypothesis claims that the ensemble average over many particles is equal to the time average over an (infinitely) long trajectory of one particle. Under the

stationarity hypothesis of the motion, we employ the ergodicity estimator $\hat{F}_\omega(n)$ from Chapter 2 [159, 175].

Figure 5.3 shows the real part of the ergodicity estimator averaged over all the trajectories in each of 14 samples. For small n, higher the packing fraction, slower the decrease of the estimator with n. However, for large n the $1/n$ scaling predicted in the case of the Brownian motion is recovered and we can safely formulate the hypothesis that ergodicity is satisfied.

5.2.6 Velocity auto-correlations

We also study the velocity auto-correlations function (VACF) which is defined as

$$C(t) = \langle \boldsymbol{v}(t) \cdot \boldsymbol{v}(0) \rangle \,, \tag{5.1}$$

where $\boldsymbol{v}(t)$ is the velocity at time t, and $\langle \cdots \rangle$ is the ensemble average. In the experimental setting, the positions are recorded with the time step $\delta = 0.04$ s, so that $t = n\delta$, and the velocity is proportional to the one-step increment: $\boldsymbol{v}(n\delta) = (\boldsymbol{x}_{n+1} - \boldsymbol{x}_n)/\delta$, with $\boldsymbol{x}_n = \boldsymbol{x}(n\delta)$. To improve statistics, we combine the time average along the trajectory of each disk and the ensemble average over many trajectories:

$$C(n\delta) = \frac{1}{M\delta^2} \sum_{m=1}^{M} \frac{1}{N_m - n - 1} \sum_{k=1}^{N_m - n - 1} \left(\Delta \boldsymbol{x}_{n+k}^{(m)} \cdot \Delta \boldsymbol{x}_k^{(m)} \right), \tag{5.2}$$

where $\Delta \boldsymbol{x}_n^{(m)} = \boldsymbol{x}_{n+1}^{(m)} - \boldsymbol{x}_n^{(m)}$ is the n-th one-step increment of the m-th disk, M is the number of disks in a sample, and N_m is the length of the m-th trajectory.

Figure 5.4 shows the normalized VACF, $C(n\delta)/C(0)$, which varies between -1 and 1, as a function of the lag time $n\delta$. For all considered samples, the VACF rapidly decreases with time and becomes close to zero for $n \geq 10$. By construction, the normalized VACF is equal to 1 at $n = 0$. Positive auto-correlations at lag time $n = 1$ can potentially be attributed to inertial effects. The negative auto-correlations observed for $n > 1$ take their root in an excess of reverse bouncing of the disks when they successively hit the trail, but not only. Since they become more pronounced when the packing fraction increases, they should also come from inter-disk collisions. In all cases, although the successive increments exhibit small but noticeable correlations, they drop very rapidly as the lag time increases. We recall that the normalized VACF for a discrete-time Brownian motion (a random walk) is 1 for $n = 0$ and 0 otherwise. Strictly speaking, the

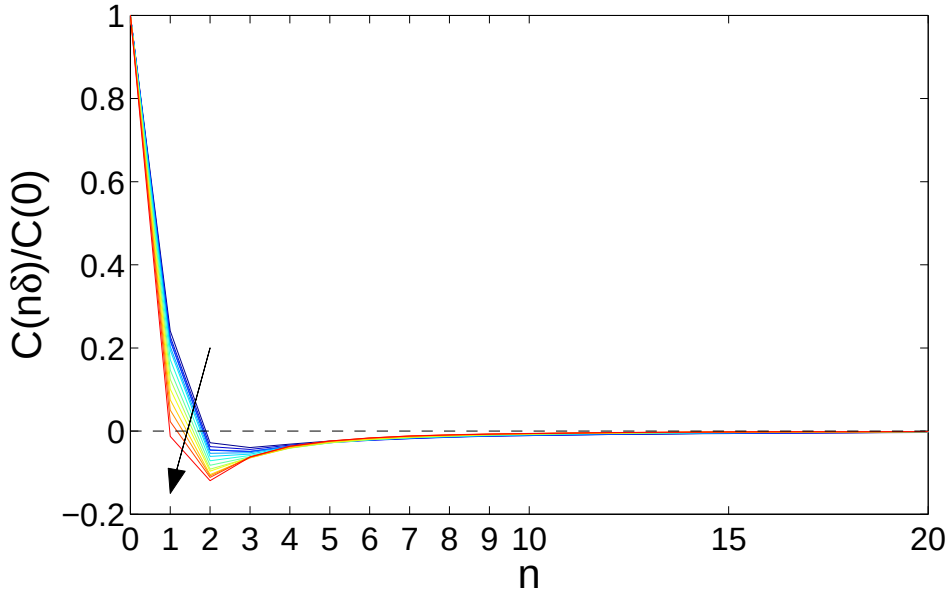


Figure 5.4: The normalized VACF, $C(n\delta)/C(0)$, as a function of the lag time $n\delta$, for 14 samples. Color changes from dark blue for the lowest packing fraction ϕ to dark red for the highest one. This change is also indicated by an arrow.

disk trajectories acquired at time step $\delta = 0.04$ s are therefore not Brownian but remain close to Brownian ones.

5.2.7 Estimation of diffusion coefficient

Now we focus on the time averaged mean square displacement, which is the most common statistical tool to probe diffusive properties of single-particle trajectories [97]. The TAMSD with the lag time n over a trajectory of length N is defined as

$$\chi_{n,N} = \frac{1}{N-n} \sum_{k=1}^{N-n} \|\boldsymbol{x}_{k+n} - \boldsymbol{x}_k\|^2. \tag{5.3}$$

If $\boldsymbol{x}_k$ are positions of planar Brownian motion with diffusion coefficient D, then the ergodicity of this process implies that

$$\chi_{n,N} \xrightarrow[N\to\infty]{} 4Dt = 4Dn\delta, \tag{5.4}$$

whereas the variance of $\chi_{n,N}$ vanishes as $N \to \infty$ [114, 220]. In other words, the TAMSD allows one to estimate the diffusion coefficient D from a *single* random trajectory, and longer

the trajectory, better the estimation.

For a fixed N, the smallest variance (and thus the best estimation) corresponds to $n = 1$, in which case $\chi_{1,N}$ is the estimator of the variance of increments. This estimator is known to be optimal for the case of Brownian motion, i.e., it is the best possible way to estimate the diffusion coefficient [62,114,116]. In practice, however, even if the studied particle is supposed to undergo Brownian motion, the acquired trajectory can be altered by various "measurement noises" such as localization error, electronic noise, drift or vibrations of the sample, post-processing errors, etc. When some of these noises are anticipated, the estimator can be adapted to provide the (nearly) optimal estimation [24, 116, 197, 198, 273]. However, the Brownian character of the studied but yet unknown process is not granted and has to be checked from the analysis of the TAMSD. In this situation, the rule of thumb consists in plotting the TAMSD versus the lag time n to first check the linear dependence and then to estimate the diffusion coefficient from the slope of the linear plot. Given the randomness of the TAMSD, this procedure can bring biases and additional statistical errors. Moreover, since fluctuations of the TAMSD grow with n (see [114,220]), the fit is often limited to small n. Figure 5.5 illustrates large fluctuations of the TAMSD estimator around the ensemble averaged TAMSD which exhibits a linear growth with n. As a consequence, an accurate estimation of the diffusion coefficient from a *single* trajectory is only possible over a narrow range of small lag times n. Note that the diffusion coefficient fitted by the ensemble average, 1.47 mm^2/s, is smaller than that estimated from the standard deviation of one-step increments, 1.83 mm^2/s (see Table 5.1). This discrepancy can be caused by eventual noises (that would affect the standard deviation of one-step increments) and auto-correlations (that would affect the TAMSD).

5.2.8 Distribution of TAMSD

One of the significant advantages of single-particle tracking is the possibility to infer information from single events, without ensemble averages. This is particularly important in microbiology because many events in a cell life are triggered by a small number of molecules. Even when many particles are tracked simultaneously, they explore different spatial regions of the cell and experience different interactions with the intracellular environment. If inferred properly, such heterogeneities may bring a much more detailed information about the cell than an ensemble average. The estimation of the diffusion coefficient from each single trajectory naturally leads

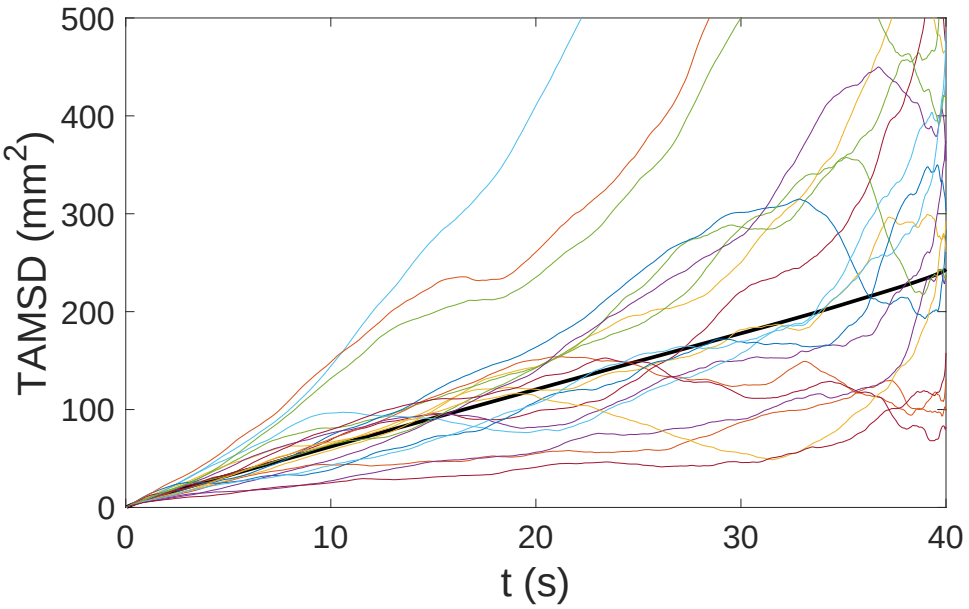


Figure 5.5: The time averaged MSD, $\chi_{n,N}$, as a function of the lag time $t = n\delta$ (with n varying from 1 to 1000), for 10 trajectories of length $N = 1001$ from the sample 1 with the lowest packing fraction. Black thick line shows the ensemble average of TAMSD over 12444 trajectories of length $N = 1001$ in this sample. The fitted diffusion coefficient from this line is 1.47 mm^2/s.

to their distribution [77, 203, 210]. However, it is important to stress that the experimentally obtained distributions include two sources of randomness: (i) the biological variability and (ii) the intrinsic randomness of the TAMSD estimator obtained from a single finite length trajectory. As a consequence, a proper biological interpretation of such distributions requires to disentangle two sources and, ideally, to remove the second one. This correction needs the knowledge of the distribution of the TAMSD estimator.

The distribution of TAMSD in the biological context was first studied via numerical simulations by Saxton [236,240]. A more general theoretical analysis of TAMSD for Gaussian processes was later performed in Refs. [6, 114, 116, 246]. We compute the distribution numerically via the inverse Fourier transform of the characteristic function of TAMSD for which the exact matrix formula was provided in Ref. [114]. This computation was shown to be fast and very accurate.

The theoretical distribution of TAMSD for Brownian motion can be compared to the empiri-

cal distribution of TAMSD obtained from the trajectories of disks. On one hand, this comparison allows one to check to which extent the acquired trajectories are close to Brownian motion. On the other hand, one can investigate in a well-controlled way the applicability of the theoretical distribution to experimental data.

Figure 5.6(a) shows the empirical distribution of TAMSD with the lag time $n = 1$ obtained by splitting each trajectory into fragments of length $N = 100$. This artificial splitting is performed to be closer to the common situation in biological applications, when the acquired trajectories are rather short. Moreover, such splitting significantly improves the statistics of the TAMSD. We compare the probability density functions of TAMSD among 14 samples and with the theoretical curves for Brownian motion. One can see notable deviations from the theoretical distribution, indicating that the acquired trajectories are not Brownian, in agreement with the analysis of Sec. 5.2.6. The two plausible reasons for the observed deviations are: (i) auto-correlations of increments at small lag times (as seen in Fig. 5.4), and (ii) small deviations from the Gaussian distribution of increments (as seen in Fig. 5.2). To check for the first reason, we plot in Fig. 5.6(b,c) the distributions of the TAMSD with larger lag times $n = 10$ and $n = 20$, at which the VACF was negligible. One gets thus a much better agreement with the theoretical distribution.

In order to check the second reason of deviations (weak non-Gaussianity), the increments of all trajectories in each sample were randomly reshuffled to fully destroy auto-correlations, and then new artificial trajectories were constructed from these increments. If the original increments were correlated Gaussian variables with the same variance, such a procedure would yield independent identically distributed Gaussian variables so that the resulting trajectories would represent Brownian motion. In this case, a perfect agreement between empirical and theoretical curves would be expected. Figure 5.6(d) shows empirical and theoretical distributions of TAMSD at the lag time $n = 1$ for reshuffled samples. The agreement is not perfect but is much better than in Fig. 5.6(a). Small residual deviations can potentially be attributed to weak non-Gaussianity of the distribution of increments.

5.2.9 Mean versus the most probable TAMSD

The non symmetric shape of the distribution of TAMSD implies that the mean value of the TAMSD is different from its mode, i.e., the most probable value or, equivalently, the position of the maximum of the PDF. This difference becomes particularly important for the analysis

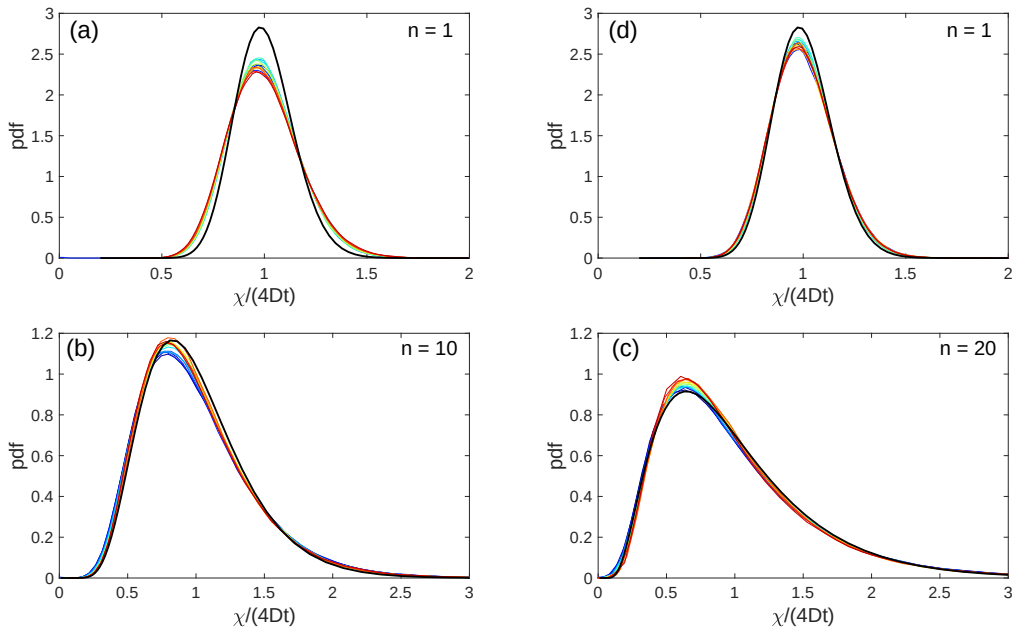


Figure 5.6: **(a,b,c)** Probability density functions of the rescaled TAMSD, $\chi_{n,N}/(4Dt)$, with $N = 100$, $t = n\delta$, and $n = 1$ (a), $n = 10$ (b), and $n = 20$ (c), for 14 samples (thin lines), and the theoretical ones for Brownian motion (thick black line). Color of thin curves changes from dark blue for the lowest packing fraction ϕ to dark red for the highest one. **(d)** Probability density functions of the rescaled TAMSD with $N = 100$, $t = n\delta$, and $n = 1$, for 14 reshuffled samples.

of single particle trajectories. When the sample of such individual trajectories is large, the empirical mean of TAMSD estimated from these trajectories is close to the expectation. In turn, when the TAMSD is estimated from few trajectories (or even from a single trajectory), it is more probable to observe a random realization near the maximum of the PDF. This issue, which was not relevant for symmetric distribution (e.g., a Gaussian distribution), may become an important bias in the analysis of TAMSD.

As discussed in Ref. [114], the distribution of TAMSD for Brownian motion is wider and more skewed for larger n/N. Moreover, the difference between the mean and the mode also grows with n/N. As suggested in [114], the distribution of TAMSD for Brownian motion and some other centered Gaussian processes (like fractional Brownian motion) can be accurately

approximated by a generalized Gamma distribution, which has a simple explicit PDF

$$p(z) = \frac{z^{\nu-1}\exp(-a/z - z/b)}{2(ab)^{\nu/2}K_\nu(2\sqrt{a/b})} \quad (z > 0), \tag{5.5}$$

with three parameters: $a \geq 0$, $b > 0$, and $\nu \in \mathbb{R}$. The moments of this distribution can be expressed as

$$\langle[\chi_{n,N}]^k\rangle = (ab)^{k/2}\frac{K_{\nu+k}(2\sqrt{a/b})}{K_\nu(2\sqrt{a/b})} \quad (k = 1, 2, 3, \ldots), \tag{5.6}$$

whereas the mode is

$$\chi_{n,N}^{\text{mode}} = \frac{\sqrt{(1-\nu)^2b^2 + 4ab} - (1-\nu)b}{2}\,. \tag{5.7}$$

For a given empirical distribution of TAMSD, the first three moments, evaluated directly from the data, can be used to calculate the parameters a, b and ν by solving numerically the system of three equations in Eqs. (5.6) for $k = 1, 2, 3$. In other words, one does not need to fit the empirical distribution in order to get this approximation.

Figure 5.7 shows the pdf of the TAMSD for the trajectories with the lowest packing fraction, with the sample length $N = 100$ and three lag times, $n = 1$, $n = 10$ and $n = 20$ (shown by symbols). From these empirical data, we evaluated the first three moments and calculated the parameters a, b and ν of the generalized gamma distribution (shown by lines). The excellent agreement validates the use of this theoretical approximation even for experimental trajectories.

5.3 Passive diffusion in the cytoplasm

In this section, we report an universal, but non-Gaussian motion type in the cellular interior, which is independent of the presence of the cytoskeleton components. We alter the cytoskeleton in a controlled manner and analyze the statistical properties of nano-tracer step-width distributions revealing non-Gaussian features for four distinct cytoskeleton states, which include and exclude the actin cortex and the microtubule network. The central result is a single universal scaling of these four increment probability densities with exponential tails, independent of cytoskeleton components present in the cell. This indicates an ubiquitous dynamic feature of intracellular transport in the cytoplasm. We are able to assign the observed anomalous transport features to spatio-temporal heterogeneities within the cytoplasmic matter and we find that the cytoskeleton components are responsible for regulating the efficiency of intracellular motion without changing the unique transport feature.

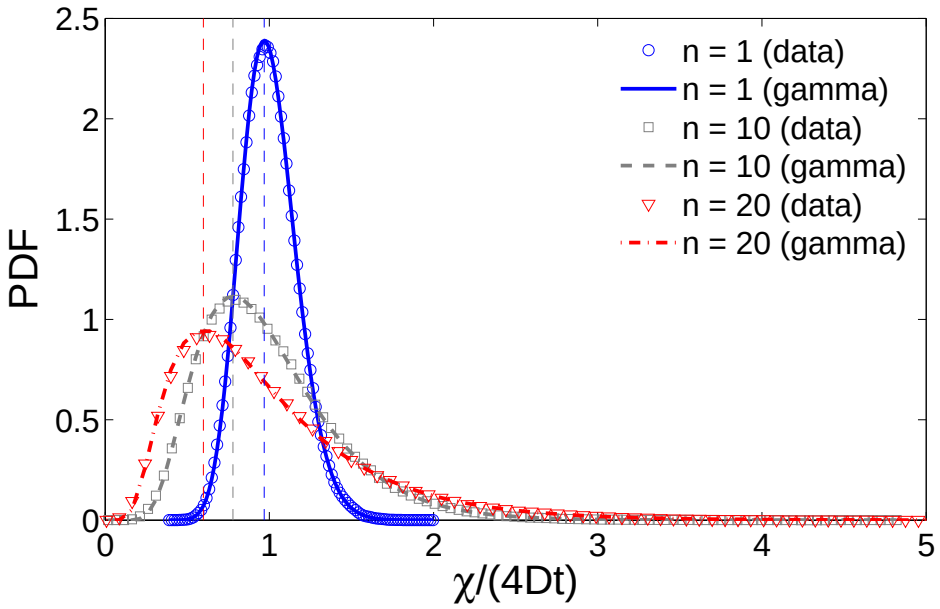


Figure 5.7: Probability density functions of the rescaled TAMSD, $\chi_{n,N}/(4Dt)$, for the sample with the lowest packing fraction, with $N = 100$, $t = n\delta$, and $n = 1$ (blue circle), $n = 10$ (gray squares), and $n = 20$ (red triangles). Lines show the generalized gamma distribution $p(z)$ from Eq. (5.5) whose parameters were obtained from the first three moments (see the text): $a = 3.27$, $b = 0.03$, $\nu = 28.13$ ($n = 1$), $a = 1.94$, $b = 0.27$, $\nu = 1.34$ ($n = 10$), and $a = 1.03$, $b = 0.58$, $\nu = 0.31$ ($n = 20$). The modes of these distributions are shown by vertical dashed lines: 0.97 ($n = 1$), 0.78 ($n = 10$), and 0.60 ($n = 20$), whereas the mean is fixed to be 1 by rescaling.

5.3.1 Presentation of the experiment

The purpose of the second experiment is to investigate the influence of cytoskeleton components on passive, cytoplasm driven, intracellular transport of particles with a diameter of 150 nm. In this respect, *Dictyostelium discoideum* cells have been prepared in four different cytoskeleton states (Fig. 5.10a): untreated wild type (WT), depolymerized microtubules (noMT), depolymerized actin cortex (noAct), and without cytoskeleton by combination of both depolymerizations (noCyt) as *D. discoideum* cells exhibit no stress fibers or intermediate filaments and are capable of surviving both cytoskeleton depolymerizations. For all experiments cell-mediated, bio-motor regulated transport (e.g. along microtubules), has been effectively excluded by using tracer par-

ticles that cannot bind to cellular components and thus reflect only interior cell dynamics. These particles that are not enclosed by vesicles have been injected directly into the cytosol and thus reflect the passive dynamics of the cytoplasmic space exclusively. As obtained by single-particle tracking, we analyzed the probability densities $P(r,t)$ of absolute magnitudes of one-dimensional increments r for different lag times Δ, based on more than $320,000$ data points of nano-particle trajectories (insets of Fig. 5.10 b-e).

5.3.1.1 Data description

Single-particle tracking was performed in Dictyostelium discoideum cells prepared in four different cytoskeleton states: untreated wild type (WT), depolymerized actin cortex (noAct), depolymerized microtubules (noMT), and combination of both depolymerisations (noCyt). An overview of the number of data points for each state is presented in Table 5.2.

Cytoskeleton state	Total	WT	noAct	noMT	NoCyt
Number of cells	68	16	18	18	16
Number of trajectories	149	33	54	38	24
Number of data points	327,444	76,435	80,533	88,302	82,174

Table 5.2: Additional information for the data sets of intracellular tracer trajectories for all four different cytoskeleton states: untreated wild type (WT), depolymerized actin cortex (noAct), depolymerized microtubules (noMT), and combination of both depolymerizations (noCyt).
The data acquisition has been performed on a microscope with spatial resolution of 54nm per pixel. The tracers, of diameter 150nm, within D. discoideum cells were imaged every $\delta = 49$ms adjusting the focal plane manually. The z coordinate is not recorded so the trajectories are two-dimensional.

All statistical analyses are applied to two-dimensional single-particle trajectories, i.e. the successive positions $\{x_m(n\delta), y_m(n\delta)\}$ $(n = 1, , N_m)$ of the m-th tracer acquired at equal time intervals of duration δ (here N_m is the number of points of the $m-$th trajectory).

5.3.2 Ergodicity

Before further analysis, we demonstrate the ergodic properties of the trajectories, justifying the use of time averaged quantities in the rest of the study. First, we use the ergodicity estimator $\hat{F}_\omega(\Delta, t_{exp})$ developed in Chapter 2, measuring the independence of increments at lag-time Δ. Figure 5.8 demonstrates that $\hat{F}_\omega(\Delta, t_{exp})$ applied to each trajectory vanish individually, as $1/\Delta$, similarly to Brownian motion. We conclude that the ergodicity hypothesis is not rejected. Consequently, we assume that time averaged quantities converge to their ensemble averages as the length of the trajectory increases, i.e. they are representative measures of physical quantities associated to the trajectories.

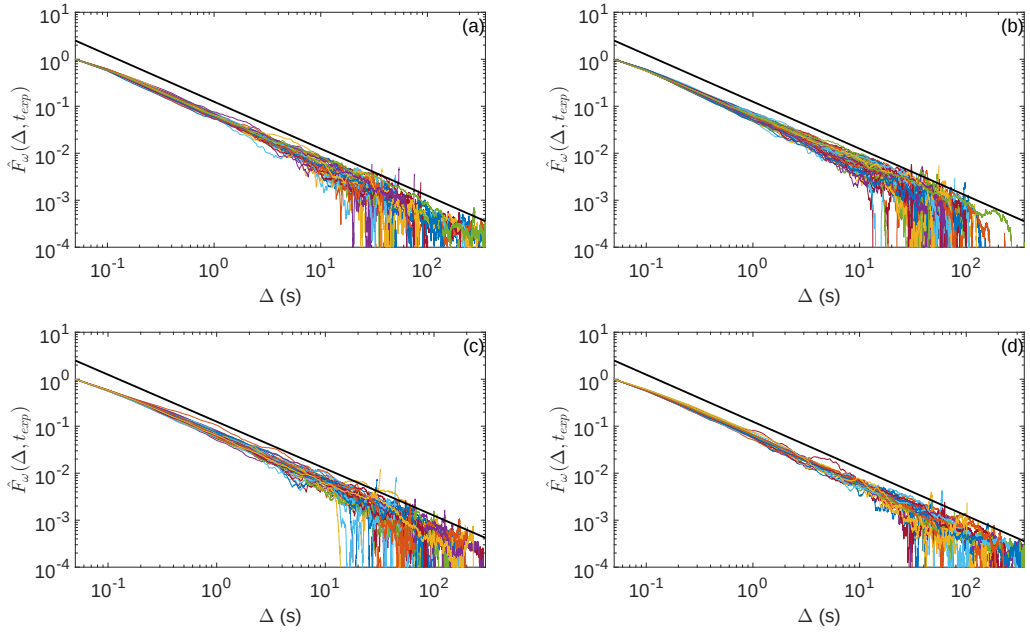


Figure 5.8: The ergodicity estimator $\hat{F}_\omega(\Delta, t_{exp})$ applied to each trajectory of duration t_{exp} as function of lag time $\Delta(s)$. The results are displayed for each cytoskeleton state: wild type **(a)**, depolymerized actin **(b)**, disrupted microtubules **(c)**, and deprived of both actin and microtubules **(d)**. Black lines indicates the $\propto 1/\Delta$ decay of Brownian motion.

Then we inquire into understanding how increasing the length of trajectories reduces of the variance of the TA MSD. For this purpose we proceed by measuring the decay of the ergodicity breaking parameter (EB)(See Chapter 1). Figure 5.9 shows that the fluctuations of variance

decay as $\propto 1/t_{exp}$ in each cytoskeleton state. In all cases, there is a transition near $t_{exp} = 2s$ slowing down the decay to zero, due to a change in the factor in front of $1/t_{exp}$.

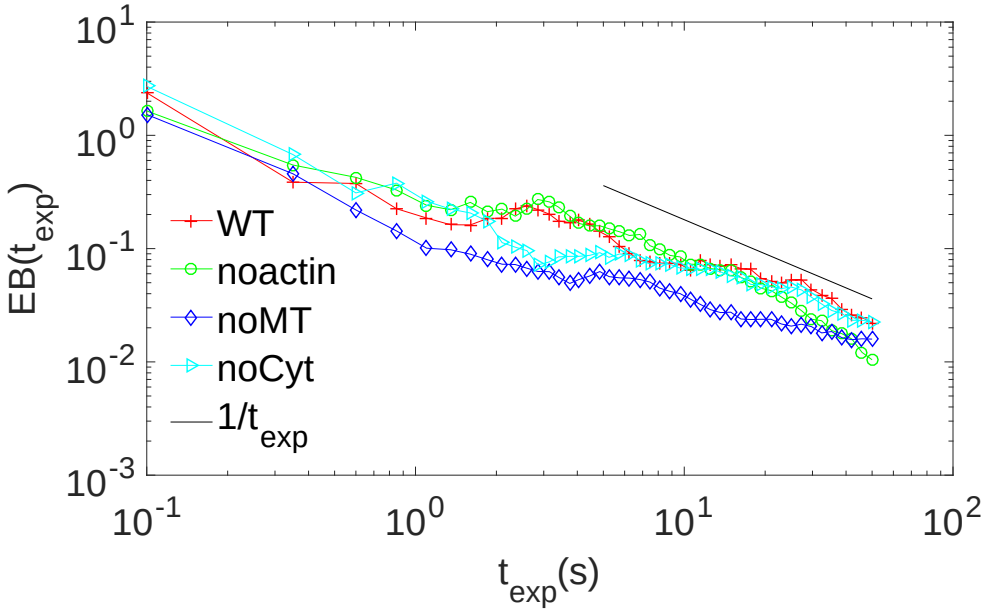


Figure 5.9: Ergodicity breaking parameter EB as a function of the trajectory length t_{exp}. The test is applied to each cytoskeleton state: Wild Type **(red crossed lines)**, Depolymerized actin **(green circles)**, Disrupted microtubules **(dark blue squares)**, and deprived of both actin and microtubules **(cyan triangles)**. Black line denotes the $\propto 1/t_{exp}$ decay for Brownian motion.

5.3.3 Probability distribution

In this section we study the probability density function of absolute increments.

5.3.3.1 Departure from Gaussianity

First, for each trajectory in all states, we test the hypothesis that the distribution of increments is Gaussian. It is not straightforward to verify this hypothesis due to a limited length of individual trajectories that prohibits an accurate computation of the probability density of increments for each single trajectory. For this reason, we resort to statistical tests of Gaussianity which can be applied to individual trajectories. Among various statistical tests, we chose the Anderson-Darling and the Shapiro-Wilk tests that were shown to be the most efficient for testing

Gaussianity on small samples [207,243]. For each cytoskeleton state, we provide in Tabla 5.3 the fraction of the trajectories, for which the hypothesis of a Gaussian distribution of increments is rejected (with the p-value below 5%). Qualitatively speaking, higher fractions mean higher number of a non-Gaussian distribution. Both tests suggest that distributions of increments of an individual trajectory cannot be considered as Gaussian because the percentage of trajectories where the hypothesis is rejected is significantly larger than the 5% (the threshold that could be obtained by chance from a Gaussian distribution).

% Rejected (p-value< 5%)	Wild Type	No Microtubules	No Actin	No Cytoskeleton
Anderson-Darling test	79%	50%	72.2%	44.9%
Shapiro-Wilk test	100%	60.6%	37.5%	66.3%

Table 5.3: The % of trajectories where the hypothesis H_0 that the increments are Gaussian is rejected according to the Anderson-Darling and the Shapiro-Wilk tests of Gaussianity applied to the increments of each individual trajectory.

The different fractions in two tests originate from the distinct features of an empirical distribution that are examined by the tests: the Shapiro-Wilk test [243] is based on a measure of discrepancy between theoretical and sample based variance, whereas the Anderson-Darling test relies on a measure of a distance between the distribution of increments obtained from the data and the theoretical one.

In summary, individual trajectories present non-Gaussian features testifying for the complex dynamics underlying the tracer's motion in the medium. Aware of the non-Gaussian character of our trajectories, we analyze more extensively the probability density and the van Hove self-correlation function used to that purpose.

5.3.3.2 The van Hove self-correlation function

In stochastic theory, the probability distribution function of increments, also known as the van Hove self-correlation function [144] and the averaged propagator, characterizes the increments $\vec{r}$ of a stochastic process $\vec{r}(t)$,

$$P_s(\vec{r},t) = \langle \delta(\vec{r}(t) - \vec{r}(0) - \vec{r}) \rangle \tag{5.8}$$

where $\langle \ldots \rangle$ denotes the expectation, i.e, the average over all trajectories of the process until time t, and $\delta(\ldots)$ is the Dirac distribution. This definition is closely related to the propagator $P(\vec{r}, t|\vec{r}_0, 0)$ that describes the probability density of finding the particle at position $\vec{r}$ at time t if it was started from $\vec{r}_0$ at time $t = 0$:

$$P_s(\vec{r}, t) = \int_0^\infty d\vec{r}_0 \rho(\vec{r}_0) P(\vec{r} + \vec{r}_0, t|\vec{r}_0, 0) \tag{5.9}$$

where $\rho(\vec{r}_0)$ is the initial density of particles. For diffusion in a homogeneous medium $P(\vec{r} + \vec{r}_0, t|\vec{r}_0, 0) = P(\vec{r}, t|0, 0) = P_s(\vec{r}, t)$ independently of $\rho(\vec{r}_0)$. For instance, this is the case for simple Brownian motion, which has the Gaussian propagator and

$$P_s(\vec{r}) = (4\pi D t)^{-d/2} \exp(-|\vec{r}|^2/(4Dt)) \tag{5.10}$$

where d is the space dimension. In turn, the cytoplasm exhibits spatial heterogeneities so that the propagator and the probability density function of increments are in general not equivalent. The tracers exploring different regions inside the cell (e.g. near the nucleus and near the plasma membrane) may have distinct dynamics. In that case $P_s(\vec{r}, t)$ represents an average over those realizations. In an experiment, the stochastic process is not known, and one aims at inferring its properties from the measured trajectories. In this case, the expectation is replaced by an ensemble average over M independent trajectories $\vec{r}_1(t), \vec{r}_2(t), \ldots, \vec{r}_M(t)$:

$$P_s(\vec{r}, t) = \frac{1}{M} \sum_{m=1}^{M} \delta(\vec{r}_m(t) - \vec{r}_m(0) - \vec{r}) \tag{5.11}$$

where the index m enumerates different trajectories. In practice, one discretizes the space of increments $\vec{r}$ by small boxes and counts how many times the observed increment $\vec{r}_m(t) - \vec{r}_m(0)$ falls into each box (see Sec. 5.3.3.3). However, a statistically accurate computation of such empirical two or three-dimensional probability density would require a very large number of increments that is still challenging for experimental in vivo data. For this reason, for an isotropic process, it is convenient to analyze the radial van Hove function to improve the statistics:

$$P(r, t) = \left(\frac{2\pi^{d/2}}{\Gamma(d/2)} \right) r^{d-1} P_s(|\vec{r}|, t) \tag{5.12}$$

where the factor in parentheses is the surface area of the d-dimensional sphere accounting for angular directions. Alternatively, one can consider the probability density of the one-dimensional projections of increments that for an isotropic process is equivalent to the radial density function.

In our case, both probability densities yield the same results, therefore we adopted the latter option as the projections on X and Y coordinates can be merged to further improve the statistics.

5.3.3.3 Application to the experiment

The density $P(r,t)$ of absolute magnitudes of one-dimensional increments r at different lag times t have been calculated, based on more than 320,000 data points (meaning around 80,000 for each cytoskeleton state) of two-dimensional single particle trajectories. The density $P(r,t)$ was evaluated at discrete points $r = k\rho$ (with a chosen bin size ρ) and $t = \Delta\delta$ as

$$P(k\rho, \Delta\delta) = \frac{p_k(\Delta\delta)}{\rho p(\Delta\delta)}, \tag{5.13}$$

where $p_k(t)$ is the number of times that the absolute one-dimensional increment $|x_m(q\delta + \Delta\delta) - x_m(q\delta)|$ or $|y_m(q\delta + \Delta\delta) - y_m(q\delta)|$ belongs to the interval $(k\rho, (k+1)\rho)$ and the normalization factor $p(\Delta\delta)$ is the sum of all $p_k(\Delta\delta)$. Here, one uses the increments for all tracers in a cytoskeleton state (i.e. $m = 1, \ldots, M$) and for all possible shifts along each trajectory (i.e. $q = 1, \ldots, N_m - \Delta$). In other words, the probability density of increments includes the ensemble average over all trajectories in a state (e.g. the WT case) and the time average along each trajectory. The time average was crucial to improve the statistics of increments. For the same purpose, the increments along X and Y coordinates were merged. From all trajectories in each of four cytoskeleton states (see Table 5.2), we had $152,804$, $160,958$, $176,528$, and $164,300$ one-step increments for the WT, noAct, noMT, and noCyt cases, respectively. Based on these increments, histograms were produced for each cytoskeleton state and at different lag times, from δ to 100δ (Fig. 5.10).

These probability densities (Fig. 5.10 b-e) start with a Gaussian-like regime for small increments and exhibit an exponential tail for larger increments. For both cases of deprived actin (noAct, noCyt) an additional second exponential tail emerges for larger values of increments. Here, the absence of the dense actin network by F-actin depolymerization leads to a broader increment distribution (Fig. 5.10c) due to less restricted particle motion, as compared to the WT state (Fig. 5.10b). In contrast, single-particle trajectories in cells lacking microtubules (noMT) yield a confined track distribution where the corresponding probability density (Fig. 5.10d) behaves similarly to that of the WT state hence exhibiting smaller increments. In the noMT

state, the cell is set to a state of pure crowding, where the cytoplasm is more viscous, lacking the fluidifying effect of the microtubules. Depolymerising both actin cortex and microtubules comprises these effects (Fig. 5.10e): the cell misses both the confinement by the actin cortex and the active stirring by microtubule motion. This results in a constrained particle motion with slightly larger increments than the noMT state, in particular, at longer lag times. This probability density analysis reveals the opposing functions of the two cytoskeleton components in a living cell: while the actin cortex confines the motion within the cell, the actively driven dynamics of the microtubules fluidifies the cytoplasm, fighting against crowding.

5.3.4 Origin of non-Gaussian distribution

So far, transport processes with non-Gaussian increments were observed in many complex systems (see Chapter 1). These experimental findings have been rationalized either by

1. Sample-based variability: The construction of the distribution of increments by concatenating normally distributed increments from trajectories with different diffusion coefficient would result in a superposition of Gaussian distributions which is not Gaussian.

2. Rare, independently occurring strong events, e.g. kicks from nearly passing MT, cargos, actin waves, that would result in more probable large displacements. Due to this reason, the enhanced tails make the distribution non-Gaussian.

3. Spatial or temporal heterogeneities of the medium that change the transport properties over time (e.g. non-constant diffusion coefficient), the increments are not identically distributed breaking the resulting in non-Gaussian distribution at least a short lag-times.

Here, we investigate which of these reasons is applicable to the eukaryotic intracellular medium, starting with sample-based variability.

5.3.4.1 Sample-based variability

In order to reduce sample-based cell-related variability (see Item 1), the one-step increments of each trajectory were rescaled by their standard deviation. The empirical standard deviation of the m-th trajectory has been calculated by averaging the standard deviation from the two

coordinates σ_m^x and σ_m^y so

$$\sigma_m = \frac{\sigma_m^x + \sigma_m^y}{2}. \tag{5.14}$$

Figure 5.11 shows the rescaled distributions of all four cytoskeleton states at three lag times: δ, 10δ, and 100δ, where δ is the one-step duration of 49 ms. The probability densities collapse for all states to a single curve at the smallest lag times, whereas they are paired into two groups at the longest lag times: the WT and noAct states exhibit larger increments caused by the sweeping motion of the microtubules, and the noMT and noCyt states show smaller increments arising from passive intracellular motion without the influence of the microtubules. Furthermore, this rescaling eliminates the second exponential tails (Fig. 5.11) in the noAct and noCyt states, due to depolymerisation of the actin cortex. Thus, these second exponential tails in the increment probability densities can be attributed to a sample variability among different cells also displayed by the larger spread in one-step increment standard deviations between individual cells in both actin deprived cases (see Table 5.5). Therefore, the vanishing of the second exponential elucidates the important role of the actin cortex unifying the cellular interior for well-regulated transport processes.

In all cytoskeleton states for the rescaled data, the transition from Gaussian-like-distributed increments to an exponential tail remains present. This occurrence of an exponential tail even in the noCyt case, without any dynamic cytoskeleton components, evidences against the second option of rare, independently occurring strong events (see Item 2), since all active dynamics have been eliminated.

5.3.4.2 Correlation of increments

To address the third cause for non-Gaussian behavior of increments distribution, a spatially heterogeneous medium (see Item 3), the auto-correlation function was calculated (Fig. 5.12) increments, was computed for each trajectory ($m = 1, \ldots, M$) as

$$U_m(\Delta) = \frac{1}{2}\left(\frac{U_m^x(\Delta)}{U_m^x(0)} + \frac{U_m^y(\Delta)}{U_m^y(0)}\right) \tag{5.15}$$

where $U_m^x(n)$ is the TA VACF (see Chapter 1) of one-step increments along X coordinate (similarly $U_m^y(n)$ for the Y coordinate). Positive correlations over time scales of about 2 seconds in both the WT and the noAct cases arise from microtubule induced swinging motions that indirectly disturb the tracers. Without microtubules, in the noMT and noCyt cases, there are

only very short-ranged negative correlations on a sub 100 ms time scale. This result cannot explain the observed exponential tail of increment distributions, at least for the noMT and noCyt cases. In fact, the distribution of increments would be unavoidably Gaussian at large lag times by the central limit theorem, if the one-step increments were independent or weakly correlated and identically distributed.

The auto-correlation function, however, does not capture all possible correlations of increments. Even if orientations of increments are uncorrelated (see Fig. 5.11), their absolute values can still remain correlated.

To confirm that the observed non-Gaussian behavior arises from the absolute value correlations, the rescaled one-step increments of each experimental trajectory have been randomly re-shuffled (Fig. 5.13) to fully destroy any correlation. As expected, the reshuffled increment probability densities collapse for all four cytoskeleton states and are Gaussian at all lag times. A plausible scenario for getting correlated absolute values of increments is based on a changing effective diffusivity along the trajectory. This arises either by a slowly evolving environment around the tracer or by the exploration of a quasi-static but spatially heterogeneous cytoplasm, or both. These two physically different scenarios correspond to random motion with either time- or space-dependent diffusion coefficients, or again both. Thus, the Δ-step increments, composed of Δ successive one-step increments with variable diffusion coefficients, become correlated.

These findings illustrate that this new way of analysis yields additional aspects to the dynamics within the cell interior in comparison to earlier studies based on mean-squared displacement or auto-correlations. To our knowledge, this presents the first experimental evidence of medium induced spatio-temporal correlations influencing intracellular transport. We now support our analysis by calculating the autocorrelation of the squared increments in Sec. 5.3.4.3.

5.3.4.3 Correlation of squared increments

The knowledge of the mean square displacement or the velocity autocorrelation function, which are both related to the covariance of increments, fully determines a centered Gaussian process. For this reason, most former works were almost exclusively focused on these basic characteristics. However, such a characterization is not complete for non-Gaussian processes, in particular, for heterogeneous diffusions. In fact, a particle diffusing in a heterogeneous medium experiences distinct local diffusivities on its way. When the equilibration of a particle with its local environ-

ment occurs on a time scale much shorter than the diffusion characteristic time, the i-th one-step increment can formally be represented as $\sigma_i \xi_i$, where ξ_i is a normalized random fluctuation (e.g., thermal noise) and σ_i its amplitude (representing the local diffusivity). This representation is inspired from a general stochastic Langevin equation,

$$dX_t = \mu(X_t, t)dt + \sigma(X_t, t)dW_t. \tag{5.16}$$

Setting the drift term to zero, $\mu(X_t, t) = 0$, one gets an infinitesimal increment as the product of space and time-dependent volatility $\sigma(X_t, t)$ and a normalized Gaussian fluctuation dW_t (white noise). Since the local diffusivity can change in time and space, σ_i depends on a random position of the particle and can thus be understood as a random variable. Even though σ_i and ξ_i are independent in the Itô stochastic convention, the autocorrelation structure of these variables is in general very complicated. Nevertheless, we will argue that the autocorrelation of squared increments can bring some evidence on the heterogeneity of diffusivities. To illustrate our arguments, we resort to a simplified model, in which fluctuations ξ_i are centered Gaussian variables with unit variance, which are independent from the amplitudes σ_i. In this setting, the autocorrelation of increments is factored into autocorrelations of ξ_i and σ_i,

$$\langle \sigma_i \xi_i \sigma_j \xi_j \rangle - \langle \sigma_i \xi_i \rangle \langle \sigma_j \xi_j \rangle = \langle \xi_i \xi_j \rangle \langle \sigma_i \sigma_j \rangle. \tag{5.17}$$

If the local diffusivities (the amplitudes σ_i) were constant, this would be simply the autocorrelation of fluctuations ξ_i. In general, however, the VACF combines the persistence (or anti-persistence) of fluctuations along the trajectory (the first factor) and the diffusivity heterogeneities (the second factor). For this reason, it is instructive to consider the autocorrelation structure of squared increments:

$$W_{ij} = \langle\, (\sigma_i \xi_i)^2 (\sigma_j \xi_j)^2 \,\rangle - \langle\, (\sigma_i \xi_i)^2 \,\rangle \langle\, (\sigma_j \xi_j)^2 \,\rangle. \tag{5.18}$$

The independence of ξ_i and σ_i and the Gaussian distribution of ξ_i imply

$$W_{ij} = \langle \sigma_i^2 \sigma_j^2 \rangle (1 + 2\langle \xi_i \xi_j \rangle^2) - \langle \sigma_i^2 \rangle^2. \tag{5.19}$$

Since the autocorrelation $\langle \xi_i \xi_j \rangle$ of fluctuations is less than 1 and decays when the lag $|i - j|$ increases, the term $\langle \xi_i \xi_j \rangle^2$ can be neglected at large lags, so that W_{ij} essentially reflects the autocorrelation of the local diffusivities (squared amplitudes σ_i^2), i.e., their heterogeneities

probed by the random trajectory. We emphasize again that this discussion is limited to the simplified model and relies on the strong assumption of independence of fluctuations and their amplitudes. Even though this assumption is not satisfied in general and thus the interpretation of the autocorrelation of squared increments is not so straightforward, this autocorrelation can still bring some valuable information and may evidence diffusivity heterogeneities. For this reason, we evaluate this quantity on the experimental trajectories by setting

$$\xi^{2*}_{\Delta}(t) = (\, dx_{\Delta}(t) - \langle dx_{\Delta}(t) \rangle \,)^2 - \langle \, (\, dx_{\Delta}(t) - \langle dx_{\Delta}(t) \rangle \,)^2 \, \rangle, \tag{5.20}$$

with $dx_{\Delta}(t) = x_{t+\Delta} - x_t$. Here we consider an increment $dx_{\Delta}(t)$ at lag time Δ, subtract its empirical mean value $\langle dx_{\Delta}(t) \rangle$, square the difference, and also subtract the mean value of the resulting centered squared increment $\xi^{2*}_{\Delta}(t)$. The time averaged autocorrelation of squared increments averaged over different realizations (trajectories) is then

$$W(\Delta) = \left\langle \frac{1}{N-\Delta} \sum_{i=1}^{N-\Delta} \xi^{2*}_{\Delta}(t) \xi^{2*}_{\Delta}(t+i) \right\rangle_M \tag{5.21}$$

The autocorrelation is then $W(\Delta)/W(0)$, as for the VACF. Figure 5.14 shows the decay of the autocorrelation of squared increments with the lag time Δ for four cytoskeleton states. In all cases, the autocorrelations are positive and slowly decreasing with the lag time. When the cytoskeleton elements are progressively removed, the local diffusivity autocorrelations are getting lower, from the WT and noAct states to noMT and noCyt states. This qualitative observation agrees with the expected reduction of heterogeneities in these cytoskeleton states.

5.3.4.4 Scaling form

Performing an additional rescaling of the n-step increments by their mean absolute increment $\langle r \rangle_t$ at lag time $t = n\delta$, the probability densities for all cytoskeleton states superimpose and become independent of the lag time (Fig. 5.15 for each cytoskeleton state).

Thus, the probability density of absolute increments exhibits a universal scaling,

$$P(r,t) = r^{-1} p(r^*), \quad r^* = \frac{r}{\langle r \rangle_t} \tag{5.22}$$

independently of the lag time Δ and whether the motion is influenced by active cytoskeleton components or not. The inverse increment r in Eq. (5.22) accounts for the dimensionality of the probability density. Thus, the transition of the increment distributions from a Gaussian-like

regime to an exponential tail is an intrinsic property of the underlying dynamics, and the shape of the probability density is captured by the master curve p. To rationalize this behavior of the master curve p, we propose a simple two-parameter distributed model of Gaussian increments with random standard deviations σ. An explicit form of the master curve p is deduced (see Sec. 5.3.5) by assuming a superposition of Gaussian where their standard deviations σ is distributed according to the χ-distribution:

$$p(r) = \frac{(\alpha r^*)^{\nu+1/2} K_{\nu-1/2}(\alpha r^*)}{\sqrt{\pi} 2^{\nu-3/2} \Gamma(\nu)}, \quad \alpha = \frac{2\Gamma(\nu+1/2)}{\sqrt{\pi}\Gamma(\nu)} \tag{5.23}$$

where $K_{\nu-1/2}$ is the modified Bessel function of the second kind, $\Gamma(\nu)$ denotes the Gamma function, and the shape parameter ν quantifies the spread of standard deviations ν. The resulting probability densities $P(r,t)$ reproduce both a plateau at small increments and an exponential tail at large increments, in perfect agreement with experimental data for all cytoskeleton states and all lag times (see Fig. 5.15).

5.3.5 Superstatistical description

When the tracers explore a heterogeneous or time-evolving medium, changes in the effective diffusivity can strongly affect the probability density of increments. In the conventional case, each tracer is assumed to follow Brownian motion with random diffusion coefficient D. This model can describe

1. the effect of the ensemble average over multiple tracers, each following Brownian motion with a fixed diffusivity which changes from one tracer to another (e.g. due to tracer size polydispersity or regional variations of the medium viscosity)

2. a single tracer exploring spatially heterogeneous or time-varying medium whose increments along the trajectory have variable amplitudes and thus affecting the increment probability density function.

In our case, Gaussianity test for individual trajectories have already excluded the first interpretation (see Sec. 5.3.3.1), that lets the heterogeneities in the medium as a plausible explanation. The variability is modeled by the distribution $p_t(D)$ of diffusion coefficients. The observed statistics of increments from an ensemble of such tracers is the average over all random realizations

of D:

$$P_s(\vec{r},t) = \int_0^\infty dD p_t(D) \frac{1}{(4\pi Dt)^{d/2}} \exp(-|\vec{r}|^2/4Dt), \tag{5.24}$$

where $p_t(D)$ is the time-dependent distribution of diffusion coefficients D. At first thought, the introduction of the time-dependent probability density $p_t(\sigma)$ does not simplify the description as compared to the probability density of increments $P_s(\vec{r},t)$. However, we will propose a scaling form of the density, $p_t(\sigma)$, in which the time dependence will be expressed only through the scale parameter. The probability density $p_t(D)$ characterising the heterogeneity of the medium is not known but may be inferred from fitting the probability density function of increments by multiple Gaussians [202, 211, 267] (that would correspond to approximating $p_t(\sigma)$ by a sum of Dirac delta peaks), exponentials or stretched-exponential functions [278, 279]. This method suffers from numerous parameters that need further to be interpreted. We therefore choose an alternative approach by modeling the probability density $p_t(\sigma)$ by a phenomenological distribution. Without any prior knowledge about the statistical properties of σ, one can choose any convenient distribution and check whether the resulting probability density of increments fits the data. If one chooses a one-sided Gaussian distribution,

$$p_t(\sigma) = \frac{2}{\lambda_t\sqrt{2\pi}} \exp\left(-\sigma^2/(2\lambda_t^2)\right) \tag{5.25}$$

with a single parameter λ_t, the probability density function of one-dimensional increments would be $2/(\pi\lambda_t)K_0(r/\lambda_t)$. This probability density function logarithmically diverges at small increments and thus is not suitable for fitting experimental data showing a plateau at $r \approx 0$. The divergence is caused by the choice (5.25) of the density $p_t(\sigma)$ that allows arbitrarily small standard deviations. To prohibit such values, one can multiply the exponential function in Eq. (5.25) by a power law vanishing at $r = 0$:

$$p_t(\sigma) = \frac{2^{1-\nu_t}}{\lambda_t\Gamma(\nu_t)}(\sigma/\lambda_t)^{2\nu_t-1} \exp\left(-\sigma^2/(2\lambda_t^2)\right) \tag{5.26}$$

The scale parameter $\lambda_t > 0$ and the shape parameter $\nu_t \geq 1/2$ of this χ-distribution can be related to the moments of the random variable σ:

$$\langle\sigma\rangle = \lambda_t\sqrt{2}\frac{\Gamma(\nu_t+1/2)}{\Gamma(\nu_t)} \tag{5.27}$$

$$\langle\sigma^2\rangle = 2\nu_t\lambda_t^2 \tag{5.28}$$

while the mode of the distribution (i.e. the most probable value) is equal to

$$\sigma_m = \lambda_t \sqrt{2\nu_t - 1}. \tag{5.29}$$

The ratio

$$\frac{\langle r_t \rangle}{\sqrt{\langle r_t^2 \rangle / 2}} = \frac{2}{\sqrt{\pi}} \frac{\Gamma(\nu_t + 1/2)}{\sqrt{\nu_t}\Gamma(\nu_t)}, \tag{5.30}$$

is independent of λ_t and might potentially be inverted numerically to extract ν_t and then determine the second parameter λ_t from the above relations. It is worth noting however that the ratio (5.30) only weakly depends on ν_t, for instance, when ν_t ranges from 1 to 5, this ratio varies from 1 to 1.1. The numerical inversion will thus be sensitive to eventual deviations of empirical data from the theoretical fit, for instance, due to the presence of the second exponential tail in the probability density of increments. At the same time, this observation tells us that the mean absolute increment $\langle r_t \rangle$ and the mean square displacement $\langle r_t^2 \rangle$ behave very similarly.

Figure 5.16 shows the time dependence of the parameters λ_t and ν_t, obtained by fitting empirical increment probability density function at different lag times by the theoretical fit model of Eq. (5.26), by a standard nonlinear Levenberg-Marquardt least-squares algorithm in Matlab (lsqcurvefit).

The scale parameter λ_t behaves similarly to the mean increment $\langle r_t \rangle$ (see Fig. 5.17), as expected from Eq. (5.27).

In turn, the shape parameter ν_t reaches a more or less constant value at long times. Its fluctuations can be attributed to the fact that the fit formula (5.27) weakly depends on ν_t so that its estimation is thus rather sensitive to noise. Finally, larger values at short lag time suggests a narrower distribution $p_t(\sigma)$ of standard deviations that is consistent with the fact that during shorter time a tracer explores a smaller region and thus experiences less heterogeneities.

We conclude that, although we have originally admitted a time dependence of both parameters of the χ-distribution (5.26), the shape parameter ν_t is found to be asymptotically constant at long times. As a consequence, the increment probability density exhibits the long-time asymptotic scaling in Eq. (5.22). with the universal scaling function

$$p(r) = \frac{(\alpha r^*)^{\nu+1/2} K_{\nu-1/2}(\alpha r^*)}{\sqrt{\pi} 2^{\nu-3/2} \Gamma(\nu)}, \quad \alpha = \frac{2\Gamma(\nu + 1/2)}{\sqrt{\pi}\Gamma(\nu)} \tag{5.31}$$

which is deduced by performing the integration in Eq. (5.24) with the probability $p_t(\sigma)$ from Eq. (5.26).

5.3.6 Cell-based variability

In order to demonstrate the cell related variability, we performed the separate analysis of acquired trajectories for each individual cell. When several trajectories were acquired in the same cell, their increments were analyzed together, as though they came from a single trajectory. Table 5.4 summarizes the estimated standard deviation of one-step increments for each cell. One can see that these values change substantially from cell to cell.

	WT		noactin		noMT		noCyt	
n	N	σ	N	σ	N	σ	N	σ
1	7000	8.8	6146	27.3	4984	11.7	6997	14.1
2	2271	13.5	1095	18.2	5835	15.8	3500	10.9
3	2154	13.1	1617	26.5	5669	8.6	3995	16.3
4	3869	11.8	4350	23.3	6393	14.2	4668	15.0
5	3593	11.3	1571	25.9	6776	16.7	959	24.1
6	6716	7.6	5625	17.9	3782	13.4	7000	16.4
7	6678	9.7	3460	13.8	6331	9.8	7000	9.9
8	7000	12.7	4925	16.2	3875	8.8	2088	14.4
9	1356	14.7	6732	15.0	5416	13.5	2092	18.2
10	5411	16.0	6395	14.3	1300	7.7	4455	10.6
11	736	11.5	4852	14.3	5787	13.7	7000	16.6
12	6973	12.9	1212	14.1	2215	14.3	7000	16.3
13	7000	7.9	4511	21.3	5433	15.8	7000	10.6
14	845	8.6	4413	13.6	2442	11.8	5001	14.4
15	6858	14.1	5492	16.5	3998	10.9	6423	12.7
16	6976	11.4	4998	22.2	6309	15.8	6996	11.5
17			5136	13.0	5947	10.6		
18			7000	18.7	5811	9.8		

Table 5.4: n index of each cell in the first column, for each cytoskeleton state, four columns present: the number of increments N in one (or few) trajectory acquired in this cell; the standard deviation of one-step increments (of lag-time $\delta = 49ms$), in nm.

The distribution of these cell-specific standard deviations can be characterized by the mean and standard deviation for four cytoskeleton states (Table 5.5). Note that the broadness of the distribution of cell specific standard deviations, estimated as the relative standard deviation (standard deviation divided mean), remains comparable in four cases. Note also that the standard deviation of increments at lag time $\delta = 49$ ms of a spherical tracer of diameter $d = 150$ nm in an aquatic solution at ambient temperature would be $\sqrt{2D\delta} = 536$ nm, where $D = k_BT/(6\pi a\eta) = 2.93 \pm 103\text{nm}^2\text{ms}^{-1}$ is the diffusion coefficient according to the Stokes-Einstein relation. This estimate gives an idea of how the intracellular crowding affects the tracer motion.

Cytoskeleton state	Mean and standard deviation (nm)	Relative standard deviation
WT	11.6 ± 2.5	0.22
noAct	18.4 ± 4.7	0.26
noMT	12.5 ± 3.0	0.24
noCyt	14.4 ± 3.6	0.25

Table 5.5: The mean, standard deviation and the relative standard deviation of cell-specific standard deviations obtained from Table 5.4.

5.4 Conclusion

In this chapter, from the statistical analysis of single trajectories, we were able to clarify the motion type in each experiment. A separate conclusion for both studies is given below.

5.4.1 Conclusion on the vibrated disks

We proposed a macroscopic realization of planar Brownian motion by vertically vibrated disks. We performed a systematic statistical analysis of many random trajectories of individual disks. The distribution of one-step increments was shown to be almost Gaussian. Since small deviations at large increments increase with the disk packing fraction, they were attributed to inter-disk collisions. The velocity auto-correlation function was positive at the lag time $n = 1$ and took negative values at $n > 1$ that rapidly vanish with n. We also analyzed the behavior of the

time averaged mean square displacement as a function of the lag time, and its fluctuations from one trajectory to another. We compared the empirical and theoretical distributions of TAMSD and revealed the sensitivity of this distribution at small lag times to eventual auto-correlations and weak non-Gaussianity. We also verified that the empirical distribution can be accurately approximated by the generalized Gamma distribution. Finally, we discussed distinctions between the mean TAMSD and the mode of its distribution. These well-controlled experimental data can serve for validating statistical tools developed for the analysis of single-particle trajectories in microbiology.

5.4.2 Conclusion on passive intracellular transport

We conclude that the dynamics-related differences between the four cytoskeleton states are captured solely by the mean absolute increments $\langle r \rangle_t$, shown in Fig. 5.17 as a function of the lag time t. In agreement with the above analysis of auto-correlations, $\langle r \rangle_t$ exhibits superdiffusive behavior for the WT and noAct cases and diffusive behavior for the other noMT and noCyt cases. Therefore, the presence of cytoskeleton components and their change in ATP-induced activity is necessary to enable cell viability as the underlying dynamics itself are not capable of sustaining the non-active transport within a cell. Thus, microtubules and the actin cortex regulate the intracellular transport like a gear shift based on a single universal transport mechanism. In summary, the analysis of absolute increment probability densities from single particle tracking has revealed (i) the homogenizing role of the actin network unifying intracellular transport, (ii) spatio-temporal heterogeneities as the reason for the non-Gaussian transport features in living cells that present (iii) a universal scaling of the increment probability densities independent of cytoskeleton components. These results pave the way for a more specific analysis of intracellular transport features based on spatio-temporal medium fluctuations also in different cell types.

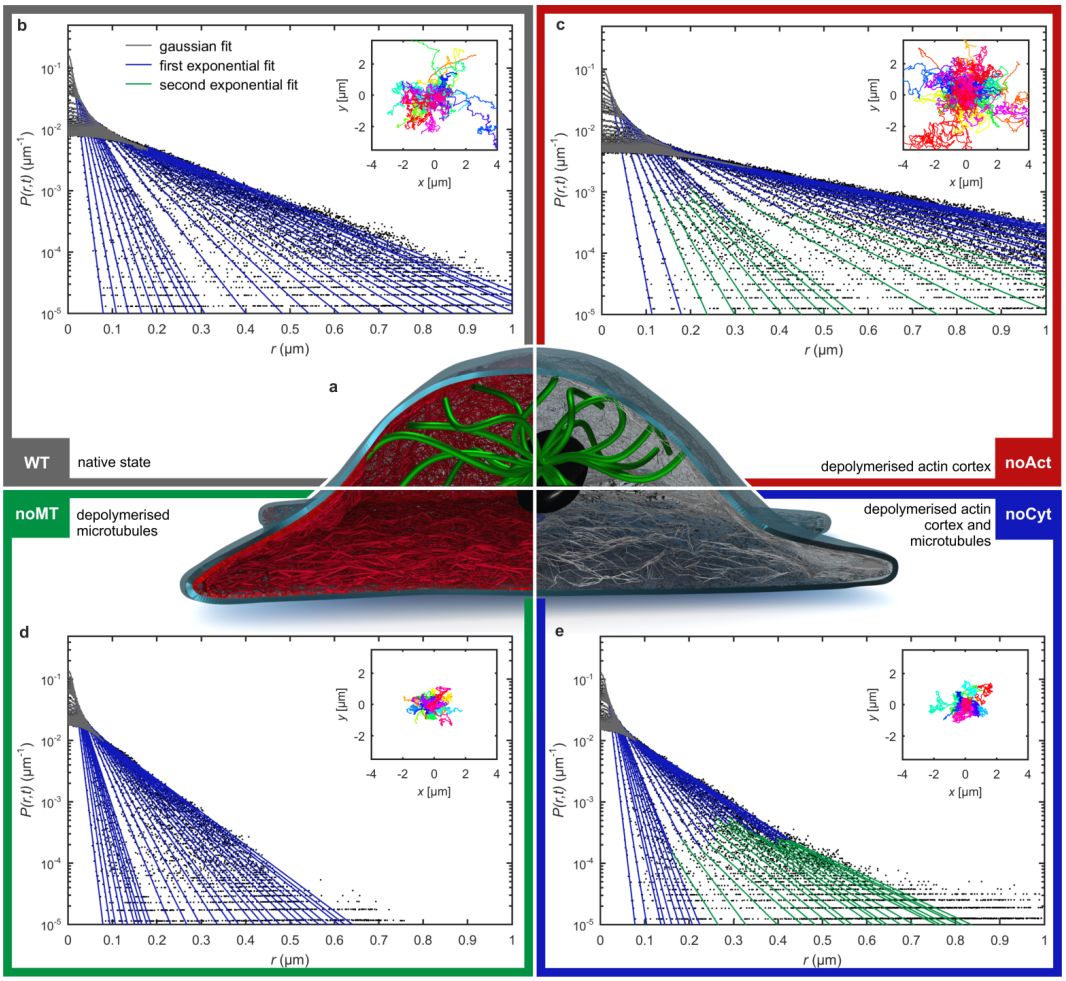


Figure 5.10: Probability density of increments of four different cytoskeleton states: **(a)** illustration of a cell with partly depolymerized cytoskeleton components (grey: membrane, black: nucleus, green: microtubule organizing center with microtubules, red: actin cortex; **(b-e)** probability densities of absolute magnitudes of one-dimensional projected increments, $P(r,t)$, for lag times $t = \delta, 2\delta, \ldots, 10\delta, 20\delta, \ldots, 100\delta$ with the inverse frame rate $\delta = 49$ ms for four different cytoskeleton states: **(b)** wild type (WT), **(c)** depolymerized actin cortex (noAct), **(d)** depolymerized microtubules (noMT) and **(e)** both depolymerizations (noCyt); solid lines show the fit of a Gaussian function (grey) for small increments followed by an exponential fit (blue) and for some lag times in the actin deprived cases (noAct and noCyt) a second exponential fit (green); insets present the acquired trajectories of particle motion displayed in different colors for each trajectory.

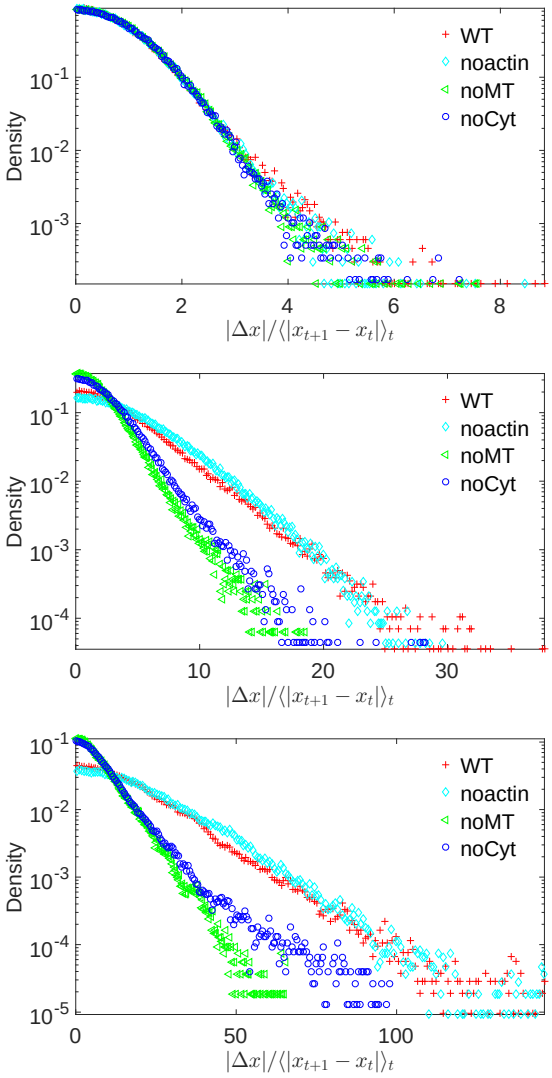


Figure 5.11: Van hove correlation fuction at time 1δ **(Top)**, 10δ **(middle)**, and 100δ **(Bottom)**, where the increments of the trajectories are rescaled by their standard deviation calculated at lag-time $\Delta = 1$. Each cytoskeletal state is displayed: Wild Type **(red crossed lines)**, Depolymerized actin **(green circles)**, Disrupted microtubules **(dark blue squares)**, and deprived of both actin and microtubules **(cyan triangles)**.

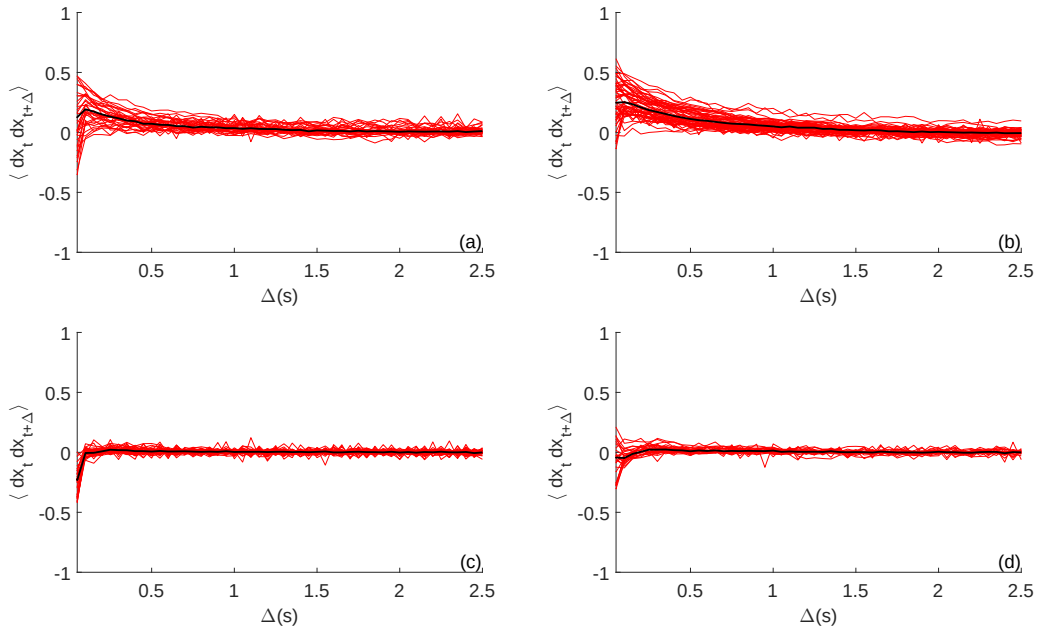


Figure 5.12: Time-averaged velocity auto-correlation function for different cytoskeleton states: Wild Type **(a)**, depolymerised microtubules **(b)**, depolymerised actin cortex **(c)**, and and deprived of both actin and microtubules **(d)**, black lines correspond to the mean value obtained by averaging the velocity auto-correlations for each trajectory in the respective cytoskeleton state, red lines marks the individual TA VACF.

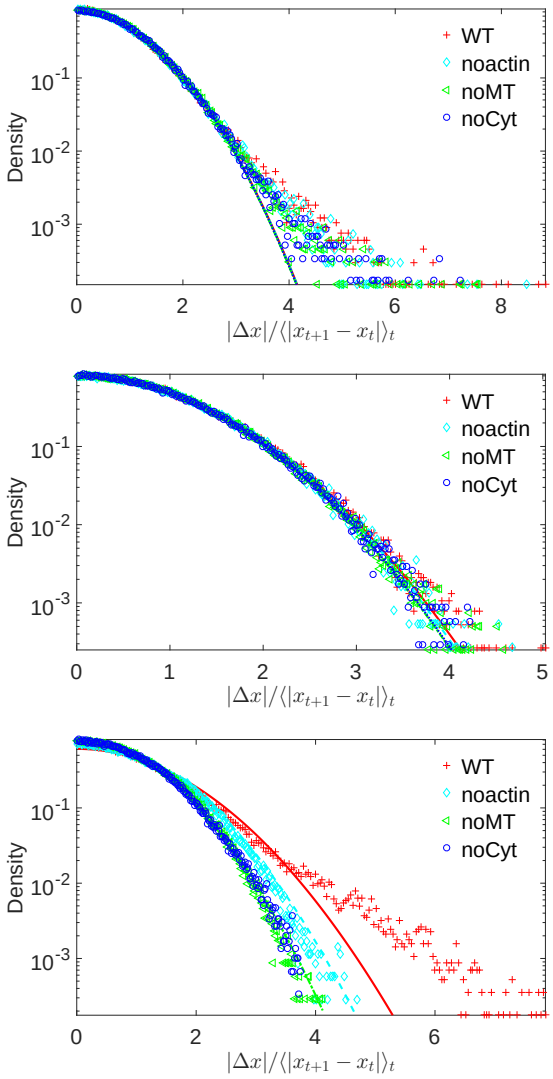


Figure 5.13: Van hove correlation function at lag-time 1δ **(Top)**, 10δ and **(Bottom)** 100δ **(middle)** where the increments of the trajectories are shuffled to destroy any correlation and rescaled by their standard deviation calculated at lag-time $\Delta = 1$. Each cytoskeleton state is displayed: Wild type **(red crossed lines)**, depolymerized actin **(green circles)**, disrupted microtubules **(dark blue squares)**, and deprived of both actin and microtubules **(cyan triangles)**. Solid lines correspond to Gaussian fits in each state.

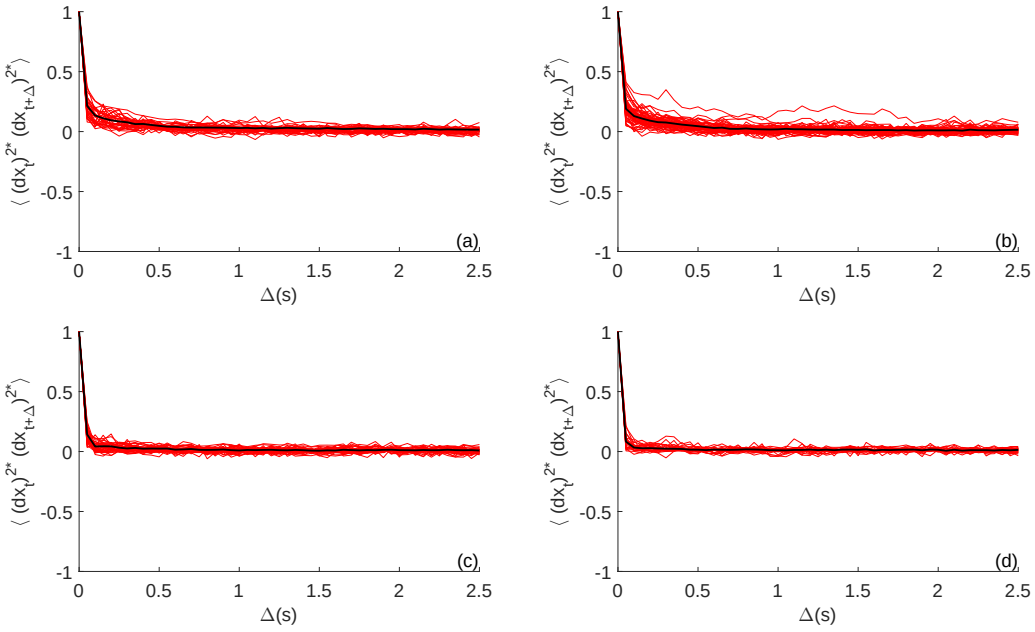


Figure 5.14: Time-averaged autocorrelation function of squared increments for different cytoskeleton states: wild type **(a)**, depolymerized microtubules **(b)**, depolymerized actin cortex **(c)**, and deprived of both actin and microtubules **(d)**, black lines correspond to the mean value obtained by averaging the autocorrelation for each trajectory in the respective cytoskeleton state, red lines marks the individual TA VACF.

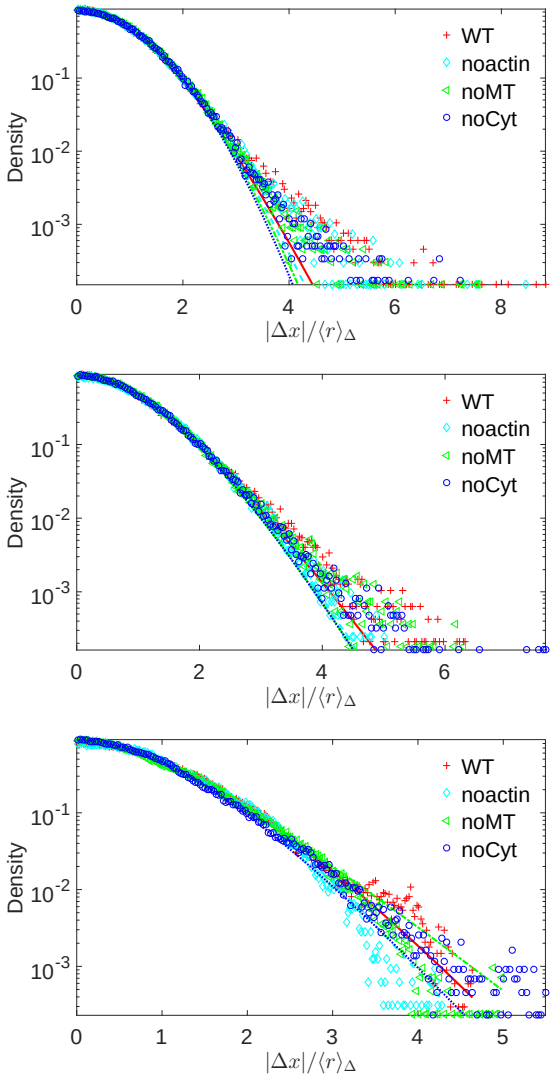


Figure 5.15: Van hove correlation fuction at lag-time **(Top)** 1δ, **(middle)** 10δ and **(Bottom)** 100δ where the increments at lag-time Δ are rescaled by their mean absolute displacement $\langle r\rangle_\Delta$. Each cytoskeletal state is displayed **(red crossed lines)** Wild Type **(green circles)** Depolymerized actin **(dark blue squares)** Disrupted microtubules and **(cyan triangles)** deprived of both actin and microtubules. Solid lines correspond to Gaussian fits in each state.

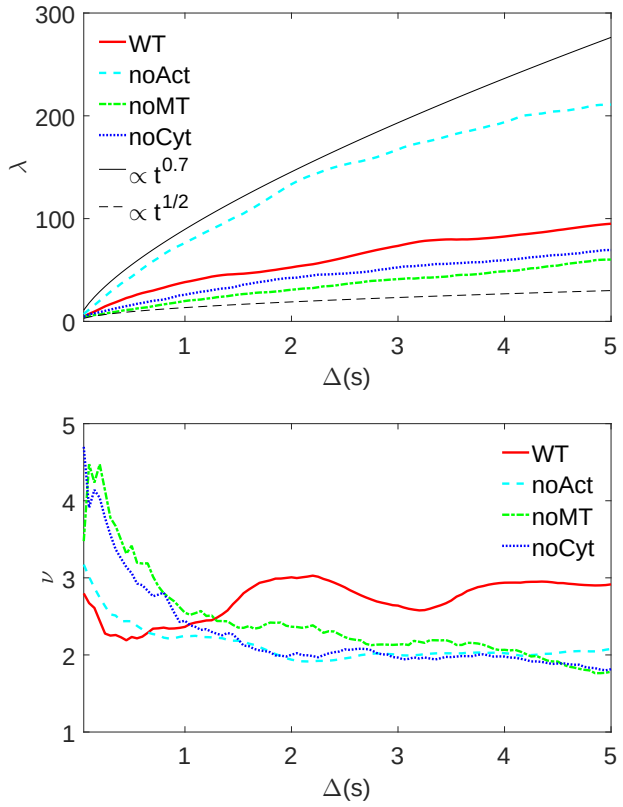


Figure 5.16: Parameters λ **(top)** and ν **(bottom)** fitted at each lag-time Δ for four cytoskeleton states: Wild type **(red line)**, depolymerized actin **(cyan dashed line)**, disrupted microtubules **(dash-dotted green line)**, and deprived of both actin and microtubules **(dotted blue line)**. **black line** (resp. **dotted black line**) is visual guideline $\propto \Delta^{0.7}(s)$ (resp. $\propto \Delta^{0.5}(s)$).

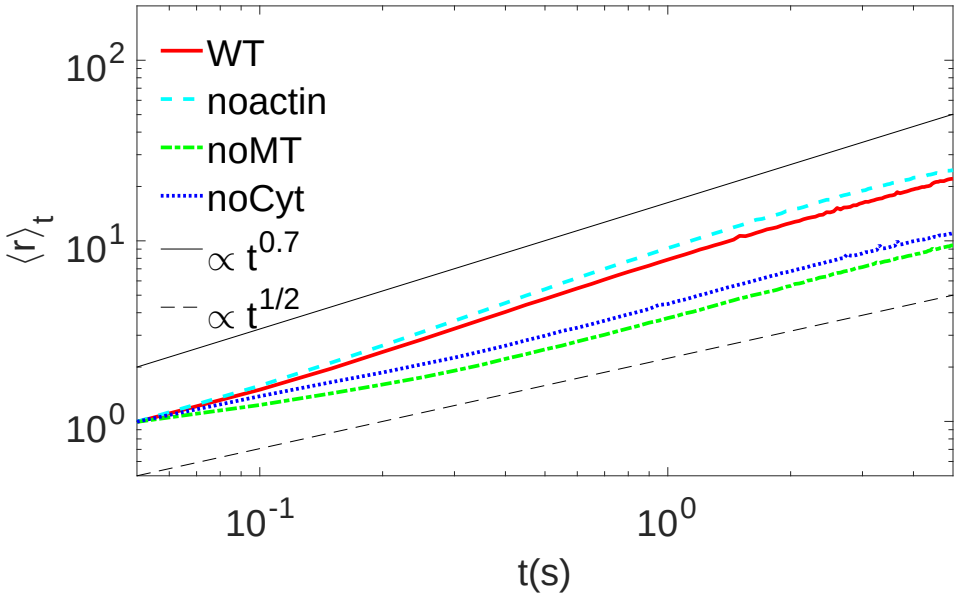


Figure 5.17: Mean absolute increment $\langle r \rangle_t$ for each cytoskeleton state **(red line)** Wild type **(green dashed line)**, depolymerized actin **(dash-dotted blue line)**, disrupted microtubules and **(black dotted line)** deprived of both actin and microtubules. **Black line** (resp. **black dashed lines**) is visual guideline $\propto \Delta^{0.7}(s)$ (resp. $\propto \Delta^{0.5}(s)$).

Chapter 6

Conclusion

During this thesis we analyzed different problems related to the analysis of a single realization of an unknown process. In this chapter, we discuss the future research directions that are suggested by this thesis.

Ergodicity

In the second chapter, the goal was to develop a methodology for assessing the exchangeability of time and ensemble averages without having access to an ensemble of trajectories. For this purpose, we proposed an estimator that can be applied to a single trajectory in order to detect weak ergodicity breaking.

Our estimator partly removes the bias of the former estimator by Magdziarz and Weron [175] due to finite size of the trajectory. By rescaling the increments of trajectories, we removed the dependence of the estimator on the level of fluctuations. In that way, our estimator shows the same distribution for any Brownian motion of fixed duration independently of its diffusion coefficient.

A functional applied to a random process is itself random. In order to show that our estimator predicts the expected behavior for three typical models of anomalous diffusion (detailed in Chapter 1), we calculated the mean estimator for these different models. The mean estimator was calculated analytically for fractional Brownian motion and shown to vanish as the lag-time increases. For the case of diffusion on fractals the mean estimator obtained from simulations confirms the known ergodic behavior of this model. After a detailed analytical analysis, we

showed that in contrast with its predecessor, our estimator does not vanish for CTRW even for an infinitely long trajectory.

We applied the estimator to two sets of experimental data. The first dataset has been recorded from the motion of mRNA in *E.coli* cell and shows no ergodicity breaking. In turn, the motion in Kv2.1 potassium channels revealed weak ergodicity breaking. Both of these results have been previously obtained from an ensemble of trajectories, but here we made the statement stronger by establishing it for each trajectory individually.

Magdziarz and Weron have mathematically proven that their estimator, based on one-step increments $x(n+1)-x(n)$, vanishes for any ergodic stationary infinitely divisible (SID) process [175]. By using the long-time increments of the process $x(n+\Delta)-x(n)$, our estimator was shown to detect ergodicity breaking for the CTRW which is not a stationary infinitely divisible process. This is one of the main results of this chapter.

However, how the estimator behaves in general for a non SID ergodic process remains an open question. In future research, we will investigate this question in order to know to which extent one can apply our estimator to trajectories with non-stationary increments. Another difficult question concerns the introduction of a significance level. Because the estimator applied to a finite trajectory will never exactly vanish, it is important to set a confidence interval above which the trajectory is considered as nonergodic. This significance level could allow an automatic testing of the ergodicity hypothesis. The group of Weron found a path to this end using the concept of ϵ-ergodicity [168–170]. They take in account the finiteness of the trajectory length and devise what should be the length of a trajectory in order to detect with a degree of certainty the ergodicity of the trajectory. However, in order to predict the desired length of the trajectory for a safe assessment, one has to specify a model on which the analytical prediction will be made. In practice, in the direction to detect ergodicity with a controlled degree of certainty, the nature of the process needs to be known, and for the nature of the process to be known, the ergodicity hypothesis needs to be verified. In future work, it would be interesting to find a way to break this circle.

Finally, the single trajectory analysis raises questions that go beyond ergodicity. Mathematically, ergodicity is defined in the limit of infinitely long trajectories. However in practice, all trajectories have finite length. Thus, the knowledge of the behavior in the long-time limit is not sufficient to ensure a precise measurement based on time averaged quantities.

The important instruction that provides the validity of the ergodicity hypothesis is that as the trajectory length increases, the variance of time-averaged quantities vanishes and the distribution of time-averaged quantities converges to the ensemble average. Ergodicity does not tell how fast the variance vanishes but at least ensures that it will vanish. A striking example when this result is not sufficient is provided in Chapter 4. The non-Gaussian diffusion model displays stationary increments at long times and it is ergodic, as also revealed by our estimator. However, the large distribution of these increments due to fluctuations of diffusivity makes much more slower the vanishing of the variance of the TA MSD. For a diffusivity correlation time $\tau = 100$ (arbitrary units), and the parameter $\nu = 1$, the TA MSD calculated with trajectories of length 1000 steps presents the same variance as that of a Brownian motion with 10 steps. To obtain the same precision as in homogeneous diffusion, the length of trajectories obeying this non-Gaussian diffusion model should be recorded on much longer times. As the estimator is based on the time-averaged Fourier transform of the trajectory, it contains more information than ergodicity alone. Future work will be dedicated to relate the decay of the estimator to the variance of time-averaged quantities.

Intermittent processes

In Chapter 2 we introduced a new method for detecting different phases of an intermittent motion. The salient point and also the main difficulty of that approach is that we do not make an a priori assumption on the underlying mechanism of the different phases of motion. For this purpose, we applied a geometric functional along a trajectory that is not specific to a given statistical property but is sensitive to a wide range of temporal changes in the statistical properties. We used as functionals two quantities derived from the convex hull: its diameter and its volume. We tested the method successfully on six models of intermittent processes presenting different intermittent properties (drift, diffusion coefficient, mean-reversion, distribution, and dimensionality) without changing any parameter of the method. Due to the integral-like nature of the convex hull, the results were shown to be robust to noise compared to other differential-like estimators (e.g. variance functional).

This chapter showed that functionals based on geometrical features of the trajectory are very powerful and can be useful in a wide range of applications.

Several further steps could be made in order to enlarge the impact of the method. First the window size (21 points) that we have chosen works well in the examples because it is sufficiently large to average out the noise and small enough to be reactive to changes in the dynamics. However, this window size is not optimal in general. Thus a question would be "how to choose appropriately the window size without prior knowledge on the process?".

In the same vein, the threshold value used for discriminating between the two processes has been chosen to be the empirical mean of the functional considered along the trajectory. We have no proof that this is the optimal choice, but we expect that it is close to optimal when the proportions of each phases are of the same order. We showed that the quality of the estimation decreases when the proportion gets more and more unequal, the empirical mean as a threshold is no longer optimal. Finding a proper threshold based on the trajectories themselves would be very interesting.

Perhaps a solution would be to take some assumptions from which one could deduce optimal choices of the lag-time and threshold value. But in this case, it would require to properly describe the statistics of the convex hull applied to a random process. So far, most of the theoretical knowledge concerns the mean volume or perimeter of the convex hull, but almost nothing is known analytically on their probability density. Moreover, we are not aware of any results concerning the diameter of the convex hull.

Our method presupposes that the process is intermittent. In the perspective of automatizing the procedure, it is crucial to develop a test that could decide beforehand whether the trajectory is intermittent or not. For now, we can calculate statistical properties associated to each sub trajectory and verify that they come from two or more well separated group in the space of all reachable statistical properties.

Non-Gaussian diffusion

Non-Gaussian diffusion is the manifestation of heterogeneities that affect the dynamic properties of the medium. In the model we presented in Chapter 4, non-Gaussian diffusion is modeled by a time-dependent diffusivity. We make a description based on a coupled Langevin equation where the diffusivity is fluctuating, with diffusivity of diffusivity σ, around an average value $\bar{D}$ with a diffusivity correlation time τ. This model can reproduce the propagator with exponential tails

with either peaked of smooth behavior near zero. We showed that the shape is determined by the parameter $\nu = \frac{\bar{D}}{\sigma^2\tau}$. The parameter ν close to zero means strong disorder characterized by an accumulation of time intervals where diffusivity is close to zero, so the distribution of displacements is peaked near zero. When fluctuations of diffusivity are smaller $\nu > 1$, diffusivity cannot approach arbitratily close to zero, so the distribution of displacements is smooth at the center. The parameter τ controls the correlations of diffusivity. Diffusivity does not evolve much on a range $t < \tau$, so increasing τ makes time averaged quantities slower converge to the ensemble average value since the process slower explores all possible diffusivities. An interesting tool for measuring the fluctuations of diffusivity is the autocorrelation of the squared increments of the process. It can be analytically calculated from the Langevin equation on diffusivity and directly compared to experimental data.

The correlations of diffusivity are limited to an exponential decay because of the harmonic trapping of the diffusivity. This unchangeable correlation structure may be limiting to some experimental cases. For example the measurement in Chapter 5 of tracers in the cytoplasm may display power-law correlated square increments. From a more general point of view, the quest to develop a continuous strictly positive process with a prescribed correlation structure is appealing in itself.

Meanwhile, we could also interrogate other interpretations of non-Gaussian diffusion. The effect of spatially dependent diffusivity is not easy to distinguish against time dependent one. Some work is needed in order to clearly state in which situation the diffusivity is space or time dependent inside the cell. It would also be helpful to understand to which extent spatial and time dependent diffusivity looks the same. In this case it could easier to treat some space-dependent diffusivity problems by approximating them by time-dependent ones.

In future work, one may look for a precise physical mechanism that is responsible for the fluctuations interpreted as diffusivity. The diffusion coefficient remains an average quantity that does not mean much without specifying an underlying mechanism.

Finally, in order to understand the impact of non-Gaussian diffusion induced by the dynamic heterogeneities on chemical reactions at the cellular level, the knowledge of the model is not sufficient. The reaction kinetics are governed by the first passage of a reactive particle to a target site. These quantities necessitate another type of description of the dynamics. Our model can then be formulated in term of a Fokker-Planck equation that is an appropriate equation for

studying first-passage problems.

Experimental evidence

In the experiment with vibrated disks of diameter 4mm, we observed a realization of Brownian motion at the macroscopic scale. We were also able to measure the effect of the packing fraction on the dynamic of the disks. The packing fraction was modified by adding new disks on the vibrating plate. At short time, negative correlations become more and more pronounced as the packing fraction increases, but the MSD remains diffusive even at rather high surface fraction. The disks crowding the plate do not induce anomalous diffusion. However the distribution of increments is affected by higher packing fraction. The Gaussian distribution of increments presents small deviations from Gaussianity on the tails. In a future experiment it would be interesting to be able to compare the collective motion of the disks with a single disk moving on the plate. In this way we could distinguish which effect is intrinsic to the system itself and which is due to the local shocks with other disks.

In the experiment in the cytoplasm, we observed the motion of a passive tracer of diameter 150nm in *Dictyostelium discoideum* cells. The cells have been prepared in four different states : Wild type (WT) compared to depolymerized actin (NoAct), disrupted microtubules (noMT), or both treatments (noCyt). On the basis of the statistical analysis developed in this study, we can state that the tracer in a medium deprived of actin displays superdiffusion which we attributed to the sweeping motion of the microtubules that constantly mix the cell content. This mechanism is a good way to fight against the macromolecular crowding inside the cell. Meanwhile, the actin network slows down the diffusivity but homogenizes the cell. Moreover, we detected a non-Gaussian distribution of increments that is independent of the cytoskeleton state. The shape of the distribution is similar to that of our model in Chapter 4 with parameter $\nu > 1$ (mild disorder). The superstatistical approach used to describe the distribution corresponds to the short-time limit of our diffusing diffusivity model. However two characteristics mismatch between this experiment and our model. First the autocorrelation of squared increments seem to decays as a power law rather than exponentially as in our model. Second, the non-Gaussian parameter obtained from experimental trajectories do not vanish during the time of the experiment, it remains mostly constant. A possible explanation would be that the correlation time of diffusivity

can be significantly larger than the trajectory duration so the increments of the trajectories are not stationary so the distribution do not converge to a Gaussian distribution. The same argument would be true for space-dependent diffusivity: if the distribution does not become Gaussian, it means that the time to explore the whole heterogeneity of the cell is significantly large than the trajectory duration.

Appendix A

Résumé en Français

Cette thèse est dédiée á l'analyse et la modélisation d'expériences où la position d'un traceur dans le milieu cellulaire est enregistrée au cours du temps. Il s'agit de pouvoir de retirer le maximum d'information à partir d'une seule trajectoire observée expérimentalement. L'enjeu principal consiste à identifier les mécanismes de transport sous-jacents au mouvement observé. La difficulté de cette tâche réside dans l'analyse de trajectoires individuelles, qui requiert de développer de nouveaux outils d'analyse statistique. Dans le premier chapitre, un aperçu est donné de la grande variété des dynamiques observables dans le milieu cellulaire. Notamment, une revue de différents modèles de diffusion anormale et non-Gaussienne est réalisée. Dans le second chapitre, un test est proposé afin de révéler la rupture d'ergodicité faible à partir dune trajectoire unique. Cest une généralisation de lapproche de M. Magdziarz et A. Weron basée sur la fonction caractéristique du processus moyennée au cours du temps. Ce nouvel estimateur est capable didentifier la rupture dergodicité de la marche aléatoire à temps continu où les temps d'attente sont distribués selon une loi puissance. Par le calcul de la moyenne de lestimateur pour plusieurs modèles typiques de sous diffusion, l'applicabilité de la méthode est démontrée. Dans le troisième chapitre, un algorithme est proposé afin reconnaître à partir d'une seule trajectoire les différentes phases d'un processus intermittent (e.g. le transport actif/passif à l'intérieur des cellules, etc.). Ce test suppose que le processus alterne entre deux phases distinctes mais ne nécessite aucune hypothèse sur la dynamique propre dans chacune des phases. Les changements de phase sont capturés par le calcul de quantités associées à l'enveloppe convexe locale (volume, diamètre) évaluées au long de la trajectoire. Il est montré que cet algorithme est efficace pour distinguer

les états dune large classe de processus intermittents (6 modèles testés). De plus, cet algorithme est robuste à de forts niveaux de bruit en raison de la nature intégrale de l'enveloppe convexe. Dans le quatrième chapitre, un modèle de diffusion dans un milieu hétérogène où le coefficient de diffusion évolue aléatoirement est introduit et résolu analytiquement. La densité de probabilité des déplacements présente des queues exponentielles et converge vers une Gaussienne au temps long. Ce modèle généralise les approches précédentes et permet ainsi d'étudier en détail les hétérogénéités dynamiques. En particulier, il est montré que ces hétérogénéités peuvent affecter de manière drastique la précision de mesures effectuées sur une trajectoire par des moyennes temporelles. Dans le dernier chapitre, les méthodes danalyses de trajectoires individuelles sont utilisées pour étudier deux expériences. La première analyse effectuée révèle que les traceurs explorant le cytoplasme montrent que la densité de probabilité des déplacements présente des queues exponentielles sur des temps plus longs que la seconde. Ce comportement est indépendant de la présence de microtubules ou du réseau dactine dans la cellule. Les trajectoires observées présentent donc des fluctuations de diffusivité témoignant pour la première fois de la présence dhétérogénéités dynamiques au sein du cytoplasme. La seconde analyse traite une expérience dans laquelle un ensemble de disques de 4mm de diamètre a été vibré verticalement sur une plaque, induisant un mouvement aléatoire des disques. Par une analyse statistique approfondie, il est démontré que cette expérience est proche d'une réalisation macroscopique d'un mouvement Brownien. Cependant les densités de probabilité des déplacements des disques présentent des déviations par rapport à la Gaussienne qui sont interprétées comme le résultat des chocs inter-disque. Dans la conclusion, les limites des approches adoptées ainsi que les futures pistes de recherches ouvertes par ces travaux sont discutées en détail.

Appendix B

Properties of random processes and how to test them

An unknown process can be characterized through its statistical properties. A good model for a random process should reproduce most of its statistical properties. In this section, we discuss the main statistical properties of random processes and the relevant methodologies to test them.

B.1 Markov property

Processes verifying the Markov property have no memory effect. The probability of an event X_{n+1} at the $(n+1)^{\text{th}}$ step only depends on X_n:

$$P(X_{n+1}|X_n, X_{n-1}, \ldots, X_1) = P(X_{n+1}|X_n) \tag{B.1}$$

Examples of processes verifying the Markov hypothesis are the Brownian motion and the diffusion in a fractal medium.

B.1.1 Ljung-Box Test

There are different way to test the Markov hypothesis [80–82, 90], we present here the Ljung-Box Test [167] which test the hypothesis H_0 that a timeseries with Gaussian increments has no serial correlation except due to randomness. The hypothesis H_1 is that increments are serially

correlated up to the lag-time k. It is based on the statistics

$$Q = N(N+2)\sum_{k=1}^{\tau} \frac{\{Y_t Y_{t+k}\}^2}{N-\tau} \tag{B.2}$$

where N is the length of the time series, τ is the maximal lag-time for which the sample autocorrelation $\{Y_t Y_{t+k}\}$ can be evaluated.

When H_0 is valid, the distribution of the statistic Q is $\chi^2(\tau)$ distributed. For a given significance level α, H_0 is rejected if the statistic $Q > \chi^2_{1-\alpha,\tau}$.

B.2 Stationarity

B.2.1 Strong stationarity

The strong stationarity states that the joint probability of events is time invariant

$$P(x_{t_1+\tau}, x_{t_2+\tau}, \ldots, x_{t_n+\tau}) = P(x_{t_1}, x_{t_2}, \ldots, x_{t_n}). \tag{B.3}$$

As a consequence, any statistical quantity derived from it is time invariant such as the moments of the distribution.

B.2.2 Weak stationarity

The weak, or wide-sense stationarity, requires that only the first moment and the autocovariance are time invariant. Then, the statistics of the process between two times t_1, t_2 depends only on the time difference $t_2 - t_1$. As an example, the covariance of a weak stationary process reads

$$\langle X_{t_1} X_{t_2} \rangle = K(|t_2 - t_1|). \tag{B.4}$$

This means that the systems is independent of its initial conditions. However, this assumption does not hold anymore for aging systems such as the CTRW with power law waiting times for which $\langle X_{t_1} X_{t_2} \rangle = K(t_2/t_1)$ such that the process depends of all its history [44, 163]. Strong stationarity implies the weak one. The situation where the only time invariant moment of the process is the first one (resp. second one) is refered to as first-order (resp. second-order) stationarity. Some methodologies were proposed to test first-order stationarity of the mean based on the logarithm of the Fourier transform of the process [219] and also for second-order stationarity using wavelet analysis [204].

B.2.3 Unit root

The concept of unit root is widely used in the time series analysis. Consider the autoregressive process of order k

$$y_n = \rho y_{n-i} + \eta_n \tag{B.5}$$

where η_n are IID variable with zero mean and variance σ^2.

Then its characteristic polynomial is

$$Q(m) = \sum_{i=0}^{k} a_i m^{p-i} \tag{B.6}$$

with $a_0 = 1$.

If the equation $Q(m) = 0$ has a root $m = 1$ of order d, then the process is "integrated of order d" (i.e. differentiating the process d times give a stationary time-series).

If all the roots are smaller than one, the process is stationary while it is called explosive when a root is larger than 1.

Dickey-Fuller test [71] propose to test the hypothesis of the presence of a unit root (i.e. $\rho = 1$). For this purpose one needs to calculate the statistic

$$\hat{\rho} = \frac{\sum_{n=1}^{N-1} y_n y_{n+1}}{\sum_{n=1}^{N} y_n^2} \tag{B.7}$$

Several models can be tested for the presence of unit root:

- Unit root $y_n = \delta y_{n-1} + \eta_n$
- Unit root with drift $y_n = \mu + \delta y_{n-1} + \eta_n$
- Unit root with drift and linear trend $y_n = \mu + \beta n + \delta y_{n-1} + \eta_n$

Each time, the corresponding distribution of $\hat{\rho}$ can be calculated analytically, resulting in an analytical form for a significance level α.

B.2.4 Trend stationarity

A time series is trend stationary if it can be written in the form

$$Y(t) = f(t) + \eta_t \tag{B.8}$$

with $f(t)$ a deterministic function and η_t a stationary random process. The fractional Brownian motion has trend stationary increments. But also processes with drift can verify this property. The Kwiatkowski-Phillips-Schmidt-Shin test (KPSS) [157] is aimed to test the null hypothesis of the process being trend-stationary against the alternative that is has unit root.

$$y_t = \nu t + r_t + \epsilon_t, \tag{B.9}$$

with ν the coefficient of the linear trend, ϵ_t a stationary error and r_t a random walk defined by $r_t = r_{t-1} + \eta_t$ with η_t IID random variables with variance σ_ϵ^2. Define e_n be the n^{th} residual from the linear fitting of y_t. Then the partial sum process is defined as $S_n = \sum_{k=1}^{n} e_n$. In this case, the statistic

$$LM = \frac{1}{\sigma_\epsilon} \sum_{n=1}^{N} S_n^2 \tag{B.10}$$

B.3 Ergodicity

B.3.1 Statement of the problem

In many experimental cases, only a few trajectories are accessible making the use of an ensemble average (over the probability space) impossible. In this case, applying time average along a trajectory is an alternative way to identify and to characterize the process.

For this purpose, it is necessary for the system to verify the ergodicity hypothesis. For an ergodic system, the statistics over an ensemble of realizations of the process at a given time is equal to the time average over a trajectory of infinite length. This is know as the Birkhoff ergodic theorem

$$\lim_{N\to\infty} \sum_{k=1}^{N} g[Y(k)] = \langle g[Y(0)] \rangle, \tag{B.11}$$

where g(.) an integrable deterministic function.

The ergodicity hypothesis is a fundamental basis for the equilibrium kinetic theory of gazes. This is the theory describing macroscopic statitical properties of an ensemble of particles at equilibrium. Boltzmann invoked ergodicity in order to derive the equipartition energy theorem [28].

It states that a particle, as time $t \to \infty$, will sample the whole phase space of the system and thus it is equivalent to look at one particle or at an instantaneous picture of independent

particle distributed on the whole phase space. A rigorous definition of ergodicity for dynamical systems is given in Sec. B.3.2.

However a variety of systems do not satisfy this hypothesis. A first class is encountered in the situation where the phase space is not simply connex, so that a portion of the phase space is not accessible to an individual particle. Thus, the average of a quantity calculated along the trajectory cannot lead to a description of the complete system. This is referred to as *strong* ergodicity breaking which is simply impossible to identify from a single trajectory.

A second class is the *weak* ergodicity breaking where the phase space is simply connex, but the dynamics of the object is slowed down and the phase space cannot be fully explored even in the limit of an infinitely long trajectory. The only way to sample correctly systems breaking the ergodicity hypothesis is to average over an ensemble of particles.

Cells are intrinsically out-of-equilibrium as there is a continuous consumption of energy in order to ensure its functioning.

The ergodicity breaking has been thoroughly investigated for various processes, including anomalous diffusions (see reviews [45,176,194] and references therein). Ito proved [132] that for a stationary Gaussian process, a sufficient condition is that the VACF verifies: $\lim_{\tau\to\infty}\langle X_t X_{t+\tau}\rangle = 0$. This results has been further generalized by Khinchin's theorem that relates ergodicity to the long-time vanishing of the VACF of a stationary process without requiring the process to be Gaussian [43, 151, 162].

Among the three models of anomalous diffusion discussed in Sec. 1.2, only the CTRW is nonergodic.

B.3.2 Definition of ergodicity for dynamical systems

Consider an infinitely divisible stationary process (SID), that is, a process $X(n)$ whose distribution can be written as the sum of independent identically distributed (IID) random variables. Under this definition, $X(n)$ can be represented as a probability measure in space (Ω, B), where Ω is the phase space of functions $f : \mathbb{N} \to \mathbb{R}$ and B is the σ-algebra of events. The dynamical system with the probability space (Ω, B, P) is associated to the transformation $S : \Omega \to \Omega$, where $S(f(n)) = f(n+1)$ completely describes the temporal evolution of the process. The dynamical system (Ω, B, P, S) is ergodic if for each invariant set $A \in B$ (i.e. a set C is invariant if for each trajectory $X(t)$, it verifies $X(t) \in C \Rightarrow X(t') \in C,\ \forall\ t' > t$), one has $P(A) = 0$ or

$P(A) = 1$, there is no invariant set except for a constant process. From these definitions, the mathematical formulation of the mixing and ergodicity conditions reads respectively

$$\lim_{n\to\infty} \mathcal{P}[A \cap S^n(B)] = \mathcal{P}(A)\mathcal{P}(B) \tag{B.12}$$

and

$$\lim_{n\to\infty} \frac{1}{n-1} \sum_{k=0}^{n-1} \mathcal{P}[A \cap S^k(B)] = \mathcal{P}(A)\mathcal{P}(B) \tag{B.13}$$

Eq. (B.12) expresses the asymptotic independence of A and B under the transformation S. According to Eq. (B.13), the mixing property implies ergodicity, as ergodicity is the time average of the first condition [173].

B.3.3 Ergodicity Breaking Parameter

The ergodicity breaking (EB) parameter characterizes the normalized variance of the time averaged mean squared displacement (TAMSD) $\bar{\delta^2}(\Delta, t_{exp}))^2$ evaluated with a lag-time Δ on trajectory of duration t_{exp} (see Sec. B.4). It was introduced to quantify deviations from ergodicity in numerous models of anomalous diffusion [43, 45, 122, 194, 196] . The ergodicity breaking parameter $EB(\Delta, t_{exp})$ as a function of the experiment duration t_{exp} (i.e. the trajectory length) evaluated with a time-lag Δ is

$$EB(\Delta, t_{exp}) = \frac{\langle(\bar{\delta^2}(\Delta, t_{exp}))^2\rangle}{\langle\bar{\delta^2}(\Delta, t_{exp})\rangle^2} - 1. \tag{B.14}$$

For an ergodic process, $\lim_{t_{exp}\to\infty} EB(\Delta, t_{exp}) = 0$ for any Δ, meaning that for a fixed Δ, the distribution of TAMSD converges to a Dirac delta distribution with $\bar{\delta^2}(\Delta, t_{exp} \to \infty) = \langle X^2(\Delta)\rangle$. Since both the VACF and EB parameter rely on ensemble averages, many trajectories are needed to reveal ergodicity breaking.

B.4 Time Averaged moments

Once ergodicity and stationarity of a process are verified, one can measure moments from a single trajectory which behaves similarly to their ensemble averaged counterparts, in average. A widely used quantity is the time-averaged mean squared displacement TA MSD, which in fact should be call TA SD because the "Mean" has no signification for a single trajectory. For a

trajectory of duration t_{exp} where the position has been recorded N times at regular interval of duration δ ($t_{exp} = N\delta$), the TA MSD at a lag-time Δ is

$$\bar{\delta^2}(\Delta, t_{exp}) = \frac{1}{N-\Delta} \sum_{k=0}^{N-\Delta} (x(k+\Delta) - x(k))^2 \tag{B.15}$$

The ensemble average TAMSD has been studied for many anomalous diffusion models (see [**?**, 150, 194]. But the distribution of the TAMSD is only known for Gaussian process [114, 115], they can be approximated by a generalized Gamma distribution. The properties fo the TAMSD in confined domain were studied [140].

In the case of nonergodic CTRW with waiting times distributed as $\propto t^{-\alpha}$, the TAMSD does not behave as the MSD but shows a linear dependence on the lag-time Δ

$$\langle \bar{\delta^2}(\Delta, t_{exp}) \rangle \propto \frac{\Delta}{t_{exp}^{1-\alpha}} \tag{B.16}$$

The TA MSD has been proposed in order to determine size of confinement zones in [221] from the TAMSD.

The time averaged velocity autocorrelation function (TA VACF) is

$$X_t \bar{X}_{t+\Delta} = \frac{1}{N-\Delta} \sum_{k=1}^{N-\Delta} Y(k+\Delta) Y(k) \tag{B.17}$$

Note that for $\Delta = 0$ the empirical variance of the process is recovered.

B.5 Testing the distribution of increments

B.5.1 General goodness-of-fit tests

Goodness-of-fit tests question the hypothesis that fluctuations around the theoretical fit are just due to intrinsic variability of data set against that the fitting distribution is not appropriate.

The test is operated in three steps

- to fit the data with the desired distribution.
- to construct a probability distribution of the distance between a cumulative distribution, obtained from simulated random data set with the fitted parameters, and the corresponding theoretical model that is chosen for fitting.

- to evaluate the probability for the empirical data for the distance using the simulated cumulative distributions.
- to reject the hypothesis if the probability is smaller than the p-value chosen

For this purpose, one has to choose a distance measure between two distributions. A general measure is

$$\Omega^2 = n \int_{-\infty}^{\infty} \left(F_n(x) - F(x)\right)^2 \omega(x) dx, \tag{B.18}$$

where n is the number of data points, $F_n(f)$ is the empirical cumulative distribution, $F(x)$ is the theoretical one and $\omega(x)$ is a weighting function. The weight $\omega(x)$ is chosen to accentuate different parts of the distribution. For $\omega(x) = 1$ the Von Cràmer statistic is obtained but with $\omega(x) = F(x)\left(1 - F(x)\right)$ the Anderson-Darling statistics is often preferred as it gives more weight to the tails of the distribution. In particular, the AD was shown to be one of the best statistics for testing Gaussianity [207, 291].

B.5.2 Shapiro-wilk test

The Shapiro-Wilk statistics [243], test the hypothesis that a vector $x_1, \ldots, x_N$ with first k moments $m_1, \ldots, m_k$ with expression

$$m_k = \int_{-\infty}^{\infty} x^k \frac{dF(x)}{dx} \tag{B.19}$$

and covariance matrix C is distributed according to a Gaussian distribution. The test statistic is

$$W = \frac{\left(\sum_{i=1}^{N} a_i \bar{x_i}\right)^2}{\sum_{i=1}^{N} (x_i - \{x_i\})^2}, \tag{B.20}$$

where $\bar{x_i}$ is the i-th smaller value of the sample and the vector

$$(a_1, \ldots, a_n) = \frac{m^T C^{-1}}{\sqrt{m^T C^{-1} C^{-1} m}} \tag{B.21}$$

Then the statistics has to be compared with tables in [243]. It was shown that this is the most efficient for testing Gaussianity on small samples [207], while Anderson-Darling test comes in second position.

B.6 Bayesian Inference

Bayesian methods based on prior knowledge on the motion are popular [1, 11, 30, 128, 230]. The prior knowledge is represented by the probability $P(\boldsymbol{x}_1, \ldots, \boldsymbol{x}_N \mid a_1, \ldots, a_K)$ of observing a trajectory $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_N$ with N points, given the parameters $a_1, \ldots, a_K$ of the model. In other words, one defines a specific model-based functional of the trajectory to "process" the observed data. Once the model is chosen, its parameters $a_1, \ldots, a_K$ can be found by maximizing the likelihood of the observed trajectory according to the Bayesian rules.

B.7 Identifying Intermittent Processes

Different statistical methods have been developed to detect change points of an intermittent stochastic process. Other attempts have been dedicated to the use of "hidden Markov model" where the trajectory is randomly swhitching between a finite number of states of motion with given exchange rates. It is possible to consider switching between two diffusion coefficient [63], or more [257], as well as alternating Brownian motion and harmonical trap [247]. Another solution is based on the maximum likelihood ratio which has been used to detect time dependent diffusivity of fractional Brownian motion [155]. Even more general method can be obtained for detecting change either in drift, diffusivity, or Hurst index for intermittent fractional Brownian motion [37].

The complete description of an intermittent process would ideally require to recognize each phase without prior, in order to limit human bias, and then identify the nature of each phase separately.

Appendix C

Matlab code for ergodicity testing

For completeness, we provide a short Matlab code that can be directly applied for analyzing single-particle tracking data.

```
function [E,F] = EFestimator(X,omega);
% This Matlab function implements the improved mixing and ergodicity estimators
% by Y. Lanoiselee and D. Grebenkov
% INPUT:   X - vector containing positions of the analyzed trajectory
%          omega - (optional) frequency (the default value is 2)
% OUTPUT:  E - the mixing estimator as a function of n
%          F - the ergodicity estimator as a function of n
if (nargin < 2)   omega = 2;  end    % Default value for omega
N = length(X)-1;   % Trajectory points are enumerated as X(0), ..., X(N)
X = X/std(diff(X));    % Normalization by the empirical standard deviation of displacements
for n=0:N,
  D(n+1) = sum( exp( (1i)*omega*(X(n+1:end) - X(1:end-n)) ) )/(N-n+1);
end
E = D - abs(sum( exp((1i)*omega*X) ))^2/N/(N+1) + 1/N;
for n=1:N,   F(n) = sum( E(2:n+1) )/n;   end
F(N+1) = NaN;   % The last point is not defined
end   % end of the function
```

Appendix D

Some properties of the convex hull

The convex hull of points of a random trajectory is a highly nontrivial geometric functional of this trajectory. Here we summarize several rigorous results for both Brownian motion and simple random walk.

Takacs showed that the mean perimeter of the convex hull of standard Brownian motion W_t in the plane is

$$\mathbb{E}\{\mathrm{vol}_1(\mathrm{Conv}(W_t))\} = \sqrt{16\pi Dt}, \tag{D.1}$$

whereas El Bachir obtained the mean area [179, 260]

$$\mathbb{E}\{\mathrm{vol}_2(\mathrm{Conv}(W_t))\} = \pi Dt. \tag{D.2}$$

More general results for a standard Brownian motion W_t in $\mathbb{R}^d$ read [84]

$$\mathbb{E}\{\mathrm{vol}_d(\mathrm{Conv}(W_t))\} = \frac{(\pi Dt)^{d/2}}{\Gamma(d/2+1)^2}, \tag{D.3}$$

$$\mathbb{E}\{\mathrm{vol}_{d-1}(\mathrm{Conv}(W_t))\} = \frac{2(4\pi Dt)^{(d-1)/2}}{\Gamma(d)}. \tag{D.4}$$

Extensions to a set of n Brownian paths have been provided in [179]. The perimeter of the planar convex hull in the case of confinement to a semi-infinite domain presents nontrivial behavior as a function of distance from the starting point to the boundary [56].

These formulas are exact for continuous Brownian motion. However, experimental trajectories are discretized so that Eqs. (D.3, D.4) can only be valid asymptotically when the number of points in the convex hull is large. In turn, our LCH estimators are based on a relatively small number of points, and Eqs. (D.3, D.4) are thus not applicable. Spitzer and Widom considered

a two-dimensional discrete random walk, modeled as a sum of independent random variables in the complex plane, and derived the average perimeter of the related convex hull [254],

$$\mathbb{E}\{L_N\} = 2\sum_{k=1}^{N} \frac{\mathbb{E}\{|x_k + iy_k|\}}{k} \tag{D.5}$$

(here $x_k + iy_k$ is the position of the walker in the complex plane after k steps), with the concentration inequality [250]

$$\mathbb{P}\{|L_N - \mathbb{E}\{L_N\}| \geq \varepsilon\} \leq 2\exp\left(-\frac{\varepsilon^2}{8\pi^2 N}\right). \tag{D.6}$$

Some other properties were reported in [250]. The asymptotic behavior of the mean perimeter and the mean area of the convex hull over planar random walks was investigated in [119]. The obtained formulas are even applicable to a moderate number of jumps that makes them suitable for the analysis of the local convex hulls. The distributions of the area and of the perimeter of the convex hull were computed numerically by a sophisticated large-deviation approach in [57]. The properties of the convex hull of Gaussian samples and of d-dimensional fractional Brownian motion were analyzed in [64, 65].

Bibliography

[1] R. P. Adams and D. J. MacKay, *Bayesian online changepoint detection*,ArXiv preprint arXiv:0710.3742 (2007).

[2] B. Alberts, A. Johnson, J. Lewis, M. Raff, K. Roberts and P. Walter, *Molecular BIology of the Cell*, Garland Science, 2008.

[3] S. Alexander and R. Orbach, *Density of states on fractals*, Jour. Phys. Lett., **43** (17), L-625-L-631 (1982).

[4] D. J. Amit and N. Brunel, *Model of global spontaneous activity and local structured activity during delay periods in the cerebral cortex*, Cereb Cortex. **7**, 237-252 (1997).

[5] D. J. Amit and N. Brunel, *Dynamics of a recurrent network of spiking neurons before and following learning*, Network: Comput. Neural Syst. **8**, 373-404 (1997).

[6] A. Andreanov and D. S. Grebenkov, *Time-averaged MSD of Brownian motion*, J. Stat. Mech. P07001 (2012).

[7] D. Arcizet, B. Meier, E. Sackmann, J. O. Rädler, and D. Heinrich, *Temporal Analysis of Active and Passive Transport in Living Cells*, Phys. Rev. Lett. **101**, 248103 (2008).

[8] M. P. Backlund, R. Joyner, and W. E. Moerner, *Chromosomal locus tracking with proper accounting of static and dynamic errors*, Phys. Rev. E **91**, 062716 (2015).

[9] D. S. Banks and C. Fradin, *Anomalous Diffusion of Proteins Due to Molecular Crowding* Biophys. J. **89**, 2960 (2005).

[10] C. B. Barber, D. P. Dobkin, and H. T. Huhdanpaa, *The Quickhull Algorithm for Convex Hulls*, ACM Trans. Math. Software **22**, 469-483 (1996).

[11] D. Barber, A. T. Cemgil, and S. Chiappa, *Bayesian Time Series Models*, Cambridge University Press, 1-26 (2011).

[12] E. Barkai, R. Metzler, and J. Klafter, *From continuous time random walks to the fractional Fokker-Planck equation*, Phys. Rev. E **61**, 132 (2000).

[13] E. Barkai, Y. Garini, and R. Metzler, *Strange kinetics of single molecules in living cells*, Phys. Today, **65**, 2935 (2012).

[14] C. Beck and E. G. D. Cohen, *Superstatistics*, Physica A, **322**, 267-275 (2003).

[15] C. Beck, E. G. D. Cohen, and H. L. Swinney, *From time series to superstatistics*, Phys. Rev. E, **72**, 056133 (2005).

[16] G. Bel and E. Barkai, *Weak Ergodicity Breaking in the Continuous-Time Random Walk*, Phys. Rev. Lett. **94**, 240602 (2005).

[17] O. Bénichou, D. S. Grebenkov, P. Levitz, C. Loverdo, and R. Voituriez, *Optimal Reaction Time for Surface-Mediated Diffusion*, Phys. Rev. Lett. **105**, 150606 (2010).

[18] O. BV́enichou, D. S. Grebenkov, P. Levitz, C. Loverdo, and R. Voituriez, *Mean first-passage time of surface-mediated diffusion in spherical domains*, J. Stat. Phys. **142**, 657-685 (2011).

[19] O. Bénichou, C. Loverdo, M. Moreau, and R. Voituriez, *Intermittent search strategies*, Rev. Mod. Phys. **83**, 81-109 (2011).

[20] O. Bénichou and R. Voituriez, *From first-passage times of random walks in confinement to geometry-controlled kinetics*, Phys. Rep. **539**, 225-284 (2014).

[21] H. C. Berg and D. A. Brown, *Chemotaxis in Escherichia coli analysed by three-dimensional tracking*, Nature **239**, 500-504 (1972).

[22] O. G. Berg, R. B. Winter, and P. H. von Hippel, *Diffusion-driven mechanisms of protein translocation on nucleic acids. 1. Models and theory*, Biochemistry **20**, 6929 (1981).

[23] H. C. Berg, *E. coli in Motion*, (Springer-Verlag, New York, 2004).

[24] A. J. Berglund, *Statistics of camera-based single-particle tracking*, Phys. Rev. E **82**, 011917 (2010).

[25] O. J. N. Bertrand D. K. Fygenson and O. A. Saleh, *Active, motor-driven mechanics in a DNA gel*, Proc. Natl. Acad. Sci. USA, **109**, 17342-17347 (2012).

[26] E. Bertseva, D. S. Grebenkov, P. Schmidhauser, S. Gribkova, S. Jeney, and L. Forró, *Optical Trapping Microrheology in Cultured Human Cells*, Eur. Phys. J. E **35**, 63 (2012).

[27] R. Bidaux, J. Chave, and R. Vocka, *Finite time and asymptotic behaviour of the maximal excursion of a random walk*, J. Phys. A: Math. Gen. **32**, 5009-5016 (1999).

[28] Vorlesungen ber Gastheorie. Leipzig: J. A. Barth. (1898). OCLC 01712811
French reedition: Leçon sur la théorie des gaz, Éditions Jacques Gabay, ISBN 2-87647-004-7

[29] P. Bongrand, *Ligand-receptor interactions*, Rep. Prog. Phys. **62**, 921-968 (1999).

[30] P. J. Bosch, J. S. Kanger, and V. Subramaniam, *Classification of dynamical diffusion states in single molecule tracking microscopy*, Biophys. J. **107**, 588-598 (2014).

[31] G. Bottazzi, *On the Pareto Type III distribution*, LEM Working Paper Series 2007/07 (2007)
Online: `http://www.lem.sssup.it/WPLem/files/2007-07.pdf`

[32] J.-P. Bouchaud, and A. Georges, *Anomalous diffusion in disordered media: Statistical mechanisms, models and physical applications*, Phys. Rep., **195**, 4-5, 127-293 (1990).

[33] C. P. Brangwynne, G. H. Koenderink, F. C. MacKintosh, and D. A. Weitz, *Intracellular transport by active diffusion*, Trends Cell Biol. **19**, 423-427 (2009).

[34] C. Bräuchle, D. C. Lamb, and J. Michaelis (Eds.), *Single Particle Tracking and Single Molecule Energy Transfer*, (Wiley-VCH, Weinheim, 2010).

[35] P. C. Bressloff and J. M. Newby, *Stochastic models of intracellular transport*, Rev. Mod. Phys. **85**, 135-196 (2013).

[36] V. Briane, C. Kervrann, and M. Vimond, *Statistical analysis of particle trajectories in living cells*, Phys. Res. E **97**, 062121 (2018).

[37] V. Briane, C. Kervrann, and M. Vimond, *A Sequential Algorithm to Detect Diffusion Switching along Intracellular Particle Trajectories*, arXiv:1804.04977 [stat.ME] (2018).

[38] X. Brokmann, J.-P. Hermier, G. Messin, P. Desbiolles, J.-P. Bouchaud, and M. Dahan, *Statistical Aging and Nonergodicity in the Fluorescence of Single Nanocrystals*, Phys. Rev. Lett. **90**, 120601 (2003).

[39] I. Bronshtein, et al., *Loss of lamin A function increases chromatin dynamics in the nuclear interior*. Nat. com. **6**, 8044 (2015).

[40] R. Brown, *A brief account of microscopical observations made in the months of June, July and August, 1827, on the particles contained in the pollen of plants; and on the general existence of active molecules in organic and inorganic bodies*, Edinburgh New Phil. J. **5**, 358-371 (1828).

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

[42] K. Burnecki, E. Kepten, J. Janczura, I. Bronshtein, Y. Garini, and A. Weron, *Universal algorithm for identification of fractional Brownian motion. A case of telomere subdiffusion*, Biophys. J. **103**, 1839-1847 (2012).

[43] S. Burov, J. H. Jeon, R. Metzler, and E. Barkai, *Single particle tracking in systems showing anomalous diffusion: the role of weak ergodicity breaking*, Phys. Chem. Chem. Phys., **13**, 1800-1812 (2010).

[44] S. Burov, R. Metzler, and E. Barkai, *Aging and nonergodicity beyond the Khinchin theorem*, Proc. Nat. Acad. Sci **107**, 13228-13233 (2010).

[45] S. Burov, J.-H. Jeon, R. Metzler, and E. Barkai,*Single particle tracking in systems showing anomalous diffusion: the role of weak ergodicity breaking*, Phys. Chem. Chem. Phys. **13**, 1800-1812 (2011).

[46] C. Calderon, *Detection of Subtle Dynamical Changes Induced by Unresolved 'Conformational Coordinates' in Single-Molecule Trajectories via Goodness-of-Fit Tests*, J. Phys. Chem. B **114**, 3242-3253 (2010).

[47] A. Caspi, R. Granek, and M. Elbaum, *Enhanced diffusion in active intracellular transport.*, Phys. Rev. Lett. **85**, 56555658 (2000).

[48] P. Chaudhuri, L. Berthier and W. Kob, *Universal nature of particle displacements close to glass and jamming transitions*, Phys. Rev. Lett., **99**, 060604 (2007).

[49] A. V. Chechkin, F. Seno, R. Metzler, and I. M. Sokolov, *Brownian yet Non-Gaussian Diffusion: From Superstatistics to Subordination of Diffusing Diffusivities*, Phys. Rev. X, **7**, 021002 (2017).

[50] L. Chen, *Stochastic Mean and Stochastic Volatility A Three-Factor Model of the Term Structure of Interest Rates and Its Application to the Pricing of Interest Rate Derivatives*, Fin. Mark., Instit. Instru., **5**, 1-88 (1996).

[51] W. Chen and K. To, *Unusual diffusion in a quasi-two-dimensional granular gas*, Phys. Rev. E **80**, 061305 (2009).

[52] B. Chen and Y. Hong, *Testing for the Markov Property in Time Series*, Eco. Theo. **28**, 130-178 (2012).

[53] A. G. Cherstvy and R. Metzler,*Anomalous Diffusion in Time-Fluctuating Non-Stationary Diffusivity Landscapes*, Phys. Chem. Chem. Phys., **18**, 840-852 (2016).

[54] J. Choi, A. Kudrolli, R. R. Rosales, and M. Z. Bazant, *Diffusion and Mixing in Gravity-Driven Dense Granular Flows*, Phys. Rev. Lett. **92**, 174301 (2004).

[55] M. V. Chubynsky and G. W. Slater, *Diffusing Diffusivity: A Model for Anomalous, yet Brownian, Diffusion*, Phys. Rev. Lett., **113**, 098302 (2014).

[56] M. Chupeau, O. Bénichou, and S. N. Majumdar, *Convex hull of a Brownian motion in confinement*, Phys. Rev. E **91**, 050104 (2015).

[57] G. Claussen, A. K. Hartmann, and S. N. Majumdar, *Convex hulls of random walks: Large-deviation properties*, Phys. Rev. E **91**, 052104 (2015).

[58] S. Condamin, O. Bnichou, V. Tejedor, R. Voituriez and J. Klafter, *First-passage times in complex scale-invariant media*, Nature **450**, 7780 (2007).

[59] S. Condamin, V. Tejedor, R. Voituriez, O. Bénichou, and J. Klafter, *Probing microscopic origins of confined subdiffusion by first-passage observables*, Proc. Nat. Ac. Sci. USA **105**, 5675 (2008).

[60] L. Conway, D. Wood, E. Tuzel, and J. L. Ross, *Motor transport of self-assembled cargos in crowded environments.*, Proc. Nat. Ac. Sci. USA **109**, 2081420819 (2012).

[61] J. C. Cox, J. E. Ingersoll, S. A. Ross, *A Theory of the Term Structure of Interest Rates*, Econometrica, **53**, 385-408 (1985).

[62] H. Cramér, *Mathematical Methods of Statistics*,(Princeton University Press, 1946).

[63] R. Das, C. W. Cairo, D. Coombs, *A Hidden Markov Model for Single Particle Tracks Quantifies Dynamic Interactions between LFA-1 and the Actin Cytoskeleton*, PLOS Comp. Biol. **5** (11): e1000556 (2009).

[64] Y. Davydov, *On convex hull of Gaussian samples*, Lit. Math. Jour. **51**, 171-179 (2011).

[65] Y. Davydov, *On convex hull of d-dimensional fractional Brownian motion*, Stat. Prob. Let. **82**, 37-39 (2012).

[66] M. de Berg, O. Cheong, M. van Kreveld, and M. Overmars, *Computational Geometry: Algorithms and Applications*, Third ed. (Springer-Verlag, Berlin-Heidelberg, 2008).

[67] P.-G. de Gennes, *Introduction to Polymer Dynamics*, (Cambridge University Press, Cambridge, 1990).

[68] J. Deseigne, S. Léonard, O. Dauchot, and H. Chaté, *Vibrated polar disks: spontaneous motion, binary collisions, and collective dynamics*, Soft Matter **8**, 5629 (2012).

[69] M. A. Desposito, C. Pallavicini, V. Levi, and L. Bruno, *Active transport in complex media: Relationship between persistence and superdiffusion*, Physica A **390**, 1026-1032 (2011).

[70] D. Dhar and R. Ramaswamy, *Classical diffusion on Eden Trees*, Phys. Rev. Lett., **54**, 1346-1349 (1985).

[71] D. A. Dickey and W. A. Fuller, *Distribution of the Estimators for Autoregressive Time Series with a Unit Root*, Jour. Amer. Stat. Asso., **74**, 366a, 427-431 (1979).

[72] T. Dieker, *Simulation of fractional Brownian motion*, M.Sc. thesis (2004). http://www.columbia.edu/ ad3217//fbm/thesis.pdf

[73] J. A. Dix and A. S. Verkman, *Crowding Effects on Diffusion in Solutions and Cells*, Ann. Rev. Biophys. **37**, 247-263 (2008).

[74] M. Doi and S. F. Edwards, *The Theory of Polymer Dynamics*, (Clarendon Press, 1986).

[75] A. A. Drăgulescu and V. M. Yakovenko, *Return of the Heston volatility model*, Quant. Fin., **2**, 443-453 (2011).

[76] D. Dufresne D, *The integrated square-root process*, *Working Paper, University of Melbourne* (2001).

[77] M. H. G. Duits, Y. Li, S. A. Vanapalli, and F. Mugele, *Mapping of spatiotemporal heterogeneous particle dynamics in living cells*, Phys. Rev. E **79**, 051910 (2009).

[78] B. Duplantier, *Le mouvement Brownien, "divers et ondoyant"*, Séminaire Poincaré **1**,155-212 (2005).

[79] A. Dupont, et al., *Single Particle Tracking in Living Cells. Is the Third Dimension Worth It?*, Biophys J **104**, 650a (2013).

[80] J. Durbin and G. S. Watson, *Testing for Serial Correlation in Least Squares Regression, I*, Biometrika, **37** (34): 409428 (1950).

[81] J. Durbin and G. S. Watson, *Testing for Serial Correlation in Least Squares Regression, II*, Biometrika, **38**, (12), 159179 (1951).

[82] J. Durbin and G. S. Watson, *Testing for Serial Correlation in Least Squares Regression, III*, Biometrika, **58**,1, 1-19 (1971).

[83] A. Einstein, *On the motion of small particles suspended in liquids at rest required by the molecular-kinetic theory of heat*, Ann. Phys. **17**, 549 (1905).

[84] R. Eldan, *Volumetric properties of the convex hull of an n-dimensional Brownian motion*, Elec. Jour. Prob. **19**, 1-34 (2014).

[85] R. J. Ellis, *Macromolecular crowding - obvious but underappreciated*, Trends Bio. Chem. Sci. **26**, 597-604 (2001).

[86] R. J. Ellis ans A. P. Minton, *Cell biology: join the crowd.*, Nature **425**, (6953):27-8 2003.

[87] D. Ernst, J. K öhler, and M. Weiss, *Probing the type of anomalous diffusion with single-particle tracking*, Phys. Chem. Chem. Phys. **16**, 7686 (2014).

[88] N. Fakhri, et al., *High-resolution mapping of intracellular fluctuations using carbon nanotubes.*, Science **344**, 10311035 (2014).

[89] K. Falconer, *Fractal Geometry: Mathematical Foundations and Applications 2nd ed.*, Wiley (2003).

[90] R. W. Farebrother, *Algorithm AS 153: Pan's Procedure for the Tail Probabilities of the Durbin-Watson Statistic* Jour. Roy. Stat. Soci., **29**, 2, 224-227 (1980).

[91] J. Feder, *Fractals*, Plenum Press, New York (1988).

[92] W. Feller, *Two singular diffusion problems*, Ann. Math. **54**, 173-182 (1951).

[93] E. Fieremans, D. S. Novikov, J. H. Jensen, J. A. Helpern, *Monte Carlo study of a two-compartment exchange model of diffusion*, NMR Biomed. **23**, 711-724 (2010).

[94] É. Fodor, D. S. Grebenkov, P. Visco and F. Van Wijland (2015).*Generalized Langevin equation with hydrodynamic backflow: Equilibrium properties.*, Phys. A, **422**, 107112 (2010).

[95] E. Fodor, H. Hayakawa, P. Visco, and F. van Wijland, *Active cage model of glassy dynamics*, Phys. Rev. E **94**, 012610 (2016).

[96] E. Frey and K. Kroy, *Brownian motion: a paradigm of soft matter and biological physics*, Ann. Phys. (Leipzig) **14**, 20 (2005).

[97] N. Gal, D. Lechtman-Goldstein, and D. Weihs, *Particle tracking in living cells: A review of the mean square displacement method and beyond*, Rheol. Acta **52**, 425-443 (2013).

[98] M. Galanti, D. Fanelli, and F. Piazza, *Conformation-controlled binding kinetics of antibodies*, Scient. Rep. **6**, 18976 (2016).

[99] X. Gan, D. Waxman, *Singular solution of the Feller diffusion equation via a spectral decomposition*, Phys .Rev. E, **91**, 012123 (2015).

[100] Y. Gefen, A. Aharony, and S. Alexander, *Anomalous Diffusion on Percolating Clusters*, Phys. Rev. Lett. **50**, 77-80 (1983).

[101] W. M. Getz and C. C. Wilmers, *A local nearest neighbor convex-hull construction of home ranges and utilization distribution*, Ecography **27**, 03835 (2004).

[102] A. K. Ghosh, A. G. Cherstvy and R. Metzler, *Non-universal tracer diffusion in crowded media of non-inert obstacles*, Phys. Chem. Chem. Phys., **17** 1847 (2015).

[103] S. K. Ghosh, A. G. Cherstvy, D. S. Grebenkov and R. Metzler, *Anomalous, non-Gaussian tracer diffusion in heterogeneously crowded environments*, New. J. Phys., **18**, 013027 (2016).

[104] R. Ghrist, *Barcodes: The persistent topology of data*, Bull. Amer. Math. Soc. **45**, 61-75 (2008).

[105] M. Giona, H.E. Roman, *Fractional diffusion equation on fractals: one-dimensional case and asymptotic behaviour*, J. Phys. A: Math. Gen. **25**, 2093 (1992).

[106] I. Golding and E. C. Cox, *Physical Nature of Bacterial Cytoplasm*, Phys. Rev. Lett. **96**, 098102 (2006).

[107] L. S. B. Goldstein and Z. Yang, *Microtubule-Based Transport Systems in Neurons: The Roles of Kinesins and Dyneins*, Ann. Rev. Neur. **23**, 39-71 (2000).

[108] J. P. Gollub, J. Clarke, M. Gharib, B. Lane and O. N. Mesquita, *Fluctuations and transport in a stirred fluid with a mean gradient*, Phys. Rev. Lett., **67**, 3507-3510 (1991).

[109] I. Goychuk, *Viscoelastic Subdiffusion: Generalized Langevin Equation Approach*, Adv. Chem. Phys., **150**, 187253 (2012).

[110] I. S. Gradshteyn and I. M. Ryzhik, *Table of Integrals, Series, and Products*, (Academic Press, 1980).

[111] M. E. Grady, E. Parrish, M. A. Caporizzo, S. C. Seeger,R. J. Composto, and D. M. Eckmann, *Intracellular nanoparticle dynamics affected by cytoskeletal integrity*, Soft Matter, **13**, 1873-1880 (2017).

[112] D. S. Grebenkov, *NMR Survey of Reflected Brownian Motion*, Rev. Mod. Phys. 79, 1077-1137 (2007).

[113] D. S. Grebenkov, M. P. Ciamarra, M. Nicodemi, and A. Coniglio, *Flow, Ordering and Jamming of Sheared Granular Suspensions*, Phys. Rev. Lett. **100**, 078001 (2008).

[114] D. S. Grebenkov, *Probability Distribution of the Time-Averaged Mean-Square Displacement of a Gaussian Process*, Phys. Rev. E **84**, 031124 (2011).

[115] D. G. Grebenkov, *Time-averaged quadratic functionals of a Gaussian process*, Phys.Rev. E **83**, 061117 (2011).

[116] D. S. Grebenkov, *Optimal and sub-optimal quadratic forms for non-centered Gaussian processes*, Phys. Rev. E **88**, 032140 (2013).

[117] D. S. Grebenkov, M. Vahabi, E. Bertseva, L. Forro, and S. Jeney, *Hydrodynamic and subdiffusive motion of tracers in a viscoelastic medium*, Phys. Rev. E **88**, 040701R (2013).

[118] D. S. Grebenkov, *First exit times of harmonically trapped particles: a didactic review*, J. Phys. A **48**, 013001 (2015).

[119] D. S. Grebenkov, Y. Lanoiselée, and S. N. Majumdar, *Mean perimeter and mean area of the convex hull over planar random walks*, J. Stat. Mech. 103203 (2017).

[120] S. Havlin, D. Ben Avraham,*Diffusion in random media*, Adv. Phys. **36** (3775), 695798 (1987).

[121] S. Havlin and D. Ben-Avraham, *Diffusion in disordered media*, Adv. Phys., **36**, 187-292 (2002).

[122] Y. He, S. Burov, R. Metzler, and E. Barkai, *Random time-scale invariant diffusion and transport coefficients*, Phys. Rev. Lett., **101**, 058101 (2008).

[123] W. He, H. Song, Y. Su, L. Geng, B. J. Ackerson, H. B. Peng and P. Tong, *Dynamic heterogeneity and non-Gaussian statistics for acetylcholine receptors on live cell membrane*, Nat. Com., **7**, 11701 (2016).

[124] E. Heinsalu, M. Patriarca, I. Goychuk, and P. Hänggi, *Use and Abuse of a Fractional Fokker-Planck Dynamics for Time-Dependent Driving*, Phys. Rev. Lett., **99** 120602 (2007).

[125] E. Hernández-García, M. A. Rodríguez, L. Pesquera and M. San Miguel, *Transport properties for random walks in disordered one-dimensional media: Perturbative calculation around the effective-medium approximation*, Phys. Rev. A, **41**, 4562 (1990).

[126] S. L. Heston, *A Closed-Form Solution for Options with Stochastic Volatility with Applications to Bond and Currency Options*, Rev. Fin. Studies., **6**, 327-343 (1993).

[127] D. J. Higham, *An Algorithmic Introduction to Numerical Simulation of Stochastic Differential Equations*, SIAM Rev., **43**, 525-546 (2001).

[128] K. Hinsen and G. R. Kneller, *Communication: A multiscale Bayesian inference approach to analyzing subdiffusion in particle trajectories*, J. Chem. Phys. **145**, 151101 (2016).

[129] F. Höfling and T. Franosch, *Anomalous transport in the crowded world of biological cells*, Rep. Progr. Phys. **76**, 046602 (2013).

[130] D. Holcman and Z. Schuss, *100 years after Smoluchowski: stochastic processes in cell biology*, J. Phys. A: Math. Theor. **50**, 093002 (2017).

[131] J. C. Hull, *Options, futures and other derivatives*, (7th Ed., Pearson Prentice Hall, Upper Saddle River, NJ, 2009).

[132] K. Itô, Stochastic integral, Proc. Imp. Acad. **20**, 5455 (1944).

[133] D. J. Jacobs, S. Mukherjee, and H. Nakanishi, *Diffusion on a DLA cluster in two and three dimensions.*, J. Phys. A : Math. Gen., **27**, 13, 4341-4350 (1994).

[134] H. M. Jaeger, S. R. Nagel, and R. P. Behringer, *Granular solids, liquids, and gases*, Rev. Mod. Phys. **68**, 1259 (1996).

[135] R. Jain and K. L. Sebastian, *Diffusion in a Crowded,Rearranging Environment*, J. Phys. Chem. B, **120**, 3988-3992 (2016).

[136] R. Jain and K. L. Sebastian, *Lévy flight with absorption: A model for diffusing diffusivity with long tails*, Phys. Rev. E, **95**, 032135 (2017).

[137] R. Jain and K. L. Sebastian,*Diffusing diffusivity: a new derivation and comparison with simulations*, J. Chem. Sci., **126**, 929-937 (2017).

[138] J. Janczura and A. Weron, *Ergodicity testing for anomalous diffusion: Small sample statistics*, J. Chem. Phys. **142**, 144103 (2015).

[139] J. H. Jensen, J. A. Helpern, A. Ramani, H. Lu, K. Kaczynski, *Diffusional Kurtosis imaging: the quantification of non-Gaussian water diffusion by means of magnetic resonance imagning*, Magn. Reson. Med., **53**, 1432-1440 (2005).

[140] J.-H. Jeon and R. Metzler, *Fractional Brownian motion and motion governed by the fractional Langevin equation in confined geometries*, Phys. Rev. E **81**, 021103 (2010).

[141] J.-H. Jeon, V. Tejedor, S. Burov, E. Barkai, C. Selhuber-Unkel, K. Berg-Sorensen, L. Oddershede, and R. Metzler, *In Vivo Anomalous Diffusion and Weak Ergodicity Breaking of Lipid Granules*, Phys. Rev. Lett. **106**, 048103 (2011).

[142] J.-H. Jeon, A. V. Chechkin and R. Metzler, *Scaled Brownian motion: a paradoxical process with a time dependent diffusivity for the description of anomalous diffusion*, Phys. Chem. Chem. Phys. **16**, 15811-15817 (2014).

[143] J.-H. Jeon, M. Javanainen, H. Martinez-Seara, R. Metzler, and I. Vattulainen, *Protein Crowding in Lipid Bilayers Gives Rise to Non-Gaussian Anomalous Lateral Diffusion of Phospholipids and Proteins*, Phys. Rev. X, **6**, 021006 (2016).

[144] A. Kahana, G. Kenan, M. Feingold, M. Elbaum, and R. Granek, *Active transport on disordered microtubule networks: the generalized random velocity model.*, Phys. Rev. E **78**, 51912 (2008).

[145] J. Kärger, *NMR self-diffusion studies in heterogeneous systems*, Adv. Coll. Int. Sci., **23**, 129-148 (1985).

[146] E. A. Katrukha, M. Mikhaylova, H. X. van Brakel, P. M. van Bergen en Henegouwen, A. Akhmanova, C. C. Hoogenraad, and L. C. Kapitein, *Probing cytoskeletal modulation of passive and active intracellular dynamics using nanobody-functionalized quantum dots*, Nat. Comm. **8**, 14772 (2017).

[147] D. Kenwright, A. Harrison, T. Waigh, P. Woodman, and V. J. Allan, *First-passage-probability analysis of active transport in live cells*, Phys. Rev. E **86**, 031910 (2012).

[148] E. Kepten, I. Bronshtein, and Y. Garini, *Ergodicity convergence test suggests telomere motion obeys fractional dynamics*, Phys. Rev. E **83**, 041919 (2011).

[149] E. Kepten, I. Bronshtein, and Y. Garini, *Improved estimation of anomalous diffusion exponents in single-particle tracking experiments*, Phys. Rev. E **87**, 052713 (2013).

[150] E. Kepten, A. Weron, G. Sikora, K. Burnecki, and Y. Garini, *Guidelines for the Fitting of Anomalous Diffusion Mean Square Displacement Graphs from Single Particle Tracking Experiments*, PLoSONE **10**, e0117722 (2015).

[151] A. I. Khinchin, *Mathematical Foundations of Statistical Mechanics* (Dover, New York, 1949).

[152] F. Kindermann, A. Dechant , M. Hohmann, T. Lausch, D. Mayer, F. Schmidt, E. Lutz, and L. Widera, *Nonergodic diffusion of single atoms in a periodic potential*, Nat. Phys., **13**, 137-141 (2016).

[153] J. Klafter and I. M. Sokolov, *First Steps in Random Walks: From Tools to Applications* (Oxford University Press, 2011).

[154] A. N. Kolmogorov, *The Wiener spiral and some other interesting curves in Hilbert space* , Dokl. Akad. Nauk SSSR., **26-2** ,115118 (1940).

[155] D. Kucharczyk, A. Wyłomańska, and G. Sikora, *Variance change point detection for fractional Brownian motion based on the likelihood ratio test.*, Phys. A: Stat. Mech. and Appl. **490**, 439-450 (2018).

[156] S. C. Kuo and M. P. Sheetz, *Force of single kinesin molecules measured with optical tweezers*, Science **260**, 232-234 (1993).

[157] D. Kwiatkowski, P. C. B. Phillips, P. Schmidt and Y. Shin, *Testing the null hypothesis of stationarity against the alternative of a unit root*, Jour. Eco., **54**, 159-178 (1992).

[158] P. Langevin *Sur la théorie de mouvement brownien*, C. R. Hebd. Seances Acad. Sci. **146**, 530 (1908).

[159] Y. Lanoiselée and D. S. Grebenkov,*Revealing nonergodic dynamics in living cells from a single particle trajectory*, Phys. Rev. E **93**, 052146 (2016).

[160] Y. Lanoiselée and D. S. Grebenkov, *Unravelling intermittent features in single particle trajectories by a local convex hull method*, Phys. Rev. E **96**, 022144 (2017).

[161] Y. Lanoiselée, G. Briand, O. Dauchot, and D. S. Grebenkov,*Statistical analysis of random trajectories of vibrated disks: towards a macroscopic realization of Brownian motion*, eprint arXiv:1712.00770.

[162] L. C. Lapas, R. Morgado, M. H. Vainstein, J. M. Rubi, and F. A. Oliveira, *Khinchin Theorem and Anomalous Diffusion*, Phys. Rev. Lett. **101**, 230602 (2008).

[163] N. Leibovich, and E. Barkai, *Aging Wiener-Khinchin Theorem.*, Phys. Rev. Lett. **115**,(8):080602 (2015).

[164] A. Le Méhauté, R. R. Nigmatullin and L. Nivanen, *Flèches du temps et géométrie fractale*, Hermes edition (1990). ISBN 2-86601-682-3

[165] K. C. Leptos, J. S. Guasto, J. P. Gollub, A. I. Pesci and R. E. Goldstein, *Dynamics of enhanced tracer diffusion in suspensions of swimming eukaryotic microorganisms*, Phys. Rev. Let. **103**, 198103 (2009).

[166] P. Levitz, *Random flights in confining interfacial systems*, J. Phys. Condens. Matter **17**, S4059 (2005).

[167] G. M. Ljung; G. E. P. Box . *On a Measure of a Lack of Fit in Time Series Models*, Biometrika, **65** (2): 297303 (1978).

[168] H. Loch, J. Janczura, A. Weron, *Ergodicity testing using an analytical formula for a dynamical functional of alpha-stable autoregressive fractionally integrated moving average processes.*, Phys. Rev. E **93**, 043317 (2016).

[169] H. Loch-Olszewska, G. Sikora, J. Janczura, A. Weron, *Identifying ergodicity breaking for fractional anomalous diffusion: Criteria for minimal trajectory length.*, Phys. Rev. E **94**, 052136 (2016).

[170] H. Loch-Olszewska, J. Szwabiński, *Detection of ϵ-ergodicity breaking in experimental data A study of the dynamical functional sensibility.*, Jour. Chem. Phys. **148**, 204105 (2018).

[171] C. Loverdo, O. Bénichou, R. Voituriez, A. Biebricher, I. Bonnet, and P. Desbiolles, *Quantifying Hopping and Jumping in Facilitated Diffusion of DNA-Binding Proteins*, Phys. Rev. Lett. **102**, 188101 (2009).

[172] A. Lubelski, I. M. Sokolov, and J. Klafter, *Nonergodicity Mimics Inhomogeneity in Single Particle Tracking*, Phys. Rev. Lett. **100**, 250602 (2008).

[173] M. Magdziarz, *Correlation cascades, ergodicity properties and long memory of infinitely divisible processes*, Stoch. Proc. App. **119**, 3416-3434 (2009).

[174] M. Magdziarz, A. Weron, K. Burnecki, and J. Klafter, *Fractional brownian motion versus the continuous-time random walk: A simple test for subdiffusive dynamics*, Phys. Rev. Let. **103**, 180602 (2009).

[175] M. Magdziarz and A. Weron, *Anomalous diffusion: Testing ergodicity breaking in experimental data*, Phys. Rev E **84**, 051138 (2011).

[176] M. Magdziarz and A. Weron, *Ergodic properties of anomalous diffusion processes*, Ann. Phys. **326**, 2431-2443 (2011).

[177] F. Mainardi, Y. Luchko, and G. Pagnini, *The fundamental solution of the space-time fractional diffusion equation*, Frac. Calc. Appl. Anal., **4**, 2, 153-192 (2001).

[178] F. Mainardi, G. Pagnini, R. K. Saxena, *Fox H functions in fractional diffusion*, Jour. Comp. App. Math., **178**, 12, 321-331 (2005).

[179] S. N. Majumdar, A. Comtet, and J. Randon-Furling, *Random Convex Hulls and Extreme Value Statistics*, J. Stat. Phys. **138**, 1-55 (2010).

[180] B. B. Mandelbrot and J. W. Van Ness, *Fractional Brownian motions, fractional noises and applications*, SIAM Rev. **10** 422-437 (1968).

[181] C. Manzo, J. a. Torreno-Pina, P. Massignan, G. J. Lapeyre, M. Lewenstein, and M. F. Garcia Parajo, *Weak Ergodicity Breaking of Receptor Motion in Living Cells Stemming from Random Diffusivity*, Phys. Rev. X, **5**, 14 (2015).

[182] G. Margolin and E. Barkai, *Nonergodicity of Blinking Nanocrystals and Other Levy-Walk Processes*, Phys. Rev. Lett. **94**, 080601 (2005).

[183] G. Margolin and E. Barkai, *Nonergodicity of a Time Series Obeying Levy Statistics*, J. Stat. Phys. **122**, 137-167 (2006).

[184] P. Massignan, C. Manzo, J. A. Torreno-Pina, M. F. Garcia-Parajo, M. Lewenstein, and G. J. Lapeyre, *Nonergodic Subdiffusion from Brownian Motion in an Inhomogeneous Medium*, Phys. Rev. Lett. **112**, 150603 (2014).

[185] J.-B. Masson, P. Dionne, C. Salvatico, M. Renner, C. G. Specht, A. Triller, and M. Dahan, *Mapping the energy and diffusion landscapes of membrane proteins at the cell surface using high-density single-molecule imaging and Bayesian inference: Application to the multiscale dynamics of glycine receptors in the neuronal membrane*, Biophys. J. **106**, 74-83 (2014).

[186] L. Mazzucato, A. Fontanini, and G. La Camera, *Dynamics of multistable states during ongoing and evoked cortical activity*, J. Neurosci. **35**, 8214-8231 (2015).

[187] J. J. McCauley, *Stochastic calculus and differential equations for physics and finance*, Cambridge University Press (2013).

[188] Y. Meroz, I. M. Sokolov, and J. Klafter, *Subdiffusion of mixed origins: When ergodicity and nonergodicity coexist*, Phys. Rev. E **81**, 010101(R) (2010).

[189] Y. Meroz, I. M. Sokolov, and J. Klafter, *Test for determining a subdiffusive model in ergodic systems from single trajectories*, Phys. Rev. Let. **110**, 090601 (2013).

[190] Y. Meroz and I. M. Sokolov, *A toolbox for determining subdiffusive mechanisms.* Phys. Rep., **573**, 129 (2015).

[191] R Metzler, WG Glckle, TF Nonnenmacher,*Fractional model equation for anomalous diffusion*, Phys. A Stat. Mech. App **211** (1), 13-24 (1994).

[192] R. Metzler and J. Klafter, *The random walk guide to anomalous diffusion: A fractional dynamics approach*, Phys. Rep. **339**, 1-77 (2000).

[193] R. Metzler, V. Tejedor, J.-H. Jeon, Y. He, W. H. Deng, S. Burov, and E. Barkai, *Analysis of Single Particle Trajectories: From Normal to Anomalous Diffusion*, Acta Phys. Pol. B **40**, 1315-1331 (2009).

[194] R. Metzler, J. H. Jeon, A. G. Cherstvy, and E. Barkai, *Anomalous diffusion models and their properties: non-stationarity, non-ergodicity, and ageing at the centenary of single particle tracking*, Phys. Chem. Chem. Phys., **16**, 24128-24164 (2014).

[195] R. Metzler, *Gaussianity Fair: The Riddle of Anomalous yet Non-Gaussian Diffusion*, Biophys. J., **112**, 413-447 (2017).

[196] M. Schwarzl, A. Godec, and R. Metzler, *Quantifying non-ergodicity of anomalous diffusion with higher order moments*, Sci. Rep., **7**, 3878 (2017).

[197] X. Michalet, *Mean square displacement analysis of single-particle trajectories with localization error: Brownian motion in an isotropic medium*, Phys. Rev. E **82**, 041914 (2010).

[198] X. Michalet and A. J. Berglund, *Optimal diffusion coefficient estimation in single-particle tracking*, Phys. Rev. E **85**, 061916 (2012).

[199] A. Michelman-Ribeiro, D. Mazza, T. Rosales, T. J. Stasevich, H. Boukari, V. Rishi, C. Vinson, J. R. Knutson, and J. G. McNally, *Direct Measurement of Association and Dissociation Rates of DNA Binding in Live Cells*, Biophys. J. **97**, 337 (2009).

[200] A. P. Minton, *The Influence of Macromolecular Crowding and Macromolecular Confinement on Biochemical Reactions in Physiological Media*, Jour. Biol. Chem. **276**, 14, 1057710580 (2001).

[201] E. W. Montroll, and G. H. Weiss, *Random Walks on Lattices. II*, Jour. Math. Phys. **6**, 167 (1965).

[202] T. Moschakis, A. Lazaridou and C. G. Biliaderis, *Using particle tracking to probe the local dynamics of barley β-glucan solutions upon gelation*, J Coll. Inter. Sci., **375**, 5059 (2012).

[203] A. Nandi, D. Heinrich, and B. Lindner, *Distributions of diffusion measures from a local mean-square displacement analysis*, Phys. Rev. E **86**, 021926 (2012).

[204] G. Nason, *A test for second-order stationarity and approximate confidence intervals for localized autocovariances for locally stationary time series.*, Jour. Roy. Stat. Soc. B, **75**, 879-904 (2013).

[205] B. Nienhuis, *Critical Behavior of Two-Dimensional Spin Models and Charge Asymmetry in the Coulomb Gas*, J. Stat. Phys. **34**, 731 (1984).

[206] F. Normant and C. Tricot, *Method for evaluating the fractal dimension of curves using convex hulls*, Phys. Rev. A **43**, 6518-6525 (1991).

[207] M. R. Nornadiah, B. W. Yap, *Power comparisons of Shapiro-Wilk, Kolmogorov-Smirnov, Lilliefors and Anderson-Darling tests*, Jour. Stat. Mod. Anal., **2**, 1, 21-33, (2011).

[208] J. O'Rourke, *Computational Geometry in C*, 2nd Ed. (Cambridge Tracts in Theoretical Computer Science, Cambridge University Press, Cambridge, 1998).

[209] A. V. Orpe and A. Kudrolli, *Velocity correlations in dense granular flows observed with internal imaging*, Phys. Rev. Lett., **98**, 238001 (2007).

[210] M. Otten, A. Nandi, D. Arcizet, M. Gorelashvili, B. Lindner, and D. Heinrich, *Local Motion Analysis Reveals Impact of the Dynamic Cytoskeleton on Intracellular Subdiffusion*, Biophys. J. **102**, 758-767 (2012).

[211] Palacci, J., Cottin-Bizonne, C., Ybert, C. and Bocquet, L., "Sedimentation and effective temperature of active colloidal suspensions." Phys Rev Lett **105**, 88304 (2010).

[212] D. Panja, *Generalized Langevin equation formulation for anomalous polymer dynamics*, J. Stat. Mech., L02001 (2010).

[213] J. Perrin, *Les Atomes*, Librairie Félix Alcan, Paris (1913).

[214] J. Perrin,L'agitation moléculaire et le mouvement brownien, Compt. Rendus Herbo. Seances Acad. Sci. Paris **146**, 967 (1908).

[215] J. Perrin, *Mouvement brownien et réalité moléculaire*, Ann. Chim. Phys. **18**, 1-114 (1909).

[216] O. Peters, *Optimal leverage from non-ergodicity*, Quant. Finance **11**, 1593-1602 (2011).

[217] O. Peters and W. Klein, *Ergodicity Breaking in Geometric Brownian Motion*, Phys. Rev. Lett. **110**, 100603 (2013).

[218] J. M. Porrá, K. G. Wang, J. Masoliver, _em Generalized Langevin equations: Anomalous diffusion and probability distributions, Phys.Rev. E, **53**, 5872 (1996).

[219] M.B. Priestley, and T. Subba Rao, *A Test for Non-Stationarity of Time-Series.*, Jour. Roy. Stat. Soc. B, **31**, 140-149 (1969).

[220] H. Qian, M. P. Sheetz, and E. L. Elson, *Single particle tracking. Analysis of diffusion and flow in two-dimensional systems*, Biophys. J. **60**, 910-921 (1991).

[221] V. Rajani, G. Carrero, D. E. Golan, G. De Vries, and C. W. Cairo,*Analysis of Molecular Diffusion by First-Passage Time Variance Identifies the Size of Confinement Zones*, Biophys. J. **100**, 1463 (2011).

[222] R. Rammal, G. Toulouse., *Random walks on fractal structures and percolation clusters*, Jour. Phys. Let., **44** (1), 13-22 (1983).

[223] J. Randon-Furling, S. N. Majumdar, and A. Comtet, *Convex hull of N planar Brownian motions: Exact results and an application to ecology*, Phys. Rev. Let. **103**, 140602 (2009).

[224] G. Rangarajan and M.Ding, *First passage time distribution for anomalous diffusion*,Phys. Lett. A **273**, 5-6, 322-330 (2000).

[225] J. F. Reverey, et al., *Superdiffusion dominates intracellular particle motion in the supercrowded cytoplasm of pathogenic Acanthamoeba castellanii.*, Sci. Rep. **5**, 11690 (2015).

[226] P. H. Richter and M. Eigen, *Diffusion controlled reaction rates in spheroidal geometry: Application to repressor-operator association and membrane bound enzymes*, Biophys. Chem. **2**, 255-263 (1974).

[227] G. Rivas, and A. P. Minton, *Macromolecular Crowding In Vitro, In Vivo, and In Between.*, Trends bioc. sci. **41**, 970981 (2016).

[228] F. Rojo, C. E. Budde Jr., H. S. Wio, and C. E. Budde, *Enhanced transport through desorption-mediated diffusion*, Phys. Rev. E **87**, 012115 (2013).

[229] F. Rouyer and N. Menon, *Velocity fluctuations in a homogeneous 2D granular gas in steady state*, Phys. Rev. Lett., **85**, 3676-3679 (2000).

[230] E. Ruggieri and M. Antonellis, *An exact approach to Bayesian sequential change point detection*, Comput. Stat. Data Anal. **97**, 71-86 (2016).

[231] J.-F. Rupprecht, O. Bénichou, D. S. Grebenkov, and R. Voituriez, *Kinetics of active surface-mediated diffusion in spherically symmetric domains*, J. Stat. Phys. **147**, 891-918 (2012).

[232] E. Sackmann, F. Keber, and D. Heinrich, *Physics of Cellular Movements*, Ann. Rev. Conden. Matt. Phys. **1**, 257-276 (2010).

[233] S. Sadegh, J. L. Higgins, P. C. Mannion, M. M. Tamkun, and D. Krapf, *Plasma Membrane is Compartmentalized by a Self-Similar Cortical Actin Meshwork*, Phys. Rev. X **7**, 011031 (2017).

[234] G. Samorodnitsky and M. S. Taqqu, *Stable Non-Gaussian Random Processes: Stochastic Models with Infinite Variance* (Chapman and Hall, New York, 1994).

[235] N. Sarnanta and R. Chakrabarti, *Tracer diffusion in a sea of polymers with binding zones: mobile vs. frozen traps*, Soft Matter, **12**, 8554 (2016).

[236] M. J. Saxton, *Lateral diffusion in an archipelago. Single-particle diffusion*, Biophys. J. **64**, 1766-1780 (1993).

[237] M. Saxton, *Single-particle tracking: models of directed transport*, Biophys. J. **67**, 2110-2119 (1994).

[238] M. J. Saxton, *Anomalous diffusion due to binding: a Monte Carlo study*, Biophys. J. **70**, 1250-1262 (1996).

[239] M. J. Saxton and K. Jacobson, *Single-particle tracking: Applications to Membrane Dynamics*, Annu. Rev. Biophys. Biomol. Struct. **26**, 373-399 (1997).

[240] M. J. Saxton, *Single-particle tracking: the distribution of diffusion coefficients*, Biophys. J. **72**, 1744-1753 (1997).

[241] C. Scalliet, A. Gnoli, A. Puglisi, and A. Vulpiani, *Cages and Anomalous Diffusion in Vibrated Dense Granular Media*, Phys. Rev. Lett. **114**, 198001 (2015).

[242] M. Scholz et al., *Cycling State that Can Lead to Glassy Dynamics in Intracellular Transport.*, Phys. Rev. X 6 (2016).

[243] S. S. Shapiro, and M. B. Wilk, *An Analysis of Variance Test for Normality (Complete Samples)*, Biometrika, **52**, 3/4, 591-611 (1965).

[244] B. O´Shaughnessy and I. Procaccia, *Analytical solutions for diffusion on the fractal objects*, Phys. Rev. Lett., **54** (5), 455458 (1985).

[245] D. Shoup and A. Szabo, *Role of diffusion in ligand binding to macromolecules and cell-bound receptors*, Biophys. J. **40**, 33 (1982).

[246] G. Sikora, M. Teuerle, A. Wyłomańska, and D. S. Grebenkov, *Statistical properties of the anomalous scaling exponent estimator based on time-averaged mean-square displacement*, Phys. Rev. E, **96**, 022132 (2017).

[247] P. J. Slator, N. Burroughs, *A Hidden Markov Model for Detecting Confinement in Single Particle Tracking Trajectories*, (2018) doi: https://doi.org/10.1101/275107.

[248] J. Ślęzak and A. Weron, *From physical linear systems to discrete-time series. A guide for analysis of the sampled experimental data*, Phys. Rev. E **91**, 053302 (2015).

[249] M. Von Smoluchowski, *Zur kinetischen Theorie der Brownschen Molekularbewegung und der Suspensionen*, Ann. Phys. **326**, 756 (1906).

[250] T. L. Snyder and J. M. Steele, *Convex hulls of random walks*,Proc. Am. Math. Soc. **117**, 1165-1173 (1993).

[251] I. M. Sokolov, *Models of anomalous diffusion in crowded environments*, Soft Matter **8**, 9043-9052 (2012).

[252] I. M. Sokolov, *What is the alternative to the AlexanderOrbach relation?*, J. Phys. A: Math. Theor., **49**, 095003 (2016).

[253] M. Spanner, F. Höfling, S. C. Kapfer, K. R. Mecke, G. E. Schröder-Turk, and T. Franosch, *Splitting of the universality class of anomalous transport in crowded media*, Phys. Rev. Lett., **116**, 060601 (2016).

[254] F. Spitzer and H. Widom, *The circumference of a convex polygon*, Proc. Amer. Math. Soc. **12**, 506-509 (1961).

[255] D. Stauffer and A. Aharony, *Introduction to Percolation Theory* 2nd Ed. (London: Taylor & Francis, 2003).

[256] B. Stuhrmann, E. Soares, M. Silva, M. Depken, F. C. MacKintosh and G. H. Koenderink, *Nonequilibrium fluctuations of a remodeling in vitro cytoskeleton*, Phys. Rev. E, **86**, 15 (2012).

[257] T. Sungkaworn, M.-L. Jobin, K. Burnecki, A. Weron, M. J. Lohse, and D. Calebiro, *Single-molecule imaging reveals receptorG protein interactions at cell surface hot spots*, Nature **550**, 543547 (2017).

[258] J. Szymanski and M. Weiss, *Elucidating the Origin of Anomalous Diffusion in Crowded Fluids*, Phys. Rev. Lett. **103**, 038102 (2009).

[259] J. Tailleur and M. E. Cates, *Statistical mechanics of interacting run-and-tumble bacteria*, Phys. Rev. Let. **100**, 218103 (2008).

[260] L. Takács, *Expected perimeter length*, Amer. Math. Month. **87**, 5-6 (1980).

[261] V. Tejedor, O. B enichou, R. Voituriez, R. Jungmann, F. Simmel, C. Selhuber-unkel, L. B. Oddershede, and R.Metzler, *Quantitative analysis of single particle trajectories: mean maximal excursion method*, Biophys. J. **98**, 1364 (2010).

[262] F. Thiel, F. Flegel, and I. M. Sokolov, *Disentangling sources of anomalous diffusion*, Phys. Rev. Lett., **111**, 010601 (2013).

[263] F. Thiel and I. M. Sokolov, *Scaled Brownian motion as a mean-field model for continuous-time random walks*, Phys. Rev. E **89**, 012115 (2014).

[264] J. Tóthová, G. Vasziová, L. Glod, and V. Lisá, *Langevin theory of anomalous Brownian motion made simple.* Eur. Jour. Phys., **32(3)**, 645655 (2011).

[265] I. M. Tolić-Nørrelykke, E.-L. Munteanu, G. Thon, L. Oddershede, and K. Berg-Sørensen, *Anomalous Diffusion in Living Yeast Cells*, Phys. Rev. Lett. **93**, 078102 (2004).

[266] J. A. Torreno-Pina, B. M. Castro, C. Manzo, S. I. Buschow, A. Cambi, and M. F. Garcia-Parajo, *Enhanced receptor-clathrin interactions induced by N-glycan-mediated membrane micropatterning*, Proc. Nat. Acad. Sci. USA **111**, 11037-11042 (2014).

[267] T. Toyota, D. A. Head, C. F. Schmidt and D. Mizuno, *Non-Gaussian athermal fluctuations in active gels*, Soft Matter, **7**, 3234 (2011).

[268] C. Tricot, *Courbes et dimension fractale*, Springer (1999).

[269] S. Türkcan, A. Alexandrou, and J.-B. Masson, *A Bayesian Inference Scheme to Extract Diffusivity and Potential Fields from Confined Single-Molecule Trajectories*, Biophys. J. **102**, 2288-2298 (2012).

[270] S. Türkcan and J.-B. Masson, *Bayesian decision tree for the classification of the mode of motion in single-molecule trajectories* PloS ONE **8**, 0082799 (2013).

[271] R. D. Vale, *The Molecular Motor Toolbox for Intracellular Transport*, Cell **112**, 467480 (2003).

[272] A. Valle, M. A. Rodriguez and L. Pesquerra, *Diffusion in a medum with space-correlated disorder*, Phys. Rev. A, **43**, 948 (1991).

[273] C. L. Vestergaard, P. C. Blainey, and H. Flyvbjerg, *Optimal estimation of diffusion coefficients from single-particle trajectories*, Phys. Rev. E **89**, 022726 (2014).

[274] G. Voisinne, A. Alexandrou, and J.-B. Masson, Biophys. J. **98**, 596 (2010).

[275] P. H. von Hippel and O. G. Berg, *Facilitated target location in biological systems*, J. Biol. Chem. **264**, 675-678 (1989).

[276] K. G. Wang, *Long-time-correlation effects and biased anomalous diffusion*, Phys.Rev. A, **45**, 833 (1992).

[277] K. Wang, J. Masoliver, and J. M. Porra, *Generalized Langevin equations: Anomalous diffusion and probability distributions*, Phys. Rev. E, **53(6)**, 58725881 (1996).

[278] B. Wang, S. M. Anthony, S. C. Bae and S. Granick, *Anomalous yet Brownian*, Proc. Nat. Acad. Sci., **106**, 15160-15164, (2009).

[279] B. Wang, J. Kuo, S. C. Bae and Granick, *When Brownian diffusion is not Gaussian*, Nat. Mater. **11**, 481-485 (2012).

[280] A. V. Weigel, B. Simon, M. M. Tamkun, and D. Krapf, *Ergodic and nonergodic processes coexist in the plasma membrane as observed by single-molecule tracking*, Proc. Nat. Acad. Sci. USA **108**, 6438-6443 (2011).

[281] M. Weiss, M. Elsner, F. Kartberg, and T. Nilsson, *Anomalous subdiffusion is a measure for cytoplasmic crowding in living cells.*, Biophys. J., **87**, 35 1824 (2004).

[282] D. Ernst, M. Hellmann, J. Khler, and M. Weiss, *Fractional Brownian motion in crowded fluids*, Soft Matter **8**, 4886 (2012).

[283] M. Weiss, *Single-particle tracking data reveal anticorrelated fractional Brownian motion in crowded fluids*, Phys. Rev. E **88**, 010101 (2013).

[284] A. Weron and M. Magdziarz, 2009, *Anomalous diffusion and semimartingales*, Eur. Phys. Lett., **86**, 60010 (2009).

[285] C. Wilhelm, *Out-of-Equilibrium Microrheology inside Living Cells*, Phys. Rev. Lett. **101**, 028101 (2008).

[286] D. Wirtz, *Particle-Tracking Microrheology of Living Cells: Principles and Applications*, Ann. Rev. Biophys. **38**, 301-326 (2009).

[287] T. A. Witten and L. M. Sander, *Diffusion Limited Aggregation, a kinetic critical phenomenon*, Phys. Rev. Lett. **47**, 1400 (1981).

[288] I. Y. Wong, M. L. Gardel, D. R. Reichman, E. R. Weeks, M. T. Valentine, A. R. Bausch, and D. A. Weitz, *Anomalous Diffusion Probes Microstructure Dynamics of Entangled F-Actin Networks*, Phys. Rev. Lett. **92**, 178101 (2004).

[289] B. J. Worton, *A convex hull based estimator of homerange size*, Biometrics **51**, 1206-1215 (1995).

[290] W. Wyss, *The fractional diffusion equation*, Jour. Math. Phys. **27**, 2782 (1986).

[291] B. W. Yap and C. H. Sim,*Comparison of various types of normality tests.*, Jour. Stat. Comp. Sim., **81**, 2141-2155, (2011).

[292] S. B. Yuste and Katja Lindenberg, *Comment on 'Mean first passage time for anomalous diffusion"*, Phys. Rev. E 69, 033101 (2004).

[293] R. Zwanzig and A. Szabo, *Time dependent rate of diffusion-influenced ligand binding to receptors on cell surfaces*, Biophys. J. **60**, 671 (1991).

Titre : Révéler les mécanismes de transport à partir d'une trajectoire individuelle dans les cellules vivantes

Mots clés : Analyse de trajectoires individuelles, diffusion anormale, processus intermittents, diffusion non-Gaussienne

Résumé : De nombreuses expériences visant à comprendre les mécanismes de transport dans le milieu intracellulaire reposent sur le suivi de traceurs individuels. Cette thèse vise à déduire des informations sur le milieu cellulaire, le traceur observé et leurs interactions à partir de l'analyse statistique d'une trajectoire. Afin de révéler la non-ergodicité à partir d'une seule trajectoire, un estimateur basé sur la fonctionnelle dynamique est proposé. Aussi, un algorithme est introduit pour reconnaître les différents états d'un processus intermittent. Par l'analyse des données expérimentales, nous révélons que les traceurs explorant le cytoplasme montrent une distribution non Gaussienne des déplacements, témoignant des hétérogénéités de diffusivité dans le cytoplasme. Pour rationaliser ces résultats, un modèle où la diffusivité est un processus stochastique est analysé. Les outils d'analyse son finalement utilisés pour analyser le mouvement aléatoire de disques posés sur un plateau vibrant.

Title : Revealing the transport mechanisms from a single trajectory in living cells

Keywords : Single trajectory analysis, anomalous diffusion, intermittent processes, non-Gaussian diffusion

Abstract : Many experiments designed to understand the transport mechanisms in the intracellular environment rely on the recording the motion of individual tracers. This thesis aims to deduce information on the cellular environment, the observed tracer and their interactions from the statistical analysis of a trajectory. In order to reveal ergodicity breaking from a single trajectory, an estimator based on dynamical functional is proposed. Also, an algorithm is introduced to recognize the different states of an intermittent process. Through the analysis of experimental data, we revealed that tracers exploring the cytoplasm show a non-Gaussian distribution of displacements, reflecting the heterogeneity of diffusivity in the cytoplasm. To rationalize these results, a model where diffusivity is a stochastic process is analyzed. The single trajectory analysis tools are finally used to analyze the random motion of disks placed on a vibrating plate.

Université Paris-Saclay
Espace Technologique / Immeuble Discovery
Route de l'Orme aux Merisiers RD 128 / 91190 Saint-Aubin, France