# Disentangling Sources of Anomalous Diffusion

Dissertation
zur Erlangung des akademischen Grades

doctor rerum naturalium
(Dr. rer. nat.)

im Fach Physik, Spezialisierung:
*Theoretische Physik*

eingereicht an der
Mathematisch-Naturwissenschaftlichen Fakultät
der Humboldt-Universität zu Berlin

von
Herrn M.Sc. Felix Thiel,

Präsidentin/Präsident der Humboldt-Universität zu Berlin:
Prof. Dr. Jan-Hendrik Olbertz

Dekanin/Dekan der Mathematisch-Naturwissenschaftlichen Fakultät:
Prof. Dr. Elmar Kulke

---

Gutachter/-innen:
1. Prof. Dr. Igor M. Sokolov
2. Prof. Domingos H. U. Marchetti
3. Prof. Dr. Günter Radons

Tag der mündlichen Prüfung: 24. September 2015

*Auch dies sind Sprüche Salomos; die Männer Hiskias, des Königs von Juda, haben sie gesammelt. Es ist Gottes Ehre, eine Sache zu verbergen; aber der Könige Ehre ist es, eine Sache zu erforschen.*

Spr. 25,1–2, Luther-Übersetzung 1984

**Zusammenfassung:** Zufällige Bewegungen wie Diffusion sind ein allgegenwärtiges Phänomen, anzufinden nicht nur in der Physik. Das Hauptobjekt von Diffusionsmodellen ist oft die mittlere quadratische Verschiebung eines Teilchens, welche für sogenannte normal-diffusive Prozesse linear mit der Zeit anwächst. Anomale Diffusion bezeichnet Prozesse, für welche sie nicht-linear wächst; ein wichtiges Beispiel ist die Bewegung großer Moleküle in biologischen Zellen. Erscheinungen wie schwache Ergodizitätsbrechung sind ebenfalls bei anomaler Diffusion zu finden, und es gibt viele mathematische Modelle zu ihrer Beschreibung. Oft ist es schwierig für ein bestimmtes Experiment das „richtige“Modell, d.h. die physikalische Ursache der Anomalie, zu finden. Eine Methode zur Trennung oder Identifikation der physikalischen Ursachen wird also dringend benötigt.

In dieser Arbeit stellten wir uns diesem Problem. Zuerst betrachteten wir ein recht allgemeines Modell zur Diffusion in ungeordneten Medien. Mithilfe der Netzwerktheorie trennten wir zwei Mechanismen, nämlich energetische und strukturelle Unordnung, welche beide zu anomaler Diffusion führen. Diese Klassen wurden dann in die Sprache der stochastischen Prozesse übertragen. Das erlaubte uns eine einfache Methode, die des fundamentalen Momentes, zu formulieren. Jene Methode ist in der Lage die energetischen und strukturellen Anteile eines Diffusionsprozesses voneinander zu trennen. Zuletzt behandelten wir Ergodizität und Ergodizitätsbrechung aus der Sicht der energetischen und strukturellen Unordnung.

**Abstract:** Random motion, in particular diffusion, is a ubiquitous phenomenon that is encountered not only in physics. The main object of a diffusion model is usually the mean squared displacement (msd) of a particle, which for so-called normal diffusion grows linearly in time. Anomalous diffusion denotes processes, in which the msd grows non-linearly; an important example is the motion of large molecules in biological cells. Many interesting properties like weak ergodicity breaking are connected to anomalous diffusion, and there are many mathematical models exhibiting anomalous behaviour. Given an experiment, it is often difficult to decide, what is the “correct” model, i.e. the physical cause for the anomaly. Therefore, a method capable of separation and identification of different physical mechanisms is urgently required.

This thesis approached the mentioned issue. First of all, we considered a quite general model for diffusion in disordered media. We used some network theory to distinguish two physical mechanisms – energetic and structural disorder. Both cause anomalous diffusion. Those classes of disorder were then translated into the language of stochastic processes. This put ourselves in position to propose a simple method, the fundamental moment, that is capable of separating the energetic and structural components of a diffusion process. At last, we discussed ergodicity and ergodicity breaking from the point of view of energetic and structural disorder.

# Contents

# 0. Tables

## 0.1. Table of abbreviations

For aesthetic reasons, abbreviations are only capitalized, when the original word starts with a capital. This is the case in the beginning of a sentence or with personal names.

| Abbreviation | Meaning |
|---|---|
| msd | mean squared displacement |
| frap | fluorescence recovery after photo-bleaching |
| fcs | fluorescence correlation spectroscopy |
| pdf | probability density function |
| spt | single particle tracking |
| tamsd | time averaged mean squared displacement |
| eamsd | ensemble averaged mean squared displacement |
| ctrw | continuous time random walk |
| fBm | fractional Brownian motion |
| sfBm | subordinated fractional Brownian motion |

## 0.2. Table of important symbols

Vectors are denoted with bold symbols. Random variables are denoted by capital Latin letters. Absolute positions are denoted by Latin letters, e.g. $\boldsymbol{x}$ or $t$, but displacements or changes are given by the corresponding Greek letters, i.e. $\boldsymbol{\xi}$ or $\tau$. The mathematical object of a Brownian motion is called a "Wiener process" here and denoted with $\boldsymbol{B}(t)$. It is the Gaussian process with almost surely continuous trajectories and $\langle \boldsymbol{B}(t)\, \boldsymbol{B}(t') \rangle = \min(t, t')$. The physical dimension of $\boldsymbol{B}(t)$ is the square root of its arguments' dimension. Function arguments are separated by a semicolon. When the function is symmetric with respect to interchange of its arguments, they are separated with a comma. Definitions are denoted with $:=$. The right-hand side of the symbol is the definition. The symbol $\sim$ means "proportional to, i.e. equal up to a factor. The symbol $\propto$ means "asymptotically equal to". Diffusivities, denoted with $K_{\gamma,\gamma'}$, have the physical dimension of $\mathrm{m}^{\gamma}/\mathrm{sec}^{\gamma'}$.

| Symbol | Meaning |
|---|---|
| $\boldsymbol{x}$, $\boldsymbol{\xi}$ | position in space, and a spacial displacement or translation |
| $t$, $\tau$ | epoch in physical time and a corresponding time difference |
| $u$, $\omega$ | epoch in operational time, and a corresponding time difference |
| $T$ | maximum measurement time |
| $d$ | dimension of Euclidean space |
| $\Omega$ | state space, usually a simple cubic lattice $\mathbb{Z}^d$ or the real space $\mathbb{R}^d$ |
| $T(\Omega)$ | the set of all possible translations of $\Omega$, usually the same as $\Omega$ |
| $\rho(\boldsymbol{x};t)$ | concentration profile or the pdf |
| $\eta(\boldsymbol{x};t)$ | activity or $\rho(\boldsymbol{x};t)\,/\rho_{\mathrm{eq}}(\boldsymbol{x})$ |
| $a$ | lattice constant |
| $K_{\gamma,\gamma'}$ | anomalous diffusivity, the subscripts indicate the physical dimension of $\mathrm{m}^\gamma/\mathrm{sec}^{\gamma'}$ |
| $\triangle$, $\boldsymbol{\nabla}$, $\boldsymbol{\nabla}^\dagger$ | Laplacian, usually a *lattice* Laplacian, incidence operator of a graph and its adjoint, sometimes gradient or divergence |
| $\beta$ | inverse temperature |
| $\lambda$, $w$ | transition rate and equilibrium transition rates |
| $A$ | adjacency matrix of the transition graph |
| $e^{-\beta\mathscr{F}}$ | Boltzmann factor |
| $\boldsymbol{X}(t)$ | position of a tracer particle, a stochastic process of mixed disorder |
| $\boldsymbol{Y}(t)$ | process of pure structural disorder, i.e. with stationary increments |
| $U(t)$ | operational time, i.e. a non-decreasing process |
| $\boldsymbol{B}(t)$ | Wiener process |
| $\Delta\boldsymbol{X}(t;t+\tau)$ | $\boldsymbol{X}(t+\tau)-\boldsymbol{X}(t)$, displacement of $\boldsymbol{X}$ in the time interval $[t,t+\tau]$ |
| $\nu$ | exponent of structural disorder, i.e. $\langle\lvert\boldsymbol{Y}(t)\rvert\rangle\sim t^\nu$ |
| $\alpha$ | exponent of energetic disorder, i.e. $\langle U(t)\rangle\sim t^\alpha$ |
| $\mathbb{P}\{\cdot\}$ | probability of an event |
| $\mathbb{P}\{\cdot\lvert:\}$, $\mathbb{E}[\cdot\lvert:]$ | conditional probability and conditional average |
| $S_Z(z)$ | survival function, i.e. $S_Z(z):=\mathbb{P}\{Z\geq z\}$ |
| $\langle\cdot\rangle$ | ensemble average |
| $[\cdot]_p$ | $p$-quantile, i.e. $p=S_Z([Z]_p)$ |
| $\mathrm{Var}[\cdot]$, $\mathrm{Cov}[\cdot,:]$ | variance and covariance, $\mathrm{Cov}[Z,Z']:=\langle ZZ'\rangle-\langle Z\rangle\langle Z'\rangle$, $\mathrm{Var}[Z]:=\mathrm{Cov}[Z,Z]$ |
| $\mathbb{E}[\cdot]$ | average over disorder |
| $\overline{O}_T$ | time average of $O$, see Eq.(4.1) |
| $J_O(T)$ | ensemble heterogeneity of $O$, see Eq.(4.4) |
| $o(\cdot)$ | small Landau symbol, $f=o(g)$, if and only if $f(x)\,/g(x)\to 0$ |
| $O(\cdot)$ | big Landau symbol, $f=O(g)$, if and only if $f(x)\,/g(x)\to c$ |

# 1. Introduction

## 1.1. A small historical overview

A small history of random motion usually starts with Robert Brown who reported in 1827 on the erratic motion of pollen in water under the microscope. Although he was not the first to observe such motion, the most prominent object in the theory of random motion, the Brownian motion, is named after him. In those early days the question was whether the motion was a biological or physical phenomenon, and already in 1784 and 1819 Ingenhousz and Bywater, respectively, argued (independently) that the motion was also present for inorganic particles, thus a physical phenomenon. In the next years, several experimenters devoted their work to this phenomenon, trying to elucidate the cause of this motion as well as its dependence on parameters like temperature or viscosity of the fluid or the size of the solved particles.[1]

The first mathematical description is due to Fick, who experimented with sodium chloride solutions, [4]. From the assumption that differences in the concentration $\rho$ of the salt induce a diffusion flux, he derived the diffusion equation

$$\frac{\mathrm{d}\rho}{\mathrm{d}t} = K_{2,1} \sum_{j=1}^{d} \frac{\mathrm{d}^2\rho}{\mathrm{d}x_j^2}.$$

$d$ is the dimension of Euclidean space. The diffusion constant $K_{2,1}$ has the physical dimension $\mathrm{m}^2/\mathrm{sec}$.[2] He validated the equation with a series of careful experiments and proceeded to research the diffusion through membranes taken from some animal's bladder. His assumption on the diffusion flux and the resulting equation presented here are now known as Fick's first and Fick's second law. Although the transport process of solved salt in water was not well investigated back then, the governing transport equation had already been well known as Fourier's law of thermal conduction. This equation appears abundantly when considering transport processes like diffusion, heat or charge transport.

Fick helped himself explaining the phenomenon with the notion of atoms and molecules. He imagined an array of forces that acts between atoms of the same and of different kind and rather than thinking of a random motion, he thought about a complex but deterministic motion of the latter.[3] It was Einstein who brought the concept of random walks into the discussion, as he considered the problem of a solution

[1] The information presented here was compiled from [1], the introduction of [2], and [3].

[2] The subscripts of the diffusivity shall indicate the physical dimension.

[3] Regarding the fact that the solved particles considered by Fick are much smaller than those of the prior mentioned authors, this conception is not surprising, but rather natural.

of small spherically shaped particles in 1905, [5]. From the concept of small random displacements or forces acting on the particle, he re-derived the diffusion equation. He bridged the gap between the description via concentration and the stochastic process, which describes the trajectory of the particle. In fact the random process he considered was the Brownian motion, named after prior mentioned botanist Robert Brown. Later Smoluchowski, [6], formulated the problem completely in terms of a random walk, and Langevin proposed a classical equation of motion with a random force term, [7], which is now known as a stochastic differential equation. The discussion about the Brownian motion, i.e. the erratic motion of small solved particles, now circled around the molecular-kinetic theory, hence the existence of atoms. The ideas of Einstein and Smoluchowski were tested systematically by Perrin, [8], and led to the establishment of the atomistic hypothesis.

Random concepts also yielded fruits in other fields like electric circuit theory to describe electrical noise by Schottky in 1918, [9]. Also, already five years before Einstein, Bachelier analysed fluctuations in stock prices, [10]. Therefore, he was the first to apply the Brownian motion process.

## 1.2. Random motion is abundant

Apparently there is a deep connection between random processes, like the motion of solved particles or stock prices, and certain partial differential equations, like Fick's second law. This allows us to describe random trajectories by such equations and vice versa it allows us to interpret such equations in terms of a randomly moving agent. Consequently, the investigation of transport equations and processes bears not only fruits in the physical sciences, but in any discipline where a certain "random motion paradigm" may be applied. Loosely spoken this paradigm states that:

*Stuff changes randomly.*

It can be used regardless of how chance enters the description. It may enter by inherently stochastic dynamics, by not knowing the exact dynamics – like in statistical mechanics – or even by not caring about the exact dynamics. Thus, the same tools used to describe a sodium chloride solution are used nowadays to model travelling people [11, 12], spreading epidemics [13, 14], the motion of animals [15–34], and the evolution of financial stocks [35–39]. All these topics caught the interest of physicists that are familiar with the necessary mathematical tools. When describing natural phenomena, the random motion paradigm applies to the motion of particles in all sorts of environments, for example in solid matter, see [40] and references therein, porous media [41–43], or in soft matter, like in soils [44], in polymer matrices, [45, 46], visco-elastic or other complex media [3, 47–51], or in general disordered media [40, 52–56]. In particular the motion of large molecules like proteins in biological cells is a recent field of study, [49, 57–73]. In polymer science one may describe the motion of polymers in polymer melts [74–77], the translocation of polymers through small pores [78, 79], or the coiling of polymers [80–82]. The most classical example of a random

walk in polymer science is of course the end-to-end-distance of a polymer chain. Also there is tracer diffusion in turbulence [83–86], active particles [87], or the motion of electrical charges or excitons [52, 88–90]. In physical chemistry, the diffusion process governs the speed of a chemical reaction of dilute agents. Therefore, understanding the agents' motion is essential, [6, 91–93]. More abstractly, the (unknown) micro-state of a complex system changes over time and can be described in terms of a random walk in state space. This interpretation is applied among other situations in glasses and large polymers, [94–104].

In all those fields, a random description successfully describes the phenomenon at hand. Whenever this is the case, the same family of mathematical tools is applied. One may argue whether those tools are particularly versatile, their field of application most certainly is. Notwithstanding that these tools can be applied almost blindly with great success, care has to be taken when doing so. Whenever it is not clear to the modeller, where and how chance enters his description, he may propose a model that fits his data but still gives wrong predictions. Even more severe, the underlying (non-)physical conception may be wrong. Such misconceptions can propagate and lead to confusion, when not unveiled by other experiments. The diffusion of large molecules in biological cells or other complex environments is an area of contemporary research, where there is a lot of discussion about the correct laws governing the motion. This will be the main motivation for this thesis. But before turning to this topic, we discuss some of the experimental techniques and accessible observables that are currently in use.

## 1.3. A theorist's overview over experimental techniques – commonly used quantities

In the first experiments on Brownian motion, it was tried to measure the velocity of the diffusing particles. This was not a very fruitful approach as a Brownian particle usually does not have a well defined velocity. The work of Einstein and Smoluchowski directed the attention to the *displacement*, rather than the velocity, of the particle.[4] This quantity was exactly what Perrin measured.[5] He took stroboscopic pictures of a well prepared ensemble of small rubber balls. It was then possible to obtain histograms of the displacements and to compute their variance, the *mean squared displacement* (msd). According to Einstein, and also Fick's theory, when $\boldsymbol{X}(t)$ denotes the position of a particle at time $t$ and $\Delta\boldsymbol{X}(t;\tau) := \boldsymbol{X}(t+\tau) - \boldsymbol{X}(t)$ the displacement, then the mean squared displacement is given by

$$\left\langle \Delta\boldsymbol{X}^2(t;t+\tau)\right\rangle = 2dK_{2,1}\tau.$$

This linear relation between msd and time arises wherever the diffusion equation appears. Since it was the first relation to be investigated, processes that exhibit such

[4] According to [1].

[5] Again according to [1], Perrin did not do it with his first intention.

behaviour are called *normal diffusive*. The square root of the msd defines a typical transport length for a lag time $\tau$.

Although the msd still is in the focus of investigation, experimental techniques have improved drastically. This is due to advanced labelling techniques using green fluorescent proteins or quantum dots, more powerful microscopes as well as cameras and computer based image processing. Nowadays, the spatial resolution is in the nanometre to sub-micrometre range. We will shortly present three techniques here. Of course there are much more, but the three will be used to introduce three of the main objects discussed in this thesis.

**Fluorescence recovery after photo-bleaching:** First of all, there are the fluorescence recovery after photo-bleaching (frap) experiments, [62]. An ensemble of tracer particles is marked with fluorescents. A short and strong laser beam deactivates all markers in a certain area and one then observes fluorescent particles diffusing back into the bleached region. In this experiment, the overall fluorescence intensity is recorded. The so obtained recovery curve is fitted to a theoretical prediction and the msd may be extracted from the data.

**Fluorescence correlation spectroscopy:** In these fluorescence correlation spectroscopy (fcs) experiments the whole fluorescence intensity is recorded, [105]. By exploiting the fact that the local intensity is proportional to the concentration of fluorescent, one can connect the fluctuation of intensity to the fluctuation of tracer concentration. This allows to measure not only the msd but also spatial and temporal correlations, hence, to probe the propagator.

**Single particle tracking:** In single particle tracking experiments (spt), individual particles are marked, their trajectories are recorded with a high resolution, which allows to reconstruct individual paths, [69, 106]. Obviously this method produces the most information, namely the trajectory $\{\boldsymbol{X}(t)\}_{t\in[0,T]}$ in the measurement interval $[0,T]$, but it is also the most complex in preparation. With the rise of single particle tracking it became possible and necessary to not only think about ensemble properties, but also about time averaged quantities obtained from a single trajectory. Hence, it is nowadays possible to address problems like ergodicity, i.e. the equivalence between time and ensemble averaging procedures, in complex environments.

Let us quickly discuss, what the propagator is. Modern diffusion models are usually formulated either in terms of a stochastic process, i.e. by making assumptions on the properties of the trajectories, or alternatively in terms of a partial or integro-differential equation that describes the evolution of the concentration. These equations are generalizations of Fick's second law. One can find the Greens' function $G(\boldsymbol{x}; t_0+\tau|\boldsymbol{x}_0; t_0)$ of such an equation, which then describes the evolution of concentration initially distributed in one position, namely in $\boldsymbol{x}_0$. In the stochastic picture this means that we fix

the initial point of the trajectory to $\boldsymbol{X}(t_0) = \boldsymbol{x}_0$, and consequently the Greens' function or propagator can be interpreted as the conditional probability of displacement. Another name for this quantity is also transition probability, fundamental solution or heat kernel. Integrals of the form

$$\int_\Omega \mathrm{d}\boldsymbol{x} \int_\Omega \mathrm{d}\boldsymbol{x}_0 \, I(\boldsymbol{x}) \, I(\boldsymbol{x}_0) \, G(\boldsymbol{x}; t_0 + \tau | \boldsymbol{x}_0; t_0)$$

bear temporal correlation functions of the position. Those are measured in fcs, where $I(\boldsymbol{x})$ is a laser intensity profile. The propagator often is the starting point of a theory. The other approach – to directly describe the trajectories' properties – is either performed by writing a stochastic differential equation, a Langevin equation, or by directly enumerating all important properties. A Langevin equations is, roughly spoken, an equation of motion with a random force, the so called noise term. When the properties of the noise are known, the msd, correlation functions or the propagator can be derived. Both approaches, starting with an equation for the propagator or with an equation for the trajectories are equivalent only under certain conditions.

Over time, the experimental techniques became more and more involved and enabled access to more intricate observables, from the msd, via correlation functions to the whole trajectory. Still, the msd and its relatives remain the most important observables. Another member of this family is the *time averaged mean squared displacement* (tamsd):

$$\overline{\Delta \boldsymbol{X}^2(\tau)}_T := \frac{1}{T - \tau} \int_0^{T-\tau} \mathrm{d}t \, \Delta \boldsymbol{X}^2(t; t + \tau) \, .$$

In contrast to the before discussed *ensemble averaged mean squared displacement* (eamsd). In time averages, one averages over subsequent time intervals of one trajectory, instead of averaging over an ensemble of different tracer's trajectories. In the same way as the eamsd fluctuates around its unknown true value for finite ensembles, the tamsd exhibits fluctuations for finite measurement times. The single particle tracking method puts the experimenter in the position to verify the ergodic hypothesis, which states the equivalence of both methods.

Over time, systems with a much richer behaviour than just a linear time dependence of the msd were discussed. Those are usually coined *anomalous diffusion*. In the next section we will discuss the traits of anomalous diffusion.

## 1.4. Features of anomalous diffusion

### 1.4.1. Non-linearity of the msd

A linear behaviour of the msd appears often but it is not universal. It is sufficient to take inertia effects of the diffusing particles into account, in order to generate other

than linear behaviour. One then observes that the diffusivity $K_{2,1}$ introduced above is not constant, but rather depends on the measurement and averaging details, on the time $t$ when the observation is started, or on the lag time $\tau$, for which the msd is evaluated. One of the first to address this problem, was Richardson in 1926, [84], who observed that the measured diffusivities of substances in atmospheric flows differ by nine orders of magnitude. To overcome this problem, he proposed an equation with a length-scale dependent diffusivity, i.e.

$$\frac{\mathrm{d}\rho}{\mathrm{d}t} = K_{2/3,1}\frac{\mathrm{d}}{\mathrm{d}x}\left|x\right|^{\frac{4}{3}}\frac{\mathrm{d}\rho}{\mathrm{d}x}$$

here presented in one dimension. The constant $K_{2/3,1}$ does not have the dimensions of a diffusivity, but rather $\mathrm{m}^{2/3}/\mathrm{sec}$. Therefore, it is hence an anomalous diffusivity. The msd in Richardson's model is given by

$$\left\langle \Delta \boldsymbol{X}^2(0;\tau)\right\rangle = cK_{2/3,1}^3\tau^3,$$

with some numerical constant $c$. It is obvious that a diffusion constant defined by $\langle \Delta \boldsymbol{X}^2(0;\tau)\rangle/(2d\tau)$ is not a constant.

An alternative approach by Batchelor for the same problem involved an a-priori time dependent diffusivity. He proposed:

$$\frac{\mathrm{d}\rho}{\mathrm{d}t} = K_{2,3}t^2\frac{\mathrm{d}^2\rho}{\mathrm{d}x^2},$$

with the anomalous diffusivity $K_{2,3}$ of dimension $[K_{2,3}] = \mathrm{m}^2/\mathrm{sec}^3$, which bears the same scaling of the msd:

$$\left\langle \Delta \boldsymbol{X}^2(0;\tau)\right\rangle = 2K_{2,3}\tau^3.$$

When the msd does not scale linearly, like it does here, one speaks about anomalous diffusion. In many physical fields, most prominently turbulence and transport in disordered systems, a non-linear time dependence of the msd can be observed. Usually the result is fitted to a power law

$$\left\langle \Delta \boldsymbol{X}^2(0;\tau)\right\rangle = 2dK_{2,\gamma}\tau^\gamma$$

for some $\gamma \neq 1$. The anomalous diffusivity $K_{2,\gamma}$ has to have the physical dimension $\mathrm{m}^2/\mathrm{sec}^\gamma$. If $\gamma$ is smaller than unity, i.e. if the spread is slower than in the normal case, one speaks of *sub-diffusion*. If it is larger the motion is called *super-diffusive*. The special case $\gamma = 2$, when length scale and time scale grow proportionally is termed *ballistic* diffusion. Other relations, for example a logarithmic time dependence, are also possible. Furthermore, there may be different time regimes, in which different scaling relations hold, which are connected by crossover sections. Usually, simple mathematical models are applied to only a few, or just one of those scaling regimes.

Diffusion in atmospheric flows [83–86] is usually superdiffusive. Apart from that, anomalous diffusion also occurs in other fields. In polymer science, a long polymer

chain can be interpreted as a random walk or diffusion process. When the number of monomers is taken as time variable, the end-to-end-distance of the chain scales normally as long as the chain is allowed to overlap with itself. If the repulsive interaction between the monomers is taken into account, the mean squared end-to-end-distance scales super-diffusively, [53, 80]. However, when considering the actual motion of a monomer part of a long chain, the process is subdiffusive, [80, 81]. The rest of the chain drags the monomer back to its previous location, inducing negative correlations in the motion. The same holds for the translocation of a polymer chain through a small pore, [78, 79]. In this situation, the position variable is given by the number of monomers that already passed the pore. It exhibits subdiffusion by the same reason as before. Similar is the process of zipping or unzipping of two DNA strands, [82].

Diffusion in disordered media can be either normal or anomalous, [40, 53], depending on the strength of disorder. Examples for such systems are charge carrier transport in disordered semiconductors [90], the diffusion of water molecules in a polymer matrix [45, 46], or the motion in porous media. Transport in such systems is usually subdiffusive, when the disorder is strong enough. The problem of diffusion in porous media can be mapped to the electrical conduction of an insulator-conductor mixture, [45, 52, 89]. Other subdiffusive physical systems are related to transport in visco-elastic environments, [3, 47–51].

The dynamics of of glasses, polymer conformations, or any other complex system can often be understood in terms of a random walk in state or configuration space, [94–96, 98–101]. In such situations, a classification in terms of the msd makes little sense, for the variable can not be interpreted as a position. Although one may apply the same array of stochastic methods, a classification of the motion in terms of the msd makes little sense. Here one is interested in relaxation behaviour, f.e. how long the system is able to remain its macroscopic state.

Superdiffusive motion appears less often in pure physics, since it requires additional energy coming from an internal or external reservoir. Hence, it is often called active motion. Some particles are able to catalyse a chemical reaction and use it for propulsion, [87, 107]. Superdiffusion is ubiquitous in the motion of living beings. Under the assumption that they move randomly, one finds superdiffusive laws in the motion of bacteria, [17, 21, 26–28, 34, 44], larger animals [23, 25, 30, 31, 33, 108–110], or humans [11–13], for further review see [18, 29]. Emphasis lies especially on the topics of epidemic spreading [14, 111] and random search theory [15, 16, 24, 30, 112, 113].

Anomalous diffusion also occurs in finance, [35–39], where one tries to model stock prices or option prices with stochastic processes.

Apparently, anomalous diffusion has a wide range of applications. But there are also other features besides the non-linear growth of the msd.

### 1.4.2. Weak ergodicity breaking

The examination of the msd provides a first and crude division in normal, slow and fast processes. Furthermore it allows us to define and quantify a transport coefficient,

namely some anomalous diffusivity

$$K_{\gamma,\gamma'} := \frac{\langle |\Delta \boldsymbol{X}(t; t+\tau)|^{\gamma} \rangle}{2d\tau^{\gamma'}},$$

with anomalous physical dimension $\mathrm{m}^{\gamma}/\mathrm{sec}^{\gamma'}$. The measurement procedure provides the details of this definition, for instance which kind of average has to be chosen. Yet, this transport coefficient may not be robust with respect to the chosen measurement technique, for instance it can change with $t$. Then it is not well defined.

One particular problem concerns ergodicity. The ergodic hypothesis is one of the foundations of statistical mechanics and states the equivalence of time and ensemble averages. Boltzmann and Maxwell were the first to use it. They argued that the evolution of a system in phase space carries it close to every region accessible on a certain energy level. By this conception, one can argue that an infinitely long phase space trajectory samples an energy surface thoroughly, so that an average over the time evolution can be replaced by an average over all states in the energy surface. In the early 1930's Birkhoff and von Neumann were able to prove the hypothesis rigorously for non-autonomous dynamical systems and for non-autonomous Hamiltonian systems respectively, [114, 115]. Those results gave birth to mathematical ergodic theory. The topic of ergodicity was reapplied to statistical mechanics in the book of Khinchin, [116]. The various ergodic theorems state the equivalence of time and ensemble averages in various probabilistic senses. Time averages are still random variables from a mathematical point of view. When the measurement interval becomes infinite, they can converge to the ensemble average value for every trajectory, on average or such that their deviations vanish. The usually mentioned precondition for ergodicity is that the whole state space is accessible. If this is not the case, the system is trapped in the accessible domain and a time average will converge to an ensemble average that is conditional to the starting domain. Such a situation is called ergodicity breaking.

Normal diffusive processes show perfect ergodic behaviour and in usual diffusion models the state space is not separated in multiple domains. Thus, it was rather surprising to see that despite of this connectedness, some processes of anomalous diffusion show a non-ergodic behaviour. In particular, there can be a disparity of the time and ensemble averages, [117–119], the distribution of the system does not approach the Boltzmann-Gibbs-distribution, [93, 120], and furthermore the time averaged transport coefficients exhibit universal fluctuations, [117, 121, 122]. A time averaged quantity is said to have universal fluctuations when it behaves randomly and this random nature remains relevant even in the limit of infinite measurement times. The contemporary collective term for these different phenomena is *weak ergodicity breaking*. In one particular model, the continuous time random walk, weak ergodicity breaking causes a linear scaling between tamsd and time, although the eamsd for this model scales subdiffusively. Therefore, more care has to be taken, when interpreting data of anomalous diffusion processes.

Weak ergodicity breaking was investigated in many theoretical models, [123–131]. It can also be observed in experiments, for instance in blinking nano-crystals [102, 132], or in the motion of large molecules inside biological cells [66, 67, 70, 72].

## 1.5. The problem of modelling intracellular transport – How can one distinguish between different models of anomalous diffusion?

Another example for anomalous diffusion is the motion of large molecules in biological cells, for instance on the cytosol or in the cell membrane. These molecules may be lipids, [66], sugars, [73], complex proteins like ion channels, DNA strands or receptors, [63, 67, 72], or other [49, 51, 58–60, 70, 71].

Their motion is usually subdiffusive, although some particles may also move superdiffusively, [59]. The superdiffusion is usually ascribed to active transport by so-called molecular motors. Those are molecules that are able to bind some cargo and drag it along protein fibres inside the cell, [133, 134]. The transport on this actin network consumes chemical energy in the form of adenosine triphosphate.

Subdiffusion on the other hand may be much more challenging to explain. The reason is that there are many factors influencing the process.

First of all, the cytosol as well as the cell membrane, are not "empty", i.e. there are many other molecules, enzymes, organelles, fibres, etc. present in the cell. This situation is usually coined *molecular crowding*, [65]. It manifests in the fact that the larger a particle, the smaller is its mobility, [73]. Large proteins can not move freely, because other things are in the way.

The environment, as seen for such a particle, is highly disordered and may not be static. Polymer solutions with such high densities are also investigated with molecular dynamics simulations and have the properties of glasses, [75–77]. The cytosol also has non-trivial fluid properties, it may be visco-elastic, [49].

In the case of proteins, it may be necessary to change the conformation in order to proceed. This change of conformation can be thermally activated or by reaching some kind of reaction site. The internal state of the protein and its dynamics is, however, not visible in a particle tracking experiment. The observer can not see any motion. In summary, the motion is affected by several factors:

- Active transport by molecular motors,
- obstruction by other parts of the cell,
- visco-elastic interaction with the environment, and
- invisible internal dynamics of the tracer particle (trapping).

Each of those mechanisms alone can be described very well by certain models. Subdiffusion caused by obstructions can be mimicked by the diffusion on a percolation cluster, visco-elastic effects can be modelled by generalized Langevin-equations and trapping effects by means of continuous time random walks. Unfortunately, when realistic situations are considered, there are unfortunately more mechanisms than just one at work simultaneously. The challenge is to identify them.

This identification is the main concern of this thesis. The zoo of anomalous diffusion models needs to be classified. The physical mechanism that is expressed in the mathematical model has to be understood so it can be matched to the experimental situation. By understanding this, one can conceive statistical tests more involved than an investigation of the msd that are able to tell different models, i.e. mechanisms apart. In the final step, a prediction of more complicated properties like ergodicity is possible based on the knowledge of the physics at work in this system. This can be considered as a very rough outline of the thesis.

Since not all models of anomalous diffusion exhibit weak ergodicity breaking, one can use its presence to make a first distinction. First passage times, i.e. the time a tracer particle needs to reach a certain target, may differ from model to model. Hence, they are a powerful tool distinguish them, [135]. Alternatively, the growth of the explored region can be examined, [136]. Furthermore, it was proposed to analyse the $p$-variation of single particle trajectories, [137]. The $p$-variation test discriminates between fractional Brownian motion and continuous time random walks, two particular models. All such tests allow for at least a binary decision between anomalous diffusion models. As already discussed in the example above, it does not have to be just one physical mechanism that needs to be identified. In particular, realistic situations may bear ergodic and non-ergodic components simultaneously, see for example [72]. There, the addition of certain drugs to the cell restored the ergodic behaviour of the tracer particles. Hence there is also a need to separate the ergodic from the non-ergodic part. A method to do so will be proposed in this thesis.

## 1.6. Aim, outline and ansatz of this thesis

The aim of this thesis is threefold. First, we need to find the physical mechanisms that lead to anomalous diffusion. As elaborated before, the most interesting situations are those, where multiple mechanisms are at work simultaneously, in particular ergodic and non-ergodic parts. Then, it is necessary to find an empirically robust method that is able to identify and separate both parts. Finally, the ergodic properties of such mixed processes can be analysed, and the separation is verified.

This programme will be tackled in the framework of a diffusion process in a disordered medium, [40, 53, 138, 139]. Although not the most general model, it is a rather realistic and rich description. Such models are generalizations of random walks on lattices and have a long tradition for instance in solid state physics, see [40]. In the mathematical community they are known as "random-resistor-networks", [140], due to the equivalence to charge transport problems. In chapter 2, we introduce these models and tackle point one: the classification of disorder. We will consider a generalization of the diffusion equation on the lattice $\Omega$,

$$\dot{\rho}(\boldsymbol{x};t) = \sum_{y\in\Omega} \left[\lambda(\boldsymbol{x};\boldsymbol{y})\,\rho(\boldsymbol{y};t) - \lambda(\boldsymbol{y};\boldsymbol{x})\,\rho(\boldsymbol{x};t)\right],$$

with a given random set of transition rates $\{\lambda\}$ between the sites of a regular lattice

$\Omega$. When the transition rates are cleverly chosen, one is able to model internal states' dynamics with sojourn or residence times per lattice site. Obstruction can be modelled with transitions that are forbidden and active transport with connections that gap a long distance.

The fixed transition rates describe a quenched (i.e. stationary) random environment. When performing an average over the random environment, that means when considering the annealed model, the anomalous diffusion becomes visible. Therefore, first the quenched model is discussed in section 2.3.1, then the annealed model in section 2.3.2. In the general case, averaging over the environment or over certain parts of the state space introduces memory effects into the annealed model that – when they are severe enough – contribute to the anomalous behaviour. In that way, visco-elastic effects can be modelled. With a suitable transformation, the annealed model can be mapped to a continuous one, albeit this is not necessary for our reasoning.

The quenched system can be understood in terms of a graph or network, [141–143], of possible transitions. We show how some results of the spectral theory of graphs correspond to the system's properties in thermodynamic equilibrium and help to understand the physical mechanism causing anomalous diffusion. By virtue of our abstract ansatz, the derived results are valid for any network system that can be described with a random master equation like the one above. We make use of the graph's Laplacian matrix, hence every other application involving this object benefits from the discussion in this thesis. However, the merit of diffusion processes is a clear physical interpretation of the state space. Other applications may lack this possibility. In our case, we obtain a classification of two different kinds of random environments, both causing anomalous diffusion, but each corresponding to another physical reason behind the anomaly.

The main object of a theory formulated in terms of a master equation is the concentration profile $\rho$ or equivalently the propagator. The single particle tracking method, on the other hand, operates with trajectories, i.e. stochastic processes. In 2.4, the description in terms of master equations will be mapped to one in terms of stochastic processes. The main tool here is the theory of generators of stochastic processes, in particular Markov-processes, [144, 145]. The physical mechanisms we will identify earlier in the chapter, translate to properties of the stochastic processes. The transition from quenched to annealed disorder in stochastic processes can only be sketched heuristically, due to the mathematical complexity of obtaining proofs. This is especially the case in the presence of long-time memory. As final result of chapter 2, we obtain a decomposition ansatz for stochastic processes that serves as a starting point for the remaining investigations. That way we are firstly able to relate the network model with possibly infinite memory to a stochastic process, and secondly we are able to theoretically identify two different components of this stochastic process.

Several stochastic processes and lattice models that are used in the rest of the thesis will also be introduced along the lines of chapter 2. The case of normal diffusion is treated in the beginning of the chapter, and along its discussion, all variables and terminology are introduced.

In chapter 3, we will demonstrate how the general ansatz on diffusion processes

developed in the first chapter, can be separated by investigation of the so-called fundamental properties. At last in chapter 4, we will discuss ergodic properties in our set-up. The thesis is summarized and concluded in the last chapter, 5.

Furthermore, the thesis is also equipped with a table of abbreviations and a table of symbols in the very beginning. The guidelines for abbreviations and symbols are also mentioned there. Some parts that would distract too much from the main discussion have been moved to the appendix. The main text refers to those sections.

# 2. Classification of disorder

## 2.1. Introduction

When making an attempt to distinguish different mechanisms that lead to anomalous diffusion, we first of all need to understand them. We need to know the mechanisms and after that we can think about how to tell them apart. This is the aim of this chapter. We will first discuss models of diffusion in disordered systems and then turn to models in terms of stochastic processes. By using properties related to thermodynamic equilibrium, we can distinguish two mechanisms, termed energetic and structural disorder.

In the beginning of this chapter, we will introduce all necessary notation and quantities using the example of normal diffusion. Then we turn to lattice models of diffusion in a disordered medium, and see how those quantities need to be generalized. A transition from quenched (in the sense of static or fixed) to annealed (in the sense of averaged) disorder allows us to determine the behaviour of the msd. More specifically, it allows us to see two ways the msd grows *not* linearly in time, hence it behaves anomalously. These two ways are then used to define the terms: Structural and energetic disorder. After that, we will show how the lattice models can be described in terms of stochastic processes, and how the terms structural and energetic disorder translate to that language.

## 2.2. Normal diffusion

The term "anomalous diffusion" implies that there is a term called normal diffusion. Normal diffusion will be presented in this section, i.e. popular models thereof and some of their properties. Alongside the presentation, some notation and basic concepts are introduced and demonstrated.

We focus mainly on two types of models: Those that describe the evolution of the concentration profile and those that specify the properties of single trajectories. The former is usually formulated as a partial differential equation or a set of rate equations, depending on whether discrete or continuous space is considered. The latter may be given by a stochastic differential equation together with reasonable assumptions on the noise, via a random walk scheme, or by directly enumerating all properties of the stochastic process. Either way, the trajectories of the tracer particle are the main object of the model, whereas in concentration models the density of the tracer particles is modelled. These descriptions are equivalent for Markovian models, but not for models with memory. The concentration description is in general not complete.

The usefulness of either one is dictated by its analytical accessibility and by the details of the experiments.

In this section, we will present both approaches in normal diffusion. We first consider a lattice model and afterwards the stochastic process defined in continuous space. In the same line, we introduce all later used notation

## 2.2.1. The binomial random walk

Consider a $d$-dimensional simple cubic lattice $\Omega$ with lattice constant $a$. The function $\rho(\boldsymbol{x};t)$ of a spacial and a temporal argument describes the mass, number or concentration of particles in site $\boldsymbol{x}$ at time $t$. Alternatively it describes the probability for one particle to be in site $\boldsymbol{x}$ at time $t$. Mass, particle number and concentration differ only in a constant factor, as long as each particle does not interact with the others. The same holds for the probability. This assumption is usually correct for low concentrations. In the rest of this text we will refer to the *concentration* $\rho(\boldsymbol{x};t)$, which is normalized to unity, i.e.

$$\sum_{\boldsymbol{x}\in\Omega} \rho(\boldsymbol{x};t) = 1. \tag{2.1}$$

We assume that at a certain epoch 0 all mass is concentrated in a single site, denoted by $\mathbf{0}$, i.e. $\rho_0(\boldsymbol{x}) := \rho(\boldsymbol{x},0) = \delta_{\boldsymbol{x},\mathbf{0}}$.

### The discrete diffusion equation

**Fick's laws and some operators from graph theory:** Pick two neighbouring sites $\boldsymbol{x}$ and $\boldsymbol{y}$ on the site and denote with $j(\boldsymbol{y};\boldsymbol{x};t)$ the current flowing from site $\boldsymbol{x}$ to $\boldsymbol{y}$ at epoch $t$. The continuity equation, or the mass preservation equation, states that the rate of change in concentration in site $\boldsymbol{x}$ is given by the current balance:

$$\dot{\rho}(\boldsymbol{x};t) = \frac{1}{a} \sum_{\boldsymbol{z}:|\boldsymbol{z}-\boldsymbol{x}|=a} j(\boldsymbol{x};\boldsymbol{z};t) =: -\boldsymbol{\nabla}^\dagger j(\boldsymbol{x};t)\,. \tag{2.2}$$

The dot denotes the partial derivative with respect to time. The symbol := denotes a definition, with the defined object on the same site as the colon. The operator $\boldsymbol{\nabla}^\dagger$ is the adjoint incidence operator of the lattice. It sums all currents flowing into the same node and acts as a lattice version of the divergence operator. Following Fick, we may assume that the current $j(\boldsymbol{y};\boldsymbol{x};t)$ is proportional to the difference in concentration:

$$j(\boldsymbol{y};\boldsymbol{x};t) = -\frac{K_{2,1}}{a}\left[\rho(\boldsymbol{x};t) - \rho(\boldsymbol{y};t)\right] =: K_{2,1}\boldsymbol{\nabla}\rho(\boldsymbol{y};\boldsymbol{x};t)\,. \tag{2.3}$$

Here, $K_{2,1}$ is the diffusivity with dimensions $[K] = \mathrm{m}^2/\mathrm{sec}$ and the incidence operator $\boldsymbol{\nabla}$ calculates the difference of a scalar quantity along one bond of the lattice. Putting

both equations together, we obtain the master equation, Fick's second law on the lattice:

$$\dot{\rho}(\boldsymbol{x};t) = \frac{K_{2,1}}{a^2} \sum_{|\boldsymbol{y}-\boldsymbol{x}|=a} [\rho(\boldsymbol{y};t) - \rho(\boldsymbol{x};t)] = -K_{2,1} \boldsymbol{\nabla}^\dagger \boldsymbol{\nabla} \rho(\boldsymbol{x};t) =: K_{2,1} \triangle \rho(\boldsymbol{x};t)\,. \quad (2.4)$$

$\triangle = -\boldsymbol{\nabla}^\dagger \boldsymbol{\nabla}$ is the so-called *lattice Laplacian*.[1] The usage of operators from graph theory and our suggestive notation allows us to symbolically restore the continuous equations.[2] More information about the incidence operators and the lattice Laplacian can be found in literature on graph theory, for instance in [141–143]. The connections to graph theory are a little deeper as we will point out in the next section. The lattice Laplacian is not explicitly time dependent and involves no memory terms. Therefore, Eq.(2.4) is local in time, i.e. the rate $\dot{\rho}$ depends only on the actual concentration profile and is independent of its prior values.

**Stationary state:** We note that Eq.(2.4) admits a stationary concentration profile, which is constant:

$$\rho_{\mathrm{st}}(\boldsymbol{x}) = \mathrm{const.} \quad (2.5)$$

Up to now, the domain $\Omega$ was not discussed. It was simply stated that it is a subset of $a\mathbb{Z}^d$. If the domain is finite, then the initial condition may not be sufficient to define the problem and additional boundary conditions are needed. Those also have to be incorporated into the Laplacian $\triangle$. Our treatment here implies that the domain is infinite (the case of so-called free diffusion) and we do not care about the boundary conditions. We only impose the normalization constraint.[3] The stationary state is only normalizable, if the domain is finite. Nevertheless, any constant function will fulfil $\triangle f = 0$.

**Transition rates and sojourn times:** Let us derive some quantities that will be needed and discussed later on. The transition rate

$$\lambda(\boldsymbol{y};\boldsymbol{x}) := \frac{K_{2,1}}{a^2} \quad (2.6)$$

[1] This operator has non-positive eigenvalues. Hence, it differs in sign from the definition of the graph Laplacian in the mathematical literature!

[2] Loosely spoken, the incidence operator is a "graph version" of the gradient, its adjoint represents the divergence. Both operators are defined for arbitrary graph topologies. Admittedly, a very suggestive notation is used here, but indeed the quantity $\boldsymbol{\nabla}\rho$ represents a current. However, only the Laplacian is uniquely defined, since the sign of the current between two sites may be chosen arbitrarily. In the product of the incidence operator and its adjoint those signs cancel ensuring uniqueness of the Laplacian.

[3] Mathematically, we are defining Eq.(2.4) on a certain subset of the Banach space $\mathcal{L}_1(\Omega)$. The first condition is that all possible values of the concentration are non-negative, i.e. $\rho(\boldsymbol{x};t) \geq 0$. The second condition is normalization, Eq.(2.1), and the last conditions are the ones on the boundary. The common choices are periodic boundary conditions, preserving translation invariance, and reflecting boundary conditions, where the diffusion flux at the boundary vanishes.

is the number of particles per unit time that move from site $\boldsymbol{x}$ to another site $\boldsymbol{y}$. When all possible escape directions are taken into account, one can compute the total number of particles leaving from site $\boldsymbol{x}$ and consequently – by taking the reciprocal – the mean sojourn time of one particle in that site:

$$\tau(\boldsymbol{x}) := \left[ \sum_{\boldsymbol{y}:|\boldsymbol{y}-\boldsymbol{x}|=a} \lambda(\boldsymbol{y};\boldsymbol{x}) \right]^{-1} = \frac{a^2}{2dK}. \tag{2.7}$$

Both quantities are homogeneous (they do not depend on the initial site $\boldsymbol{x}$). Furthermore the transition rate $\lambda(\boldsymbol{y};\boldsymbol{x})$ is isotropic (it does not depend on the target site $\boldsymbol{y}$) and symmetric (with respect to commutation $\boldsymbol{x} \leftrightarrow \boldsymbol{y}$). That means that in stationary state, i.e. when the concentration is a constant, the number of particles going from $\boldsymbol{x}$ to $\boldsymbol{y}$ is the same as the number of particles moving in reverse direction. This defines the equilibrium transition rate:

$$w(\boldsymbol{x},\boldsymbol{y}) := \lambda(\boldsymbol{x};\boldsymbol{y})\, \rho_{\mathrm{st}}(\boldsymbol{y}) = \frac{K_{2,1}}{a^2} \tag{2.8}$$

**The mean squared displacement:** The standard observable in diffusion is the mean squared displacement (msd), the variance of the particles position. It can easily be shown that the msd grows linearly in time (in the case of free diffusion). This is done by inspecting its time derivative and reordering the summation

$$\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}t} \left\langle \boldsymbol{X}^2(t) \right\rangle =& \frac{\mathrm{d}}{\mathrm{d}t} \sum_{\boldsymbol{x}\in a\mathbb{Z}^d} \boldsymbol{x}^2 \rho(\boldsymbol{x};t) = \sum_{\boldsymbol{x}\in a\mathbb{Z}^d} \boldsymbol{x}^2 \dot{\rho}(\boldsymbol{x},t) \\
=& \frac{K_{2,1}}{a^2} \sum_{j=1}^{d} \sum_{\boldsymbol{x}\in a\mathbb{Z}^d} \boldsymbol{x}^2 \left[ \rho(\boldsymbol{x}+a\boldsymbol{e}_j;t) + \rho(\boldsymbol{x}-a\boldsymbol{e}_j;t) - 2\rho(\boldsymbol{x};t) \right] \\
=& \frac{K_{2,1}}{a^2} \sum_{j=1}^{d} \sum_{\boldsymbol{x}\in a\mathbb{Z}^d} \rho(\boldsymbol{x},t) \left[ (\boldsymbol{x}-a\boldsymbol{e}_j)^2 + (\boldsymbol{x}+a\boldsymbol{e}_j)^2 - 2\boldsymbol{x}^2 \right] \\
=& \frac{K_{2,1}}{a^2} \sum_{j=1}^{d} 2a^2 \sum_{\boldsymbol{x}\in a\mathbb{Z}^d} \rho(\boldsymbol{x},t) = 2dK_{2,1}.
\end{aligned}$$

$\boldsymbol{e}_j$ is the unit vector in $j$-th direction. In the last line we have used the normalization of the concentration. The derivation is the same for any initial condition. This is due to the time and spacial translation invariance of the master equation and leads to the same law for $\left\langle \Delta\boldsymbol{X}^2(t,t+\tau) \right\rangle$. We obtain the main feature of normal diffusion: The linear time-dependence of the msd

$$\left\langle \Delta\boldsymbol{X}^2(t,t+\tau) \right\rangle = 2dK_{2,1}\tau. \tag{2.9}$$

**Solution and propagator:** With the special initial condition $\rho_0(\boldsymbol{x}) = \delta_{\boldsymbol{x},\boldsymbol{x}_0}$, the solution to Eq.(2.4) is the Greens' function or fundamental solution of the equation. We hereby call it the propagator, $G(\boldsymbol{x}; t|\boldsymbol{x}_0; 0)$. In appendix A.1, we show that it is given by:

$$G(\boldsymbol{x}; t|\boldsymbol{x}_0; 0) = e^{-\frac{t}{\tau}} \prod_{j=1}^{d} I_{[\boldsymbol{e}_j \cdot (\boldsymbol{x}-\boldsymbol{x}_0)]/a}\left(\frac{t}{d\tau}\right), \tag{2.10}$$

where $I_n(x)$ denotes a modified Bessel function of the first kind. It is more instructive, however, to discuss it in relation to the random walk, which will be done in the next subsection.

### The binomial random walk process

**Interpretation as a Markov process and first properties:** Eq.(2.4) can also be understood as the forward equation of a Markov process, [144]. Then consequently $K_{2,1}\triangle$ is the generator of the process. The formal solution to (2.4) involves the time evolution operator

$$\rho(\boldsymbol{x}; t) = \mathcal{U}(t)\, \rho_0(\boldsymbol{x}) := e^{K_{2,1} t \triangle} \rho_0(\boldsymbol{x})\,.$$

The kernel of the time evolution operator is the propagator, which can now be identified as the transition probability of the Markov process. We may now interpret the concentration profile $\rho(\boldsymbol{x}; t)$ as the probability density function (pdf) of the position $\boldsymbol{X}(t)$ of *one* tracer that moves on the lattice. This process is called the binomial random walk. Useful properties, like stationarity and independence of the increments are a consequence of the Markov property and the time independence of the generator. However, the Markovianity is not a consequence of (2.4) but rather an additional assumption. Such an interpretation of a rate equation becomes more questionable, when memory effects are involved or the generator is time-dependent, [146].

Much information on the trajectories can already be extracted from $\triangle$. This is because the current of probability (and thus the transition rates) corresponds to jumps of the random walker. The above defined mean sojourn time, see Eq.(2.7), is the average time the walker remains at the same position (hence the name). We also see that the transition rates, defined in Eq.(2.6), can be related to the probability that $\boldsymbol{X}$ jumps along one link of the lattice. This is done by simple normalization

$$p(\boldsymbol{y}; \boldsymbol{x}) := \frac{\lambda(\boldsymbol{y}; \boldsymbol{x})}{\sum\limits_{\boldsymbol{z}:|\boldsymbol{z}-\boldsymbol{x}|=a} \lambda(\boldsymbol{y}; \boldsymbol{x})} = \frac{1}{2d}. \tag{2.11}$$

These transition probabilities are – like the transition rates – homogeneous and isotropic. The jump directions are uniformly distributed among all possible directions.

**Subordination:** The separation in jumps and waiting times allows for a simple numerical scheme to generate trajectories of the process, but it is also a little bit more subtle. The reason is that every continuous-time Markov-process $\boldsymbol{X}(t)$ can be represented as a composition $\boldsymbol{Y}(U(t))$ of a discrete time Markov process (a Markov chain), $\{\boldsymbol{Y}(u)\}_{u\in\mathbb{N}}$ and a Poisson-process, $\{U(t)\}_{t\in\mathbb{R}_{0,+}}$. $U$ connects the physical time with the internal time, the number of jumps, and will be called *operational time* of the process. The process $\boldsymbol{Y}(u)$ is called uniformized Markov-chain (sometimes embedded Markov chain), see [145], and its transition matrix is given by $[e^{K\tau\triangle} - \mathbb{1}]$ in our case. The Poisson-process $U(t)$ controls the time of the jumps, and has the intensity $\tau^{-1}$, with $\tau$ being the mean sojourn time of Eq.(2.7).

The time between the jumps is exponentially distributed with an average being $\tau$ and the probability that $U(t)$ assumes the value $u$ is given by:

$$\mathbb{P}\{U(t) = u\} = e^{-\frac{t}{\tau}} \frac{\left(\frac{t}{\tau}\right)^u}{u!}.$$

On the other hand, when the tracer starts from $\boldsymbol{x}_0$, the probability to be in $\boldsymbol{x}$ after $u$ steps is given by:

$$\mathbb{P}\{\boldsymbol{Y}(u) = \boldsymbol{x}|\boldsymbol{Y}(0) = \boldsymbol{x}_0\} = \frac{1}{(2d)^u} \sum_{u=\sum_j \frac{2au_j+x_j-x_{0,j}}{a}} \frac{u!}{\prod_{j=1}^{d} u_j! \left(\frac{au_j+x_j-x_{0,j}}{a}\right)!}.$$

Here one splits the total number of jumps $u$ into the number of steps $u_j$ made in positive or negative $\boldsymbol{e}_j$ direction. Each direction is chosen with the same probability $1/(2d)$. This leads to a sum over multinomial coefficients. Consequently, we may express the propagator in terms of a subordination formula, via:

$$G(\boldsymbol{x}; t|\boldsymbol{x}_0; 0) = \sum_{u=0}^{\infty} \mathbb{P}\{\boldsymbol{Y}(u) = \boldsymbol{x}|\boldsymbol{Y}(0) = \boldsymbol{x}_0\} \mathbb{P}\{U(t) = u\}. \tag{2.12}$$

The equivalence of (2.10) and Eq.(2.12) is shown in appendix A.1.

The representation $\boldsymbol{X}(t) = \boldsymbol{Y}(U(t))$ is the first example of subordination, the composition of two stochastic processes. The uniformized chain in our example contains all spatial properties. Such an interpretation of two composed stochastic processes is only possible, when the operational time does not decrease. Only in that case causality is not violated.

The derived propagator (2.10) is still pretty cumbersome. One reason is that we formulated the problem on a lattice. When taking the continuum limit $a \to 0$, the expressions simplify a lot. In particular, the typical scaling between displacement and time becomes much more apparent. A stochastic justification for this approach is the Moivre-Laplace limit theorem. The latter states that the multinomial variable $\boldsymbol{Y}(u)$ may be approximated by a Gaussian variable for a large number of jumps. This approximation gets better the more summands are taken into account, that means it gets better for larger times. Another approach would be a coarse graining,

renormalization procedure that rescales the characteristic length scale. This procedure can be applied on the propagator directly, or on the diffusion equation, i.e. on the lattice Laplacian. It is also possible to work in Fourier space and to consider only the limit of small wave vectors in the propagator as well as in the diffusion equation. The resulting continuous space process is the Brownian motion that will be described in the next section.

## 2.2.2. Brownian motion

### The continuous diffusion equation

When the continuum limit $a \to 0$ is taken in the master equation (2.4), the lattice Laplacian becomes the usual Laplacian differential operator. We obtain the common diffusion equation in continuum:

$$\dot{\rho}(\boldsymbol{x};t) = K_{2,1} \sum_{j=1}^{d} \frac{\partial^2}{\partial x_j^2} \rho(\boldsymbol{x};t) = K_{2,1} \Delta \rho(\boldsymbol{x};t)\,. \tag{2.13}$$

Furthermore, the incidence operator can be identified with the gradient and its adjoint with the negative divergence. We recover the usual formulation of Fick's laws and the conservation of mass. Again we need to specify boundary conditions. We choose free boundary conditions in infinite space. As in the discrete case (2.13) is local in time. The stationary state is a constant concentration profile. Integration by parts gives a linear time dependence of the msd for free diffusion:

$$\left\langle \Delta \boldsymbol{X}^2(t, t+\tau) \right\rangle = 2dK\tau. \tag{2.14}$$

The propagator is easily obtained, it is a Gaussian

$$G(\boldsymbol{x};\tau|\boldsymbol{0};0) = \frac{1}{(4\pi K_{2,1}\tau)^{\frac{d}{2}}} \exp\left[-\frac{\boldsymbol{x}^2}{4K_{2,1}\tau}\right]. \tag{2.15}$$

We can see that the spacial argument $\boldsymbol{x}$ only appears in terms of a scaling variable $|\boldsymbol{x}| / \sqrt{K_{2,1}\tau}$. Hence, spacial displacement and elapsed time are not independent, but have to obey the scaling relation given in (2.14).

### The Brownian motion process

Parallel to the treatment above, we can interpret the diffusion equation in terms of a Markov process. This process is called Brownian motion. It can be uniquely defined as the Gaussian process with independent increments that are distributed according to Eq.(2.15) and that has almost surely continuous trajectories. Again, one could represent Brownian motion as a composition $\boldsymbol{Y}(U(t))$ of a Poisson process $U$ and the sum of $u$ i.i.d. Gaussian random variables, $\boldsymbol{Y}(u)$. This representation is, however, not very common as the definition of Brownian motion bears enough nice properties itself. Furthermore, since both are infinitely divisible, the Poisson process' intensity and the Gaussian random variables' variance are arbitrary. It is only necessary to fix their ratio $K_{2,1}$.

### 2.2.3. Summary

This section was devoted to describe normal diffusion. We presented a lattice model and a continuous space model and discussed properties of the stochastic processes and of the equations that govern their propagators. Both exhibit a linear behaviour of the msd, see Eq.s (2.9) and (2.14). In the rest of this thesis we will talk about normal diffusion, whenever such a linear law holds. This linear relation may only hold asymptotically for large times, but even then we will still refer to normal diffusion. Other authors will only call a Brownian motion normally diffusive and also require other properties like Gaussianity, stationarity and independence increments, or scaling of higher moments than the second, see [40]. We will not do this! We call a process normal in a certain regime of times, when the msd is asymptotically linear in that regime. Nevertheless, let us recapitulate some of the properties of those epitomes of normal diffusion.

First of all, we mention the homogeneity and the isotropicity of Eq.s (2.4) and (2.13). Normal diffusion is motion in a very symmetric and regular medium. In the lattice model, the transition rates $\lambda(\boldsymbol{y};\boldsymbol{x})$ do neither depend on the position $\boldsymbol{x}$, nor on the link $(\boldsymbol{y};\boldsymbol{x})$. Similarly, the sojourn times $\tau(\boldsymbol{x})$ do not depend on the lattice site. Furthermore, the topology of the lattice is very regular; every site has the same number of next neighbours and every link spans the same distance (namely $a$). In summary, normal diffusion takes place in a medium completely lacking any disorder. The corresponding stochastic processes have stationary and independent increments. They are Markovian and lack any memory. The governing equations are local in time and lack a dependence on the history of $\rho$.

In both cases, lattice and continuum, we can represent the stochastic process as subordinated to a Poisson process. Furthermore we can see that the pdf is (at least asymptotically) of a special scaling form, i.e. the space dependence in Eq.(2.15) appears only in the form $|\boldsymbol{x}|/\sqrt{K_{2,1}\tau}$.

We only considered diffusion in infinite space, rather than diffusion in the presence of boundaries. Here, the concentration profile will spread indefinitely and it will thin out. When the system is finite on the other hand, the concentration profile will ultimately converge to the uniform distribution, which constitutes the equilibrium distribution. The Gaussian (or binomial) profile of the concentration is a consequence of the (lattice) Laplacian and thus it is a consequence of the identification of possible micro- or meso-states and the transitions between them.

In our treatment, we completely neglected the velocity of the particle. The position of the tracer particle constitutes its whole state space. This is only possible, when the inertia of the tracer can be neglected. For example, when the medium is very viscous, or when the tracer particle is much heavier than the particles of the medium, so that collisions between them effectively do not change the tracers velocity. If one leaves this so called over-damped regime, the velocity variables must also be considered in the state space. In the continuous case, the time dependence of the msd is still linear, because the velocity decorrelates but, for short times, the tracer will move ballistically, and $\langle \boldsymbol{X}^2(t)\rangle \sim t^2$. The corresponding process is not a Brownian motion any more but

the integral of an Ornstein-Uhlenbeck process. The majority of the models treated in this thesis will however neglect inertia and velocity.

In the sequel, all mentioned properties, homogeneity, isotropy, scaling form, appearance of the Poisson process, and so on, will be subject to generalization. We will see that a linear law like (2.14) can not be erected, when disorder of a certain strength is introduced to transition rates or when different scaling forms are used for the stochastic processes. All this is going to be elucidated in the following sections.

## 2.3. Disorder in master equations

**A new master equation:** As pointed out in the last section, Brownian motion takes place in a very regular and ordered medium. In the following section, we want to investigate disordered systems. We will introduce the disorder in the rates of the master equation. In a later section, we will show how this reflects in a description via stochastic processes.

We start with a generalized version of (2.4) that describes the evolution of the concentration profile:

$$\dot{\rho}(\boldsymbol{x};t) = \triangle_\lambda \rho(\boldsymbol{x};t) := \sum_{y\in\Omega} \left[\lambda(\boldsymbol{x};\boldsymbol{y})\,\rho(\boldsymbol{y};t) - \lambda(\boldsymbol{y};\boldsymbol{x})\,\rho(\boldsymbol{x};t)\right]. \tag{2.16}$$

Before, the transition rates $\lambda(\boldsymbol{y};\boldsymbol{x})$ have been very regular. They were symmetric, homogeneous and isotropic. All those assumptions are now dropped. We only assume that they are not negative. Since all diagonal terms in above equation cancel, the diagonal entries $\lambda(\boldsymbol{x};\boldsymbol{x})$ are irrelevant and can be set to zero.

The new master equation describes diffusion in a disordered medium, [139]. The product $\lambda(\boldsymbol{y};\boldsymbol{x})\,\rho(\boldsymbol{x})$ is proportional to the number of particles moving from $\boldsymbol{x}$ to $\boldsymbol{y}$ in unit time, and is the main modelling object of the theory. (2.16) can be motivated in the following ways: Coming from irreversible thermodynamics, one may combine the continuity equation, with a linear response ansatz, similar to Fick's treatment. Following Onsager, one would assume the diffusion flux $j$ as linear in corresponding thermodynamic force, which can be identified as the gradient of the chemical potential:

$$\dot{\rho} = -\boldsymbol{\nabla}^\dagger j = -\frac{1}{T}\boldsymbol{\nabla}^\dagger L \boldsymbol{\nabla}\mu.$$

Here, $L$ is the Onsager coefficient, $T$ the temperature assumed to be constant and $\mu$ the chemical potential, which is itself a function of the concentration. When $L$ is not constant, and the chemical potential is not a trivial function of $\rho$, a linearisation of the equation bears (2.16). Another approach is to interpret (2.16) as the discretization of a Fokker-Planck-equation with a random potential. The disorder is modelled with the set of possibly random rates $\{\lambda(\boldsymbol{y};\boldsymbol{x})\}_{\boldsymbol{x},\boldsymbol{y}\in\Omega}$, which will be called the *environment* henceforth. Our knowledge of the environment or lack thereof is modelled with the probability distribution of the environment.

Master equations like (2.16) have been used intensively to study phenomena related to diffusion, see for instance [40, 53] and references therein. The disordered Laplacian $\triangle_\lambda$, in the various forms discussed below, also appears when considering networks of masses connected with ideal springs, [147, 148], and resistor networks, [89]. All those phenomena are thus formally related up to a certain point. Eq.(2.16) is also thoroughly discussed in mathematical literature, [56, 149–151], usually in connection to random walks. This relation will be exploited in the next section.

**Applicability of the model:** Of course, Eq.(2.16) is not the most general formulation. With our ansatz, we express the assumption that the process with quenched disorder is a time-homogeneous Markov process. We identify the state space of the tracer particle with its position in Euclidean space and thus we treat the so-called over-damped case. This identification is not necessary. Other mathematical features like non-linearity, time dependence of the rates or even memory effects could be included. Non-linear terms are necessary for proper description of chemical reactions. Time dependent rates are necessary, when the system is subject to external driving. One could argue that the motion in the complete state (or phase) space truly is governed by a time-homogeneous equation like the one above, but often the whole state space is either not observable or not known. For diffusions, there could be additional properties, like the velocity of the tracer, some other internal states of the tracer, or positions of other particles that are needed to completely characterize the system state. Memory effects will emerge in Eq.(2.16) when the state space is projected to a set of "relevant" variables, i.e. position variables, [152]. The master equation above is thus not applicable in those cases, explicitly not in the under-damped limit. We are "only" free to choose a proper ensemble of environments $\{\lambda(\boldsymbol{y};\boldsymbol{x})\}$.

**Programme for the rest of the section:** Let us take a little time to sketch the next steps. The transition rates $\lambda$ define the problem. For one fixed environment Eq.(2.16) describes a model with quenched disorder, since the transition rates are independent of time, i.e. the disorder is static. We are interested in the long time behaviour of the msd, which is not easily accessible in quenched disorder. Some properties of the system can be made out in quenched disorder, and we will use them to find an appropriate average over the disorder. In this averaged or annealed model the behaviour of the msd is easier to obtain. The physical meaning of the average depends on the problem. It may correspond to taking different samples of the medium or to different regions of the same infinite medium. One of the main ideas here is that a tracer particle diffusing in a disordered medium explores its surrounding. On larger length and time scales the tracer feels not the details of the medium but rather a local average. In this conception, the long-time, large-scale behaviour corresponds to an average over the largest region of the medium. In that way, the long-time regime corresponds to an homogeneous average over the disorder. The system "homogenizes" on such large scales.

In what way the average has to be taken depends on the definition of the quantities of

interest. For vividness consider the diffusion of some small particles through a porous medium like a soaked sponge. We fix the concentration of the diffusing particles on both sides of the sponge and wait until a stationary diffusion flux from one side to the other establishes. The ratio of that flux to the difference in concentration is the diffusivity.[4] When this ratio is known, we know that the diffusion is normal and is characterized by the just measured diffusivity. Now, the sponge can be replaced with a homogeneous liquid, which has the same diffusivity. With respect to this measurement protocol, the disordered medium (the sponge) and the homogeneous one are identical. A sketch of that procedure can be found in Fig. 2.1 on page 31. In the same sense we would like to replace Eq.(2.16) with (2.4) with an appropriate effective diffusivity.

Whenever the long-time behaviour is normal diffusion, we will obtain a finite value of that effective diffusivity. Although the system is disordered, the disorder is "weak" and does not lead to an anomalous scaling of the msd. However, when effective diffusivity becomes zero of infinite, the diffusion can no longer be described with a normal law, but only with an anomalous one. In this case the disorder is considered to be "strong". Successful homogenization is thus a sufficient condition for normal diffusion. Understanding, how homogenization fails, enables us to understand the emergence of anomalous diffusion.

In the next subsection, we will discuss some properties of the quenched disorder model, Eq.(2.16). These properties will then be used to formulate an appropriate annealed model and an effective diffusivity. After obtaining this effective diffusivity, we will discuss whenever this diffusivity becomes non-finite, and are thus put in the position to identify two major mechanism that lead to anomalous diffusion.

## 2.3.1. Quenched disorder

At first, we want to discuss the quenched case. This means that we fix a set of transition rates for discussion. We start with clarifying the assumptions that give well-defined transition rates.

#### Graph representation, Stationary state and equilibrium properties

**The transition graph:** A first and trivial observation is that every $\lambda(\boldsymbol{y};\boldsymbol{x})$ must be non-negative:

$$\lambda(\boldsymbol{y};\boldsymbol{x}) \geq 0. \tag{2.17}$$

Otherwise the interpretation as a rate of particles moving from one site to another is not possible. It is on the other hand very well possible that a certain rate vanishes. In that case, the corresponding transition is forbidden. In normal diffusion, this is the case whenever the distance of the sites is larger than the lattice constant. The set $\Omega$ of all lattice sites, together with the set $T(\Omega)$ of all possible transitions can be interpreted

[4] This definition of the diffusivity is analogous to the definition of the electrical conductance or resistance.

as a graph. This graph may be represented via its adjacency matrix $A(\boldsymbol{y};\boldsymbol{x})$, which evaluates to unity if $\lambda(\boldsymbol{y};\boldsymbol{x}) > 0$ and zero otherwise:

$$A(\boldsymbol{y};\boldsymbol{x}) := \mathbb{1}_{(0,\infty)}(\lambda(\boldsymbol{y};\boldsymbol{x})) = \begin{cases} 0, & \lambda(\boldsymbol{y};\boldsymbol{x}) = 0 \\ 1, & \lambda(\boldsymbol{y};\boldsymbol{x}) > 0 \end{cases}. \tag{2.18}$$

$\mathbb{1}_I(x)$ is an indicator function returning unity if $x \in I$ and zero otherwise. The so-defined *graph of transitions* is subject to several assumptions. Firstly, the graph is assumed to be fully connected. That means, every state $\boldsymbol{y}$ can be reached from every other state $\boldsymbol{x}$ with a certain succession of transitions.[5] A graph that is not fully connected either has absorbing sub-graphs, from where the tracer can not return, or it splits up into several non-communicating sub-graphs. Either way, we could consider the dynamics on the sub-graph and the treatment (up to some transient behaviour) would not change. Our second assumption is that the graph is undirected, i.e., whenever $\lambda(\boldsymbol{y};\boldsymbol{x}) > 0$ then $\lambda(\boldsymbol{x};\boldsymbol{y}) > 0$. For every transition, the reverse transition is allowed. Some of our results can be generalized to directed graphs in the quenched case.

**The thermodynamic equilibrium state:** Following Schnakenberg in [104], the graph representation alone[6] is sufficient to proof the existence and stability of a proper stationary state, $\rho_{\mathrm{st}}(\boldsymbol{x})$, for the master equation (2.16). The stationary state can explicitly be calculated via Kirchhoff's theorem, [153].[7] Schnakenberg reports that much more information can be obtained from the graph representation, but we will only need the proof of the existence and stability of the stationary state. The full argument is given in the appendix in section A.2.

The stationary state is defined by the equation:

$$0 = \triangle_\lambda \rho_{\mathrm{st}}. \tag{2.19}$$

Let $T_{l,\boldsymbol{x}}$ be one of the directed spanning trees of the transition graph,[8] and $\boldsymbol{\xi} = (\boldsymbol{z};\boldsymbol{y}) \in T_{l,\boldsymbol{x}}$ a directed link from that tree, then the stationary state solution reads

$$\rho_{\mathrm{st}}(\boldsymbol{x}) = \frac{\sum_l \prod_{\boldsymbol{\xi}\in T_{l,\boldsymbol{x}}} \lambda(\boldsymbol{\xi})}{\sum_{\boldsymbol{x}\in\Omega}\sum_l \prod_{\boldsymbol{\xi}\in T_{l,\boldsymbol{x}}} \lambda(\boldsymbol{\xi})}. \tag{2.20}$$

[5] Mathematically spoken, we require that for all $\boldsymbol{x},\boldsymbol{y} \in \Omega$ there exists a $n \in \mathbb{N}$ such that $A^n(\boldsymbol{y};\boldsymbol{x}) > 0$.

[6] To be more precise, the appearance of the Laplacian.

[7] Kirchhoff's original theorem was concerned with computing the electrical current for a given voltage profile in a resistor network. He constructed the solution consisting of rational functions of the resistances by considering spanning trees of the corresponding graph. Schnakenberg used the same technique to compute the stationary state of Eq.(2.16). In graph theory, the number of spanning trees in a graph is computed in that way and one refers to "(Kirchhoff's) matrix tree theorem", [141–143, 145].

[8] A spanning tree $T_l$ is built by removing links from a graph until no cycles are left. In the directed spanning tree $T_{l,\boldsymbol{x}}$ every link is assigned to a direction and it points towards the node $\boldsymbol{x}$.

Eq.(2.20) summarizes Kirchhoff's theorem, [153]: the stationary state solution is a normalized sum of products of transition rates that are chosen according to the directed spanning trees of the graph. As Schnakenberg noted, the usefulness of this solution depends on the topology of the graph, and thus on how easy one can determine all spanning trees.

It is apparent that the stationary state is properly normalized for finite graphs. Furthermore, it was shown in [104] that it is an attractor, i.e. $\rho(\boldsymbol{x};t) \to \rho_{\mathrm{st}}(\boldsymbol{x})$, as $t \to \infty$. When the transition graph consists only of one communicating component and every transition is reversible – as were our assumptions – all local currents along the bonds vanish in the stationary state. This is the so called *detailed balance* condition, which allows us to identify the stationary state as an equilibrium state, i.e. $\rho_{\mathrm{st}} = \rho_{\mathrm{eq}}$, with:

$$0 = j_{\mathrm{eq}}(\boldsymbol{y},\boldsymbol{x}) = \lambda(\boldsymbol{y};\boldsymbol{x})\,\rho_{\mathrm{eq}}(\boldsymbol{x}) - \lambda(\boldsymbol{x};\boldsymbol{y})\,\rho_{\mathrm{eq}}(\boldsymbol{y})\,. \tag{2.21}$$

Under these conditions the equilibrium state nowhere vanishes and we may apply the logarithm to identify the equilibrium state with some Boltzmann-distribution:

$$\rho_{\mathrm{eq}}(\boldsymbol{x}) =: N^{-1} e^{-\beta \mathscr{F}(\boldsymbol{x})}. \tag{2.22}$$

$\beta = 1/(k_B T)$ is the inverse temperature and $\mathscr{F}$ is some free energy. When the system has detailed balance, the Boltzmann factors are unique, although not necessarily normalizable. Therefore the main object of our treatment is the Boltzmann-factor.

The requirement of detailed balance means that our system has a notion of thermodynamic equilibrium. In a certain sense, we are working close to equilibrium, we can still "see" it, albeit the system may never reach it. When $\Omega$ is infinitely large it may happen that the equilibrium state $\rho_{\mathrm{eq}}$ is not normalizable and the constant $N$, appearing in Eq.(2.22), becomes zero.[9] Still, when the disordered Laplacian acts on the function of Boltzmann factors, the result is zero. We can thus say (already by looking at Eq.(2.19)) that $\rho_{\mathrm{eq}}$ is an eigenstate of $\triangle_\lambda$ with eigenvalue zero, although this eigenstate may not be a member of the set of normalizable distributions. If the transition graph is not fully connected, but splits up into several non-communicating sub-graphs, let's say $\Omega = \Omega_1 \cup \Omega_2$, then eigenvalue zero is degenerated. The reason is that both $\rho_{\mathrm{eq},i}(\boldsymbol{x}) := N_i^{-1} e^{-\beta \mathscr{F}(\boldsymbol{x})} \mathbb{1}_{\Omega_i}(\boldsymbol{x})$, for $i = 1, 2$ are orthogonal eigenstates to the eigenvalue zero. Each region has its own independent equilibrium state. Our first assumption on the transition graph assures its uniqueness.

For the following discussion, we require the detailed balance condition and it is crucial. It is our main restriction on the master equation and rules out some interesting cases. The random-force model, [53], can only be treated by the following methods when it has a scalar potential. The same is true for a non-over-damped particle subject to a deterministic force and thermal kicks. When the deterministic force is not the gradient of a potential, it is not possible to satisfy Eq.(2.21), without having $\rho_{\mathrm{eq}}(\boldsymbol{x})$. The concentration in equilibrium must vanish on all closed elementary circles of the

[9] When this is the case, we set $N = 1$ and refuse to care about the prefactor.

graph, when detailed balance is supposed to hold. As mentioned above, it also has to vanish outside of an absorbing region. Thus, it is convenient to restrict ourself to undirected transition graphs. Our second assumption assures that the equilibrium concentration is nowhere zero.

**Equilibrium transition rates:** From now on, we follow the arguments of Camboni and Sokolov [45]. Eq.(2.21) defines a quantity symmetric on the bonds, viz. the rate of particles going from $\boldsymbol{x}$ to $\boldsymbol{y}$ in equilibrium.. We call this quantity the equilibrium transition rate:

$$w(\boldsymbol{x}, \boldsymbol{y}) := \lambda(\boldsymbol{y}; \boldsymbol{x})\, e^{-\beta \mathscr{F}(\boldsymbol{x})}. \tag{2.23}$$

This decomposition was already used in [154] in the context of effective medium approximations.

The original transition rates $\lambda(\boldsymbol{y}; \boldsymbol{x})$ have been decomposed into three parts. Firstly, there is the topology of the transition graph, represented by the adjacency matrix $A(\boldsymbol{y}, \boldsymbol{x})$. Then there are the symmetric equilibrium rates $w(\boldsymbol{y}, \boldsymbol{x})$ that describe the dynamics in thermodynamic equilibrium, and finally there is the equilibrium concentration given by the Boltzmann factors. The decomposition reads:

$$\lambda(\boldsymbol{y}; \boldsymbol{x}) = A(\boldsymbol{y}, \boldsymbol{x})\, w(\boldsymbol{y}, \boldsymbol{x})\, e^{\beta \mathscr{F}(\boldsymbol{x})}. \tag{2.24}$$

In the next section, the original master equation is rewritten in terms of the new quantities.

### Master equation for activities

Using Eq.(2.24) of the last section, we have

$$\dot{\rho}(\boldsymbol{x}; t) = \sum_{y \in \Omega} A(\boldsymbol{y}, \boldsymbol{x})\, w(\boldsymbol{y}, \boldsymbol{x}) \left[ e^{\beta \mathscr{F}(\boldsymbol{y})} \rho(\boldsymbol{y}; t) - e^{\beta \mathscr{F}(\boldsymbol{x})} \rho(\boldsymbol{x}; t) \right].$$

We observe that the quantity appearing in the braces is proportional to the activity. Hence, the last equation can be formulated as an equation governing the dynamics of the activities

$$e^{-\beta \mathscr{F}(\boldsymbol{x})} \dot{\eta}(\boldsymbol{x}; t) = \sum_{y \in \Omega} A(\boldsymbol{y}, \boldsymbol{x})\, w(\boldsymbol{y}, \boldsymbol{x}) \left[ \eta(\boldsymbol{y}; t) - \eta(\boldsymbol{x}; t) \right], \tag{2.25}$$

where

$$\eta(\boldsymbol{x}; t) := \frac{\rho(\boldsymbol{x}; t)}{\rho_{\mathrm{eq}}(\boldsymbol{x})} = N e^{\beta \mathscr{F}(\boldsymbol{x})} \rho(\boldsymbol{x}; t). \tag{2.26}$$

**Decomposition of the disordered Laplacian:** Now it turns out that the action of the original operator of transition rates can be split up. The first contribution is the action of the Boltzmann-factors, given by an operator $\mathcal{B}$ that is diagonal in the usual base of the nodes $\{\boldsymbol{x}\}$:

$$\mathcal{B}f(\boldsymbol{x}) := e^{-\beta\mathscr{F}(\boldsymbol{x})} f(\boldsymbol{x})\,.$$

The other one is the action of the equilibrium rates, given by an operator $\mathcal{W}$ that is diagonal in the base of all *links* of the graph:

$$\mathcal{W}g(\boldsymbol{y};\boldsymbol{x}) := w(\boldsymbol{y},\boldsymbol{x})\, g(\boldsymbol{y};\boldsymbol{x})\,.$$

The adjacency matrix is used to define the transition graph's incidence operator that transforms concentrations into currents:[10]

$$\boldsymbol{\nabla}_A f(\boldsymbol{y};\boldsymbol{x}) = \mathrm{sign}(\boldsymbol{y};\boldsymbol{x})\, A(\boldsymbol{y};\boldsymbol{x})\, [f(\boldsymbol{y}) - f(\boldsymbol{x})] \tag{2.27}$$

The incidence operator and its adjoint define a symmetric Laplacian:

$$\triangle_{A,w} f(\boldsymbol{x}) := -\boldsymbol{\nabla}_A^\dagger \mathcal{W} \boldsymbol{\nabla}_A f(\boldsymbol{x}) = \sum_{\boldsymbol{y}\in\Omega} A(\boldsymbol{y},\boldsymbol{x})\, w(\boldsymbol{x},\boldsymbol{y})\, [f(\boldsymbol{y}) - f(\boldsymbol{x})]\,.$$

Now, we can rewrite (2.25) as

$$\mathcal{B}\dot{\eta} =: \triangle_{A,w}\eta. \tag{2.28}$$

The symmetric operator $\triangle_{A,w}$ is a perturbation of the regular Laplacian of the graph, given by $-\boldsymbol{\nabla}_A^\dagger \boldsymbol{\nabla}_A$. The operator $\mathcal{B}$ can be seen as a local change of the time scale. This decomposition of the disordered Laplacian can be found in [45], where a simple cubic topology was considered. But it also appeared earlier, e.g. in [154]. In the last text, however, $\eta$ was not identified as the activity.

The original equation (2.16) reads in the new notation

$$\dot{\rho}(\boldsymbol{x};t) = \triangle_{A,w}\mathcal{B}^{-1}\rho(\boldsymbol{x};t) = \sum_{\boldsymbol{y}\in\Omega} A(\boldsymbol{x},\boldsymbol{y})\, w(\boldsymbol{x},\boldsymbol{y}) \left[\frac{\rho(\boldsymbol{y};t)}{e^{-\beta\mathscr{F}(\boldsymbol{y})}} - \frac{\rho(\boldsymbol{x};t)}{e^{-\beta\mathscr{F}(\boldsymbol{x})}}\right]. \tag{2.29}$$

It is simply governed by the transposed operator. We can also identify:

$$\triangle_\lambda = \triangle_{A,w}\mathcal{B}^{-1}. \tag{2.30}$$

[10] The definition of the incidence operator involves assigning a positive direction to each link $(\boldsymbol{y};\boldsymbol{x})$, here expressed with the sign operator. Hence the incidence operator is not uniquely defined. The Laplacian $\triangle = -\boldsymbol{\nabla}^\dagger\boldsymbol{\nabla}$ is though, because the sign appears only squared and cancels.

**Connection to Fick's laws:** Eq.(2.28) can be understood as a generalization of Fick's laws on a graph. One starts with the continuity equation on the graph: $\dot{\rho} = -\boldsymbol{\nabla}_A^\dagger j$, but instead of assuming that the diffusion flux is given by the gradient of concentration, we take it as the gradient of the chemical potential $j(\boldsymbol{y};\boldsymbol{x}) = \beta L(\boldsymbol{y};\boldsymbol{x})\,\boldsymbol{\nabla}_A \mu$ with a bond dependent Onsager coefficient, $L$. This is the usual Onsager linear response theory with the correct thermodynamic force for diffusion: the gradient of the chemical potential. The chemical potential is the logarithm of the fugacities, which can be identified with the activities. Hence we have $\dot{\rho} = -\boldsymbol{\nabla}_A^\dagger L \boldsymbol{\nabla}_A \ln \eta$. The activities can be identified as $\eta = \mathcal{B}^{-1}\rho$, and near equilibrium they fluctuate around unity. This allows for linearising the logarithm term; we obtain $\dot{\rho} = \boldsymbol{\nabla}_A^\dagger L \boldsymbol{\nabla}(\rho/\rho_{\mathrm{eq}})$. In conclusion, we can interpret equation (2.16) and its relatives as a generalization of Fick's laws to situations near equilibrium with disordered equilibrium state, disordered Onsager coefficients (which can be identified with bond diffusivities) and complex topology.

The new master equation Eq.(2.25) is formally equivalent to the evolution equation of potential in an electrical network, and also to an elastic network, see [53, p.167 et seq.] and the appendix A.3. This analogy is exploited later on, when the effective medium diffusivity is calculated.

**Further properties of the quenched model:** When the probability law of the environment is known and suitably well behaved, it is possible to make probabilistic statements about the diffusion process in quenched disorder. The usual way is to consider a Dirichlet form defined by the Laplacian $\triangle_\lambda$. Then, some estimates with respect to the Dirichlet form translate to estimates of the propagator. Those in turn are used to proof statements about for instance the spectral dimension. But they may also be used to show convergence of the cleverly rescaled propagator to some other stochastic processes propagator. This way, one is able to show that the rescaled solution of Eq.(2.16) behaves like a Gaussian and thus the diffusion process is normal. Such scaling limits can be understood in various probabilistic ways: for almost every environment, in probability, or in distribution. However, rigorous proofs are difficult, and we will not dive deeper into this topic here. We rather investigate the annealed case and refer to [140] for an overview. The diffusion is normal, for $d$-dimensional simple cubic lattices above the percolation threshold when the rates $\lambda$ are bounded away from zero. When there is a significant number of rates that are very low, the diffusion may be subdiffusive, see [140, ch. 8.1.].

**Summary:** In this section we have generalized the discrete diffusion equation by introducing asymmetrical disorder in the transition rates. We constructed a network of possible transitions and split up the rates into three parts:

1. A topological part described by the adjacency matrix $A(\boldsymbol{y};\boldsymbol{x})$ of the transition graph,

2. the equilibrium concentration described by the Boltzmann factors $e^{-\beta\mathscr{F}}$,

3. and the symmetric equilibrium rates $w(\boldsymbol{y}, \boldsymbol{x})$ that describe the dynamics in equilibrium.

In order to do this, we assumed the following three conditions on the transition graph to hold:

1. It is fully connected, i.e. $\forall \boldsymbol{x}, \boldsymbol{y} \in \Omega \; \exists n \in \mathbb{N} : (A^n)(\boldsymbol{y}, \boldsymbol{x}) > 0$,
2. it is undirected, i.e. . $A(\boldsymbol{y}, \boldsymbol{x}) = A(\boldsymbol{x}, \boldsymbol{y})$,
3. and most importantly detailed balance, Eq.(2.21), holds.

The existence of an attracting stationary state is ensured by Kirchhoff's theorem. It uses the topology of the graph to construct the solution. The symmetrization of the rates was possible by exploiting the detailed balance condition that requires all microscopical currents to vanish in equilibrium. By virtue of detailed balance, the stationary state is an equilibrium solution. We needed to require that the transition graph is undirected to ensure that the equilibrium concentration does not vanish anywhere. Furthermore we assumed that the graph is fully connected, so that the equilibrium state is a unique eigenvector of the disordered Laplacian with eigenvalue zero. Finally we used the decomposition, to split the disordered Laplacian into a diagonal operator and a symmetric one. This lead to a restatement of the master equation in terms of activities, Eq.(2.25).

In the following section, we investigate under which conditions the disordered system may be replaced with an ordered one.

### 2.3.2. Annealed disorder

All considerations in the last section were based on the detailed knowledge of the set $\{\lambda(\boldsymbol{y}; \boldsymbol{x})\}$ of transition rates. From this knowledge we were able to determine the equilibrium properties of the sample. Although those do determine the diffusive behaviour, it may be difficult to infer the exponent of anomalous diffusion directly. This would be much easier, when the detailed and complicated environment could be replaced with an homogeneous annealed one. One could justify this approach in the following way: When a tracer particle explores the lattice $\Omega$, it also explores the set $\{\lambda(\boldsymbol{y}; \boldsymbol{x})\}$. On a certain coarse grained scale, the tracer only feels a certain spacial average of aforementioned rates. Hence, we can expect that on the largest possible scale all transition rates can be replaced with one effective rate $\lambda^* := K^*/a^2$. This renormalization procedure also implies a replacement of the transition graph's adjacency matrix $A(\boldsymbol{y}, \boldsymbol{x})$ with an effective one $A^*(\boldsymbol{y}, \boldsymbol{x})$ that hopefully has more regular properties than the old one. In fact, the procedure involves a replacement of the disordered Laplacian $\triangle_\lambda$ with an appropriate effective one $K^*\triangle^*$. Whenever this is possible, we speak about *homogenization* of the system. In that particular situation the dynamics of the disordered system at the largest scale are dominated by a homogeneous system. Consequently when it turns out that $K^*\triangle^*$ is the lattice Laplacian from Eq.(2.4), the system's dynamics are normal diffusion, except for a certain transient regime.

When will the dynamics be anomalous? Given that the system homogenizes, then the dynamics can not be normal when either the homogeneous operator $\triangle^*$ is not the lattice Laplacian, or if it is, when the corresponding diffusivity $K^*$ is not finite. Consider the latter case, the diffusivity $K^*$ can vanish, indicating a diffusion slower than normal diffusion: subdiffusion. On the other hand, when $K^*$ diverges, the process is faster than normal diffusion and it is superdiffusive.

Our intention can compared to electric conduction. Consider a block of some material between two well conducting plates at different electric potential. The gradient of potential will induce transport of electric charge from one plate to the other, an electric current flowing in the material. From a macroscopic point of view, the strength of the electric current and the difference in potential can be measured and both of them define the electrical resistance. The latter is a macroscopic and thus averaged quantity. Although at a microscopical level, the charge transport is described by Eq.(A.7) from the appendix, from a macroscopic point of view, we can not distinguish between inhomogeneous bond conductances and a homogeneous effective one. The macroscopic measurement defining the resistance is blind for microscopical fluctuations in the conductance, *if they are small*! When the material in question is an insulator or a super-conductor, $R$ will attain non-finite values, i.e. infinity for an insulator and zero for a super-conductor. The charge transport inside the medium then is anomalous. However, as long as transport is normal, we can replace the disordered material with a standard industry resistor of the same resistance and achieve the same results within the scope of the measurement procedure.

This picture can be transferred back to the diffusion problem, see Fig. 2.1. We want to investigate the diffusion of some solved particles through a complex disordered medium, for instance a sponge or some biological living cell. Place the medium between two membranes. On each site of the membrane, there is a solution with different concentration of solved particles. This induces a drop of chemical activity along the medium, which in turn results in solved particles flowing from one reservoir to the other. This diffusive current will eventually reach a stationary state and can be measured. The ration between the stationary current strength and the activity difference *in the limit of infinitely far seperated membranes* exactly defines an average diffusivity $K^*$. In [45], the authors argue that this *quasi-stationary* limit defines the exact diffusivity for free diffusion. The problem is analogue to the electrical one, when $K^*$ is finite. The sponge can safely be replaced with a homogeneous solution of the same diffusivity, but only as long it is finite!

Would we be able to find sufficient conditions for the system to homogenize to a normal diffusive one, we would at the same time know under which conditions anomalous diffusion is possible. Or rephrased: When Eq.(2.16) or Eq.(2.25) can not be replaced by Eq.(2.4), then the diffusion process is asymptotically anomalous. By identifying the different sufficient conditions for successful homogenization, we can *classify* different mechanisms that lead to anomalous diffusion. This is the primary object of the whole chapter.

Mathematically, we have to deal with some kind of coarse graining map acting on the operator $\triangle_\lambda$, and we have to investigate its limit. This is a very difficult task

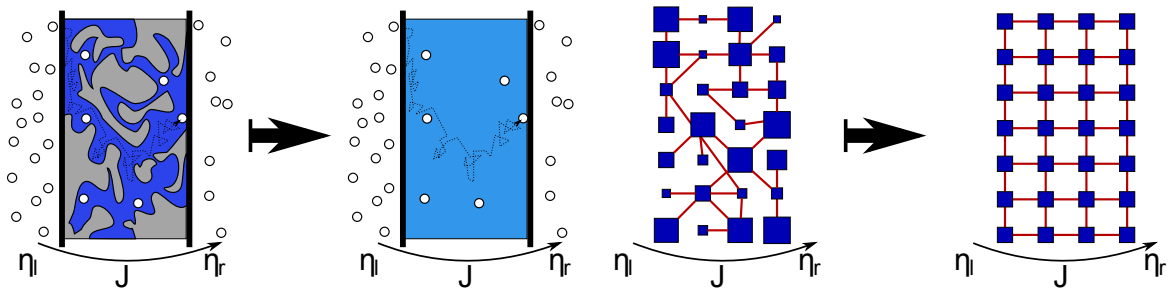



Figure 2.1.: Replacement of a disordered medium with an ordered one
A part of the disordered medium is put between two ideal membranes. On each site of the membrane the chemical activity of tracer molecules is fixed and after some time a stationary diffusion flux establishes. The ratio between stationary current $J$ and the difference in activity $\eta_l - \eta_r$ is the effective diffusivity. When this is known the disordered medium can be replaced with a homogeneous one with the same diffusivity. The homogeneous model will bear the same results with respect to this measurement as the disordered one by construction and can be considered as the annealed description. The same holds for a disordered lattice model. Inhomogeneities in topology, symmetric rates and equilibrium distance will eventually homogenize, and result in an effectively homogeneous description.

to do rigorously and goes beyond the scope of this thesis. We will treat the problem here only in a more appropriate heuristic manner. Still, there are rigorous results, see e.g. [56, 149]. The limit problem for the operator is in a certain sense equivalent to a scaling limit problem for the stochastic process. Thus the limit theorems for random walks are also related to the homogenization of a disordered medium, [150, 155]

In the following we present some of the possible ways to obtain this effective diffusivity.

### Effective medium approximation

When each link of the environment is independent of the other links, an effective medium average can be applied. This self-consistent technique was first used by Bruggemann to compute the dielectric constant of material mixtures, [156]. Later, it was also applied in the context of percolation conduction and diffusion problems, [41, 52, 89, 154, 157]. We will use the equivalence to the electrical network and relate the effective medium conductance with an effective medium diffusivity. Although it works better, when the network has independent and symmetric link properties, it was also generalized to incorporate link correlations, and asymmetry [157]. However, since we already split up $\triangle_\lambda$ into the symmetric part $\triangle_{A,w}$ and the equilibrium state operator $\mathcal{B}$, we may disregard the latter for once.

The effective medium approximation replaces the operator $\triangle_{A,w}$ with a reference operator $\triangle^*$ and an effective diffusivity in such a way that the reference resolvent is equal to the disorder averaged (annealed) one, $\mathbb{E}[(s - \triangle_{A,w})^{-1}]$. In that way the

propagator of the annealed model and of the reference model coincide and consequently their macroscopic diffusivities. The details of the derivation are given in the appendix A.4. In the simplest case, a simple cubic lattice with i.i.d. equilibrium rates on each bond, the effective medium approximation replaces each bond with the same effective bond, such that the macroscopic diffusivity remains unchanged. For topologies more difficult than the simple cubic one, we have to be more careful.

The transition graph of the reference model is chosen nice enough to be tractable. We assume that the reference operator has the following form:

$$\triangle^* f(\boldsymbol{x}) := \sum_{\boldsymbol{y}\in\Omega} A^*(\boldsymbol{x},\boldsymbol{y})\,\frac{f(\boldsymbol{y})-f(\boldsymbol{x})}{D^*(\boldsymbol{x},\boldsymbol{y})}. \tag{2.31}$$

Hence, beside the effective diffusivity, there are two more quantities to be determined: The topology of the reference transition graph given is by the adjacency matrix $A^*$, and the distance dependence is $D^*(\boldsymbol{x},\boldsymbol{y})$. Both quantities are subject to certain restrictions. The reference transition graph can not erase links that are possibly present in the original system. Hence $A^*(\boldsymbol{x},\boldsymbol{y})$ *must* be unity whenever $A(\boldsymbol{x},\boldsymbol{y})$ *could* be unity. A reasonable choice would be to construct a reference model with a regular transition graph that has a maximum connection distance $r$, i.e. take $A^*(\boldsymbol{x},\boldsymbol{y}) = \mathbb{1}_{(0,r]}(|\boldsymbol{x}-\boldsymbol{y}|)$. Depending on the possible topology of the original transition graph, this $r$ can be finite, as well as infinite. In the latter case, we say the graph has *long-range connections*. The distance factor $D^*$ on the other hand needs to match the stochastic distance dependence of the equilibrium rates $w$ on large scales. That means, when $f_{(\boldsymbol{x},\boldsymbol{y})}(A,w)$ denotes the joint pdf of $A(\boldsymbol{x},\boldsymbol{y})$ and $w(\boldsymbol{x},\boldsymbol{y})$, and when it has the following scaling form

$$\begin{aligned} f_{(\boldsymbol{x},\boldsymbol{y})}(A,w)\,\mathrm{d}A\mathrm{d}w = {}& (1-p(\boldsymbol{x},\boldsymbol{y}))\,\delta(A)\,\delta(w)\,\mathrm{d}A\mathrm{d}w \\ & + p(\boldsymbol{x},\boldsymbol{y})\,\delta(A-1)\,\bar{f}(wD(\boldsymbol{x},\boldsymbol{y}))\,D(\boldsymbol{x},\boldsymbol{y})\,\mathrm{d}A\mathrm{d}w, \end{aligned}$$

then $D^*$ is identified on large scales as $D^* = D/p$. Here $p(\boldsymbol{x},\boldsymbol{y})$ denotes the probability that $\boldsymbol{x}$ and $\boldsymbol{y}$ are connected and $\bar{f}$ is a scaling function. For simplicity, we may assume that the reference model is translationally invariant, so that $D^*$ is only a function of the displacement $\boldsymbol{\xi} = \boldsymbol{y}-\boldsymbol{x}$, and isotropic, so that $D^*$ is a function of $|\boldsymbol{\xi}|$ only.

Under these assumptions, the effective medium diffusivity is well defined and the asymptotic msd in the effective medium is given by:

$$\left\langle \boldsymbol{X}^2(t)\right\rangle = C_{\triangle^*}\mathbb{E}[AwD^*]_{\mathrm{EM}}\,t. \tag{2.32}$$

$\mathbb{E}[AwD^*]_{\mathrm{EM}}$ denotes the effective medium average of the bond diffusivity. It is defined by an implicit equation and should (with a proper choice of $D^*$) not depend on the link $(\boldsymbol{x},\boldsymbol{y})$. $C_{\triangle^*}$ is a *connection factor* that is defined by

$$C_{\triangle^*} = \sum_{\boldsymbol{\xi}\in T(\Omega)} A^*(\boldsymbol{x},\boldsymbol{x}+\boldsymbol{\xi})\,\frac{\boldsymbol{\xi}^2}{D^*(\boldsymbol{x},\boldsymbol{x}+\boldsymbol{\xi})} = \sum_{\boldsymbol{\xi}\in T(\Omega)} \mathbb{1}_{(0,r]}(|\boldsymbol{\xi}|)\,\frac{\boldsymbol{\xi}^2}{D^*(|\boldsymbol{\xi}|)} \tag{2.33}$$

for translationally invariant, isotropic systems.

In that way, we may put:

$$\mathbb{E}[\triangle_{A,w}]_{\mathrm{EM}} = \mathbb{E}[AwD^*]_{\mathrm{EM}}\,\triangle^*. \tag{2.34}$$

The connection factor arises when the msd is computed in the same way as for the binomial random walk. This equation summarizes the discussion on effective medium approximation.

### Coarse graining and projection of state space

The effective medium average performs best when applied to an environment consisting of i.i.d. symmetric transition rates. We have to expect comparably poor results, when the transition rates are either correlated or asymmetric. If the equilibrium state is disordered, that means $\mathcal{B}$ is not the identity operator, the transition rates $\lambda$ are asymmetric and correlated to each other via the Boltzmann factors. In this case one has to resort to other methods to explore the properties of the annealed model.

One method could be coarse graining. By switching to a mesoscopic scale, small scale – and thus short time – features are averaged and one expects a certain degree of homogenization in suitable ensembles of environments. In terms of the operator $\mathcal{B}$ this would mean that – at a certain scale – we may safely replace $\mathcal{B}$ by a multiplication, i.e. $\mathcal{B} = \mathbb{E}\big[e^{-\beta\mathscr{F}}\big]\,\mathbb{1}$. This requires that the fluctuations in the equilibrium state are somehow bounded and there exists an average Boltzmann factor. This would also mean that on mesoscopic scales and upwards the activities and concentrations only differ by a constant factor and hence are independent from each other.

**Coarse graining in the transition graph:** In fact the authors of [45] argue that this can be done. They consider the same model Eq.(2.16) as we do, with the restriction to only allow next neighbour transitions. From the observation that in the stationary state (in which the diffusivity is defined) the activity at $\boldsymbol{x}$ is a weighted mean of its neighbouring activities

$$\eta(\boldsymbol{x}) = \sum_{\boldsymbol{y}\in\Omega} \frac{A(\boldsymbol{x},\boldsymbol{y})\,w(\boldsymbol{x},\boldsymbol{y})}{\sum\limits_{\boldsymbol{z}\in\Omega} A(\boldsymbol{x},\boldsymbol{z})\,w(\boldsymbol{x},\boldsymbol{z})}\eta(\boldsymbol{y})$$

they conclude that the activity profile must homogenize at some mesoscopic scale. A spatial average over such mesoscopic regions is applied and results in $\rho = \mathbb{E}\big[e^{-\beta\mathscr{F}}\big]\,\eta$.

The same argument applies to our case: The activity is a weighted average over all next neighbours. This average, however, does not translate directly into a spatial average! Nodes that are next neighbours may be very far apart or very near, depending on the topology of the transition graph, and how the nodes are embedded in the Euclidean space. The set of nodes, over which the average is carried out, is a closed ball with respect to the chemical distance metric, which is the distance in the transition graph. The diameter of such a ball in Euclidean space may not grow linearly with the

chemical distance and their scaling relation essentially determines one of the fractal dimensions of the transition graph. In the presence of long-range connections it will grow faster, when there are many connections missing with respect to a simple cubic lattice, it may grow slower. The interpretation of coarse graining and a transition to a mesoscopic scale is only possible, when this scale can be chosen small with respect to the total diameter of the lattice but still averages over a sufficient number of nodes. This is possible, when there is a maximum connection distance $r$, i.e. in the absence of long-range connections. Otherwise, the activity will not homogenize on any mesoscopic scale, but on the whole lattice.

If such a homogenization is possible, the average of the activities translates into an average over the Boltzmann factors. This spacial average over the Boltzmann factors on several nodes will by virtue of the strong law of large numbers converge to the expectation of the Boltzmann factor, provided enough nodes participate in the averaging procedure. In that case we have:

$$\mathcal{B} = \mathbb{E}\left[e^{-\beta\mathscr{F}}\right] \mathbb{1} \tag{2.35}$$

on the mesoscopic scale. The expectation of the Boltzmann factors does not need to be finite. In fact, its divergence leads to a departure from normal diffusion. Expressed in a non-rigorous way, this would lead to $\dot{\rho} = 0$, indicating very slow dynamics.

**Application of more general projections:** The coarse graining procedure we just described, was applied in the transition graph, we tried to take an average over next neighbours in that topology. In the presence of long-range connections, the interpretation as an average over mesoscopic regions in Euclidean space fails. Such an average still can be applied. It belongs to a larger part of operations, namely projections, that also includes the elimination of irrelevant internal states. The original treatment of such projections is due to Zwanzig, [152], who demonstrated that disregarding parts of the full dynamics may lead to memory effects. A repetition of his treatment in Laplace domain is given in appendix A.5.

The procedure remains the same, regardless of the nature of the projection, be it a true coarse graining procedure or an elimination of irrelevant variables. The coarse grained dynamics are described by a generalized master equation, consisting of a Laplacian with a memory kernel and a new source term:

$$\dot{\rho}(\boldsymbol{x};t) = \int_0^t \mathrm{d}t'\, \mathcal{M}_{\triangle_\lambda}(t-t')\, \rho(\boldsymbol{x};t') + \mathcal{I}_{\triangle_\lambda}[\rho_0;t]\,. \tag{2.36}$$

The interpretation is clear: the diffusing particles' small scale dynamics (inside the coarse grained region) are not visible any more but still relevant. As a result, the time needed for a transition is not an exponential random variable any more. Such non-exponential transitions lead to a non-trivial memory kernel giving dynamics, which are *non-local* in time. Hence, the operator $\mathcal{M}_{\triangle_\lambda}$ does not only act on the spacial coordinate but also involves memory effects. The source term $\mathcal{I}_{DisLaplacian}$ appearing in

the equation is a part of the initial concentration that was not able to make a transition visible on larger scales. It can be eliminated when an appropriate mesoscopic initial condition is chosen.

Under certain conditions, the memory effects of $\mathcal{M}_{\triangle_\lambda}$ can be neglected. This is the case, when the particle can escape the coarse grained region in a finite time. Then $\mathbb{E}\left[e^{-\beta\mathscr{F}}\right]$ is finite and the application of Eq.(2.35) is justified on large time scales. Whenever this is not the case, the reduction to a time local equation is not possible on any time scale and we say the system has a *long-range memory*.

The exact form of the memory kernel depends of course on the details of the projection that is used. Interestingly, it is not important which is the nature of the projection. It may be a coarse graining transformation, like Machta's renormalization group, [55, 158], it may be an average over irrelevant sub-states, like considered in [40], or a real projection in Euclidean space, like the random walk on the comb, [159].

**Correlations and memory:** In principle we have to expect correlations and non-stationary behaviour when a coarse graining procedure is applied. However, the topology of the transition graph and the details of the coarse graining transformation decide whether those correlations are of temporal or spatial nature. Temporal correlations occur when the tracer particles get lost inside a mesoscopic box and are related to non-exponentially distributed transition rates.

The archetype for temporal correlations is the random walk on a comb, [159]. Here, a random walker is placed on a two dimensional lattice, but it is only allowed to make a step in $x$-direction, when it is on the $x$-axis. When we project the motion onto the $x$-axis, the random walker is trapped on the teeth of the comb. When those teeth have a finite length, there is a mean time that the walker spends on the teeth between a transition on the spine. On time scales much larger than this mean time, the diffusion on the spine will appear normal. This is not the case when the teeth have infinite length, because the return time to the spine then has an infinite mean. Under these conditions, the random walk on a comb has long-range memory, which reflects in long waiting times between transitions along the $x$-axis.

The correlations in this example are, however, only of temporal nature, since when the walker returned to the spine it will jump with equal probability to the right or to the left. The reason is that all teeth of the comb are connected in the same manner to each other. When this is not the case, i.e. when not all mesoscopic regions are connected similarly, we additionally have to expect spatial correlations. An example for such behaviour would be a particle with an additional velocity variable, for instance the Lévy-Walk, [85]. When an average over the velocity variable is applied, the particle will still show a certain persistence. I.e. it keeps its direction of motion for a certain period of time (that is the mean time to change its velocity). If that mean time is infinite, the particle "gets stuck" in that particular velocity state and moves for a long time in the same direction. This kind of long-range memory induces a persistent motion, i.e. positive correlations in successive displacements.

In any case, a long-time deviation from normal diffusive behaviour can only be

expected, when the mean sojourn time in some mesoscopic region diverges. But the sojourn times are proportional to the Boltzmann factors, as Eq.(2.56) will later on show. Using again the argument of the strong law of large numbers, the mean sojourn time is finite, when $\left\langle e^{-\beta\mathscr{F}}\right\rangle$ is finite. In that sense, the discussion of the last section boils down to Eq.(2.35).

#### An annealed description with normal behaviour

Combining both methods of the last sections results in a difficult superposition of multiple effects. In the absence of long-range memory our non-rigorous derivation bears the following annealed formula:

$$\mathbb{E}[\mathcal{B}]\,\dot{\eta} = \mathbb{E}[\triangle_{A,w}]_{\mathrm{EM}}\,\eta.$$

We identify the operator on the left-hand side with the identity times the average Boltzmann factor and the operator on the right hand side with the effective medium approximation. When going back to a description by concentrations, we obtain:

$$\dot{\rho} = \frac{\mathbb{E}[AwD^*]_{\mathrm{EM}}}{\mathbb{E}[e^{-\beta\mathscr{F}}]}\triangle^*\rho. \tag{2.37}$$

The effective coefficient of normal diffusion is now given by

$$K_{\mathrm{eff}} = \frac{1}{2d}\frac{C_{\triangle^*}\mathbb{E}[AwD^*]_{\mathrm{EM}}}{\mathbb{E}[e^{-\beta\mathscr{F}}]}. \tag{2.38}$$

It summarizes all aspects of the original model (2.16): The disorder in the equilibrium state with the average Boltzmann factor, the disorder in the equilibrium transition rates with the effective medium diffusivity, and the properties of the reference system with the connection factor of the reference Laplacian, $C_{\triangle}$. This effective diffusivity has to be compared with the one of [45]. They used a simple cubic lattice with next neighbour transitions, where $D$ is equal to $a^2$ if two nodes are neighbours and zero otherwise. This gives $C_{\triangle^*} = 2d$ and $K_{\mathrm{eff}} = a^2\mathbb{E}[Aw]_{\mathrm{EM}}/\mathbb{E}\left[e^{-\beta\mathscr{F}}\right]$. The effective diffusivity is subject to discussion in the next section. We discuss in particular the cases when it is not finite.

### 2.3.3. Characterization of structural and energetic disorder

Eq.(2.38) allows us to classify the disorder we encounter in the master equation (2.16). In the last section we have seen that the original Master equation may be replaced with an annealed equation (2.37). The annealed dynamics is normal, when the effective coefficient of normal diffusion, which is given by Eq.(2.38), is finite. That means,

$$0 < 2dK_{\mathrm{eff}} = \frac{C_{\triangle^*}\mathbb{E}[AwD^*]_{\mathrm{EM}}}{\mathbb{E}[e^{-\beta\mathscr{F}}]} < \infty \tag{2.39}$$

is a sufficient condition for normal diffusion of the annealed model. A non-finite value of $K_{\text{eff}}$ can only occur when either the enumerator or the mean in the denominator is not finite. Therefore, $K_{\text{eff}}$ becomes non-finite either through properties derived from $\mathcal{B}$ or from $\triangle_{A,w}$. Most of the discussion of [45] applies here as well. By allowing arbitrary topology in the transition graph we extend their results. When $K_{\text{eff}}$ becomes zero, we have to deal with subdiffusion, when it diverges, the system is superdiffusive. We start by discussing the denominator.

**Energetic disorder:** The divergence of the average Boltzmann factor is connected to infinite mean sojourn times, as was pointed out before. As long as the free energy is not almost everywhere infinite, it can not be zero. We reproduce an argument of [45]. Consider the site energies' pdf $p(\mathscr{F})$ and the median $\mathscr{F}_m$ of their distribution.[11] Then we have:

$$\begin{aligned}\mathbb{E}\left[e^{-\beta\mathscr{F}}\right] &= \int_{-\infty}^{\mathscr{F}_m} \mathrm{d}\mathscr{F}\, p(\mathscr{F})\, e^{-\beta\mathscr{F}} + \int_{\mathscr{F}_m}^{\infty} \mathrm{d}\mathscr{F}\, p(\mathscr{F})\, e^{-\beta\mathscr{F}} \\ &> e^{-\beta\mathscr{F}_m} \int_{-\infty}^{\mathscr{F}_m} \mathrm{d}\mathscr{F}\, p(\mathscr{F}) + 0 = \frac{1}{2} e^{-\beta\mathscr{F}_m} > 0.\end{aligned}$$

The average Boltzmann factor may, however, diverge. This is connected to a strong disorder in the equilibrium state. Whenever this is the case, $K_{\text{eff}}$ will vanish and the motion will be subdiffusive. Since the disorder is found in the equilibrium state and thus in the energy landscape, we will refer to *energetic disorder*. As already discussed in section 2.3.2, energetic disorder is connected to non-stationarity and memory effects.

### Structural disorder

Let us consider the mean in the enumerator. It consists of two connected factors, the connection factor $C$ and the effective medium bond diffusivity $\mathbb{E}[AwD^*]_{\text{EM}}$. Their product may vanish or diverge.

The connection factor is given by

$$C_{\triangle^*} = \sum_{\boldsymbol{\xi}\in T(\Omega)} \mathbb{1}_{(0,r]}(|\boldsymbol{\xi}|)\, \frac{|\boldsymbol{\xi}|^2}{D^*(|\boldsymbol{\xi}|)}.$$

$C$ can only be zero, when the reference operator $\triangle^*$ is proportional to unity operator. The reference operator has to be zero then (otherwise, probability would not be conserved), and there would be no motion at all. The set of all possible transitions, or translations respectively, $T(\Omega)$ can be identified with $\mathbb{Z}^d$ itself. Consequently, when there are no long-range connections, i.e. $r < \infty$, the series and hence the connection

[11] The median free energy is the one chosen such, that it is with probability one half larger than any other.

factor remains finite. When there are long-range connections, $C$ only remains finite, when $D^*(|\boldsymbol{\xi}|)$ grows faster than $|\boldsymbol{\xi}|^3$. Otherwise the diffusing agents will perform long-range jumps with a high enough rate to amplify the diffusive transport. This will result in superdiffusion.

The effective medium bond diffusivity $\mathbb{E}[AwD^*]_{\mathrm{EM}}$ can be infinite or zero when an inappropriate reference graph is chosen. When the "right" reference system is chosen, it is smaller than infinity. It can still be zero, when the original transition graph lacks a sufficient amount of links that the reference graph is supposed to have. This is the case at the percolation threshold or on any fractal set. These effects can be mimicked when there is strong disorder in the equilibrium rates. Then there are links in the original network that exist but are rarely traversed. Both cases correspond to *local anisotropy*. The transition probabilities are not isotropic, the diffusing agents are forced to make detours, they need a large number of jumps in order to cross a small physical (Euclidean) distance.

It is not possible that the connection factor diverges and at the same time the effective medium bond diffusivity vanishes. This would mean that we have a percolation situation on a complete graph. The percolation concentration of a fully connected graph on the other hand is zero. Therefore, non-finiteness in the connection factor and in the effective medium bond diffusivity can not cancel each other.

Both properties describe the *structure* of the medium that is the connectivity and transition between different sites. It is important to note that in our approach three kinds of information are merged. The Boolean information $A$ on whether two sites are connected at all, the distance $|\xi|$ between the sites and the rate $w$ that describes the strength of the current. All three result in a motion obeying detailed balance. Hence it is anomalous diffusion, but still happens in thermodynamic equilibrium. Since it is motion In thermodynamic equilibrium it has to be stationary. Long-range connections, as well as local anisotropy result in stationary motion. We call stationary anomalous diffusion as *structural disorder*.

With the introduction of structural and energetic disorder for lattice models, we achieved one of the main goals of this chapter. In the next subsection, some lattice models featuring energetic and/or structural disorder are introduced. After that, we will generalize the notion to stochastic processes.

### 2.3.4. Examples

Let us shortly discuss some prominent models and classify the featured disorder.

**Random trap model:** We start with the random trap model. Consider a simple cubic lattice in $d$-dimensions with only nearest neighbour connections. On each site of the lattice a random energy trap is placed, i.e. there is a negative potential $U_T(\boldsymbol{x})$, but the potential vanishes between the sites. A sketch of the energy landscape can be found in Fig. 2.2a. As a consequence the motion is isotropic, each possible transition has the same overall probability. The sojourn times, however, differ from site to site. For any

given environment $\{U_T(\boldsymbol{x})\}$, the system may be described by the master equation:

$$\dot{\rho}(\boldsymbol{x},t) = w_0 \sum_{\boldsymbol{y}:|\boldsymbol{y}-\boldsymbol{x}|=a} \left[\rho(\boldsymbol{y})\, e^{\beta U_T(\boldsymbol{y})} - \rho(\boldsymbol{x})\, e^{\beta U_T(\boldsymbol{x})}\right]. \tag{2.40}$$

The transition rates have been modelled with Arrhenius' law and $w_0$ is a constant rate that determines the speed of a single transition. The mean sojourn time per site is given by

$$\tau(\boldsymbol{x}) := \frac{1}{2dw_0} e^{-\beta U_T(\boldsymbol{x})} \tag{2.41}$$

and can be compared with Eq.(2.7). Hence, the deeper the trap, the longer the sojourn time.

The dynamics is governed by the environment of energy traps $\{U_T(\boldsymbol{x})\}$. A common choice here is taking the traps to be i.i.d. exponentially distributed with mean $-E_T$. By a simple change of variables, this leads to a Pareto law for the mean waiting times

$$p_{\tau(\boldsymbol{x})}(\tau) := \frac{2dw_0}{\beta E_T} \left(2dw_0\tau\right)^{-1-\frac{1}{\beta E_T}} \mathbb{1}_{[1,\infty)}(2dw_0\tau)\,. \tag{2.42}$$

The average Boltzmann factor is proportional to the average sojourn time, i.e. to the mean of the above distribution. However, the mean of this Pareto distribution diverges, when $\beta E_T \geq 1$. Therefore, when the traps are on average deeper than the thermal energy $\beta^{-1}$, we have a case of energetic disorder induced anomalous diffusion. In fact, one can show that for the annealed model (i.e. when the average over the trap depths is taken), the msd scales as $\langle \boldsymbol{X}^2(t)\rangle \propto t^\alpha$. The exponent $\alpha$ depends on the dimension of the lattice and the mean trap depth, see [53, 158]:

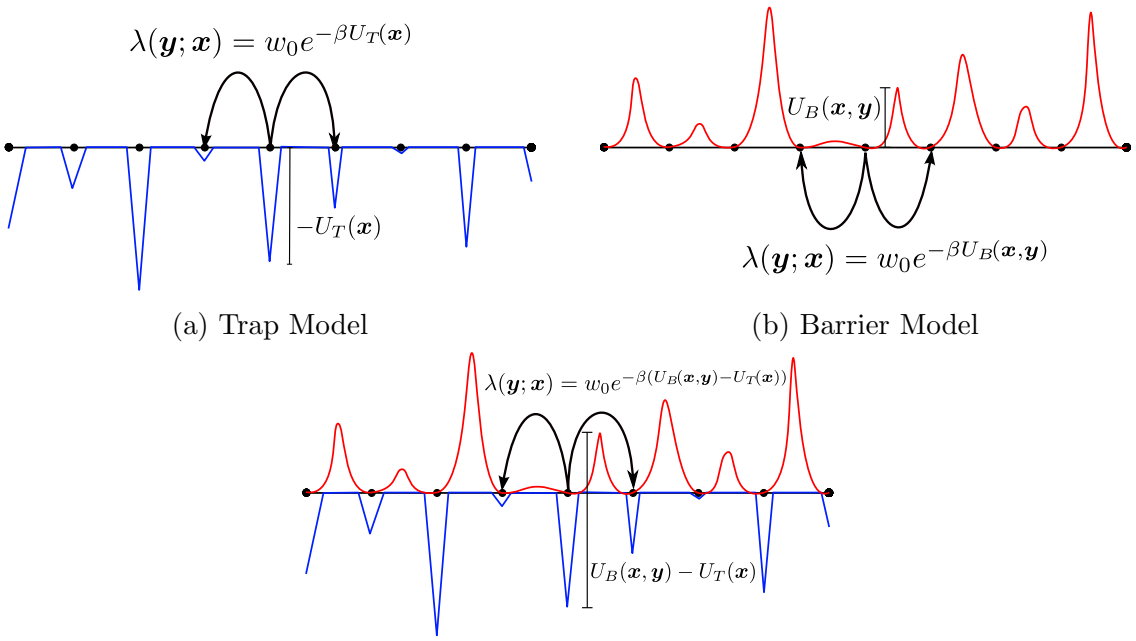
$$\alpha = \begin{cases} \dfrac{2}{2-d+d\max(1,\beta E_T)}, & d<2 \\ \dfrac{1}{\max(1,\beta E_T)}, & d>2 \end{cases} \tag{2.43}$$

and the motion may be subdiffusive. The crossover is due to the fact that in lower dimensions the random walk is recurrent, so each trap is visited several times, whereas in higher dimensions the walk is transient and each trap is only visited once. When the walker revisits a certain site, the new waiting time is drawn according to the same distribution, this leads to correlations in the waiting time sequence and the critical case $d=2$ bears additional logarithmic corrections, as $\left\langle \boldsymbol{X}^2(t)\right\rangle \propto t^\alpha \ln^{1-\alpha} t$. Interestingly, the anomalous behaviour vanishes, when the initial distribution is the equilibrium one, see [54]. This stresses the fact that energetic disorder is a non-equilibrium feature.

This so called quenched trap model has been applied in charge transport in disordered semiconductors, glassy systems and polymer dynamics, see [74, 90, 96–98, 160]. It is often employed to provide a phenomenological description of processes with a

wide range of relaxation times. There also have been analytical and numerical studies that investigate the non-stationary, i.e. ageing, features of the model, [124, 130, 131, 151, 161–167]. Since the inhomogeneous sojourn times are a feature defined on the nodes, thus independent on the topology, there are also some studies with other topologies, for instance on fractals [131]. In numerical simulations, the trap model can be generated by a simple random walk with exponentially distributed waiting times. The mean sojourn time on each site has to be chosen according to Eq.(2.41).

Let us turn to a model of structural disorder type.




(a) Trap Model

(b) Barrier Model

(c) Random potential model

Figure 2.2.: Sketches of the energy landscapes
(2.2a): A sketch of the dynamics in the quenched trap model. A particle sitting in a trap of depth $-U_T(\boldsymbol{x})$ waits an exponentially distributed time with mean $\tau(\boldsymbol{x})$ and then makes a transition with equal probability in any direction. (2.2b): A sketch of the dynamics in the random barrier model. The particle sits in site $\boldsymbol{x}$ and jumps with probability $p(\boldsymbol{y};\boldsymbol{x})$ to a neighbour site. Its sojourn time is determined by the detailed balance condition. (2.2c): A sketch of the dynamics in the random potential model. The particle at site $\boldsymbol{x}$ not only has to escape the trap of depth $-U_B(\boldsymbol{x})$ but at the same time it has to surmount the barrier of height $U_B(\boldsymbol{x},\boldsymbol{y})$. The total energy needed for a transition is the difference (remember that the trap depth is negative) of both. The barriers lead to an anisotropy in the transition probabilities, whereas the traps lead to an inhomogeneous equilibrium state.

**Random barrier model:** The set-up of the random barrier model is exactly the same as in the random trap model. In contrast to before, each site has zero potential but random, positive energy barriers $U_B(\boldsymbol{y}, \boldsymbol{x})$ are placed on each link. A sketch of this model can be found in Fig. 2.2b. The transition rates are again given by Arrhenius' law and are symmetric:

$$\lambda(\boldsymbol{y}; \boldsymbol{x}) = w(\boldsymbol{y}, \boldsymbol{x}) = w_0 e^{-\beta U_B(\boldsymbol{y}, \boldsymbol{x})}. \tag{2.44}$$

This equation can be compared to (2.8). The probabilities to jump to a neighbour site are anisotropic and given by the normalized transition rates:

$$p(\boldsymbol{y}; \boldsymbol{x}) := \frac{e^{-\beta U_B(\boldsymbol{y}, \boldsymbol{x})}}{\sum\limits_{\boldsymbol{z}:|\boldsymbol{z}-\boldsymbol{x}|=a} e^{-\beta U_B(\boldsymbol{z}, \boldsymbol{x})}}. \tag{2.45}$$

The governing master equation reads:

$$\dot{\rho}(\boldsymbol{x}, t) = w_0 \sum_{|\boldsymbol{y}-\boldsymbol{x}|=a} e^{-\beta U_B(\boldsymbol{y}, \boldsymbol{x})} \left[\rho(\boldsymbol{y}) - \rho(\boldsymbol{x})\right] \tag{2.46}$$

and the model may also show subdiffusion when the effective medium average of $e^{-\beta U_B(\boldsymbol{y}, \boldsymbol{x})}$ vanishes. When the barriers $U_B$ are distributed identically and independently this is only possible in one dimension. On the line the tracer particle will be trapped between two "high" barriers that can not be surpassed in small time. Then the msd in the annealed model can be shown to be $\left\langle \boldsymbol{X}^2(t) \right\rangle \propto t^{2\nu}$, see [53]. The exponent $\nu$ is one half in dimensions higher than one and

$$\nu = \frac{1}{1 + \max(1, \beta E_B)} \tag{2.47}$$

in one dimension. $E_B$ is the average barrier height. In this work, we choose the distribution of the barrier heights always as exponential. Again, we see that for barriers that are high compared to thermal energy, i.e. when $\beta E_B > 1$, the annealed motion is anomalous. Since the disorder is to be found in the symmetric transition rates, and the equilibrium state is the uniform one, the anomaly is due to structural disorder. In higher dimensions the tracer particle is able to find a way around those high barriers, and the diffusion will be restored to normal.

In contrast to the trap model, the probability to jump to some neighbour is anisotropic. The equilibrium distribution is however perfectly homogeneous. Hence the anomalous diffusion of the barrier model happens in thermodynamic equilibrium.

The random barrier model is a standard model for conduction in disordered media, [40, 53]. In the mathematical community the model is treated as the "random conductance model" and there is a huge body of work, see [140]. Several rigorous results, also for more complex than the here considered simple cubic topology have been derived, [56].

In numerical simulations, the barrier model is treated like a biased random walk, where the local bias is given by Eq.(2.45). The sojourn time per site is exponentially distributed and its mean is the norming factor in Eq.(2.45) times $w_0^{-1}$.

**Random Potential model:** The features of the random barrier and the random trap model can be combined to construct a model featuring energetic and structural disorder at the same time. A sketch of the energy landscape can be found in Fig. 2.2c. For a fixed environment of traps and barriers, the motion is described by the following master equation:

$$\dot{\rho}(\boldsymbol{x},t) = w_0 \sum_{\boldsymbol{y}:|\boldsymbol{y}-\boldsymbol{x}|=a} e^{-\beta U_B(\boldsymbol{y},\boldsymbol{x})} \left[\rho(\boldsymbol{y})\, e^{\beta U_T(\boldsymbol{y})} - \rho(\boldsymbol{x})\, e^{\beta U_T(\boldsymbol{x})}\right]. \tag{2.48}$$

In this work we assume the traps and barriers to be i.i.d. exponentially distributed with mean $-E_T$ for the traps and $E_B$ for the barriers. The transition probabilities only depend on the energy barriers

$$p(\boldsymbol{y};\boldsymbol{x}) := \frac{e^{-\beta U_B(\boldsymbol{y},\boldsymbol{x})}}{\sum\limits_{|\boldsymbol{z}-\boldsymbol{x}|=a} e^{-\beta U_B(\boldsymbol{z},\boldsymbol{x})}}, \tag{2.49}$$

and the mean waiting times are governed by the energy traps

$$\tau(\boldsymbol{x}) = w_0^{-1} \frac{e^{-\beta U_T(\boldsymbol{x})}}{\sum\limits_{|\boldsymbol{z}-\boldsymbol{x}|=a} e^{-\beta U_B(\boldsymbol{z},\boldsymbol{x})}}. \tag{2.50}$$

Averaging over the random landscapes bears subdiffusive motion whenever either the mean trap depth or the mean barrier height (in one dimension) are larger than the thermal energy. The msd scales as

$$\left\langle \boldsymbol{X}^2(t) \right\rangle \propto t^{2\nu\alpha} \tag{2.51}$$

with the same exponents as above.

In simulations the random potential model is treated like the barrier model. Only the sojourn time has to be adjusted by a factor $e^{-\beta U_T(\boldsymbol{x})}$ on each site.

## 2.3.5. Summary

In this section we have discussed master equation models for anomalous diffusion. The master equation describes the temporal evolution of the concentration profile. It constitutes a model of quenched disorder. The disorder is expressed by the random set of transition rates, which define the particle currents between the states. We restricted ourselves to transition rates that obey a detailed balance equation, so that there is a meaningful (although possibly not normalizable) equilibrium distribution. The disorder in the transition rates was then identified as an inhomogeneity in the equilibrium distribution, an anisotropy in the equilibrium transition rates (or transition probabilities) and a complex topology in the graph of possible transitions. By considering the activities rather than the concentration, we have been able to split the

action of the master equation operator into the action of the equilibrium distribution and the action of the equilibrium rates.

When going to an annealed description, both decouple and can be averaged independently. The equilibrium rates and the topological part can be replaced with a homogeneous and isotropic reference model by means of an effective medium approximation. The equilibrium distribution becomes uniform in a coarse grained description. In the annealed description the diffusion is normal, when the effective normal diffusivity given by

$$2dK_{\text{eff}} = \frac{C_{\triangle^*}\mathbb{E}[AwD^*]_{\text{EM}}}{\left\langle e^{-\beta\mathscr{F}}\right\rangle}$$

is neither zero nor infinity. Here $C_{\triangle^*}$ is the connection factor, describing the topology of the reference system. It may diverge in the presence of long-range connections. $\mathbb{E}[AwD^*]_{\text{EM}}$ is the effective medium bond diffusivity. It may vanish in percolation-like situations. Finally, $\left\langle e^{-\beta\mathscr{F}}\right\rangle$ is the expected Boltzmann factor, which diverges when the mean sojourn time on a lattice site does. This was the main formula of the section.

When $K_{\text{eff}}$ is not finite, we were able to define two mechanisms of anomalous diffusion: Energetic disorder, when the denominator diverges, and structural disorder, when the enumerator becomes zero or infinite. Due to energetic disorder there can only be subdiffusion in our lattice models, whereas structural disorder can induce sub- as well as superdiffusion. Both may be working at the same time. It is apparent that energetic disorder corresponds to anomalous diffusion far from equilibrium and structural disorder leads to anomalous behaviour in equilibrium. These two terms are the basis of the thesis' discussion. With the random potential model and its relatives we presented epitomes for energetic and structural disorder, as well as for a mixed situation.

In the next section, we will expand the notion of structural and energetic disorder to stochastic processes.

## 2.4. Subordinated stochastic processes

In this section, we will expand the notion of structural and energetic disorder to stochastic processes. First it will be demonstrated how the master equation Eq.(2.16) (or equivalently Eq.(2.25)) corresponds to a proper stochastic process. This gives rise to a stochastic description for the quenched disorder model. Secondly the transition to annealed disorder is performed. When this is done, we are able to extend the definition of structural and energetic disorder to stochastic processes. In the rest of the section, some processes that will appear later in the thesis will be presented and classified.

### 2.4.1. Quenched disorder

When asking for the corresponding stochastic process of some master equation, one asks about the properties of the single particles' trajectories that form the concen-

tration profile. From a mathematical point of view, stochastic processes can be characterized by their finite-dimensional distributions, i.e. the joint pdf of the random vector $(\boldsymbol{X}(t_1), \boldsymbol{X}(t_2), \ldots, \boldsymbol{X}(t_n))$ for any $n$, and certain smoothness properties.[12] For instance, the Wiener process is the unique Gaussian process with independent and stationary increments, whose correlation function is the minimum of its arguments, and that has almost surely continuous trajectories. Although such a characterization can be useful in an analytical situation, it is cumbersome when it comes to simulations or data analysis. High dimensional distributions are difficult to access, let alone continuity properties.

Random walks or jump processes may be much more fit for such purposes. Here the process is characterized by a sequence of jumps from one state to another and a sequence of waiting times in between the jumps. Those two sequences do not need need to be independent of each other and may lack nice properties like stationarity. But it turns out that such a representation is quite possible in our case. We consider equation (2.29), here repeated for convenience

$$\dot{\rho}(\boldsymbol{x};t) = \triangle_\lambda \rho(\boldsymbol{x};t) = \sum_{y\in\Omega} A(\boldsymbol{y},\boldsymbol{x})\, w(\boldsymbol{y},\boldsymbol{x}) \left[ \frac{\rho(\boldsymbol{y};t)}{e^{-\beta\mathscr{F}(\boldsymbol{y})}} - \frac{\rho(\boldsymbol{x};t)}{e^{-\beta\mathscr{F}(\boldsymbol{x})}} \right].$$

As always the initial condition is fixed to $\rho(\boldsymbol{x};0) =: \rho_0(\boldsymbol{x})$. The formal solution to the equation is given by the time evolution operator and involves the exponential of the operator $\triangle_\lambda$:

$$\rho(\boldsymbol{x};t) = \mathcal{U}_\lambda(0;t)\,\rho_0(\boldsymbol{x}) := e^{t\triangle_\lambda}\rho_0(\boldsymbol{x})\,. \tag{2.52}$$

Since $\triangle_\lambda$ does not depend explicitly on time,

$$\mathcal{U}_\lambda(t;t+\tau)\,\mathcal{U}_\lambda(0;t) = e^{(t+\tau-t)\triangle_\lambda} e^{t\triangle_\lambda} = e^{(t+\tau)\triangle_\lambda} = \mathcal{U}_\lambda(0;t+\tau)$$

holds and the family of time evolution operators forms a semi-group.[13]

When properly normalized, we may interpret the concentration profile as the pdf of a tracer particles position. We denote this stochastic process with $\boldsymbol{X}_\lambda(t)$. Consequently, the kernel of the time evolution operator is the propagator and the transition matrix of that process, i.e.

$$\langle \boldsymbol{y}|\, e^{t\triangle_\lambda}\, |\boldsymbol{x}\rangle = G_\lambda(\boldsymbol{y};t|\boldsymbol{x};0) = \mathbb{P}\{\boldsymbol{X}_\lambda(t) = \boldsymbol{y} | \boldsymbol{X}_\lambda(0) = \boldsymbol{x}\}\,. \tag{2.53}$$

By virtue of the semi-group property, the transition probability obeys the Chapman-Kolmogorov equation, hence $\boldsymbol{X}_\lambda(t)$ is (for fixed environment $\lambda$) a Markov process.

[12] This means that the processes are constructed according to Kolmogorov's extension theorem, from the finite-dimensional distributions. In that set-up it is not possible to make statements that involve the evaluation of $\boldsymbol{X}(t)$ at infinitely many epochs. Thus we can not make statements about continuity, differentiability or similar properties. Other constructions are possible.

[13] It does not form a group, since $\mathcal{U}_\lambda$ is not necessarily invertible. The time evolution of the process can not be reverted beyond the preparation epoch.

In the theory of Markov processes, $\triangle_\lambda$ is called the generator of the process [144, 145]. By our assumptions its largest eigenvalue is zero and corresponds to the stationary state (the thermodynamic equilibrium). All other eigenvalues are negative and correspond to modes in the pdf that decay over time. The spectral structure of $\triangle_\lambda$ is responsible for conservation of probability and for $\mathcal{U}_\lambda$'s stochasticity. The semi-group properties express explicitly the Markov property of the process $\boldsymbol{X}_\lambda(t)$. However, the continuous time Markov process $\boldsymbol{X}_\lambda(t)$ still depends explicitly on the environment $\lambda$ and thus on the Boltzmann factors as well as the equilibrium rates. We thus deal with a random walk *in the quenched environment*.

Some features may still be obtained at this point. Since $\boldsymbol{X}_\lambda$ is a Markov process, it can be decomposed into a uniformized Markov chain and a Poisson process of suitable intensity, see for instance [145, ch. 5]. The Poisson process then serves as an internal time clock determining the epochs of jumps, when tries to change its state. The transitions are given by the uniformized Markov chain. Although this representation is always possible when the mean interval time of the Poisson process is smaller or equal than the smallest sojourn time $\tau(\boldsymbol{x})$ found on the lattice, the representation is not optimal. A lot of the trials will not change the state of $\boldsymbol{X}$, because the sojourn times are disordered and the smallest of them dominates the Poisson process. We will choose a somewhat different approach here, and interpret $\boldsymbol{X}_\lambda$ as a jumping process. A decomposition $\boldsymbol{X}_\lambda(t) =: \boldsymbol{Y}_\lambda(U_\lambda(t))$ like for the binomial random walk is still possible.

We define a continuous time Markov process $\{\boldsymbol{Y}_\lambda(u)\}_{u\in\mathbb{R}_{0,+}}$ that jumps from site $\boldsymbol{x}$ to site $\boldsymbol{y}$ with probability:

$$p_\lambda(\boldsymbol{y};\boldsymbol{x}) := \frac{\lambda(\boldsymbol{y};\boldsymbol{x})}{\sum\limits_{\boldsymbol{z}\in\Omega}\lambda(\boldsymbol{z};\boldsymbol{x})} = \frac{A(\boldsymbol{y},\boldsymbol{x})\,w(\boldsymbol{y},\boldsymbol{x})}{\sum\limits_{\boldsymbol{z}\in\Omega}A(\boldsymbol{z},\boldsymbol{x})\,w(\boldsymbol{z},\boldsymbol{x})}. \tag{2.54}$$

This equation generalizes Eq.(2.11) and shows that $\boldsymbol{Y}_\lambda$ is fully determined by the equilibrium rates $w$ and the topology of the transition graph. Between the jumps, the process $\boldsymbol{Y}_\lambda(u)$ has to wait an exponentially distributed waiting time of mean $[\sum_{\boldsymbol{z}\in\Omega}A(\boldsymbol{z},\boldsymbol{x})\,w(\boldsymbol{z},\boldsymbol{x})]^{-1}$. The Markov-generator for $\boldsymbol{Y}_\lambda(u)$ thus is $\triangle_{A,w}$. The total time up to epoch $u$ that $\boldsymbol{Y}_\lambda(u)$ spends in a certain site $\boldsymbol{x}$ is the occupation or *local time of* $\boldsymbol{Y}_\lambda$:

$$l_{\boldsymbol{Y}_\lambda}(\boldsymbol{x};u) := \int_0^u \mathrm{d}u'\,\delta_{\boldsymbol{x},\boldsymbol{Y}_\lambda(u')}. \tag{2.55}$$

Since the random walker $\boldsymbol{Y}_\lambda(u)$ must be anywhere on the lattice, we have the identity:

$$u = \sum_{\boldsymbol{z}\in\Omega} l_{\boldsymbol{Y}_\lambda}(\boldsymbol{z};u)\,.$$

The difference between $\boldsymbol{X}_\lambda$ and $\boldsymbol{Y}_\lambda$ is only temporal. When $\boldsymbol{Y}_\lambda(u)$ reaches $\boldsymbol{x}$ and

resides there for some time, $\boldsymbol{X}_\lambda(t)$ spends the $e^{-\beta\mathscr{F}(\boldsymbol{x})}$-fold time on that site, namely:

$$\tau_\lambda(\boldsymbol{x}) = \left[\sum_{\boldsymbol{z}\in\Omega} \lambda(\boldsymbol{z};\boldsymbol{x})\right]^{-1} = \frac{e^{-\beta\mathscr{F}(\boldsymbol{x})}}{\sum\limits_{\boldsymbol{z}\in\Omega} A(\boldsymbol{z};\boldsymbol{x})\, w(\boldsymbol{z},\boldsymbol{x})}. \tag{2.56}$$

Again, this formula generalizes Eq.(2.7) for normal diffusion. At this point the Boltzmann factors enter. We see that the local time of $\boldsymbol{X}_\lambda(t)$ advances slower, when the random walker arrives at a site with a high mean sojourn time. The local times of $\boldsymbol{X}_\lambda$ and $\boldsymbol{Y}_\lambda$ only differ by a constant factor, $e^{-\beta\mathscr{F}(\boldsymbol{x})}$. When some operational time epoch $u$ is fixed, we know that the physical time $t$ already advanced to

$$\sum_{\boldsymbol{x}\in\Omega} e^{-\beta\mathscr{F}(\boldsymbol{x})} l_{\boldsymbol{Y}_\lambda}(\boldsymbol{x};u)\,.$$

The inverse of this process relates the operational time with the physical time[14]

$$U_\lambda(t) := \inf\left\{ u \geq 0 \middle| t < \sum_{\boldsymbol{x}\in\Omega} e^{-\beta\mathscr{F}(\boldsymbol{x})} l_{\boldsymbol{Y}_\lambda}(\boldsymbol{x};u) \right\}. \tag{2.57}$$

A transformation of this sort is called change of local time in the mathematical literature, see for instance [150]. The local Boltzmann factors $e^{-\beta\mathscr{F}(\boldsymbol{x})}$ are the speed measure that communicates the transformation. Since our formulation is on a lattice, the expressions expressions for the local time are pretty accessible, more care has to taken in the continuous case.[15] Nevertheless, equivalent formulations are possible. The local time thus becomes a powerful tool when proving limit theorems, [149, 150, 155].

Note that $U_\lambda$ depends on the details of the realization $\boldsymbol{Y}_\lambda$. Still, all the spacial dynamics of $\boldsymbol{X}_\lambda$ is given by $\boldsymbol{Y}_\lambda$, whereas the temporal features are given by $U_\lambda$, which only implicitly depends on $\boldsymbol{Y}_\lambda$. Let us now discuss the transition to the annealed model. This will be done in two steps, considering first the structural, then the energetic disorder part.

## 2.4.2. Annealed disorder

**Structural disorder:** Let us consider a master equation involving only structural disorder, hence the process $\boldsymbol{Y}_\lambda(t)$. Denote the pdf of that process with $\rho_{\boldsymbol{Y}_\lambda}$. Since its generator is given by $\triangle_{A,w}$, the evolution equation for the pdf takes the form:

$$\dot{\rho}_{\boldsymbol{Y}_\lambda}(\boldsymbol{x};t) = \triangle_{A,w}\rho_{\boldsymbol{Y}_\lambda}(\boldsymbol{x};t)\,.$$

[14] The alternative definition

$$U_\lambda(t) := \sup\left\{ u \geq 0 \middle| t > \sum_{\boldsymbol{x}\in\Omega} e^{-\beta\mathscr{F}(\boldsymbol{x})} l_{\boldsymbol{Y}_\lambda}(\boldsymbol{x};u) \right\}$$

differs from the original one only in its continuity properties. The original process, defined by Eq.(2.57) is continuous to the left with right limits (làdcàg). The other process is continuous to the right with left limits (càdlàg). For our discussion, this usually makes no difference.

[15] Then, the local time needs to be defined in terms of a measure.

When the effective medium approximation can be applied, we may replace $\triangle_{A,w}$ with the reference operator $K^*\triangle^*$. The such defined annealed process $\boldsymbol{Y}(t)$ is stationary and Markovian. In the presence of long-range connections, and when the distance factor $D^*$ of the reference model grows like $|\boldsymbol{x}-\boldsymbol{y}|^{1+\mu}$ with some $\mu \in (0,2)$, the reference Laplacian behaves at large scales like a fractional Laplacian. The corresponding process is a lattice version of a stable motion with index $\mu$ (which will be introduced later in the section). This is also possible when the original model had no disorder at all (i.e. when one starts with the reference system). Such motion is superdiffusive and has structural "disorder" with respect to normal diffusion. This is due to the presence of long-range connections in the transition graph.

On the other hand, we also discussed that the effective medium approximation may break down if the effective bond diffusivity vanishes or the original topology of the transition graph is complicated. The dynamics of the annealed process is in general not Markovian any more. Furthermore, it makes no sense to postulate a governing equation for the pdf, since the full hierarchy of finite dimensional distributions is needed to characterize the process. As we have seen the main feature of structural disorder is not the Markovianity of the process, but rather an homogeneous equilibrium state. Such an equilibrium state is invariant under spacial translations. Therefore any initial condition $\boldsymbol{Y}(0) = \boldsymbol{x}_0$ is sampled from equilibrium. In the same way, the dynamics needs to be invariant under temporal translations. In equilibrium, the exact time of preparation is irrelevant, only the elapsed time *since* preparation is relevant. Therefore, the *increment process*, $\Delta\boldsymbol{Y}(t;t+\tau) := \boldsymbol{Y}(t+\tau) - \boldsymbol{Y}(t)$ is stationary. This means, it does not depend on $t$. The moments of $\boldsymbol{Y}$ may exhibit a power law behaviour

$$\langle|\Delta\boldsymbol{Y}(t,t+\tau)|\rangle \sim \tau^{\nu} \tag{2.58}$$

but this is not crucial.

The stationarity of $\boldsymbol{Y}$'s increments is the main assumption on the structural disorder part. That means, we will talk about structural disorder in stochastic processes, whenever we find a process with *stationary increments*. This assumption may even hold for non-Markovian models, i.e. those that can not be represented by a proper linear master equation. One of those models, the fractional Brownian motion, will be presented later in this section.

In that sense our motivation coming from the quenched disorder master equation is somewhat stronger than the last discussion. However, with that approach we were able to see that this structural part is an interplay of the topology of the transition graph and the *equilibrium* rates. The process $\boldsymbol{Y}$ is an anomalous diffusion process in thermodynamic equilibrium, since the latter is uniform. The anomaly is caused by disorder (or memory effects) of the environment that does not change the equilibrium distribution. This is in contrast to the energetic disorder part, as we will show now.

**Energetic disorder:** We started the discussion of structural disorder disregarding the non-uniform equilibrium state. Let us again add it and discuss the changes. As

already mentioned before, the difference between $\boldsymbol{Y}_\lambda$ and $\boldsymbol{X}_\lambda$ is the local time. The disorder in the Boltzmann factors enters as a speed measure in the operational time $U_\lambda$.

$U_\lambda(t)$ is a process that can easily be understood. Without the additional factor, $U_\lambda(t)$ would just be a linear function of $t$. Since $\boldsymbol{Y}_\lambda(u)$ has to be somewhere in $\Omega$, each moment spent in some place adds to $U_\lambda$. With the introduction of the Boltzmann factors, the shape of $U_\lambda(t)$ will change to a piecewise linear one. The slope is given by $e^{-\beta\mathscr{F}}$ and changes whenever $\boldsymbol{Y}_\lambda$ makes a jump. The times between the jumps are exponentially distributed. In physical time they have spatially inhomogeneous means.

From here on it is easy to understand, what happens in the annealing procedure. All sites become equivalent, thus $U(t)$ and $\boldsymbol{Y}(u)$ decouple. When $\boldsymbol{Y}(u)$ is transient, each site is only visited once on average and $U(t)$ will become a simple renewal process. Depending on the statistics of the sojourn times $\tau(\boldsymbol{x})$ (i.e. if $\langle\tau(\boldsymbol{x})\rangle$ exists), the limit will be a Poisson process or a renewal process with heavy tailed waiting time distribution. When $\boldsymbol{Y}(u)$ is recurrent, the correlations introduced by revisiting a site have to be taken into account. This changes the statistics of the waiting time distribution, [158, 168].

After averaging over the energetic disorder, we obtain some non-decreasing process $U(t)$. One usually assumes that $U$ shows on average a power law behaviour in time

$$\langle U(t)\rangle \sim t^\alpha, \tag{2.59}$$

but in principle the only restriction on $U(t)$ is that it must be *non-decreasing*, otherwise causality is lost.

We will henceforth speak about $U(t)$ as the energetic disorder part of the overall process. Again, the formulation in stochastic processes is much more general, than the formulation in master equations. The latter allows us to understand that energetic disorder is connected to disorder in the equilibrium state and thus corresponds to anomalous diffusion far from equilibrium. When preparing the process in equilibrium and analysing the msd, it will show a linear scaling, as was shown in [54]. Energetic disorder also does not require a simple power law scaling of $U$. For instance in [169] a logarithmic time dependence arose from interaction of different tracers. One can even take $U$ as a deterministic function of time, as was done in [170] or in Batchelor's process [83, 171]. It is very important to notice that except in the Poissonian case or when $U$ is a deterministic linear function of time, $U(t)$ does not have stationary increments. And it is this non-stationarity that signalizes energetic disorder.

**Mixed disorder:** We summarize the discussions of the last paragraphs and gather all our assumptions on stochastic processes. We do not specify any more whether the state space is continuous or the motion is on a lattice, but we will interpret the state as a point in Euclidean space (and not in phase space). It is assumed that the tracer position process admits the following representation

$$\boldsymbol{X}(t) = \boldsymbol{Y}(U(t))\,. \tag{2.60}$$

Which is a composition of two stochastic processes, often referred to as a subordinated process, [144]. Here $\boldsymbol{Y}(u)$ is process with stationary increments and $U(t)$ is a non-decreasing process. $U(t)$ will be called the *operational time* in the rest of the text.

$\boldsymbol{Y}(u)$ represents the structural disorder of the process, whereas $U(t)$ stands for the energetic disorder. The power law assumptions (2.58) and (2.59) are taken just only for convenience. Other relations may be used as well. $U$ may show more complicated growth or $\boldsymbol{Y}$ may show several crossovers. The main requirements on both are that $\boldsymbol{Y}$'s increments are stationary and $U$ does not decrease! Those features may not be as apparent coming from the master equation (2.16) as they are expressed in the language of stochastic processes.

Structural disorder is absent when $\boldsymbol{Y}$ is a Brownian motion. Energetic disorder is absent when $U$ has stationary increments, i.e. when it is a Poisson process or a deterministic linear function of time.

Let us now introduce some popular stochastic processes and classify them.

## 2.4.3. Examples

**Continuous time random walks:** There are many possible variants of continuous time random walks (ctrw). The main idea is that a random walker experiences a sequence of random displacements and has to wait a random time in between its jumps. There are ctrw models with correlations between displacement and waiting times [85], and with correlations between the phases [172, 173]. In the most simple versions, the displacements and waiting times at each phase are independent on each other and on all other phases. The displacements are Gaussian random variables and the waiting times taken from a Pareto distribution (and this is exactly how to create sample trajectories numerically). After additional coarse graining, the process can be represented as the composition of a Wiener process and a renewal process

$$\boldsymbol{X}(t) = \sqrt{2K_{2,\alpha}}\boldsymbol{B}(U(t)) \tag{2.61}$$

where $K_{2,\alpha}$ is a diffusivity with dimensions $[K_{2,\alpha}] = \mathrm{m}^2/\mathrm{sec}^\alpha$, and $U(t)$ is the non-stationary renewal process that counts the number of jumps. This means that the inter-renewal times are Pareto distributed with parameter $\alpha > 0$. When $\alpha < 1$, then $\langle U(t)\rangle \sim t^\alpha$. In the marginal case $\alpha = 1$ it holds that $\langle U(t)\rangle \sim t/\log t$, otherwise $\langle U(t)\rangle \sim t$. Hence, ctrw is a model for subdiffusion, when $\alpha \leq 1$, i.e. when the waiting time distribution lacks a finite mean. The diffusion in ctrw is impaired by the presence of long periods without any motion, the random walker is transiently trapped. Furthermore, since the Wiener process is stationary, the anomaly is due to pure energetic disorder. This is not surprising, since the ctrw is a mean field model for the random trap model from section 2.3.4. When an average over the random environment of the traps is applied, for instance by some renormalization procedure [158], the limit process will be the ctrw for dimensions higher than two. For lower dimensions the recurrence of the random walk will destroy the simple renewal picture. The limit process will be a different one. In one dimension, $U$ will be a Kesten-Spitzer-process, [168]. The ctrw was originally introduced in the context of charge

carrier diffusion in disordered semiconductors [90]. Due to its analytical accessibility it was a constant object of study and generalization [38, 92, 174–178] until today. In experimental context, relatives of the ctrw are used to model diffusion processes, which exhibit transient trapping, [66, 70, 72].

**Batchelor's process:** When one properly rescales the centre of mass of a cloud of continuous time random walkers, then one obtains the Batchelor's process, often also called scaled Brownian motion. This is also a process of energetic disorder and may be represented as

$$\boldsymbol{X}(t) := \sqrt{2\alpha K_{2,\alpha}}\boldsymbol{B}(t^{\alpha}) \tag{2.62}$$

with again $\alpha > 0$. The operational time is for this particular process given by a deterministic function. The power law can be replaced with another non-decreasing function, like in [170]. As it has been mentioned and described in [179], Batchelor's process is a mean field model for ctrw. It was first proposed by G.K. Batchelor as a model for turbulent diffusion, [83]. In the original formulation, the model was stated as a diffusion equation with a time dependent diffusion coefficient

$$\dot{\rho}(\boldsymbol{x};t) = \alpha K_{2,\alpha} t^{\alpha-1} \triangle \rho(\boldsymbol{x};t)\,, \tag{2.63}$$

in opposition to Richardson's diffusion equation with a space-dependent diffusivity. Despite its prevalence when modelling frap recovery curves [62, 73, 180, 181] it remains a phenomenological model. However, Postnikov reported a physical situation, namely diffusion in layers, where Batchelor's process arises as a physical model, [171].

**Stable motions:** The stable motion is a generalization of Brownian motion, in which the independent Gaussian increments are replaced with independent $\mu$-stable ones, $\mu \in (0,2)$, [182, 183]. This results in a process whose increments' moments scale as $\langle |\Delta \boldsymbol{Y}(t;t+\tau)|^{\gamma}\rangle \sim \tau^{\gamma/\mu}$ but diverge whenever $\gamma \geq \mu$. The divergence of the distribution's moments, in particular the divergence of the mean squared displacement poses a problem for data analysis. To extract the scaling one has to resort to other measures like the quantile $[|\Delta \boldsymbol{Y}(t,t+\tau)|^{\gamma}]_p$, the value, which is smaller than a certain fraction $p$ of all realizations of the random variable $|\Delta \boldsymbol{Y}(t;t+\tau)|^{\gamma}$. It is the inverse of the survival function, which is the complement of the cumulative distribution function For the increment process, it is defined implicitly by the equation

$$p = S_{|\Delta \boldsymbol{Y}_\tau|^{\gamma}}\Big([|\Delta \boldsymbol{Y}(t,t+\tau)|^{\gamma}]_p\Big) := \mathbb{P}\Big\{|\Delta \boldsymbol{Y}(t,t+\tau)|^{\gamma} > [|\Delta \boldsymbol{Y}(t,t+\tau)|^{\gamma}]_p\Big\}\,. \tag{2.64}$$

This quantity always exists, and for a stable motion it is proportional to $\tau^{\gamma/\mu}$. Nevertheless, since stable motions have independent and stationary increments, this is superdiffusion caused by pure structural disorder. This is also apparent from the fact that the generator of this process is the fractional Laplacian, a special kind of non-local operator, [183].

Stable motions are often called Lévy flights have been a very popular model for animal and human motion, [11, 19, 22, 30, 31], where power laws in the pdfs of the displacements were found. Over the years, some of those findings had to be corrected though, [20]. They also attracted much attention in the context of random searches, [16, 24, 30, 112, 113]. In physics, the model was used to phenomenologically describe tracer diffusion in turbulent flows, [84, 86], but has been criticized for its lack of the second moment. The divergence of the second moment corresponds to an infinite velocity, which in turn translates into infinite kinetic energy. To overcome this problem, the long range displacements in Lévy flights have been compensated with long travel times, so that a finite velocity is restored. This modification of Lévy-flights was termed "Lévy walks" and was first introduced by Shlesinger and Klafter in [85] and is subject to investigation up to now, [184]. Regardless of the difficulties for the physicist, the fractional Laplacian is exactly the continuous limit of the long range reference operator considered in section 2.3.2. Thus stable motions arise when considering lattice models with long-range connections.

Stable motions are easily generated numerically by means of a random walk. Each step the position is incremented by a symmetric $\mu$-stable random variable with a certain scale factor and time is incremented by an exponential random variable of a certain mean. The ratio between the scale factor and the mean of the exponential random variable defines the anomalous diffusivity.

**Fractional Brownian motion:** Fractional Brownian motion (fBm) is a model of anomalous diffusion caused by long-range correlations. It is defined as the Weyl fractional derivative of a Wiener process

$$\boldsymbol{Y}(t) := \frac{\sqrt{2K_{2,2\nu}}}{\Gamma\left(\nu+\frac{1}{2}\right)}\left[\int_{-\infty}^{t} \mathrm{d}\boldsymbol{B}(t')\ (t-t')^{\nu+\frac{1}{2}-1} - \int_{-\infty}^{0} \mathrm{d}\boldsymbol{B}(t')\ (0-t')^{\nu+\frac{1}{2}-1}\right] \tag{2.65}$$

but it is rather characterized as a Gaussian process with the covariance function

$$\langle \boldsymbol{Y}(t_1)\,\boldsymbol{Y}(t_2)\rangle = K_{2,2\nu}\left[t_1^{2\nu}+t_2^{2\nu}-|t_1-t_2|^{2\nu}\right]. \tag{2.66}$$

$\nu$ is the so called self-similarity or Hurst index and ranges from zero to unity, [185]. In fact fBm is the only Gaussian self similar process. For $\nu < 1/2$ the motion is subdiffusive and anti-correlated, for $\nu > 1/2$ it is superdiffusive due to strong positive correlations. When the Hurst index is exactly one half, the motion reduces to Brownian motion.

For Hurst indices smaller than one half, it is used as a model for diffusion in a viscoelastic medium in particular large molecules in living cells, [51, 60, 71], or polymer translocation through a small pore, [78, 79]. Since it has stationary increments, fBm is also a process of structural disorder.

In order to numerically generate trajectories of fBm, we used a technique described

by Saxton, [180]. The following Weierstrass-Mandelbrot function

$$W(t) := \sum_{n=-\infty}^{\infty} \frac{\cos(C_n) - \cos\left(C_n + z^n \frac{2\pi t}{T}\right)}{z^{n\nu}} \tag{2.67}$$

approximates a one dimensional fractional Brownian motion on the interval $[0, T]$. Here $\{C_n\}_{n\in\mathbb{Z}}$ is a sequence of $[0, 2\pi)$-uniformly distributed random variables, and $z$ is an irrational number. The approximation gets better as $z$ approaches one. In our simulations, we use $z = \sqrt{\pi}$. The series is evaluated from $n = -10$ to 50.

**Subordinated fractional Brownian motion:** A process of mixed origins can easily be constructed by composition of fBm with the renewal process of ctrw. The so obtained subordinated fractional Brownian motion (sfBm) features anomalous diffusion with stationary and non-stationary parts:

$$\boldsymbol{X}_{\text{sfBm}} := \boldsymbol{Y}_{\text{fBm}}(U(t)) \tag{2.68}$$

Although there is no derivation from first principles, this toy model is often used when trying to understand diffusion of proteins in biological cells, [66, 70, 72]. Related processes are also used in finance models [36].

For simulation of the process it is sufficient to create trajectories of a fBm as described above, and stop them by means of a time increment as it is done in a ctrw. That means, in each step of the random walk, the position is updated according to a fBm process and time is updated by a Pareto-distributed random variable, like in ctrw.

## 2.5. Summary

In chapter 2 of the thesis, we tried to classify two physical mechanisms that are responsible for anomalous diffusion: structural and energetic disorder. After introducing the necessary notation in normal diffusion, we started with a model of diffusion in disordered media: a master equation with random rates:

$$\dot{\rho}(\boldsymbol{x}; t) := \sum_{y\in\Omega} \left[\lambda(\boldsymbol{x}; \boldsymbol{y})\, \rho(\boldsymbol{y}; t) - \lambda(\boldsymbol{y}; \boldsymbol{x})\, \rho(\boldsymbol{x}; t)\right].$$

By considering the graph of all possible transitions, using Kirchhoff's theorem and requiring a detailed balance condition, we were able to split up the random rates $\lambda(\boldsymbol{y}; \boldsymbol{x})$ into three parts

$$\lambda(\boldsymbol{y}; \boldsymbol{x}) = A(\boldsymbol{x}, \boldsymbol{y})\, w(\boldsymbol{x}, \boldsymbol{y})\, e^{-\beta \mathscr{F}(\boldsymbol{x})}.$$

The topology of the transition graph that is described by its adjacency matrix $A$ describes the possible jumps the tracer particle is allowed to make. The symmetric

equilibrium rate $w(\boldsymbol{x}, \boldsymbol{y})$ denotes the rate of particles moving from one site to another in thermodynamic equilibrium. Finally, the Boltzmann factor $e^{-\beta \mathscr{F}(\boldsymbol{x})}$ is proportional to the number of particles in $\boldsymbol{x}$ at thermodynamic equilibrium. This decomposition allowed us to rewrite the master equation in terms of the thermodynamic activities:

$$e^{-\beta \mathscr{F}(\boldsymbol{x})} \dot{\eta}(\boldsymbol{x}; t) := \sum_{y \in \Omega} A(\boldsymbol{x}, \boldsymbol{y})\, w(\boldsymbol{x}, \boldsymbol{y})\, [\eta(\boldsymbol{y}; t) - \eta(\boldsymbol{x}; t)]\,.$$

After deriving the last equation, we set out to find an appropriate annealed model, i.e. the proper way to average over the random realizations of $\lambda$. We argued that the symmetric operator, the right-hand side of the last equation, can be treated with the effective medium approximation and the equilibrium state appearing on the left-hand side will homogenize, when a suitable coarse graining transformation is applied. In the annealed description, the concentration profile thus is determined by:

$$\mathbb{E}\left[e^{-\beta \mathscr{F}}\right] \dot{\rho}(\boldsymbol{x}; t) = \sum_{\boldsymbol{y} \in \Omega} \mathbb{E}[AwD^*]_{\mathrm{EM}}\, \frac{\mathbb{1}_{(0,r]}(|\boldsymbol{x} - \boldsymbol{y}|)}{D^*(|\boldsymbol{x} - \boldsymbol{y}|)}\, [\rho(\boldsymbol{y}; t) - \rho(\boldsymbol{x}; t)]\,.$$

Within the effective medium approach, the transition graph has been replaced with a regular reference graph, with maximum connection distance $r$. $r$ is determined by the farthest possible transition of the original ensemble of transition graphs, i.e. by $A$. Large jumps are penalized with a distance factor $D^*$ that makes them less probable, and is determined by the statistical distance dependence of the equilibrium rates $w$. The correct choice of $r$ and $D^*$ makes the effective medium bond diffusivity $\mathbb{E}[AwD^*]_{\mathrm{EM}}$ well defined. The average Boltzmann factor $\mathbb{E}\left[e^{-\beta \mathscr{F}}\right]$ is well defined, when the system successfully homogenizes.

Under these conditions, we derived the effective coefficient of normal diffusion:

$$K_{\mathrm{eff}} = \frac{C_{\triangle^*} \mathbb{E}[AwD^*]_{\mathrm{EM}}}{2d \mathbb{E}[e^{-\beta \mathscr{F}}]}.$$

When it is infinite, the process is superdiffusive, when it vanishes the process is subdiffusive. Two physical mechanisms can be identified in that way: The denominator can not vanish. It can only diverge, when the equilibrium state is strongly disordered. This leads to subdiffusion and is called energetic disorder. Energetic disorder is related to diffusion far from equilibrium. When the enumerator is non-finite, on the other hand, we speak about structural disorder. The enumerator can vanish, when the effective medium bond diffusivity vanishes, which is possible in percolation-like situations (i.e. on percolation clusters or on tree-like structures with strong disorder in the equilibrium rates). When the enumerator diverges, the connection factor $C_{\triangle^*}$ is responsible. This is possible in the presence of long-range connections in the reference graph. The non-finiteness of the enumerator is connected to anomalous properties of the dynamics in thermodynamic equilibrium, hence coined structural disorder.

The formula for the effective coefficient of normal diffusion put us in the position to define two classes of anomalous diffusion mechanisms. Energetic disorder, corresponding to strong disorder in the equilibrium state, and structural disorder, corresponding

to disorder or anomalous properties in the equilibrium rates. Since the usual description of anomalous diffusion models is stochastic processes, we needed to expand the two concepts to the language of stochastic processes.

This was done by interpreting the operator in master equation with random rates as the generator of a Markov process. We discussed that the annealing procedure possibly destroys the Markovianity of the process, but other features remain. First, the structural disorder motion, happening in equilibrium, becomes a process with stationarity increments. The energetic disorder part can be seen as a change of local time and becomes a subordinator. A stochastic process of mixed, structural and energetic, disorder therefore has the following subordinated form:

$$\boldsymbol{X}(t) = \boldsymbol{Y}(U(t))\,.$$

Where $\boldsymbol{Y}$ represents the structural disorder component and is a process with stationary increments. $U$ is called the operational time. It represents the energetic disorder and is a non-decreasing process.

In that way we defined the terms energetic and structural disorder. Let us now turn to a method that is able to separate both ingredients $\boldsymbol{Y}$ and $U$ from each other.

# 3. Fundamental moment and exponent

## 3.1. Introduction

The terms energetic and structural disorder have been introduced in the last chapter. It was shown, how the anomalous behaviour could emerge in equilibrium or non-equilibrium or possibly in a mix of both situations. We were able to separate the stationary and non-stationary part with the help of a subordination ansatz

$$\boldsymbol{X}(t) = \boldsymbol{Y}(U(t))\,. \tag{3.1}$$

Where the operational time $U(t)$ is a non-decreasing process and $\boldsymbol{Y}(u)$ is a process with stationary increments. $U$ represents the energetic, non-equilibrium disorder, whereas $\boldsymbol{Y}(u)$ represents the structural disorder part giving anomalous diffusion in equilibrium.

The distinction of structural and energetic disorder is a great tool to understand different mechanisms that cause anomalous diffusion. Nevertheless, if we are not able to identify the mechanism in a realistic situation, the distinction remains worthless. Therefore we need a test that helps us to separate $\boldsymbol{Y}$ and $U$. This is also the question, whether an experiment needs to be described with an equilibrium or non-equilibrium model. There are already methods to distinguish between different models. For instance, there is the $p$-variation test that discriminates between ctrw and fBm, [137]. Furthermore an inspection of the coverage of a process distinguishes fBm from diffusion on a fractal, [136]. The investigation of the process' ergodic properties can also help to identify non-stationary models. This will be discussed in greater detail in the next chapter, i.e. chapter 4. Here we propose a method that is not only able to distinguish between energetic and structural disorder, but also to break up the anomaly into its structural and energetic part.

The rest of the chapter is structured as follows. First, we introduce our basic assumptions on the energetic and structural disorder parts. Then we present the fundamental properties and show how they separate the energetic and structural disorder. After that, we discuss these methods in several examples and close the chapter with a summary.

## 3.2. A scaling assumption

For the rest of this chapter, we refine our assumptions on the structural disorder part. Despite having stationary increments, we assume that $\boldsymbol{Y}$ exhibits a certain scaling

behaviour. In terms of the survival function, our assumption reads:

$$S_{|\Delta \boldsymbol{Y}_\omega|}(\xi) := \mathbb{P}\{|\Delta \boldsymbol{Y}(u; u+\omega)| \geq \xi\} = f\left(\frac{F(\xi)}{\omega}\right). \tag{3.2}$$

Here $f(x)$ gives the scaling form of the pdf and $F(\xi)$ is a strictly monotonously growing function that connects the operational time increment $\omega$ with the displacement $\xi$. The absence of the operational time epoch $u$ on the right hand side of the last equation expresses the stationarity of the increments. $f(x)$ is a non-increasing function with $f(0) = 1$ and $f(x \to \infty) = 0$. Since the survival function is the complement of the cumulative distribution function, we can express (3.2) in terms of the pdf of the increments:

$$\rho_{|\Delta \boldsymbol{Y}_\omega|}(\xi) = -\frac{\mathrm{d}}{\mathrm{d}\xi} S_{|\Delta \boldsymbol{Y}_\omega|}(\xi) = -\frac{1}{\omega} F'(\xi)\, f'\left(\frac{F(\xi)}{\omega}\right). \tag{3.3}$$

The function $F$ may show different power law regimes and can be as complicated as necessary, we only need to require that it is invertible. It is, however, pretty abstract. We focus on the most important candidate, which is a simple power function, $F(\xi) = \xi^{\frac{1}{\nu}}$. With this choice, the scaling assumption reads

$$\rho_{|\Delta \boldsymbol{Y}_\omega|}(\xi) = -\frac{1}{\nu} \frac{\xi^{\frac{1}{\nu}-1}}{\omega} f'\left(\frac{\xi^{\frac{1}{\nu}}}{\omega}\right). \tag{3.4}$$

From Eq.(3.4) (or alternatively Eq.(3.2)) we readily obtain the moments of $\Delta \boldsymbol{Y}$:[1]

$$\langle |\Delta \boldsymbol{Y}(u; u+\omega)|^\gamma \rangle = A_{\gamma\nu} \omega^{\gamma\nu}, \tag{3.5}$$

or alternatively

$$\langle F^\gamma(|\Delta \boldsymbol{Y}(u; u+\omega)|) \rangle = A_\gamma \omega^\gamma. \tag{3.6}$$

The constant $A_\gamma$ depends on the scaling function $f(x)$ and is not necessarily finite.

Although $U(t)$ could be any non-decreasing function, we assume that it behaves like a power law on average:

$$\langle U(t) \rangle = t^\alpha. \tag{3.7}$$

The assumptions (3.7) and (3.4) enable us to quantify the energetic disorder with $\alpha$ and the structural disorder with $\nu$. More difficult set-ups are certainly possible, we focus on this case, because it is the most lucid.

[1] The constant is given by

$$A_\gamma = -\int_0^\infty \mathrm{d}\zeta\, f'(\zeta)\, \zeta^\gamma.$$

## 3.3. Definition of the fundamental properties and first properties

We know that $\boldsymbol{X}(t)$ is the composition of two random processes. Imagine, we would know $U(t)$ and $\boldsymbol{X}(t)$ over the course of time. Then we could reconstruct $\boldsymbol{Y}$. Mathematically, this can be expressed by using conditional random variables; we have $\boldsymbol{Y}(u) = \boldsymbol{X}(U(t))\,|U(t) = u$. Unfortunately, such a decomposition is only theoretically possible. In practice, we have access to moments. Nevertheless, using the conditional average, we can compute the moments of $\boldsymbol{X}$'s displacement:[2]

$$\begin{aligned}\left\langle |\Delta \boldsymbol{X}(t;t')|^{\gamma} \right\rangle =& \left\langle \mathbb{E}\left[ |\Delta \boldsymbol{X}(t;t')|^{\gamma} \,|U(t)\right] \right\rangle = \left\langle \mathbb{E}\left[ |\Delta \boldsymbol{Y}(\Delta U(t;t'))|^{\gamma} \,|\Delta U(t;t')\right] \right\rangle \\ =& A_{\gamma\nu} \left\langle \Delta U^{\gamma\nu}(t;t') \right\rangle .\end{aligned}$$

By choosing the correct $\gamma$ this expression becomes linear in $\Delta U$ and therefore additive!

We may fix this special index, and call it the *fundamental exponent* of the process. The corresponding moment is the *fundamental moment*. The pair is defined in the following way:

*The fundamental moment is the one moment of displacement that is additive in time, i.e. it fullfills:*

$$\left\langle |\Delta \boldsymbol{X}(t_1;t_3)|^{\gamma_F} \right\rangle = \left\langle |\Delta \boldsymbol{X}(t_1;t_2)|^{\gamma_F} \right\rangle + \left\langle |\Delta \boldsymbol{X}(t_2;t_3)|^{\gamma_F} \right\rangle , \tag{3.8}$$

for $t_1 \leq t_2 \leq t_3$. *The corresponding index $\gamma_F$ is called the fundamental exponent.*

When an ensemble of trajectories of $\boldsymbol{X}$ is given, obtained from measurements or simulations, Eq.(3.8) is easily numerically inverted for $\gamma_F$. Moments are easy and quick to calculate. The drawback of our method is that an *ensemble* is needed and results are expected to get better the larger it is. Hence, it has to be modified in the single trajectory scenario. In Fig. 3.1 you can find a small sketch illustrating the additivity property of the fundamental moment.

**The fundamental exponent reveals structural disorder:** This is easy to see by using Eq.(3.5) and the definition (3.8). Evaluating the $\gamma$-th power of displacement we have:

$$\begin{aligned}0 =& \left\langle |\Delta \boldsymbol{X}(t_1;t_3)|^{\gamma} \right\rangle - \left\langle |\Delta \boldsymbol{X}(t_1;t_2)|^{\gamma} \right\rangle - \left\langle |\Delta \boldsymbol{X}(t_2;t_3)|^{\gamma} \right\rangle \\ =& A_{\gamma\nu} \left\{ \left\langle \Delta U^{\gamma\nu}(t_1;t_3) \right\rangle - \left\langle \Delta U^{\gamma\nu}(t_1;t_2) \right\rangle - \left\langle \Delta U^{\gamma\nu}(t_2;t_3) \right\rangle \right\}\end{aligned}$$

[2] Taking the conditional average is possible, when $U(t)$ is stochastically independent of $\boldsymbol{Y}$ or when it is a change of local time. In the latter case, $\boldsymbol{Y}$ and $U(t)$ are defined on the same probability space and thus $\boldsymbol{Y}$ is measurable with respect to $U$.

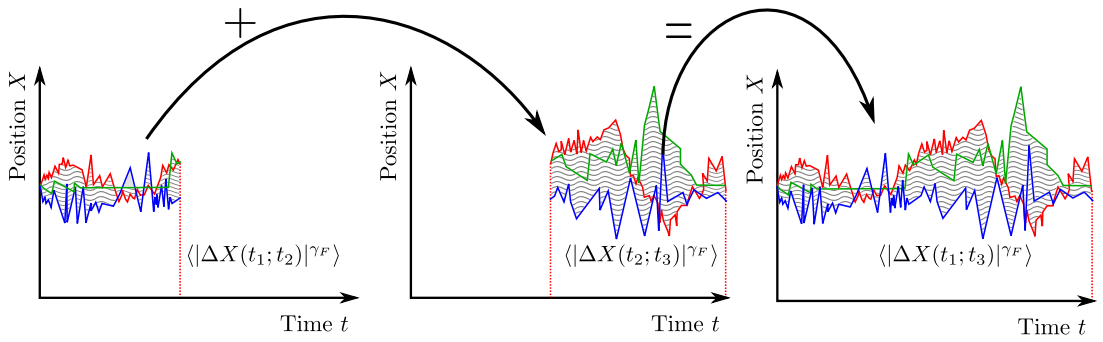



Figure 3.1.: The additivity of the fundamental moment
Computing the fundamental moment of some consecutive intervals and adding the results gives the same value as computing the fundamental moment of the composite interval.

When choosing $\gamma = \gamma_F = \frac{1}{\nu}$, we have by additivity of the expectation:

$$\begin{aligned} 0 =& \langle|\Delta \boldsymbol{X}(t_1;t_3)|^{\gamma_F}\rangle - \langle|\Delta \boldsymbol{X}(t_1;t_2)|^{\gamma_F}\rangle - \langle|\Delta \boldsymbol{X}(t_2;t_3)|^{\gamma_F}\rangle \\ =& A_1 \langle U(t_3) - U(t_1) - U(t_2) + U(t_1) - U(t_3) + U(t_2)\rangle = A_1 \langle 0\rangle . \end{aligned}$$

We recover the definition. Hence, if one is able to find the fundamental index, one also has knowledge about the scaling exponent $\nu$. In that way one reveals the structural disorder part of the process.

**The fundamental moment reveals energetic disorder:** When the fundamental exponent is already known, all moments of the subordinator are directly accessible. This is seen, by taking the same conditional average as before:

$$\langle|\boldsymbol{X}(t)|^{m\gamma_F}\rangle = A_m \langle U^m(t)\rangle . \tag{3.9}$$

This is of course only true, when the corresponding moment of $\boldsymbol{X}$ exists, i.e. when $A_m$ is finite. In particular the scaling exponent $\alpha$ of energetic disorder part can be inferred from the slope of the fundamental moment in a log-log-plot.

**Diverging fundamental moment:** When $\boldsymbol{X}$ does not possess all moments, the fundamental moment may also diverge. Still, an additivity property may hold and the scaling behaviour can be extracted by using more robust measures than moments. The fundamental exponent can be identified by looking for additivity in the quantiles:

$$[|\Delta \boldsymbol{X}(t_1;t_3)|^{\gamma_F}]_p = [|\Delta \boldsymbol{X}(t_1;t_2)|^{\gamma_F}]_p + [|\Delta \boldsymbol{X}(t_2;t_3)|^{\gamma_F}]_p . \tag{3.10}$$

To see this, consider the definition of the $p$-quantile and apply the same conditional average as before:

$$\begin{aligned} p =& \mathbb{P}\Big\{ |\Delta \boldsymbol{X}(t;t')|^{\gamma_F} \geq \left[ |\Delta \boldsymbol{X}(t;t')|^{\gamma_F} \right]_p \Big\} \\ =& \Big\langle \mathbb{P}\Big\{ |\Delta \boldsymbol{X}(t;t')| \geq \left[ |\Delta \boldsymbol{X}(t;t')|^{\gamma} \right]_p^{\frac{1}{\gamma_F}} \Big| \Delta U(t;t') \Big\} \Big\rangle \\ =& \Big\langle S_{|\Delta \boldsymbol{Y}_{\Delta U(t;t')}|} \Big( \left[ |\Delta \boldsymbol{X}(t;t')|^{\gamma_F} \right]_p^{\frac{1}{\gamma_F}} \Big) \Big\rangle \\ =& \left\langle f\left( \frac{\left[ |\Delta \boldsymbol{X}(t;t')|^{\gamma_F} \right]_p}{\Delta U(t;t')} \right) \right\rangle . \end{aligned}$$

Although this definition is implicit, one still can observe that $\left[ |\Delta \boldsymbol{X}(t_1;t_2)|^{\gamma_F} \right]_p$ scales linearly with the operational time and will thus be additive in time.

**Different scaling regions:** The unique scaling behaviour of $\boldsymbol{Y}$ is a requirement that is not always met. In some important situations, for instance in the Rouse model, there will be a crossover in the scaling behaviour. Such situations can not be captured with a single exponent $\nu$, but they can be treated with the more general ansatz via $F(\xi)$, see Eq.(3.2) and Eq.(3.3). Using equation (3.6), we can compute expectation values of all powers of $F(|\Delta \boldsymbol{X}|)$. In particular the expectation of $F(|\Delta \boldsymbol{X}|)$ scales linear in the operational time and hence is additive. We can thus define a *fundamental expectation* as the one expectation of the displacements that fulfils:

$$\langle F(|\Delta \boldsymbol{X}(t_1;t_3)|) \rangle = \langle F(|\Delta \boldsymbol{X}(t_1;t_2)|) \rangle + \langle F(|\Delta \boldsymbol{X}(t_2;t_3)|) \rangle \tag{3.11}$$

The fundamental expectation contains all information of the stationary part of the process and thus all information on the structural disorder. Its powers reveal the time dependence of the operational time, hence they reveal all energetic disorder:

$$\langle F^m(|\boldsymbol{X}(t)|) \rangle = A_m \langle U^m(t) \rangle . \tag{3.12}$$

Knowing $F$ allows us to infer $\alpha$ from the slope of $\langle F(|\boldsymbol{X}(t)|) \rangle$. Of course finding a function is much more complicated than finding a simple exponent. Especially the inversion of Eq.(3.11) is difficult. This shows the strength of the specialized ansatz $F(\xi) = \xi^{\frac{1}{\xi}}$.

**Connection to other dimensions:** Defining scaling exponents has a long history, especially in diffusion on fractals, [186]. Coming from that area, one will notice the similarity between the fundamental exponent and the walk dimension. The walk dimension relates the displacement $\xi$ that a random walker experiences with the length of the time interval $\tau$, during which the displacement occurs. The walk dimension $d_w$ is defined by $\xi^{d_w} \sim \tau$. For a normal diffusive process, this index is of course one, for a random walker on a fractal it may take different values. The fundamental moment also measures the scaling between displacement and time, so one is tempted to replace

the cumbersome definition involving some additivity properties with this easy scaling relation. This is, however, only partly true, because the fundamental moment describes the scaling between displacement $\xi$ and *operational time* $\omega$. The latter may scale with the physical time $\tau$ according to a completely different law, say $\omega \sim \tau^{\alpha}$. In the regime where those simple power law relations hold, we have

$$\tau \sim \xi^{d_w} \sim \omega^{\frac{d_w}{\gamma_F}} \sim \tau^{\frac{d_w \alpha}{\gamma_F}},$$

and thus the equality

$$\gamma_F = \alpha d_w. \tag{3.13}$$

In the absence of energetic disorder, we will have $\omega \sim \tau$ and thus the fundamental exponent coincides with the walk dimension. Their values depart however when energetic disorder is introduced, as $\alpha$ is no longer unity. We want to stress here that the fundamental exponent is *not* some modified walk dimension, since it is not defined as a scaling exponent, but as the unique index, at which the moments of displacement become an additive function in time.

**Fundamental properties of normal diffusion:** In the absence of energetic disorder the operational time scales linear with the physical time. And without structural disorder, $\nu = \frac{1}{2}$. Therefore, we have a fundamental exponent of two, and an exponent $\alpha$ equal to one. Of course the effects of energetic and structural disorder can cancel each other, when $1 = \frac{2\alpha}{\gamma_F}$, but we will not consider this case here.

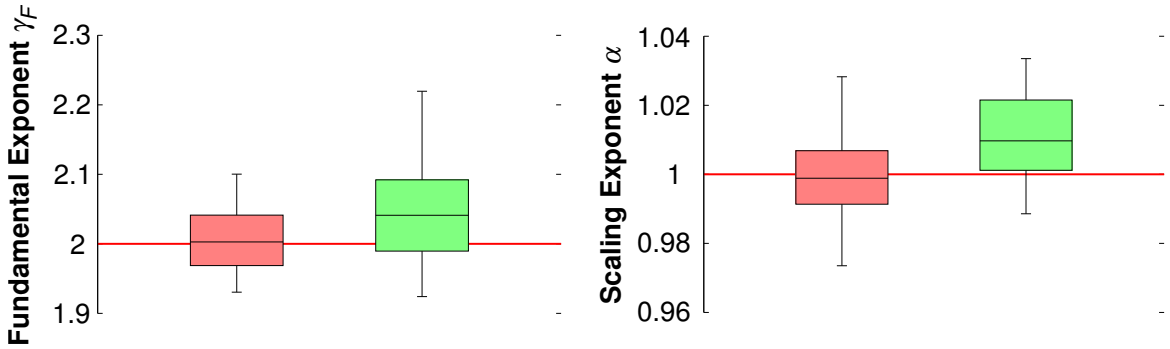
We compared the prediction $\gamma_F = 2$ with simulations of Brownian motion and the binomial random walk. For each process we generated an ensemble of 2048 trajectories in the time interval $[0, 2048]$ with unit time steps. The fundamental exponent was determined by inverting the equation

$$\Delta M(\gamma; t) := \langle |\Delta \boldsymbol{X}(0; t)|^{\gamma} \rangle - \left\langle \left|\Delta \boldsymbol{X}\left(0; \tfrac{t}{2}\right)\right|^{\gamma} \right\rangle - \left\langle \left|\Delta \boldsymbol{X}\left(\tfrac{t}{2}; t\right)\right|^{\gamma} \right\rangle = 0$$

for $\gamma = \gamma(t)$. This can be done by varying $\gamma$ until $\Delta M(\gamma; t)$ changes its sign. The zero is obtained by linear interpolation and bisection. $\gamma_F$ is the average of all $\gamma(t)$ that have been found. Often, $\Delta M(\gamma; t)$ does not change its sign in a reasonable range for $\gamma$. (This can happen with up to 70% of the time points.) When there is a change of sign, the graph of $\Delta M(\gamma; t)$ often intersects with the abscissa under a fairly high angle.

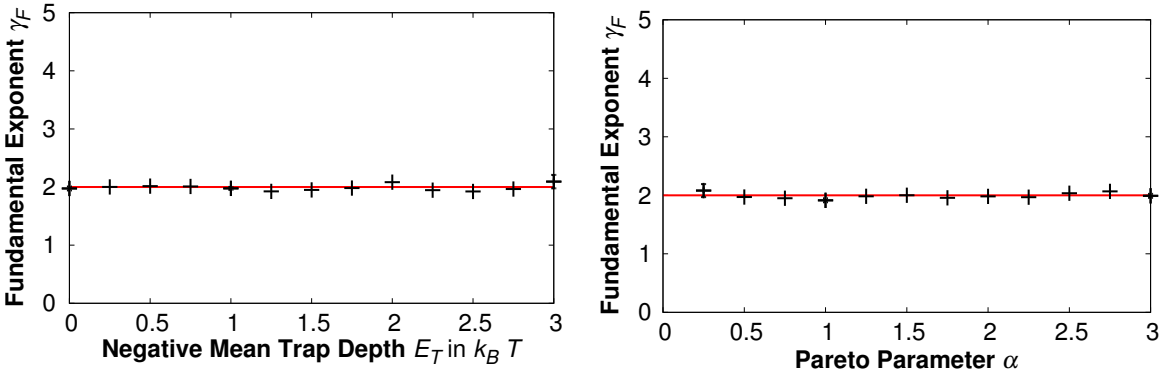
The scaling exponent $\alpha$ was determined by applying usual linear regression to the double logarithmic plot $\ln t$ versus $\ln\langle |\boldsymbol{X}(t)|^2 \rangle$.

Box plots for the fundamental exponent $\gamma_F$ are shown in Fig. 3.2a. Box plots for the exponent $\alpha$ can be found in Fig. 3.2b. The fundamental exponent is clearly two, although the fluctuations around this average value (here measured with the interquartile distance) depend on the model and increase with its complexity. The same holds for the exponent $\alpha$ which is scattered around unity, with differing but small fluctuations.




(a) The fundamental exponent (b) Scaling of the fundamental moment

Figure 3.2.: Fundamental properties of normal diffusion
The fundamental exponent is two for models of normal diffusion. Deviations from the prediction depend on the model. The same holds for the fluctuations of the exponent $\alpha$, which is unity in normal diffusion. The left box plot (red) shows the simulation results for Brownian motion, the right one (green) for the binomial random walk. For each model we evaluated 64 simulations with 2048 time steps and the same number of trajectories. The box denotes the interval that contains 50% of all data points (first and third quartile). The bar in the box is the median, larger than half of the data points. The full range of data is marked with the "whiskers" of the box. The red line is the theoretical prediction.




(a) Trap model (b) Ctrw

Figure 3.3.: Fundamental moment for the random trap model and ctrw
Every point is calculated from 2048 trajectories of length $10^6\tau$. Error bars are given for three points and are smaller than the symbols. Both models are of pure energetic disorder type, and their fundamental exponent is $\gamma_F = 2$.

## 3.4. In pure energetic disorder

We discuss models with energetic disorder, those that have been introduced in chapter 2. For those processes the fundamental moment is the msd and the anomaly is described by the exponent $\alpha$, which is the scaling between physical and operational time. The latter, $U(t)$, is the source of the anomaly, its most striking feature is the lack of stationary increments. Energetic disorder is thus closely connected with ageing, the explicit dependence on the preparation epoch. In contrast to processes of structural disorder, models with energetic disorder change over time. One of the most prominent consequences of this behaviour, the lack of ergodicity, is discussed in chapter 4.

**Uncorrelated increments:** In the absence of structural disorder the fundamental exponent is two. Thus the msd is additive in time and this immediately implies that such processes have uncorrelated increments, since

$$\begin{aligned} 0 =& \left\langle \Delta \boldsymbol{X}^2(t_1;t_3) \right\rangle - \left\langle \Delta \boldsymbol{X}^2(t_1;t_2) \right\rangle - \left\langle \Delta \boldsymbol{X}^2(t_2;t_3) \right\rangle \\ =& 2 \left\langle \Delta \boldsymbol{X}(t_1;t_2) \cdot \Delta \boldsymbol{X}(t_2;t_3) \right\rangle . \end{aligned}$$

Obviously the opposite also holds. Anomalous diffusion processes with uncorrelated increments are processes of pure energetic disorder. Hence the msd is always the fundamental moment in pure energetic disorder. To see this we simulated ctrw and the random trap model for various choices of the Pareto-parameter and mean trap depth, respectively. The result is shown in Fig. 3.3 (Fig. 3.3a for the trap model, Fig. 3.3b for ctrw). The measured fundamental exponent is always two.

**Models with multiple scaling regimes:** The msd being fundamental is true even for non-trivial scaling behaviour of the operational time. Consider two special processes

of energetic disorder. The first one being a deterministically time-changed Brownian motion:

$$\boldsymbol{X}^{(1)}(t) = \sqrt{2K_{2,1}}\boldsymbol{B}\Big(t(1+\tfrac{t}{\tau_0})\Big)\,,$$

and the second one, $\boldsymbol{X}^{(2)}(t) := \sqrt{2K_{2,1}}\boldsymbol{B}(\tilde{U}(t))$ being a ctrw with a cut-off waiting time distribution

$$\psi(t) \sim \frac{e^{-\frac{t}{\tau_2}}}{\Big(1+\frac{t}{\tau_1}\Big)^{1+\alpha}},$$

with an appropriate normalization. $\tilde{U}$ is the renewal process defined by $\psi(t)$. When $\alpha \in [0,1)$, and $0 \ll \tau_1 \ll \tau_2$, the expectation of $\tilde{U}$ shows a crossover from sub-linear to a linear time-dependence. The fundamental moment, which is the second one, will exhibit this crossover behaviour as well. The same holds for $\boldsymbol{X}^{(1)}(t)$, having a crossover from a linear to ballistic regime, when $t \gg \tau_0$. Still, as can be demonstrated in $\boldsymbol{X}^{(2)}(t)$, the msd is the fundamental one, because

$$\begin{aligned}\Big\langle |\Delta\boldsymbol{X}^{(2)}(t;t'')\,|^2\Big\rangle =& 2dK_{2,1}\Big\langle \Delta\tilde{U}(t;t'')\Big\rangle \\ =& 2dK_{2,1}\Big\langle \Delta\tilde{U}(t;t') + \Delta\tilde{U}(t';t'')\Big\rangle \\ =& \Big\langle |\Delta\boldsymbol{X}^{(2)}(t;t')\,|^2\Big\rangle + \Big\langle |\Delta\boldsymbol{X}^{(2)}(t';t'')\,|^2\Big\rangle\,.\end{aligned}$$

Let us now turn to models of structural disorder.

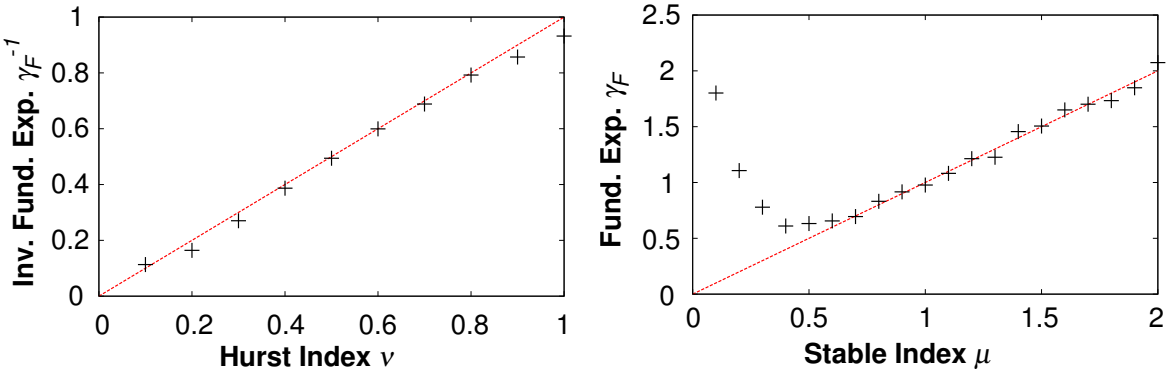
## 3.5. In pure structural disorder

The capital difference between models of energetic and structural disorder is the stationarity of the increments, which is invariance under temporal translation. Consequently $\langle|\Delta\boldsymbol{X}(0;\tau)\,|^\gamma\rangle$ and $\langle|\Delta\boldsymbol{X}(\tau;2\tau)\,|^\gamma\rangle$ will bear the same value for a process of solely structural disorder. Hence the fundamental moment is the one, which is linear in the time lag $\tau$. Let us discuss some of the models from the last chapter.

**Fractional Brownian motion:** The Hurst exponent $\nu$ of the fractional Brownian motion defines its scaling by $\langle|\boldsymbol{Y}(t)\,|^\gamma\rangle \sim t^{\gamma\nu}$. The fundamental exponent is thus given by

$$\gamma_F = \tfrac{1}{\nu}. \tag{3.14}$$

This is nicely shown in Fig. 3.4a, where the measured fundamental exponent is plotted against the Hurst exponent.




(a) FBm (b) Stable motion

Figure 3.4.: Fundamental moment for fBm and stable motions
Crosses are simulation results, the dotted line is the theoretical prediction. Every point is calculated from 1000 trajectories with 1000 time steps. (3.4a): result for fBm. (3.4b): result for stable motion. The predictions from Eq.(3.14) and Eq.(3.15) are given by red lines. Simulations follow the prediction, when the parameters do not assume extreme values.

**Stable motion:** The stable motion is an example for a process where the fundamental moment diverges. Since the $\gamma$-th moment scales as $t^{\gamma/\mu}$, the fundamental exponent is equal to the Lévy index $\mu$, i.e.

$$\gamma_F = \mu. \tag{3.15}$$

Again we compared the prediction with simulations. Here, however, we had to resort to the alternate definition via the quantile, Eq.(3.10). Inverting the equation, we produced the results of Fig. 3.4b. It works well as long as $\mu$ does not adopt extremely low values. For computing the quantile it is necessary to sort the list of displacements, which is a time-consuming process when compared to computing the moments. Therefore one should resort to determine the fundamental scaling with help of the moments, when one is sure that they do not diverge.

**Barrier model:** At last, the random barrier model is also a process of structural disorder. The scaling exponent $\nu$ is given by Eq.(2.47) and its reciprocal gives the fundamental exponent, which can be expressed in terms of the mean barrier height as

$$\gamma_F = 1 + \max(1, \beta E_B)\,. \tag{3.16}$$

This prediction also matches the simulations nicely as depicted in Fig. 3.5a. Since the results of the barrier model will be compared to those of the random potential model, Fig. 3.5a will be found in the next section.

## 3.6. In mixed disorder

The power of the fundamental properties lies in the fact that they are able to separate the energetic and structural disorder. We will demonstrate this in the random potential model, as was already done in [187] by the author. The msd of this model behaves possibly sub-diffusively as

$$\left\langle \boldsymbol{X}^2(t) \right\rangle \sim t^{2\nu\alpha},$$

where $\nu$ is calculated from the mean barrier height by Eq.(2.47) and $\alpha$ is given by Eq.(2.43).

The structural disorder parameter given by $\nu$ gives rise to the fundamental exponent as

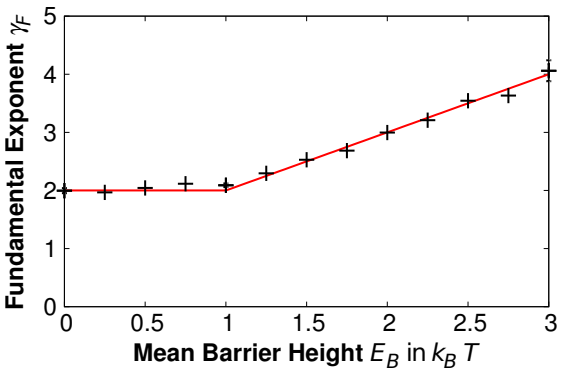
$$\gamma_F = \frac{1}{\nu} = 1 + \max(1, \beta E_B) \tag{3.17}$$

which is the same value as in the barrier model. We simulated the random potential model for different values of the mean barrier height $E_B$ and mean trap depth $E_T$ and compared the data to this prediction. The result is plotted in Fig. 3.5b. Each of the dotted lines corresponds to simulation runs with a fixed value of the mean trap depth. Apparently the latter has no influence on the value of the fundamental exponent, as all dotted lines scatter around the prediction of the barrier model. They do increase the fluctuations though. This demonstrates how the fundamental exponent is able to identify the structural disorder.

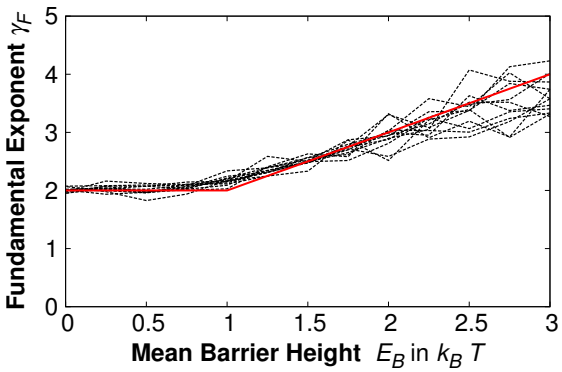
Now, we can compare the fundamental moment of the mixed model to the fundamental moment of the trap model. The time dependence of the fundamental moment gives the energetic disorder, as it scales with $t^\alpha$. This is seen in Fig. 3.5c, where we plotted the fundamental moments of the random barrier, trap and potential model against time in a double logarithmic manner. All models have a mean barrier height of $3.0\beta^{-1}$ and mean trap depth of $3.0\beta^{-1}$. Clearly the fundamental moment of the random trap and the random potential model have the same slope, since they share the same degree of energetic disorder. Here we can identify $\alpha$ as 0.5. In contrast, the fundamental moment of the random barrier model has a unit slope in the log-log-plot, showing its linear time dependence and the lack of energetic disorder.

## 3.7. Summary

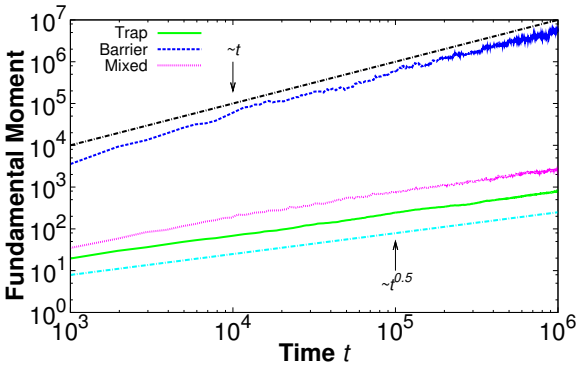
In this chapter, we proposed a method to separate the energetic and structural disorder parts of a process. We started with the subordination ansatz $\boldsymbol{X}(t) = \boldsymbol{Y}(U(t))$, and the additional but not crucial assumption that the moments of the operational time $U(t)$ as well as the moments of the increments $\Delta\boldsymbol{Y}$ behave as power laws. The scaling exponent of the operational time $\alpha$, describes the energetic disorder, the exponent $\nu$ of the stationary part describes the structural disorder.




(a) Barrier model




(b) Random potential model




(c) Time dependence

Figure 3.5.: Fundamental exponent of the random barrier and potential model (3.5a): Fundamental exponent for the random barrier model. (3.5b): Fundamental exponent for the random potential model. Every point was calculated from 2048 trajectories with $10^6$ time steps. Every dotted line corresponds to one fixed value of the mean trap depth ranging from zero to $3.0\beta^{-1}$. The latter apparently does not influence the fundamental exponent, but just the fluctuations of the result. Both predictions are the same, given by Eq.(3.16) and (3.17), respectively, since both models share the same structural disorder properties.
(3.5c): Here we show the fundamental moment of a random trap, barrier and potential model plotted versus time in a double logarithmic fashion. The parameters are $E_B = 3.0\beta^{-1}$ and $E_T = 3.0\beta^{-1}$. The fundamental moment is only influenced by the energetic disorder, and consequently the trap model and the random potential model have the same slope. The dotted lines are power laws inserted for convenience.

It was demonstrated that there is a moment of $\boldsymbol{X}$'s displacement, which is additive in time, i.e. which fulfils:

$$\langle |\Delta \boldsymbol{X}(t_1;t_3)|^{\gamma_F} \rangle = \langle |\Delta \boldsymbol{X}(t_1;t_2)|^{\gamma_F} \rangle + \langle |\Delta \boldsymbol{X}(t_2;t_3)|^{\gamma_F} \rangle \,,$$

for any three epochs $t_1 \leq t_2 \leq t_3$. This moment was called the fundamental moment and its corresponding exponent $\gamma_F$ was called the fundamental exponent. In the case, when the structural disorder part obeys a scaling law with exponent $\nu$, the fundamental exponent is given by:

$$\gamma_F = \frac{1}{\nu}.$$

The fundamental moment on the other hand is proportional to the moments of the operational time and thus its time dependence unveils the value of $\alpha$. In this way the knowledge of the fundamental properties enables one to know $\alpha$ and $\nu$ and hence to separate structural from energetic disorder.

The method can be applied to the quantiles of $\Delta \boldsymbol{X}$ when the fundamental moment itself is infinite. In that case, the equation

$$[|\Delta \boldsymbol{X}(t_1;t_3)|^{\gamma_F}]_p = [|\Delta \boldsymbol{X}(t_1;t_2)|^{\gamma_F}]_p + [|\Delta \boldsymbol{X}(t_2;t_3)|^{\gamma_F}]_p$$

defines the fundamental exponent and the time dependence of the quantiles reveals the operational time. The method also needs to be modified when the increments $\Delta \boldsymbol{Y}$ don't show a simple and unique power law scaling. Then a non-trivial scaling function $F$ needs to be found, such that the fundamental expectation

$$\langle F(|\Delta \boldsymbol{X}(t_1;t_3)|) \rangle = \langle F(|\Delta \boldsymbol{X}(t_1;t_2)|) \rangle + \langle F(|\Delta \boldsymbol{X}(t_2;t_3)|) \rangle$$

is additive. $F$ describes the non-trivial scaling between operational time and the stationary increments of the structural disorder part.

Since the fundamental exponent is defined with the additivity of a moment and not by a scaling relation alone it needs to be distinguished from the random walk dimension. The fundamental exponent is the scaling exponent between displacement of the process and the operational time, not the physical time. Both dimensions are related via

$$\gamma_F = \alpha d_w.$$

The absence of structural disorder is indicated by $\nu = 1/2$ , i.e. $\gamma_F = 2$. Absence of energetic disorder on the other hand implies a linear growth of the fundamental moment or expectation, hence $\alpha = 1$. Consequently, models with pure energetic disorder are those with uncorrelated increments, regardless of whether the energetic disorder is trivial (i.e. there exists a unique $\alpha$) or not. Pure structural disorder processes, on the other hand, are those with stationary increments and a fundamental exponent different from two.

We demonstrated the theoretical considerations numerically in ctrw and the random trap model for pure energetic disorder, in fBm, stable motion and the random barrier model for pure structural disorder, and in the random potential model for a model of mixed disorder.

We thus illustrated that the fundamental properties are indeed able to serve their purpose: the separation of structural and energetic disorder. As long as there are unique scaling laws, it is a very powerful method. It becomes unhandy, when the process in question does not possess all moments and one needs to resort to quantiles. Structural disorder with different scaling regimes also poses a problem as the fundamental exponent is not uniquely defined. Furthermore, the method relies on ensemble properties and is thus only applicable, when a sufficiently large ensemble of trajectories is accessible. When this is the case, however, it is, as we showed, perfectly fit to reveal the structural and energetic components of the process, and hence to shed light on the question whether the diffusion happens in thermodynamic equilibrium or not.

# 4. Statistical Aspects of Ergodicity

## 4.1. Introduction

As has been done before, we will introduce the problem and the notation in terms of normal diffusion. We abbreviate the increment process with $\Delta \boldsymbol{X}_\tau(t) := \Delta \boldsymbol{X}(t; t+\tau) := \boldsymbol{X}(t+\tau) - \boldsymbol{X}(t)$. In this section, $\boldsymbol{Y}(t) := \sqrt{2K_{2,1}}\boldsymbol{B}(t)$ denotes a standard Brownian motion. Definitions are given in terms of general stochastic processes $\boldsymbol{X}(t)$.

Ergodic theory is concerned with time averages of certain observables, that means, with functions of the system's state. Here we are only interested in observables of the process' increments, i.e. some function $O(\Delta \boldsymbol{X}_\tau(t))$, usually a power of its absolute value. Given a measurement in the interval $[0, T]$, the time average of $O$ is given by:

$$\overline{O(\Delta \boldsymbol{X}_\tau)}_T := \frac{1}{T-\tau} \int_0^{T-\tau} \mathrm{d}t\, O(\Delta \boldsymbol{X}_\tau(t)) \tag{4.1}$$

For fixed lag time $\tau$, the average is performed over the time $t$, so that it needs to be dropped as argument in the final expression. $\overline{O(\Delta \boldsymbol{X}_\tau)}_T$ is a functional of a stochastic process and consequently a random variable. However, under certain conditions the time average assumes a deterministic limit, when the measurement interval becomes infinite, i.e. when $T \to \infty$. In this limit, we drop the subscript $T$ and write:

$$\overline{O(\Delta \boldsymbol{X}_\tau)} := \lim_{T\to\infty} \frac{1}{T-\tau} \int_0^{T-\tau} \mathrm{d}t\, O(\Delta \boldsymbol{X}_\tau(t))\,.$$

The famous ergodic hypothesis states that the time average produces the value at thermodynamic equilibrium:

$$\overline{O(\Delta \boldsymbol{X}_\tau)} = \langle O(\Delta \boldsymbol{X}_\tau) \rangle_{\mathrm{eq}}\,. \tag{4.2}$$

The average on the right hand side is performed with respect to the distribution of $\Delta \boldsymbol{X}_\tau$ at thermodynamic equilibrium. Since the time average is a random variable, the limit has to be understood in a certain probabilistic sense, which will be made clear later.

In mathematics, ergodic theory started with the works of Birkhoff [114] and von Neumann [115] in 1931. Birkhoff proved for autonomous dynamical systems that the average time the system spent in some region of its state space is a non-vanishing constant for almost any trajectory. This constant turns out to be the stationary

measure of that region. This is essentially a limit theorem on the local times of the system. Von Neumann proved that the time averaging operator (here defined by Eq.(4.1)) of an autonomous Hamiltonian system converges weakly to a limit. This ensures that for almost any initial condition the time averages take a certain value, which can be identified as its equilibrium value.

Both formulations provide results that do not depend on the exact realization of the process, i.e. that hold for almost all trajectories or initial conditions (which fixes the probabilistic sense). The proofs use the formal time evolution operators of the dynamical system and require the "metrical transitivity" of state space. This means that the only invariant sets under the time evolution operation are the empty set and the whole state space. Hence, the system will explore its full state over the course of time and is not confined to some sub-domain. When this is not the case, the time averages will still assume a deterministic value, but this value depends on the region in state space that the system was initialized in.

Another very important but sometimes forgotten prerequisite is the stationarity of the system. In fact, the original ergodic theorems were given for time-independent systems. The notation of an equilibrium average makes little sense when $\langle O(\Delta \boldsymbol{X}_\tau(t))\rangle_{\text{eq}}$ does not assume a value independent of $t$, hence is not equilibrated.

Today, ergodic theorems are provided for a variety of stationary stochastic processes – in particular for Markovian processes – and they are a basic part of the theory of stochastic processes as well as of measure theory, [144, 145, 188]. Apart from metrical transitivity there are other indicators for ergodicity. For instance a stationary Gaussian process is ergodic, when its power spectrum (the Fourier transform of its autocorrelation function) does not contain any delta functions, see f.e. a note in [182, pp.581]. Ergodic properties of Markovian processes can be attributed to spectral properties of their generators. A spectral decomposition of the generator reveals that the time evolution operator and the time averaging operator both converge to projectors that map any initial distribution to the stationary state. Although such a spectral analysis would be in the spirit of chapter 2, it needs to be performed for the generator of the *increment process* (since we are interested in the ergodic properties of the increments). The latter is either trivial, i.e. zero, or too complicated to be treated in a general set-up. Hence, we refrain from such an investigation. Rather we will focus on the expected value of the time averages and on their fluctuations in sections 4.1.1 and 4.1.2 respectively. This is also done in recent physical literature on the subject. Another characterisation can be done in terms of correlation functions and is due to Khinchin and Birkhoff, [116]. He gave a very clear condition for ergodicity, which is related to the time averages' fluctuations and will be presented in 4.1.3. In this section, these concepts will be introduced for Brownian motion. In the remainder of the chapter, they will be discussed in anomalous diffusion processes.

### 4.1.1. Agreement of Averages

The main observable in diffusion is the displacement $\Delta \boldsymbol{X}_\tau(t)$ and its powers. For Brownian motion, the eamsd is given by

$$\left\langle |\Delta \boldsymbol{Y}_\tau(t)\,|^2 \right\rangle = 2dK_{2,1}\tau.$$

In particular, since Brownian motion has stationary increments, this value is independent of $t$ and it is only determined by the lag time. We can therefore identify the eamsd with the equilibrium average form Eq.(4.2). Let us compute the expected tamsd, this is a double – time and ensemble – average:

$$\left\langle \overline{|\Delta \boldsymbol{Y}_\tau|^2}_T \right\rangle = \frac{\langle |\Delta \boldsymbol{Y}_\tau|^2 \rangle}{T-\tau} \int_0^{T-\tau} \mathrm{d}t \; = 2dK_{2,1}\tau. \tag{4.3}$$

It bears the same value. Therefore, Brownian motion is ergodic in mean.

In the following chapter, we will present processes that violate this relation. We will also discuss when this is the case. Regardless of the coincidence of both values, one can ask about the magnitude of fluctuations around that expectation value.

### 4.1.2. Ensemble heterogeneity

Apart from just the expectation, one can try to quantify the fluctuations of time averages. This can be done by investigating the normalized variance of a time averaged observable:

$$J_O(T) := \frac{\mathrm{Var}\left[\overline{O(\Delta \boldsymbol{X}_\tau)}_T\right]}{\left\langle \overline{O(\Delta \boldsymbol{X}_\tau)}_T \right\rangle^2}. \tag{4.4}$$

This quantity measures how heterogeneous an ensemble of trajectories is with respect to the functional $O$; we call it the ensemble heterogeneity parameter. When it decays to zero for $T \to \infty$, all trajectories of infinite length appear similar with respect to $O$. For processes with stationary increments the denominator is a constant with respect to $T$, and the decay of $J_O$ is only possible when the variance of the time average in the enumerator vanishes. In such a case, a diffusivity defined in terms of the tamsd will deviate from the ensemble averaged value with relative fluctuations given by the ensemble heterogeneity. For anomalous diffusion processes with weak ergodicity breaking such a diffusivity may remain a random variable and $J_O$ will approach a finite positive value in the limit $T \to \infty$. This is the reason why $J_O$ is often called "weak ergodicity breaking parameter" in the literature. However, as we proceed to show, a non-vanishing $J_O$ is not necessary for weak ergodicity breaking. Therefore, we proposed the alternative name.

For the msd of Brownian motion, one can compute the ensemble heterogeneity exactly. We will first present the result and calculate the expression later with the use

of correlation functions. Since there are currently no other time scales than the lag time $\tau$ and the measurement time $T$, the heterogeneity parameter must be a function of a dimensionless combination of both, for instance $\tau/(T-\tau)$. The result is given by

$$J^{(\mathrm{Bm})}_{\Delta \boldsymbol{Y}^2_\tau}(T) = \begin{cases} 2 - \frac{4}{3}\left(\frac{\tau}{T-\tau}\right)^{-1} + \frac{1}{3}\left(\frac{\tau}{T-\tau}\right)^{-2}, & \tau > T-\tau \\ \frac{4}{3}\frac{\tau}{T-\tau} - \frac{1}{3}\left(\frac{\tau}{T-\tau}\right)^{2}, & \tau \leq T-\tau \end{cases}, \tag{4.5}$$

and vanishes as $T \to \infty$.

This means two things. First, since Brownian motion has stationary increments, the variance of the tamsd itself vanishes. This means that the absolute fluctuations decay, not just the relative ones. And secondly, a diffusivity defined by $K := \overline{\Delta \boldsymbol{Y}_\tau}/(2d\tau)$ is a well-defined deterministic quantity.

In the rest of the chapter, we will encounter processes, which possess a heterogeneous ensemble. In such cases the time averages bear fluctuations that are always of the same order of magnitude as their expectations. A time averaged diffusivity will then not be well-defined.

We see that not only $\langle|\overline{\Delta \boldsymbol{Y}(\tau)}_T - \langle \Delta \boldsymbol{Y}(\tau)\rangle|\rangle$ decays to zero as $T \to \infty$, but also $\langle|\overline{\Delta \boldsymbol{Y}(\tau)}_T - \langle \Delta \boldsymbol{Y}(\tau)\rangle|^2\rangle$. Thus the ergodic relationship does not only hold on average, but also in variance. Such a limit up to the $p$-th moment is called "limit in $L^p$". It fixes the probabilistic sense, in which we have to understand equation (4.2). Such a concept of ergodicity holding in $L^p$ is sometimes called "weak ergodicity" in the mathematical literature. When the ergodicity equation (4.2) holds in $L^2$, one often speaks about "ergodicity in mean square". We will do the same here.

The variance of the time averages can also be expressed in terms of correlation functions. This leads to another condition for ergodicity in mean square, Birkhoff-Khinchin theorem.

### 4.1.3. Irreversibility and Birkhoff-Khinchin Theorem

The correlation function of an observable $O$ is given by

$$C_O(t,t') := \mathrm{Cov}[O(\Delta \boldsymbol{X}_\tau(t)), O(\Delta \boldsymbol{X}_\tau(t'))]\,. \tag{4.6}$$

The correlation function is obviously symmetric with respect to interchange of arguments, and for a process with stationary increments it is a function solely of $|t-t'|$. Using this expression, we can express the variance of $O$'s time average as

$$\mathrm{Var}\left[\overline{O(\Delta \boldsymbol{X}_\tau)}_T\right] := \frac{1}{(T-\tau)^2}\int_0^{T-\tau} \mathrm{d}t \int_0^{T-\tau} \mathrm{d}t'\, C_O(t,t')\,.$$

The symmetry of the integrand can be used to half the integration domain from a square to a triangle where $t < t'$. This gives a factor two and the remaining integration

domain is equivalently parametrized by $t_m = \min(t, t')$ and $\tau_d := |t - t'|$. We obtain

$$\mathrm{Var}\left[\overline{O(\Delta \boldsymbol{X}_\tau)}_T\right] := \frac{2}{(T-\tau)^2} \int_0^{T-\tau} \mathrm{d}\tau_d \int_0^{T-\tau-\tau_d} \mathrm{d}t_m \, C_O(t_m, t_m + \tau_d) \,. \tag{4.7}$$

From here it is easy to see that, when the correlation function decays fast enough to zero, the integral will vanish. Consequently the variance of $O$'s time average vanishes as well. To see this, we consider a process with stationary increments. We use the symmetry of the integrand and apply the variable transformations $x := t_m / (T - \tau)$ and $y := \tau_d / (T - \tau)$ to bring the variance to following form

$$\mathrm{Var}\left[\overline{O(\Delta \boldsymbol{X}_\tau)}_T\right] = 2 \int_0^1 \mathrm{d}y \int_0^{1-y} \mathrm{d}x \, C_O((T-\tau)\, y) \,.$$

The inner integration is trivially performed. When attempting to take the limit $T \to \infty$, we pull it into the integral. As long as the correlation function is bounded, the dominated convergence lemma applies and the limit reads:

$$\lim_{T\to\infty} \mathrm{Var}\left[\overline{O(\Delta \boldsymbol{X}_\tau)}_T\right] = 2 \int_0^1 \mathrm{d}y \; (1-y) \lim_{T\to\infty} C_O((T-\tau)\, y) \tag{4.8}$$

and vanishes, if the limit function in the integrand is zero.

This condition, namely:

$$\lim_{\tau_d \to \infty} C_O(\tau_d) = 0 \tag{4.9}$$

is called *irreducibility* of the observable. The reasoning here is known as Birkhoff-Khinchin theorem, [116]. Summarizing, this theorem states that irreducibility is sufficient for ergodicity in mean square. This theorem has been discussed for anomalous diffusion e.g. in [128, 189].

For Brownian motion the correlation function in question is the one of $\Delta \boldsymbol{Y}_\tau^2$. Since the increment process is Gaussian, we can apply Wick's theorem. It states that for any Gaussian vector $\boldsymbol{Z} := (Z_1, Z_2, Z_3, Z_4, \ldots)$ with zero mean all higher order correlations can be expressed in terms of the two point correlations.

$$\langle Z_a Z_b Z_c Z_d \rangle = \langle Z_a Z_b \rangle \langle Z_c Z_d \rangle + \langle Z_a Z_c \rangle \langle Z_b Z_d \rangle + \langle Z_a Z_d \rangle \langle Z_b Z_c \rangle \,, \tag{4.10}$$

where $a, b, c, d$ can be any numbers from $\{1, 2, 3, 4, \ldots\}$. When this relation is applied to some Gaussian process $t \mapsto \Delta \boldsymbol{X}_\tau(t)$, we obtain:

$$C_{\Delta \boldsymbol{X}_\tau^2}(t, t') = 2 \left(\mathrm{Cov}[\Delta \boldsymbol{X}_\tau(t) \,, \Delta \boldsymbol{X}_\tau(t')]\right)^2 = 2 C^2_{\Delta \boldsymbol{X}_\tau}(t, t') \,, \tag{4.11}$$

which is a relation between the correlation function of the increments and the one of the squared increments.

In the case of Brownian motion, this function is given by:

$$C_{\Delta \boldsymbol{Y}_\tau}(t,t') := 2dK_{2,1}\left[\tau - |t-t'|\right] \mathbb{1}_{[0,\tau)}(|t-t'|)\,. \tag{4.12}$$

The indicator function ensures that it vanishes whenever $|t-t'| > \tau$. Hence the msd of Brownian motion is irreducible.

Going the way back, by plugging Eq.(4.12) into Eq.(4.11), plugging the result into Eq.(4.7) and performing the integral gives:

$$\begin{aligned}
\mathrm{Var}\left[\overline{|\Delta \boldsymbol{Y}_\tau|^2}_T\right] =& 4\left(\frac{2dK_{2,1}}{T-\tau}\right)^2 \int_0^{T-\tau} \mathrm{d}\tau_d \int_0^{T-\tau-\tau_d} \mathrm{d}t_m \, \left[\tau-\tau_d\right]^2 \mathbb{1}_{[0,\tau)}(\tau_d) \\
=& \left(\frac{2dK_{2,1}}{T-\tau}\right)^2 \Big\{ 4\left(\tau \wedge T - \tau\right)\tau^2\left(T-\tau\right) - \left(\tau \wedge T - \tau\right)^4 \\
&+ \left(\tau \wedge T - \tau\right)^2\left(2\tau^2 - T\tau\right) + \frac{4}{3}\left(\tau \wedge T - \tau\right)^3\left(T+\tau\right) \Big\}\,.
\end{aligned}$$

Here we abbreviated $t \wedge t' := \min(t,t')$. Dividing this expression by (4.3) gives Eq.(4.5)

### 4.1.4. Weak ergodicity breaking

The topic of "weakly breaking ergodicity" was already raised. The phenomenon is understood as a non-coincidence of time and ensemble averages (hence a breaking of weak ergodicity in the $L^1$-sense), in situations with a connected state space. By just inspecting metrical transitivity, one would expect an ergodic behaviour and is surprised when it is not present. This phenomenon appears in non-stationary processes, when the requirements of classical ergodic theory are not met. It is usually quantified by the disparity of time and ensemble averages, i.e. a violation of (4.2), and in terms of the ensemble heterogeneity. The latter is also called the weak ergodicity breaking parameter and a non-vanishing limit value is interpreted as a criterion for weak ergodicity breaking. We will show in the rest of the chapter, however, that both features are not equivalent, hence we proposed a different name.

Such an analysis has already been performed for several models [102, 118, 122, 125, 130, 131, 190–192], in particular ctrw [117, 120, 121] and fBm [126, 193]. We will try to lead the same discussion for anomalous diffusion models of mixed origins, although this is not always possible in its whole generality. Other quantities have also been proposed for ergodicity breaking parameters, for instance limit values of response functions [123], generalizations of correlation functions [194] or some special values of characteristic functions of the process (the dynamical functional) [129]. Weak ergodicity breaking can also be observed experimentally, for instance in fluorescent nano-crystals [132] and in diffusion in biological cells [67, 70, 72].

In the rest of the chapter, we will discuss the presented concepts first in structural, then in energetic and finally in mixed disorder. We will close with a summary. In the rest of the chapter $\boldsymbol{Y}$ will again denote a process of pure structural disorder.

## 4.2. In pure structural disorder

### 4.2.1. General considerations

The main assumption of pure structural disorder processes is the stationarity of their increments, i.e. $\Delta \boldsymbol{Y}_\tau(t) = \Delta \boldsymbol{Y}_\tau$. This ensures that the ensemble averages of any power of the displacement are stationary and can be interpreted as equilibrium averages. Let us investigate the consequences for time averages.

**Agreement of the averages:** The stationarity of the displacements leads to an ergodic behaviour in mean value.

$$\left\langle \overline{O(\Delta \boldsymbol{Y}_\tau)}_T \right\rangle = \frac{\langle O(\Delta \boldsymbol{Y}_\tau)\rangle}{T-\tau} \int_0^{T-\tau} \mathrm{d}t \; = \langle O(\Delta \boldsymbol{Y}_\tau)\rangle \,. \tag{4.13}$$

This holds in particular for any power of the displacement $|\Delta \boldsymbol{Y}_\tau|^\gamma$.

**Asymptotic properties of the ensemble heterogeneity:** Since the expectations of the time averages do not depend on the time $t$, the ensemble heterogeneity parameter and the variance of the corresponding time average are proportional. Another consequence of stationarity is that the correlation function of any observable does only depend on the time difference $\tau_d := |t - t'|$. Thus, when computing the variance of some time average using Eq.(4.7), the integral can directly be simplified by using symmetry and a variable transformation to $t_m := \min(t, t')$ and $\tau_d := |t - t'|$.

Let us evaluate the expression for the ensemble heterogeneity asymptotically. To do so, we assume that the correlation function $C_O(\tau_d)$ is continuous and bounded, especially for small arguments. This implies that the variance of $O$ is finite. The second assumption is that the correlation function behaves asymptotically like

$$C_O(\tau_d) = C_O^{(\infty)} + \mathcal{O}\big(\tau_d^{-\epsilon}\big) \,, \tag{4.14}$$

for large arguments $\tau_d$. $\epsilon$ is supposed to be some real number. The ensemble heterogeneity is given by

$$J_O(T) = \frac{2}{(T-\tau)^2 \langle O(\Delta \boldsymbol{Y}_\tau)\rangle^2} \int_0^{T-\tau} \mathrm{d}\tau_d \; (T - \tau - \tau_d)\, C_O(\tau_d) \,. \tag{4.15}$$

The remaining integral in Eq.(4.15) is split up at some arbitrary but small time $\tau_c \ll T$. In the first part, which is integrated from 0 to $\tau_c$, we may use the mean value theorem to find a constant $C_O^*$ such, that

$$\int_0^{\tau_c} \mathrm{d}\tau_d \, \frac{T-\tau-\tau_d}{(T-\tau)^2} C_O(\tau_d) = \frac{C_O^*}{(T-\tau)^2} \int_0^{\tau_c} \mathrm{d}\tau_d \; (T - \tau - \tau_d) = \mathcal{O}\big(T^{-1}\big) \,.$$

The second part is treated with the asymptotic expansion. The leading term of this integral is half of the limit value $C_O^{(\infty)}$. Depending on the value of $\epsilon$, the next term is either of order $T^{-\epsilon}$, of order $T^{-1}$ or it has a logarithmic correction. We may summarize this short discussion:

$$J_O(T) = \frac{C_O^{(\infty)}}{\langle O(\Delta \boldsymbol{Y}_\tau)\rangle^2} + \begin{cases} \mathcal{O}\big(T^{-\epsilon}\big)\,, & \text{for } \epsilon < 1 \\ \mathcal{O}\big(T^{-1}\ln T\big)\,, & \text{for } \epsilon = 1\,. \\ \mathcal{O}\big(T^{-1}\big)\,, & \text{for } \epsilon > 1 \end{cases} \tag{4.16}$$

And it is apparent that the ensemble is homogeneous if and only if the power correlator decays to zero. This is irreversibility, this is the asymptotic decorrelation of the increments and we computed the first correction to Birkhoff-Khinchin theorem.

**Processes with independent increments:** A simple application of the latter formula is processes with independent increments. Here the correlation functions have to be zero whenever $\tau_d > \tau$ and thus the ensemble heterogeneity parameter decays like $T^{-1}$, but only when the correlation function does not diverge at $\tau_d = 0$! The latter is the case, when the variance of $O$ diverges.

**Gaussian processes:** For a Gaussian process, the correlation function of the msd can be replaced by the correlation function of the increments. As discussed in the introduction, this is due to Wick's theorem applied to the increment process. However, another expression in terms of the eamsd can be found. Some algebraic manipulations and exploitation of stationarity bears:

$$C_{\Delta \boldsymbol{Y}_\tau}(\tau_d) = \frac{1}{2}\left\{\left\langle \Delta \boldsymbol{Y}^2_{\tau+\tau_d}\right\rangle + \left\langle \Delta \boldsymbol{Y}^2_{|\tau-\tau_d|}\right\rangle - 2\left\langle \Delta \boldsymbol{Y}^2_{\tau_d}\right\rangle\right\}. \tag{4.17}$$

When we assume – as usual – that $\langle \Delta \boldsymbol{Y}^2_\tau\rangle \propto 2dK_{2,2\nu}\tau^{2\nu}$ for large $\tau$, and expand Eq.(4.17) for large $\tau_d$, we can identify the exponent $\epsilon$ of decay of the correlator as $\epsilon = 4 - 4\nu$. We obtain the ensemble heterogeneity of a Gaussian process of structural disorder:

$$J_{\Delta \boldsymbol{Y}^2_\tau}(T) = \begin{cases} \mathcal{O}\big(T^{-1}\big)\,, & \text{for } \nu < \dfrac{3}{4} \\ \mathcal{O}\big(T^{-1}\ln T\big)\,, & \text{for } \nu = \dfrac{3}{4}\,. \\ \mathcal{O}\big(T^{4\nu-4}\big)\,, & \text{for } \nu > \dfrac{3}{4} \end{cases} \tag{4.18}$$

Hence, the fluctuations die out when the diffusion is asymptotically sub-ballistic, i.e. $\nu < 1$. But we also have two interesting special cases. First, there is the super-ballistic behaviour when $\nu > 1$. Then, the ensemble heterogeneity parameter will grow to infinity and beyond. The time averages fluctuate strongly when compared to their expectations. However when the process is ballistic in the asymptotic limit, i.e.

$\nu = 1$, then the correlation functions approach a constant as $\tau_d \to \infty$ and so does $J$. The ensemble heterogeneity approaches the value two:

$$J_{\Delta \boldsymbol{Y}_\tau^2}(T) = \frac{4\left(2dK_{2,2}\tau^2\right)^2}{(T-\tau)^2\left\langle \Delta \boldsymbol{Y}_\tau^2 \right\rangle} \int_{\tau_c}^{T-\tau} \mathrm{d}\tau_d \, (T-\tau-\tau_d) + \mathcal{O}\big(T^{-1}\big) = 2 + \mathcal{O}\big(T^{-1}\big) \,. \quad (4.19)$$

This means that for an asymptotically ballistic Gaussian process, the fluctuations of the mean squared displacement never die out. The diffusivity or equivalently the velocity measured via time averages are thus ill-defined, because they are inherently random.

Let us apply the general considerations to some of the introduced processes.

## 4.2.2. Examples

**Fractional Brownian motion:** The first example to present is fractional Brownian motion. It is a stationary process with mean squared displacement given by

$$\left\langle |\Delta \boldsymbol{Y}_\tau|^2(\tau) \right\rangle = 2dK_{2,2\nu}\tau^{2\nu}.$$

The Hurst parameter $\nu$ is taken from $(0,1)$. For this Gaussian process, Eq.(4.18) applies immediately. In [126], the pre-factors have been computed. They are given by

$$J^{(\mathrm{fBm})}_{\Delta \boldsymbol{Y}_\tau^2}(T) = \begin{cases} \dfrac{\tau}{T}\displaystyle\int_0^\infty \mathrm{d}z \, \left[(z+1)^{2\nu} + |z-1|^{2\nu} - 2z^{2\nu}\right]^2, & \text{for } \nu \in (0, \dfrac{3}{4}) \\ \dfrac{9}{16}\dfrac{\tau}{T}\ln\dfrac{\tau}{T}, & \text{for } \nu = \dfrac{3}{4} \\ \dfrac{4\nu^2(2\nu-1)}{(4\nu-3)(4\nu-2)}\left(\dfrac{\tau}{T}\right)^{4-4\nu}, & \text{for } \nu \in (\dfrac{3}{4}, 1) \end{cases} . \quad (4.20)$$

The integral in the equation converges to some constant value. The special case of ballistic motion does not apply for fBm, since the process is not defined for $\nu = 1$. For $\nu = \frac{1}{2}$ the Brownian motion case is restored.

**Stable motions:** Stable motions are an example for processes with independent increments, and consequently have an homogeneous ensemble. Nevertheless, care has to be taken that a valid observable is chosen. When $\mu$ is the stable index of the motion, $\langle |\Delta \boldsymbol{Y}(t; t+\tau)\,|^\gamma \rangle$ diverges, whenever $\gamma \geq \mu$. However choosing $\gamma < \mu/2$ ensures that $|\Delta \boldsymbol{Y}(t; t+\tau)\,|^\gamma$ has a finite variance. Our argument applies for such a choice of $\gamma$ and $J$ decays to zero as $T^{-1}$.

Let us now turn to processes with additional energetic disorder.

## 4.3. In energetic and mixed disorder

### 4.3.1. General considerations

When introducing energetic disorder, we lose the stationary-increments property of the whole process $\boldsymbol{X}(t) = \boldsymbol{Y}(U(t))$. This complicates the discussion and makes general arguments messier. The general plan is to use the conditional relation $\boldsymbol{Y} = \boldsymbol{X}|U$ and the properties of $\boldsymbol{Y}$.

**No agreement of the averages:** Lacking stationarity, there is no longer an equilibrium value of the msd. It is thus not apparent at all why an ergodic equation like

$$\frac{1}{T-\tau}\int_0^{T-\tau} \mathrm{d}t' \; \langle O(\Delta \boldsymbol{X}_\tau(t'))\rangle = \langle O(\Delta \boldsymbol{X}_\tau(t))\rangle$$

is supposed to hold. When a time average is applied, one averages over the time $t$ that explicitly appears on the right hand side of this equation. Above equation can thus only be true for one, but not all times $t$. In fact the above equation implies that $\langle O(\Delta \boldsymbol{X}_\tau(t))\rangle = \langle O(\Delta \boldsymbol{X}_\tau(t'))\rangle$ for any pair of epochs except at the origin. This can be seen by multiplying the equation by $T-\tau$ and taking a derivative with respect to $T-\tau$.

This does on the other hand not mean that the process $\boldsymbol{X}(t)$ has a heterogeneous ensemble. It is still possible that the fluctuations of the time average decay to zero; one has to normalize them properly, though.

**A Birkhoff-Khinchin argument:** The ensemble heterogeneity is not any more proportional to the variance of the time average. Hence, even when the latter vanishes, the absolute fluctuations of the time average can grow. The ensemble is homogeneous, when the variance grows slower than the square of the expected time average. This can be summarized in a generalized version of equation (4.9). Let $C_O(t,t')$ be the correlation function of $O(\Delta \boldsymbol{X}_\tau(t))$. It is not any more a mere function of the time difference, but still we can rewrite the ensemble heterogeneity in the way we did before as

$$J_O(T) = 2\int_0^1 \mathrm{d}y \int_0^{1-y} \mathrm{d}x \, \frac{C_O((T-\tau)\,x, (T-\tau)\,(x+y))}{\left\langle \overline{O(\Delta \boldsymbol{X}_\tau)}_T \right\rangle^2}.$$

Under the same assumptions as in Birkhoff-Khinchin theorem, namely continuity and boundedness of the integrand, limit and integration can be interchanged. A sufficient condition for an homogeneous ensemble is given when

$$\lim_{T\to\infty} \frac{C_O(xT, yT)}{\left\langle \overline{O(\Delta \boldsymbol{X}_\tau)}_T \right\rangle^2} = 0 \tag{4.21}$$

holds for almost every[1] $x, y \in [0,1]$.

Contrary to the arguments of Birkhoff-Khinchin theorem, an homogeneous ensemble does not imply ergodicity here. The process does not possess stationary increments, therefore the heterogeneity parameter is not proportional to the variance of the time average. Although the relative fluctuations of the time average die out, the absolute fluctuations may not.

**Processes subordinated to Gaussian ones:** When trying to use conditional expressions for the correlation function, the expressions get quite confusing. In general, since $\Delta U_\tau(t)$ is not stationary, the function

$$\left\langle \Delta \boldsymbol{X}^2(t, t+\tau)\, \Delta \boldsymbol{X}^2(t'; t'+\tau) \,|U\right\rangle$$

is a function of four arguments. The expressions become simpler, when $\boldsymbol{Y}(u)$ is Gaussian. Again using Wick's theorem, Eq.(4.10), we have

$$\begin{aligned}\left\langle \Delta \boldsymbol{X}^2(t_1,t_2)\, \Delta \boldsymbol{X}^2(t_3,t_4) \,|U\right\rangle =& \left\langle \Delta \boldsymbol{Y}^2_{\Delta U(t_1,t_2)}|U\right\rangle \left\langle \Delta \boldsymbol{Y}^2_{\Delta U(t_3,t_4)}|U\right\rangle \\ &+\frac{1}{2}\left\{\left\langle \Delta \boldsymbol{Y}^2_{\Delta U(t_1,t_3)}|U\right\rangle + \left\langle \Delta \boldsymbol{Y}^2_{\Delta U(t_2,t_4)}|U\right\rangle \right. \\ &\left. - \left\langle \Delta \boldsymbol{Y}^2_{\Delta U(t_1,t_4)}|U\right\rangle - \left\langle \Delta \boldsymbol{Y}^2_{|\Delta U(t_2,t_3)|}|U\right\rangle \right\}^2 .\end{aligned}$$

We abbreviate $g(\omega) := \left\langle \Delta \boldsymbol{Y}^2_\omega \right\rangle$ and write

$$\begin{aligned}K_\tau(t,t';U) :=& \frac{1}{2}\left\{g(|\Delta U(t,t')|) + g(|\Delta U(t+\tau,t'+\tau)|)\right. \\ &\left. - g(|\Delta U(t,t'+\tau)|) - g(|\Delta U(t+\tau,t')|)\right\}^2 . \end{aligned} \tag{4.22}$$

$K_\tau$ is a conditional average, therefore a random variable. We can express the correlation function of $\boldsymbol{X}$'s square displacement in terms of $g$ and $K_\tau$ as

$$C_{\Delta \boldsymbol{X}^2_\tau}(t,t') = C_{g(\Delta U_\tau)}(t,t') + \left\langle K_\tau(t,t';U)\right\rangle . \tag{4.23}$$

This expression has to be compared with the square of Eq.(4.17), which gives the equivalent expression for pure structural disorder. We can make out two differences: The first being the influence of $\Delta U$'s non-stationarity that modifies the correlation function of $\Delta \boldsymbol{Y}$, the second being the randomness of $U$ that introduces a whole new term. It is the covariance of some function of $U$'s increments – here identified with its correlation function – which vanishes when $U$ is a deterministic function of time.

When we apply the standard manipulations to the heterogeneity parameter, we obtain

$$J_{\Delta \boldsymbol{X}^2_\tau}(T) = \frac{2\int_0^{T-\tau} \mathrm{d}\tau_d \int_0^{T-\tau-\tau_d} \mathrm{d}t_m \left[C_{g(\Delta U_\tau)}(t_m, t_m+\tau_d) + \left\langle K_\tau(t_m, t_m+\tau_d; U)\right\rangle\right]}{\left\{\int_0^{T-\tau} \mathrm{d}t \left\langle g(\Delta U_\tau(t))\right\rangle\right\}^2}.$$

[1] I.e. for almost every point with respect to the Lebesgue-measure on $[0,1]\times[0,1]$.

When comparing the first of the summands to the definition of $J$, we observe that it is an heterogeneity parameter itself. Consequently, we may write:

$$J_{\Delta \boldsymbol{X}_\tau^2}(T) = J_{g(\Delta U_\tau)}(T) + 2\frac{\int_0^{T-\tau} \mathrm{d}\tau_d \int_0^{T-\tau-\tau_d} \mathrm{d}t_m \, \langle K_\tau(t_m, t_m+\tau_d; U)\rangle}{\left\{\int_0^{T-\tau} \mathrm{d}t \, \langle g(\Delta U_\tau(t))\rangle\right\}^2}. \tag{4.24}$$

The most striking difference to Eq.(4.15) is the appearance of the extra heterogeneity of the operational time. When the latter is deterministic, or when it fulfils Eq.(4.21), it does not contribute to the heterogeneity of $\boldsymbol{X}$. Without the subordination this term has cancelled, but now it constitutes the contribution of the operational time's correlation. The second difference is that $K_\tau$ has a much more intricate form than before, as has just been discussed.

$U$ is a monotonous process, i.e. $t < t'$ implies $U(t) \leq U(t')$. Therefore, when $\boldsymbol{Y}$ is a process with independent increments, $K_\tau(t, t+\tau_d; U)$ has to vanish for large arguments $\tau_d > \tau$. The reason is that $U$ respects causality. The integral of $K_\tau$ will be cut off at $\tau_d = \tau$ and thus is less significant in the limit $T \to \infty$. However, a Gaussian process with stationary and independent increments can already be identified as Brownian motion. Hence there is no structural disorder and the energetic disorder must be the cause of a finite ensemble heterogeneity. On the other hand, when the operational time $U(t)$ has independent increments, the first part will be less significant in the limit, since the covariance vanishes, when $\tau_d > \tau$.

Loosely spoken, the ensemble heterogeneity feeds of two contributions, the heterogeneity of $g(\Delta U_\tau)$ and the correlations of $\boldsymbol{Y}$ expressed in $K_\tau$. To be more precise, a calculation is needed and we therefore turn to some examples.

## 4.3.2. Examples

**Batchelor's process:** For Batchelor's process, or scaled Brownian motion, we can refine Eq.s(4.24) and (4.22) with the operational time $U(t) = t^\alpha$ and with $g(u) = 2dK_{2,\alpha}u$. Due to the deterministic nature of the of the operational time, $J_{g(\Delta U_\tau)}$ vanishes identically and the ensemble heterogeneity parameter can be written as

$$\begin{aligned} J^{(\mathrm{Bat})}_{\Delta \boldsymbol{X}_\tau^2}(T) =& \frac{4d^2K_{2,\alpha}^2 \int_0^{T-\tau} \mathrm{d}\tau_d \int_0^{T-\tau-\tau_d} \mathrm{d}t_m \, \left[(t_m+\tau)^\alpha - (t_m+\tau_d)^\alpha\right]^2 \mathbb{1}_{[0,\tau]}(\tau_d)}{\left\{2dK_{2,\alpha}\int_0^{T-\tau} \mathrm{d}t \, t^\alpha\right\}^2} \\ =& (\alpha+1)^2 \int_0^{\tau} \mathrm{d}\tau_d \int_0^{T-\tau-\tau_d} \mathrm{d}t_m \, \frac{\left[(t_m+\tau)^\alpha - (t_m+\tau_d)^\alpha\right]^2}{(T-\tau)^{2\alpha+2}}. \end{aligned}$$

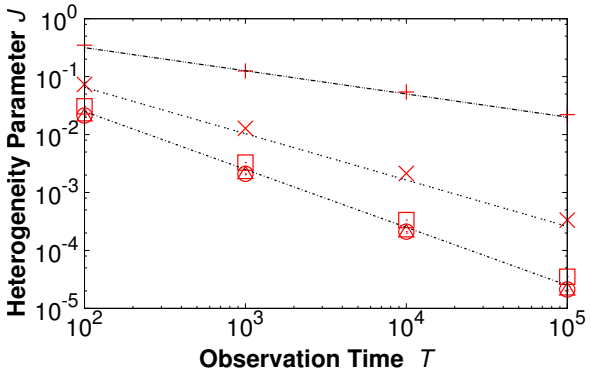



Figure 4.1.: Scaling of Batchelor's process' ensemble heterogeneity
$J$ for scaled Brownian motion was simulated for different values of $\alpha$ and $T$. The $\alpha$ values are 0.2 (+), 0.4 (×), 0.6 (⊡), 0.8 (⊙), and 1.0 (△). Each point is calculated from 5000 trajectories. The given power laws are (from top to bottom line) $T^{-0.4}$, $T^{-0.8}$ , and $T^{-1.0}$. They show that the heterogeneity decays as predicted by Eq.(4.25).

The whole problem was already discussed by the author in [179], where also the integral was shown to be

$$J^{(\mathrm{Bat})}_{\Delta \boldsymbol{X}^2_\tau}(T) \propto \begin{cases} 4\int_0^1 \mathrm{d}y \int_0^\infty \mathrm{d}x \, [(x+1)^\alpha - (x+y)^\alpha]^2 \left(\frac{\tau}{T}\right)^{2\alpha}, & \alpha < \frac{1}{2} \\ \frac{1}{3}\frac{\tau}{T}\ln\left(\frac{T}{\tau}\right), & \alpha = \frac{1}{2} \\ \frac{4\alpha^2}{3\,(2\alpha-1)}\frac{\tau}{T}, & \alpha > \frac{1}{2} \end{cases} . \tag{4.25}$$

The double integral converges in the given parameter regime and can be expressed in terms of a beta function. The decay of the heterogeneity parameter with measurement time is compared to simulations in Fig. 4.1. The slopes of the graphs are in good accordance to our predictions.

This shows that Batchelor's process has an homogeneous ensemble and the fluctuations of the time averages die out when $T$ becomes large. One would thus expect ergodic behaviour, but this is not the case, because

$$\left\langle \overline{|\Delta \boldsymbol{X}_\tau|^2}_T \right\rangle = \frac{2dK_{2,\alpha}}{\alpha+1}\frac{T^{\alpha+1} - \tau^{\alpha+1} - (T-\tau)^{\alpha+1}}{T-\tau} = \frac{2dK_{2\alpha}}{T^{1-\alpha}}\tau + \mathcal{O}\big(T^{-1}\big) . \tag{4.26}$$

Most remarkably, the time average grows asymptotically linearly in the lag time $\tau$, whereas the eamsd grows like:

$$\left\langle \Delta \boldsymbol{X}^2_\tau(t) \right\rangle = 2dK_{2,\alpha}\left[(t+\tau)^\alpha - t^\alpha\right] \propto \begin{cases} \frac{2\alpha dK_{2,\alpha}}{t^{1-\alpha}}\tau, & t \gg \tau \\ 2dK_{2,\alpha}\tau^\alpha, & \tau \ll t \end{cases} .$$

Hence, tamsd and eamsd exhibit the same scaling, but only in a certain regime and with a difference in the transport coefficients.

The diffusivity that is defined by the time averages depends explicitly on the measurement time $T$ and decays (or grows) as a power law. The reason is the non-stationarity of the process's increments. We want to stress here that ensemble heterogeneity and ergodicity are different properties of the process. Homogeneity of the ensemble implies ergodicity, when the process has stationary increments, but not necessarily when it has not. When the process has non-stationary increments, but has an homogeneous ensemble, the latter only ensures that the time averaged quantities are well defined deterministic objects (in our case we measure either zero or infinity, due to the lack of boundaries).

Let us now turn to a process with a random subordinator, where the heterogeneity of the operational time plays a role.

**Continuous time random walks:** The ensemble heterogeneity of ctrw was already investigated in [117], although in a different manner. The main observation in ctrw, very similar to Batchelor's process, is the linear growth of the double averaged mean squared displacement:

$$\left\langle \overline{|\Delta \boldsymbol{X}_\tau|^2}_T \right\rangle = \frac{2dK_{2,\alpha}}{\Gamma(2+\alpha)} \frac{T^{\alpha+1} - \tau^{\alpha+1} - (T-\tau)^{\alpha+1}}{T-\tau} = \frac{2dK_{2,\alpha}}{\Gamma(1+\alpha)\, T^{1-\alpha}} \tau + \mathcal{O}\big(T^{-\alpha}\big) \,. \tag{4.27}$$

As already discussed, this is a consequence of the increments' non-stationarity. However, contrary to the last example, the ensemble heterogeneity is not zero, rather it is given by:

$$J^{(\mathrm{ctrw})}_{\Delta \boldsymbol{X}^2_\tau}(T) = 2\frac{\Gamma^2(1+\alpha)}{\Gamma(1+2\alpha)} - 1 + \mathcal{O}\big(T^{-\alpha}\big) \,. \tag{4.28}$$

A comparison of this equation with simulation data is given in Fig. 4.2. In sub-figure 4.2b, we also present the decay of the error with growing measurement times (in a slightly more general set-up), which fulfils our expectation. The correct data set is the one with $\nu = 1/2$. The simulation values for $J$ are good, whenever $\alpha$ is large. For smaller values the very slow decay of the error ($T^{-\alpha}$!) renders the simulated data useless.

As a consequence of the heterogeneity, the time averaged diffusivity remains a random character and differs from trajectory to trajectory. This behaviour is in contrast to Batchelor's process and results from the random character of the operational time. Such universal (i.e. present in the asymptotic limit) fluctuations in the transport coefficients have already been discussed in [117, 118, 122]. In these publications one can find histograms of the transport coefficients that show a very broad shape.

Let us shortly discuss how to obtain Eq.(4.28). Following our approach, we have to compute (or estimate) all integrals in Eq.(4.24). This was done in [195] by the author,

to where we refer to for details. The ctrw can be represented as a subordinated Brownian motion, i.e. $\boldsymbol{X}(t) = \sqrt{2K_{2,\alpha}}\boldsymbol{B}(U(t))$, where the operational time is a renewal process with heavy tailed waiting time distribution. This relation allows us to identify $g(\omega) = 2dK_{2,\alpha}\omega$ in Eq.(4.24) and in Eq.(4.22). Using the expressions for the moments of $\Delta U_\tau(t)$ that are given in [195], we can compute the denominator of Eq.(4.24) as

$$\left\{ \int_0^{T-\tau} \mathrm{d}t \, \langle \Delta U_\tau(t) \rangle \right\}^2 = \frac{\tau^2 T^{2\alpha}}{\Gamma^2(1+\alpha)} + o\big(T^{2\alpha-1}\big) \, .$$

The integral over $\Delta U$'s covariance is

$$2 \int_0^{T-\tau} \mathrm{d}\tau_d \int_0^{T-\tau-\tau_d} \mathrm{d}t \, \mathrm{Cov}[\Delta U_\tau(t) \, , \Delta U_\tau(t)] = \tau^2 T^{2\alpha} \frac{2\Gamma^2(1+\alpha) - \Gamma(1+2\alpha)}{\Gamma(1+2\alpha)\,\Gamma^2(1+\alpha)} + o\big(T^{2\alpha}\big) \, .$$

The $K_\tau$-functional is given by

$$K_\tau(t, t+\tau_d; U) = 8dK_{2,\alpha}^2 \Delta U^2(t+\tau_d; t+\tau) \, \mathbb{1}_{[0,\tau]}(\tau_d)$$

and the its integral is $o(T^\alpha)$. Putting everything together gives Eq.(4.28). Noteworthy is the fact that the $K_\tau$-functional does not contribute in the asymptotic limit. It is the heterogeneity of $U$ that is responsible for the heterogeneity of the ensemble.

**The heterogeneity of the fundamental moment:** We close the discussion of ctrw with reviewing the method of He et al. to obtain $J$, [117]. Their approach is based on the scaling relation

$$\overline{|\Delta \boldsymbol{X}_\tau|^2}_T \sim \frac{1}{T} U(T)$$

that leads with a proper rescaling to

$$\frac{\overline{|\Delta \boldsymbol{X}_\tau|^2}_T}{\left\langle \overline{|\Delta \boldsymbol{X}_\tau|^2}_T \right\rangle} \sim \frac{U(T)}{\langle U(T) \rangle} .$$

Thus, the distribution of the rescaled time averaged mean squared displacement is determined by the distribution of the subordinator. In the case of ctrw this distribution can be expressed in terms of the Lévy-distribution with pdf $l_{\alpha,1}(x)$:

$$p_{\overline{|\Delta \boldsymbol{X}_\tau|^2}_T / \langle \overline{|\Delta \boldsymbol{X}_\tau|^2}_T \rangle}(\xi) = \frac{1}{\alpha \xi} \sqrt[\alpha]{\frac{\Gamma(1+\alpha)}{\xi}} \, l_{\alpha,1}\left( \sqrt[\alpha]{\frac{\Gamma(1+\alpha)}{\xi}} \right) . \tag{4.29}$$

It is, due to the scaling property of $U$ independent of $T$, whence the limit $T \to \infty$ is trivial. The ensemble heterogeneity parameter is the variance of the dimensionless time

average and can be computed from above equation, bearing (4.28). The main point of the argument is the scaling relation, which is plausible but lacks rigour. The approach of [117] does not provide an estimation of the error term as we did. It does provide the whole distribution of the renormalized time averages, though. Furthermore, Eq.(4.29) is very successful in predicting the simulations as can be seen in [117]. Their approach works by virtue of the linear scaling with operational time; because they investigated the time average of the *fundamental moment*. Consequently, we expect Eq.(4.29) to hold in the presence of structural disorder, when the mean squared displacement is replaced with the fundamental one. The rescaled time averages of the fundamental moment therefore have the same distribution as the rescaled operational time has.

**Subordinated fractional Brownian motion:** The last example is a process of true mixed disorder, the subordinated fractional Brownian motion. The process has been investigated by the author in [195], where also the ensemble heterogeneity has been calculated. The calculation follows the same line as for the ctrw but the intermediate results have to be adjusted. First of all, for fBm we have $\langle \Delta \boldsymbol{Y}_\omega^2 \rangle = g(u) = 2dK_{2,2\nu\alpha}\omega^{2\nu}$. Hence, the $K_\tau$-functional takes a much more complicated form than in ctrw:

$$\begin{aligned} K_\tau(t, t+\tau_d; U) =& \frac{\left(2dK_{2,2\nu\alpha}\right)^2}{2} \left\{ |\Delta U(t; t+\tau_d)|^{2\nu} + |\Delta U(t+\tau; t+\tau_d+\tau)|^{2\nu} \right. \\ & \left. - |\Delta U(t; t+\tau_d+\tau)|^{2\nu} - |\Delta U(t+\tau; t+\tau_d)|^{2\nu} \right\}^2 . \end{aligned}$$

It is easy to verify that the ctrw-result is restored, when the Hurst index $\nu$ is set to one half. In that particular case, the whole function vanishes when $\tau_d > \tau$. However, the integral over $K_\tau$ can still be shown to be $\mathcal{O}(T^\alpha)$. What remains to do is to calculate the heterogeneity of $\Delta U_\tau^{2\nu}$. Each part can be estimated separately, giving

$$\begin{aligned} & \int_0^{T-\tau} \mathrm{d}\tau_d \int_0^{T-\tau-\tau_d} \mathrm{d}t_m \, \left\langle \Delta U_\tau^{2\nu}(t_m) \, \Delta U_\tau^{2\nu}(t_m+\tau_d) \right\rangle \\ =& \frac{\Gamma^2(1+2\nu)}{\Gamma^2(2+\alpha\,(2\nu-1))} \frac{\tau^{2+2\alpha(2\nu-1)} T^{2\alpha}}{\Gamma(1+2\alpha)} + o\big(T^{2\alpha}\big) , \end{aligned}$$

and

$$\left\{ \int_0^{T-\tau} \mathrm{d}t \, \left\langle \Delta U_\tau^{2\nu}(t) \right\rangle \right\}^2 = \frac{\Gamma^2(1+2\nu)}{\Gamma^2(2+\alpha\,(2\nu-1))} \frac{\tau^{2+2\alpha(2\nu-1)} T^{2\alpha}}{\Gamma^2(1+\alpha)} + \mathcal{O}\big(T^{2\alpha-1}\big) .$$

This results in

$$J_{\Delta \boldsymbol{X}_\tau^2}^{(\mathrm{sfBm})}(T) = J_{|\Delta U_\tau|^{2\nu}}(T) + \mathcal{O}\big(T^{-\alpha}\big) = 2\frac{\Gamma^2(1+\alpha)}{\Gamma(1+2\alpha)} - 1 + \mathcal{O}\big(T^{-\alpha}\big) , \tag{4.30}$$

the same result as in ctrw. Apparently the ensemble heterogeneity is determined solely by the heterogeneity of the operational time. Although the $K_\tau$-functional has a broad

support (in contrast to the small support $\tau_d < \tau$ in the Brownian motion case), its integral is still of smaller order than the heterogeneity of $U$. In this particular model, the structural disorder has no impact on the ensemble heterogeneity parameter, it is only determined by the randomness of the non-stationary increments of the subordinator. Equation (4.30) can be compared with simulations, which is done in Fig. 4.2. In sub-figure 4.2a we compared the asymptotic value with simulations. The result is good as long as the values of $\alpha$ and $\nu$ are reasonable. If they are not, i.e. $\alpha$ is close to zero, or $\nu$ is far from one half, the results become very bad, which is a consequence of the slow decay of the error term ($\mathcal{O}(T^{-\alpha})$). In sub-figure 4.2b, we thus plotted the residual against the measurement time $T$, they show a power law decay like $T^{-\alpha}$ when the Hurst index is smaller than 0.9. In the case $\nu = 0.9$, the decay is even slower, a behaviour that is not captured by the estimations of [195], which we reproduced here. In this regime it may even be questionable to decide from the simulations whether the asymptotic value is correct or not, since the decay is so slow. However, simulations with long measurement times and short lag times become are very costly.

## 4.4. Summary

In this chapter, we investigated certain properties of stochastic processes related to ergodicity. An observable is called ergodic, when its time average bears the same value as an ensemble average in thermodynamic equilibrium, in a certain sense, i.e.

$$\overline{O(\Delta \boldsymbol{X}_\tau)} = \langle O(\Delta \boldsymbol{X}_\tau) \rangle_{\mathrm{eq}} .$$

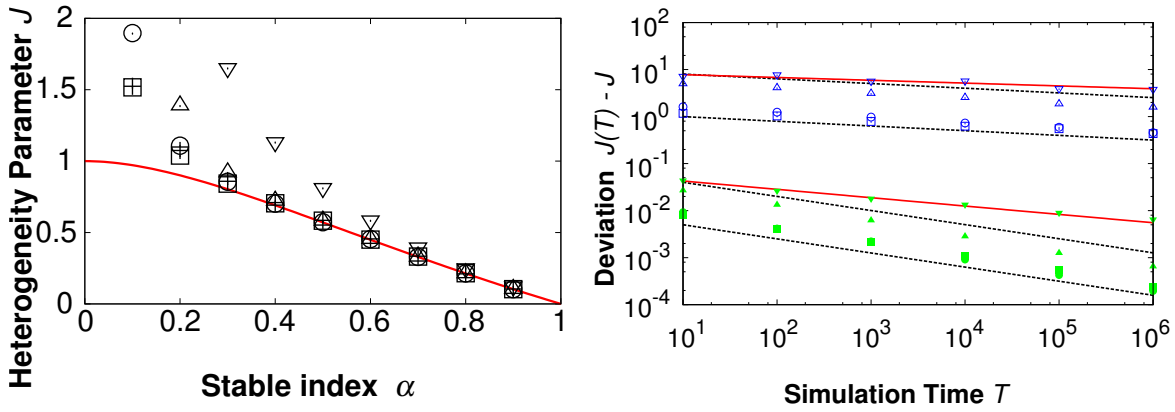
This relation holds for processes with stationary increments, and without a separation of state space.

We here investigated the validity of the ergodic relation for some powers of the absolute value of displacement in mean and in mean square. Furthermore, we investigated the ensemble heterogeneity or weak ergodicity breaking parameter

$$J_O(T) := \frac{\mathrm{Var}\left[\overline{O(\Delta \boldsymbol{X}_\tau)}_T\right]}{\left\langle \overline{O(\Delta \boldsymbol{X}_\tau)}_T \right\rangle^2},$$

which measures how homogeneous an ensemble of trajectories appears with respect to the observable $O$. We said that processes, for which $J_O$ decays to zero in the limit of infinite measurement times $T$, have an homogeneous ensemble. We expressed the ensemble heterogeneity in terms of the correlation function of $C_O(t, t')$, defined by the covariance of the observable evaluated at two epochs $t$ and $t'$. Birkhoff-Khinchin theorem states that for a stationary observable, the decay of this correlation function with growing distance $|t - t'|$ is sufficient for ergodicity in mean square. This property, the decay of the correlation function, is called irreversibility.

We discussed the related phenomenon of weak ergodicity breaking, which denotes a departure of the ergodic relation in processes without separation of state space.




(a) Simulated values

(b) Scaling with measurement time

Figure 4.2.: Ensemble heterogeneity for sfBm

(4.2a): This plot shows simulated values of sfBm's ensemble heterogeneity. Each symbol corresponds to a different value of the Hurst parameter: $\nu = 0.1$ (+), 0.3 (□), 0.5 (⊙), 0.7 (△), 0.9 (▽) Every point is an average of 15 simulations with 4096 trajectories, with $T = 10^6$ and $\tau = 2^{-17}T$. The prediction, Eq.(4.30), is given by the solid red line. The agreement is good for large $\alpha$ and gets worse for smaller values. The ctrw case is $\nu = 0.5$ (⊙).

(4.2b): Here we plotted the deviation of the simulated values from the prediction over the measurement time. The deviation decays as a power law $T^{-\alpha}$ except for $\nu = 0.9$, where the decay is even slower. Hurst indices are 0.1 (□), 0.3 (⊙), 0.7 (△), and 0.9 (▽). $\alpha$ is given by 0.1 (empty symbols) and 0.3 (full symbols). The symbols for $\alpha = 0.3$ are shifted for lucidity. For comparison, dotted lines with slopes $-0.1$ (top) and $-0.3$ (bottom) are given. Solid lines are fits for $\nu = 0.9$ and have the slopes $-0.18 \pm 0.01$ (bottom line, $\alpha = 0.3$) and $-0.06 \pm 0.01$ (top line, $\alpha = 0.1$). Each symbol is an average of 25 simulations with $10^3$ trajectories.

Another effect that may go hand in hand with the violation of the ergodic relation is the inherent randomness of diffusivities measured via time averages. In that case, the transport coefficient exhibits universal fluctuations and the corresponding ensemble heterogeneity $J_O$ approaches a non-zero value even in the limit $T \to \infty$.

We started the treatment with processes with pure structural disorder. Such processes have stationary increments, and are ergodic in mean. Also, we showed that irreversibility implies that $J_O(T)$ decays to zero. Consequently, processes with independent increments are ergodic in mean square, whenever the mean square of the corresponding observable is finite.

For Gaussian processes we used Wick's theorem to express the correlation function of the square displacement in terms of the correlation function of the simple displacements. We inferred the asymptotic behaviour of the ensemble heterogeneity. It decays for $T \to \infty$ as long as the process is asymptotically sub-ballistic. In the ballistic case, $J_O$ will approach the value two. In the super-ballistic case, it will diverge as $T \to \infty$. The discussion was exemplified in fBm and stable motions.

After that, we turned to processes with additional energetic disorder. It was shown that the lack of stationary increments implies a violation of ergodicity in mean. Nevertheless, in an attempt to apply an argument á la Birkhoff-Khinchin, we found a sufficient condition for the asymptotic decay of the heterogeneity parameter, namely:

$$\lim_{T\to\infty} \frac{C_O(xT, yT)}{\left\langle \overline{O(\Delta \boldsymbol{X}_\tau)}_T \right\rangle^2} = 0.$$

When the structural disorder component $\boldsymbol{Y}(u)$ is a Gaussian process, we were able, by again using Wick's theorem to rewrite the heterogeneity parameter as

$$J_{\Delta \boldsymbol{X}^2_\tau}(T) = J_{g(\Delta U_\tau)}(T) + 2\frac{\int_0^{T-\tau} \mathrm{d}\tau_d \int_0^{T-\tau-\tau_d} \mathrm{d}t_m \, \langle K_\tau(t_m, t_m+\tau_d; U)\rangle}{\left\{ \int_0^{T-\tau} \mathrm{d}t \, \langle g(\Delta U_\tau(t))\rangle \right\}^2}.$$

Here the function $g(\omega) := \langle \Delta \boldsymbol{Y}^2_\omega \rangle$ abbreviates the dependence of $\boldsymbol{Y}$'s increments on the operational time and $K_\tau$ is a random function connected to the correlation function of $\boldsymbol{Y}$:

$$\begin{aligned} K_\tau(t, t'; U) :=& \frac{1}{2} \{ g(|\Delta U(t, t')|) + g(|\Delta U(t+\tau, t'+\tau)|) - \\ & - g(|\Delta U(t, t'+\tau)|) - g(|\Delta U(t+\tau, t')|) \}^2 . \end{aligned}$$

Hence it turned out that there is a direct contribution of the operational time and a contribution of $\boldsymbol{Y}$'s correlations.

We proceeded to discuss examples, and began with Batchelor's process. This process was not ergodic in mean, but had an homogeneous ensemble. The next process we considered was ctrw, for which it was shown that it is non-ergodic in mean, but this

time with a heterogeneous ensemble. The major contribution to the heterogeneity parameter was due to the operational time's heterogeneity, since the structural part is Brownian motion and has independent increments.

After discussing ctrw, we reviewed the method of [117] to calculate $J$ and remarked that it relies only on a linear relation between msd and operational time. Hence, the method also applies to the fundamental moment, even in the presence of structural disorder. The time average of the fundamental moment has the same probability law as the operational time.

In the end we investigated sfBm, a process of mixed disorder. Here again it turned out that only the operational time's heterogeneity determines the heterogeneity of the process. Consequently, the ensemble heterogeneity of sfBm is the same as the one of ctrw. We compared the analytical results with simulations. For Batchelor's process the results coincide with the theory quite good. For ctrw and sfBm, the simulations agree with the theory in reasonable parameter regimes. This means that for moderate values of $\nu$ the heterogeneity exhibited large but predicted deviations. For large Hurst parameters however, the deviations are even larger than the bounds given by our theory.

Nevertheless, alone by comparison of Batchelor's process and ctrw, we showed that universal fluctuations in time averaged transport coefficients do not go hand in hand with a violation of the ergodic hypothesis. Both processes have non-stationary increments, and thus violate (4.2). However only in ctrw, which has a *random* operational time, we observe fluctuations in the time averaged diffusivity, indicated by a non-vanishing value of the ensemble heterogeneity. Interestingly, the exact same value is recovered in sfBm, indicating (although without a rigorous proof) that universal fluctuations in time averages are a consequence of a random operational time. A mild structural disorder (in the sense of a sub-ballistic scaling of the structural part) as well as mild energetic disorder (in the sense of a non-random operational time) seems not to be sufficient for observing random transport coefficients. On the other hand, the situation of a process, which is subordinated to an asymptotically ballistic Gaussian process – which alone possesses an heterogeneous ensemble – still needs to be investigated.

# 5. Summary and Conclusion

We now summarize the thesis. It was introduced with chapter 1. We have started with the theoretical description of Fick, who found the same equation for the dispersal of sodium chloride in water as Fourier found for the transport of heat. We have mentioned some further theoretical treatment of diffusion or Brownian motion by Einstein, Smoluchowski, Langevin and others, who established several different approaches to the problem. Also the experimental works of Perrin were brought up. They lead to the establishment of the atomistic conception.

We continued to argue that the models that describe the motion of microscopic particles can be phenomenologically employed whenever there are processes that can be understood as randomly changing. This, so we said, enables the modeller to describe phenomena not only in physics, but also in economics, biology, and health sciences. One of the recent fields in physics is the motion of large molecules in biological cells.

In contemporary experiments, tracer particles are marked with fluorescents and tracked either in compound or individually. This enables to measure the mean squared displacement (msd), or certain correlation functions derived from the propagator.

In the earliest experiments and theories, it was shown that the msd grows linearly with time. Later, it was discovered that there are situations, in which the msd exhibits a non-linear growth. Those were coined anomalous diffusion in contrast to a normal linear behaviour. Among others, anomalous diffusion appears in turbulence, polymer science, disordered systems, finance, animal motion and biophysics. A feature often appearing in anomalous diffusion processes is the so-called weak ergodicity breaking, which is the non-coincidence of time and ensemble averages, when the system's state space consists of only one component. Often when weak ergodicity breaking appears, time averaged transport coefficients like the diffusivity exhibit random fluctuations that do not vanish, even in the limit of infinitely long measurement intervals.

The main motivation of this thesis was the diffusion of large molecules in biological cells, which is anomalous and may exhibit weak ergodicity breaking. In this particular example, many effects are superposed and generate the anomalous behaviour together. The particles motion is obstructed by other components of the cell, there are visco-elastic interactions with its surroundings, there may be internal states of the particle that change its mobility, and it may be subject to active transport by molecular motors. Each effect itself can be described by a particular mathematical model, but the challenge is to understand their superposition. Furthermore, when data was collected, it is not a trivial problem to infer the correct mathematical model that explains the data.

The aim of this thesis was to provide a method that is able distinguish different models of anomalous diffusion. In that way, not only the correct mathematical de-

scription is identified, but also the underlying physical mechanism that is reflected by the mathematics.

In chapter 2 of the thesis, we tried to classify two physical mechanisms that are responsible for anomalous diffusion: structural and energetic disorder. After introducing the necessary notation in normal diffusion, we started with a model of diffusion in disordered media: a master equation with random rates. Such a model is able to reproduce many of the aforementioned effects.

For each fixed set of random rates, we identified three parts. First, there was the graph of all possible transitions and its corresponding adjacency matrix. Second, we used Kirchhoff's theorem to construct the stationary distribution on this graph and argued that it is a thermodynamic equilibrium. This equilibrium state was proportional to the Boltzmann factors, which were the second ingredient. The last part were the equilibrium transition rates, the rate of particles moving from one site to another in the equilibrium state. By virtue of a detailed balance condition those turned out to be symmetric.

This decomposition was used to find an appropriate annealed description of the disordered system. By a combination of effective medium and coarse graining arguments, we replaced the Boltzmann factors with their average, the transition graph with a regular "reference" graph and the equilibrium rates with an effective medium bond diffusivity. The reference graph had to be chosen such, that it connects all those sites that may have been connected in the original graph. Furthermore, the strength of the reference connections had to be chosen according to the stochastic spatial dependence of the equilibrium rates. Only then the effective medium bond diffusivity could assume a finite value.

We used the annealed model to derive the effective coefficient of normal diffusion given by

$$2dK_{\text{eff}} = \frac{C_{\triangle^*}\mathbb{E}[AwD^*]_{\text{EM}}}{\mathbb{E}[e^{-\beta\mathscr{F}}]}.$$

When it is infinite, the process is superdiffusive, when it vanishes the process is subdiffusive. Two physical mechanisms can be identified in that way: Energetic disorder, when the denominator is not finite and structural disorder, when the enumerator is not finite. Both can be present simultaneously. Structural disorder appears, when the connection factor $C_{\triangle^*}$ is infinite or the effective medium bond diffusivity vanishes. This is possible, when there are long-range connections or in situations similar to percolation. Energetic disorder, on the other hand, is possible when the average Boltzmann factor diverges.

We proceeded with transferring the terms energetic and structural disorder to the language of stochastic processes. We first started to interpret the master equation with random rates as the forward equation of a Markov process. It was argued that the Markov property may get lost in the annealing procedure, but that in pure structural disorder another feature remains, namely the stationarity of increments. Furthermore, energetic disorder was connected to a change of local time and hence to a subordination

procedure. A stochastic process of mixed, structural and energetic, disorder therefore was said to have the following subordinated form:

$$\boldsymbol{X}(t) = \boldsymbol{Y}(U(t))\,.$$

Where $\boldsymbol{Y}$ represents the structural disorder component and is a process with stationary increments. $U$ is called the operational time. It represents the energetic disorder and is a non-decreasing process. This finished chapter 2.

This subordination ansatz was analysed in more detail in chapter 3. There we tried to find a method that is able to separate $\boldsymbol{Y}$ from the operational time $U$. Only when this is possible, the theoretical distinction made in the before chapter, has practical value. This was indeed possible under the additional assumption that the structural disorder component has a scaling form. Then it is possible to find the fundamental moment of displacement, which is additive in time, i.e. which fulfils:

$$\langle|\Delta\boldsymbol{X}(t_1;t_3)|^{\gamma_F}\rangle = \langle|\Delta\boldsymbol{X}(t_1;t_2)|^{\gamma_F}\rangle + \langle|\Delta\boldsymbol{X}(t_2;t_3)|^{\gamma_F}\rangle\,,$$

for any $t_1 \leq t_2 \leq t_3$. The corresponding exponent $\gamma_F$ is called the fundamental exponent and the reciprocal of the scaling exponent of the structural disorder component: $\gamma_F = 1/\nu$. We showed that the fundamental moment reveals the time dependence of the operational time, regardless of how complicated that time dependence is.

It was demonstrated that the fundamental exponent can also be inferred from the quantiles of the increments, when the fundamental exponent does not exist. Furthermore, when the structural disorder component has a complicated scaling form, there exists a fundamental expectation that is additive in time. The discussion was supported with several numerical simulations of processes with pure and mixed disorder.

In the final chapter, namely chapter 4, we turned to the phenomenon of weak ergodicity breaking. The ergodicity of the msd and other averages of the displacement was discussed in terms of several statistics. Namely, in terms of the expected time averages, the absolute and relative variance of the time average, and in terms of correlation functions. When the expected time average has been equal to its ensemble average in the limit of long measurement intervals, we said the observable was ergodic in mean. If additionally, the time average's variance vanished, we said that the observable is ergodic in mean square. Furthermore, the relative variance is a measure of how similar different trajectories of the ensemble appear when compared with respect to the observable in question. This relative variance was defined by:

$$J_O(T) := \frac{\mathrm{Var}\left[\overline{O(\Delta\boldsymbol{X}_\tau(t))}_T\right]}{\left\langle\overline{O(\Delta\boldsymbol{X}_\tau(t))}_T\right\rangle^2},$$

and we called it the ensemble heterogeneity, although it is usually called weak ergodicity breaking parameter in the literature. When $J$ vanishes asymptotically, we said that the ensemble is homogeneous, it is heterogeneous otherwise. $J$ quantifies the fluctuation of the time averages in the asymptotic limit. A heterogeneous ensemble hence is connected to the universal fluctuation phenomena in the transport coefficients.

We discussed first processes of pure structural disorder, and then those with additional energetic disorder. In processes of pure structural disorder, any observable of the displacement is ergodic in mean, since the increments are stationary. Furthermore, the ensemble heterogeneity parameter is proportional to the variance of the time averages and Birkhoff-Khinchin theorem ensures that an homogeneous ensemble is a sufficient condition for ergodicity in mean square. In Gaussian processes, we expressed $J$ for the msd in terms of the displacements correlation function by using Wick's theorem. That way we showed that every sub-ballistic process has an homogeneous ensemble, a ballistic Gaussian process has a heterogeneous ensemble and finally that $J$ diverges for any super-ballistic Gaussian process.

In the presence of energetic disorder, the ergodicity in mean breaks down. Since the state space in our models consists of one communicating component, this is weak ergodicity breaking. General considerations lead us to a Birkhoff-Khinchin-like sufficient condition for a homogeneous ensemble. When the structural disorder component is a Gaussian process, we were able to use Wick's theorem once again, and to split up the heterogeneity parameter into two parts:

$$J_{\Delta \boldsymbol{X}_\tau^2}(T) = J_{g(\Delta U_\tau)}(T) + 2\frac{\int_0^{T-\tau} \mathrm{d}\tau_d \int_0^{T-\tau-\tau_d} \mathrm{d}t \ \langle K_\tau(t_1, t_2; U)\rangle}{\left\{\int_0^{T-\tau} \mathrm{d}t \ \langle g(\Delta U_\tau(t))\rangle\right\}^2}.$$

The first of those two parts was the ensemble heterogeneity of the energetic disorder component. The second was connected to the correlation function of the structural disorder component. Hence, we saw that the operational time's heterogeneity is a direct contribution to the compound process' heterogeneity.

We discussed a few special processes. First of all Batchelor's process, which breaks ergodicity in mean, but has an homogeneous ensemble. Another process, ctrw, has a heterogeneous ensemble and still breaks ergodicity in mean. This showed that universal fluctuations expressed by a non-vanishing heterogeneity parameter and weak ergodicity breaking don't always go hand in hand. Hence we refrained from the notion of a weak ergodicity breaking parameter.

We reviewed the method of [117] to compute $J$ for ctrw and found that the time average of the fundamental moment has to have the same distribution as the operational time.

Finally, we turned to sfBm, as a candidate for mixed disorder. As it was the case in ctrw, the heterogeneity of the operational time process dominated the heterogeneity parameter of the whole process, bearing the same result as in ctrw. The discussion was supported with numerical simulations of sfBm and Batchelor's process.

We conclude the thesis now. Models for diffusion or random motion can be applied in many fields of science. Sometimes it is necessary to transcend bare description and prediction of data, by rather understanding the responsible mechanisms. For physical systems exhibiting random motion, we described two of such mechanisms, energetic and structural disorder. Both types are expressed in the subordination ansatz, which

is a great a-priori tool, but leaves a lot of open questions. In particular, we gave no rigorous statements or limit theorems to support our homogenization arguments. Although interesting, such questions currently lie beyond the skills of the author.

The fundamental properties are an excellent and practicable tool to separate the energetic and structural disorder components. They enable us to determine the physical mechanisms at work in real-life situations, and to understand better transport phenomena in complex environments. We hope that they are soon used for that purpose. But some more theoretical work needs to be done, too. For instance, the situation when the structural disorder component exhibits crossover needs to be investigated in greater detail. Furthermore, the fundamental properties need to be generalized to single particle situations. Such a technique would essentially be a more intricate version of the $p$-variation test.

The analysis of weak ergodicity breaking is a very active field of research, to which we could only give a small contribution. Nevertheless, we were able to shed a little light on the connection between stationarity of the process' increments and the heterogeneity of its ensemble.

We especially hope that the techniques described in this thesis turn out to be useful when investigating the diffusion of large molecules in biological cells.

# A. Appendices

## A.1. The propagator of the discrete diffusion equation

We apply a standard method of solving Eq.(2.4) with free boundary conditions and concentrated initial condition. This solution can be found throughout the literature, for example in [40, p.269 et seq.] or in [196, p.41]. We start with applying Fourier and Laplace transform to (2.4), obtaining

$$s\rho(\boldsymbol{k};s) - \rho_0(\boldsymbol{k}) = \frac{2K_{2,1}}{a^2}\sum_{j=1}^{d}\left[\cos(a\boldsymbol{k}\cdot\boldsymbol{e}_j)-1\right]\rho(\boldsymbol{k};s)\,,$$

where

$$\rho(\boldsymbol{k};s) := \int_0^\infty \mathrm{d}t \sum_{\boldsymbol{x}\in a\mathbb{Z}^d} e^{\imath\boldsymbol{k}\cdot\boldsymbol{x}-st}\rho(\boldsymbol{x};t)$$

and analogous for $\rho_0(\boldsymbol{k})$. We may identify $2K_{2,1}/a^2 = d/\tau$ from Eq.(2.7), and solve the equation for $\rho(\boldsymbol{k};s)$:

$$\rho(\boldsymbol{k};s) = \frac{\tau\rho_0(\boldsymbol{k})}{s\tau + 1 - \frac{1}{d}\sum_{j=1}^{d}\cos(ak_j)}, \tag{A.1}$$

where $k_j$ is the $j$-th component of $\boldsymbol{k}$. We obtain the propagator, when the initial condition is a Kronecker delta, i.e. when $\rho_0(\boldsymbol{k}) = e^{\imath\boldsymbol{k}\cdot\boldsymbol{x}_0}$. For this initial condition, the solution can be obtained in original space. When first inverting the Laplace transform, an exponential is obtained. Inverting the Fourier transform afterwards gives the following product of integrals:

$$G(\boldsymbol{x};t|\boldsymbol{x}_0;0) = e^{-\frac{t}{\tau}}\prod_{j=1}^{d}\left\{\frac{a}{2\pi}\int_{-\frac{\pi}{a}}^{\frac{\pi}{a}}\mathrm{d}k_j\, e^{-\imath k_j(x_j-x_{0,j})+\frac{t}{d\tau}\cos(ak_j)}\right\}.$$

The expression in the brackets can be identified with a modified Bessel function of the first kind (use [197, p.376, Eq. 9.6.19])

$$G(\boldsymbol{x};t|\boldsymbol{x}_0;0) = e^{-\frac{t}{\tau}}\prod_{j=1}^{d} I_{(x_j-x_{0,j})/a}\left(\frac{t}{d\tau}\right). \tag{A.2}$$

This is Eq.(2.10) of the main text.

Now a series expansion, [197, p.375, Eq.9.6.10], is applied to the modified Bessel function. After the expansion, the Cauchy product of the series is computed and the expression is multiplied with a unit factor $u!/u!$:

$$\begin{aligned}
G(\boldsymbol{x};t|\boldsymbol{x}_0;0) =& e^{-\frac{t}{\tau}} \prod_{j=1}^{d} I_{(x_j-x_{0,j})/a}\left(\frac{t}{d\tau}\right) \\
=& e^{-\frac{t}{\tau}} \prod_{j=1}^{d} \sum_{u_j=0}^{\infty} \frac{\left(\frac{t}{2d\tau}\right)^{2u_j+\frac{(x_j-x_{0,j})}{a}}}{u_j!\left(\frac{au_j+x_j-x_{0,j}}{a}\right)!} \\
=& e^{-\frac{t}{\tau}} \sum_{u=0}^{\infty} \sum_{u=\sum_j \frac{2au_j+x_j-x_{0,j}}{a}} \prod_{j=1}^{d} \frac{\left(\frac{t}{2d\tau}\right)^{2u_j+\frac{(x_j-x_{0,j})}{a}}}{u_j!\left(u_j+\frac{(x_j-x_{0,j})}{a}\right)!} \\
=& e^{-\frac{t}{\tau}} \sum_{u=0}^{\infty} \left(\frac{t}{2d\tau}\right)^{u} \sum_{u=\sum_j \frac{2au_j+x_j-x_{0,j}}{a}} \frac{1}{\prod_{j=1}^{d} u_j!\left(\frac{au_j+x_j-x_{0,j}}{a}\right)!} \\
=& e^{-\frac{t}{\tau}} \sum_{u=0}^{\infty} \frac{1}{u!}\left(\frac{t}{2d\tau}\right)^{u} \sum_{u=\sum_j \frac{2au_j+x_j-x_{0,j}}{a}} \frac{u!}{\prod_{j=1}^{d} u_j!\left(\frac{au_j+x_j-x_{0,j}}{a}\right)!}.
\end{aligned}$$

The fraction in the last line is a multinomial coefficient, a generalization of the binomial coefficient. We obtain:

$$G(\boldsymbol{x};t|\boldsymbol{x}_0;0) = e^{-\frac{t}{\tau}} \sum_{u=0}^{\infty} \frac{1}{u!}\left(\frac{t}{\tau}\right)^{u} \frac{1}{(2d)^u} \sum_{u=\sum_j \frac{2au_j+x_j-x_{0,j}}{a}} \frac{u!}{\prod_{j=1}^{d} u_j!\left(\frac{2au_j+x_j-x_{0,j}}{a}\right)!}$$

and may identify the probability that a Poisson process $U(t)$ assumes the value $m$ at time $t$

$$\mathbb{P}\{U(t)=u\} = e^{-\frac{t}{\tau}} \frac{1}{u!}\left(\frac{t}{\tau}\right)^{u},$$

and the probability that a discrete time random walker $\boldsymbol{Y}(u)$ on $a\mathbb{Z}^d$ that started at $\boldsymbol{x}_0$, is in $\boldsymbol{x}$ after $u$ steps:

$$\mathbb{P}\{\boldsymbol{Y}(u)=\boldsymbol{x}|\boldsymbol{Y}(0)=\boldsymbol{x}_0\} = \frac{1}{(2d)^u} \sum_{u=\sum_j \frac{2au_j+x_j-x_{0,j}}{a}} \frac{u!}{\prod_{j=1}^{d} u_j!\left(\frac{2au_j+x_j-x_{0,j}}{a}\right)!}.$$

This way we obtain:

$$G(\boldsymbol{x};t|\boldsymbol{x}_0;0) = \sum_{u=0}^{\infty} \mathbb{P}\{\boldsymbol{Y}(u)=\boldsymbol{x}|\boldsymbol{Y}(0)=\boldsymbol{x}_0\}\,\mathbb{P}\{U(t)=u\}\,, \tag{A.3}$$

which is Eq.(2.12)

## A.2. Existence and stability of the stationary state in quenched disorder master equations

We start with equation (2.16) of the main text

$$\dot{\rho}(\boldsymbol{x};t) = \sum_{y\in\Omega} \left[\lambda(\boldsymbol{x};\boldsymbol{y})\,\rho(\boldsymbol{y};t) - \lambda(\boldsymbol{y};\boldsymbol{x})\,\rho(\boldsymbol{x};t)\right] =: \triangle_\lambda \rho(\boldsymbol{x};t)$$

and reproduce the results of Kirchhoff [153] and Schnakenberg [104]. We assume the environment, i.e. the set of rates $\{\lambda\}$, as fixed. As already suggested in the main text, the environment defines a graph with the set of nodes $\Omega$. A link is placed between $\boldsymbol{x}$ and $\boldsymbol{y}$ when $\lambda(\boldsymbol{y};\boldsymbol{x}) > 0$ or $\lambda(\boldsymbol{x};\boldsymbol{y}) > 0$. This condition is a little different from the one given in the main text, but more close to Schnakenberg's treatment. We denote the (directed) bonds with Greek letters, i.e. $(\boldsymbol{y};\boldsymbol{x}) =: \boldsymbol{\xi}$.

We set out to find the stationary state $\rho_{\rm st}$ that is defined by $\triangle_\lambda \rho_{\rm st} = 0$ and eventually some boundary conditions, for free diffusion this is the normalization condition. As already noted by Kirchhoff, it can be determined by examining the spanning trees of the graph. Those are sub-graphs that span all of $\Omega$ but contain no loops, they are countable. The directed spanning tree $T_{l,\boldsymbol{x}}$ is the $l$-th spanning tree, in which all links are pointing towards the node $\boldsymbol{x}$.

According to Kirchhoff's theorem, the stationary state is given by a sum of products of the rates:

$$\rho_{\rm st}(\boldsymbol{x}) = \frac{\sum_l \prod_{\boldsymbol{\xi}\in T_{l,\boldsymbol{x}}} \lambda(\boldsymbol{\xi})}{\sum_{\boldsymbol{x}\in\Omega}\sum_l \prod_{\boldsymbol{\xi}\in T_{l,\boldsymbol{x}}} \lambda(\boldsymbol{\xi})}. \tag{A.4}$$

This is Eq.(2.20) of the main text. We see immediately that $\rho_{\rm st}$ is positive and properly normalized for any finite graph. Furthermore it is apparent that $\rho_{\rm st}$ vanishes on a sub-domain when the only links reaching outside of that sub-domain are *leaving*. In that case, one factor $\lambda(\boldsymbol{\xi})$ pointing inside the sub-domain would be zero and the whole product vanishes. $\rho_{\rm st}$ splits up into several linear independent states, when the transition graph consists of multiple non-communicating sub-graphs. To have a unique stationary solution, we need to require that there is only one communicating component.

The stationary state is an attractor. To see this, one needs to consider the Kullback-Leibler divergence[1] of any concentration profile $\rho$ obeying (2.16)

$$K[\rho(t)\,;\rho_{\rm st}] := \sum_{\boldsymbol{x}\in\Omega} \rho(\boldsymbol{x};t)\ln\frac{\rho(\boldsymbol{x};t)}{\rho_{\rm st}(\boldsymbol{x})}. \tag{A.5}$$

[1] When $\rho(\boldsymbol{x}) = 0$, we put $\rho(\boldsymbol{x})\ln\frac{\rho(\boldsymbol{x})}{\rho_{\rm st}(\boldsymbol{x})} = 0$.

We can show that this is a Lyapunov function. To see this note first that $K[\rho;\rho_{\mathrm{st}}] \geq 0$, due to Gibbs inequality:

$$K[\rho;\rho_{\mathrm{st}}] = -\sum_{\boldsymbol{x}\in\Omega} \rho(\boldsymbol{x}) \ln \frac{\rho_{\mathrm{st}}(\boldsymbol{x})}{\rho(\boldsymbol{x})} \geq \sum_{\boldsymbol{x}\in\Omega} \rho(\boldsymbol{x}) \left[1 - \frac{\rho_{\mathrm{st}}(\boldsymbol{x})}{\rho(\boldsymbol{x})}\right] = \sum_{\boldsymbol{x}\in\Omega} [\rho(\boldsymbol{x}) - \rho_{\mathrm{st}}(\boldsymbol{x})] = 0.$$

We used $\ln x \leq x - 1$ and the normalization of $\rho$ and $\rho_{\mathrm{st}}$. Equality holds only when $\rho = \rho_{\mathrm{st}}$. Furthermore $K$ is positive definite, because $\frac{\delta^2 K[\rho;\rho_{\mathrm{st}}]}{\delta\rho(\boldsymbol{x})\delta\rho(\boldsymbol{y})}|_{\rho=\rho_{\mathrm{st}}} = \delta_{\boldsymbol{x},\boldsymbol{y}} \frac{1}{\rho_{\mathrm{st}}(\boldsymbol{x})} \geq 0$. For $K$ to be a Lyapunov function, it remains to show that its total time derivative is smaller or equal to zero:

$$\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}t} K[\rho(t)\,;\rho_{\mathrm{st}}] =& \sum_{\boldsymbol{x}\in\Omega} \dot{\rho}(\boldsymbol{x};t) \ln \frac{\rho(\boldsymbol{x};t)}{\rho_{\mathrm{st}}(\boldsymbol{x})} + \frac{\mathrm{d}}{\mathrm{d}t} \sum_{\boldsymbol{x}\in\Omega} \rho(\boldsymbol{x};t) \\
=& \sum_{\boldsymbol{x},\boldsymbol{y}\in\Omega} \ln \frac{\rho(\boldsymbol{x})}{\rho_{\mathrm{st}}(\boldsymbol{x})} \left[\lambda(\boldsymbol{x};\boldsymbol{y})\,\rho(\boldsymbol{y}) - \lambda(\boldsymbol{y};\boldsymbol{x})\,\rho(\boldsymbol{x})\right] + \frac{\mathrm{d}}{\mathrm{d}t} 1 \\
=& \sum_{\boldsymbol{x},\boldsymbol{y}\in\Omega} \lambda(\boldsymbol{x};\boldsymbol{y})\,\rho(\boldsymbol{y}) \ln \frac{\rho(\boldsymbol{x})\,\rho_{\mathrm{st}}(\boldsymbol{y})}{\rho(\boldsymbol{y})\,\rho_{\mathrm{st}}(\boldsymbol{x})} \\
=& \sum_{\boldsymbol{x},\boldsymbol{y}\in\Omega} \left[\lambda(\boldsymbol{x};\boldsymbol{y}) - \delta_{\boldsymbol{x},\boldsymbol{y}} \sum_{\boldsymbol{z}\in\Omega} \lambda(\boldsymbol{z};\boldsymbol{x})\right] \rho(\boldsymbol{y}) \ln \frac{\rho(\boldsymbol{x})\,\rho_{\mathrm{st}}(\boldsymbol{y})}{\rho(\boldsymbol{y})\,\rho_{\mathrm{st}}(\boldsymbol{x})} \\
\leq& \sum_{\boldsymbol{x},\boldsymbol{y}\in\Omega} \left[\lambda(\boldsymbol{x};\boldsymbol{y}) - \delta_{\boldsymbol{x},\boldsymbol{y}} \sum_{\boldsymbol{z}\in\Omega} \lambda(\boldsymbol{z};\boldsymbol{x})\right] \rho(\boldsymbol{y}) \left[\frac{\rho(\boldsymbol{x})\,\rho_{\mathrm{st}}(\boldsymbol{y})}{\rho(\boldsymbol{y})\,\rho_{\mathrm{st}}(\boldsymbol{x})} - 1\right] \\
=& \sum_{\boldsymbol{x}\in\Omega} \frac{\rho(\boldsymbol{x})}{\rho_{\mathrm{st}}(\boldsymbol{x})} \triangle_\lambda \rho_{\mathrm{st}}(\boldsymbol{x}) - \sum_{\boldsymbol{y}\in\Omega} \rho(\boldsymbol{y}) \sum_{\boldsymbol{x}\in\Omega} \lambda(\boldsymbol{x};\boldsymbol{y}) + \sum_{\boldsymbol{y}\in\Omega} \rho(\boldsymbol{y}) \sum_{\boldsymbol{z}\in\Omega} \lambda(\boldsymbol{z};\boldsymbol{y}) \\
=& 0
\end{aligned}$$

In the fourth line we added a term on the diagonal. This is allowed, since all summands on the diagonal $\boldsymbol{x} = \boldsymbol{y}$ vanish. Hence, The Kullback-Leibler divergence is a proper Lyapunov function and consequently, $\rho(\boldsymbol{x};t) \to \rho_{\mathrm{st}}$, as $t \to \infty$. The stationary state is a stable attractor.

Last but not least, the stationary state is not a thermodynamic equilibrium, when there are unbalanced microscopical fluxes, i.e.:

$$\lambda(\boldsymbol{y};\boldsymbol{x})\,\rho_{\mathrm{st}}(\boldsymbol{x}) = \lambda(\boldsymbol{x};\boldsymbol{y})\,\rho_{\mathrm{st}}(\boldsymbol{y})\,.$$

To fulfil this detailed balance condition, we may require that $\lambda(\boldsymbol{y};\boldsymbol{x}) > 0 \Leftrightarrow \lambda(\boldsymbol{x};\boldsymbol{y}) > 0$. In that case we can identify the stationary state $\rho_{\mathrm{st}}$ with an equilibrium state $\rho_{\mathrm{eq}}$.

## A.3. Analogy with electrical and elastic networks

Consider an electric network, that is a set of nodes $\boldsymbol{x}, \boldsymbol{y} \in \Omega$ that are connected to each other by some conductances $\sigma(\boldsymbol{y},\boldsymbol{x})$ and to the ground with a capacitance $c(\boldsymbol{x})$. Let

$u(\boldsymbol{x};t)$ denote the electric potential at node $\boldsymbol{x}$ and epoch $t$, then $u$ behaves according to

$$c(\boldsymbol{x})\,\dot{u}(\boldsymbol{x};t) = \sum_{\boldsymbol{y}} A(\boldsymbol{y},\boldsymbol{x})\,\sigma(\boldsymbol{y},\boldsymbol{x})\,[u(\boldsymbol{y};t) - u(\boldsymbol{x};t)]\,. \tag{A.6}$$

The charge $q(\boldsymbol{x};t) := c(\boldsymbol{x})\,u(\boldsymbol{x};t)$ on the other hand, behaves according to

$$\dot{q}(\boldsymbol{x};t) = \sum_{\boldsymbol{y}} A(\boldsymbol{y},\boldsymbol{x})\,\sigma(\boldsymbol{y},\boldsymbol{x})\left[\frac{q(\boldsymbol{y};t)}{c(\boldsymbol{y})} - \frac{q(\boldsymbol{x};t)}{c(\boldsymbol{x})}\right]. \tag{A.7}$$

The last equations can be identified with Eq.s(2.16) and (2.25) of the main text, when taking $u = \eta$, $q = \rho$, $w = \sigma$ and $c$ as the Boltzmann factors. As was pointed out in the supplementary material to [45], the identification of potential and activities is only an approximation (more correctly we have to identify the chemical and electrical potential). This leads to a different behaviour of both models when the concentration gradients are very large, and $\mu(\boldsymbol{y}) - \mu(\boldsymbol{x}) = \log\eta(\boldsymbol{y}) - \log\eta(\boldsymbol{x}) \approx \eta(\boldsymbol{y}) - \eta(\boldsymbol{x})$ is not a good approximation any more.

Master equation operators admitting an decomposition into a product of a symmetric and a diagonal operator also occur in elastic networks. Here, each node $\boldsymbol{x}$ of the graph is identified with a point mass $m(\boldsymbol{x})$. The links are identified with ideal springs connecting the masses. When the displacement of the mass at $\boldsymbol{x}$ at time $t$ is denoted by $\boldsymbol{r}(\boldsymbol{x};t)$, and a damping term is included, we obtain the following equation of motion:

$$m(\boldsymbol{x})\,\ddot{\boldsymbol{r}}(\boldsymbol{x};t) + \gamma\dot{\boldsymbol{r}}(\boldsymbol{x};t) = \sum_{\boldsymbol{y}} A(\boldsymbol{y},\boldsymbol{x})\,k(\boldsymbol{y},\boldsymbol{x})\,[\boldsymbol{r}(\boldsymbol{y};t) - \boldsymbol{r}(\boldsymbol{x};t)]\,.$$

We can identify $\mathcal{B}$ with the masses and $\triangle_{A,w}$ with the operator involving the spring constants $k$. When the damping is either large or small, one of the terms on the left hand side of the equation can be neglected. In the purely elastic limit ($\gamma = 0$), we recover the same operator $\mathcal{B}\triangle_{A,w}$ that governs the evolution of the activities. The dynamics of the displacements however is vastly different from the activities' dynamics, because they belong to different function classes ($\boldsymbol{r}$ can be negative and is not bounded) and their equations involve a different time derivative. On the other hand, the evolution operators share the same spectral properties.

## A.4. Effective medium approximation

In this section we explain Eq.(2.34) of the main text. The original idea is due to Bruggemann, [156], and is presented nicely in [89]. The latter article presented two approaches that both are reproduced here. The second approach is also inspired by [157].

## A.4.1. Electrical network approach

One way to tackle the effective medium approximation is to interpret the diffusion problem as a random resistor network, see Eq.(A.6) with $c(\boldsymbol{x}) = 1$. Then, finding the effective diffusivity is equivalent to finding the effective conductivity of the network. The main idea is to consider a reference network with fixed conductances $\sigma^*$. Then the conductance of one link is replaced with the original one $\sigma$ and after that, the change of the potential drop along the link is calculated. $\sigma^*$ is chosen such, that the average change in the potential drop (when averaged with respect to the random conductances) is zero.

Let us be more specific and isolate two nodes along a link of the reference network. The total resistance between both nodes in the effective medium is given by $R^*$. We replace the effective medium conductance $\sigma^*$ along the bond with the original one given by $\sigma$. The total conductance between the nodes changes from $(R^*)^{-1}$ to $(R^*)^{-1} - \sigma^* + \sigma$. This changes the original potential difference along the link from its value in the effective medium, $\Delta u^*$, to a new value $\Delta u$.

Let us introduce a current $i_0$ at one end of the link and extract it at the other end. The strength of the current is chosen such, that it compensates the change in the voltage drop. I.e., we apply

$$i_0 = \Delta u^* \left(\sigma^* - \sigma\right).$$

By knowing $i_0$, we can compute the change of the voltage drop along the link using the total resistance of the altered network

$$\Delta u - \Delta u^* = \frac{i_0}{(R^*)^{-1} - \sigma^* + \sigma} = \Delta u^* \frac{\sigma^* - \sigma}{(R^*)^{-1} - \sigma^* + \sigma}.$$

The effective medium condition states that this difference in voltages vanishes on the average regardless of the original voltage drop $\Delta u^*$. Averaging over the random bond conductances and identifying them with the transition rates (i.e. $\sigma = Aw$ and $\sigma^* = A^*K^*/D^*$) gives

$$0 \stackrel{!}{=} \mathbb{E}\left[\frac{\frac{A^*K^*}{D^*} - Aw}{(R^*)^{-1} - \frac{A^*K^*}{D^*} + Aw}\right], \tag{A.8}$$

where the average is performed with respect to the joint distribution of $A$ and $w$.

## A.4.2. Operator approach

**Pseudo-diagonalization:** Another approach is to compare the propagator or resolvent of the original and the reference graph. We compare $\triangle_{A,w}$ to some reference operator $K^*\triangle^*$, which is at first sight arbitrary. The reference operator is given by Eq.(2.31) of the main text and a first observation is that it can be brought into a

pseudo-diagonal form. We now use a quantum-mechanical notation, $\langle \boldsymbol{x}|\rho\rangle := \rho(\boldsymbol{x})$, abbreviate the links[2] with $\boldsymbol{\xi} := (\boldsymbol{x}, \boldsymbol{y}) \in T(\Omega)$ and define a link state via:

$$|(\boldsymbol{x}, \boldsymbol{y})\rangle := \frac{1}{\sqrt{2}} \left[|\boldsymbol{x}\rangle - |\boldsymbol{y}\rangle\right]. \tag{A.9}$$

Using the link states we write

$$\triangle_{A,w} = \sum_{\boldsymbol{x},\boldsymbol{y}\in\Omega} A(\boldsymbol{x}, \boldsymbol{y})\, w(\boldsymbol{x}, \boldsymbol{y}) \left[|\boldsymbol{x}\rangle\langle\boldsymbol{y}| - |\boldsymbol{x}\rangle\langle\boldsymbol{x}|\right] = \sum_{\boldsymbol{\xi}\in T(\Omega)} \left[-2A(\boldsymbol{\xi})\, w(\boldsymbol{\xi})\right] |\boldsymbol{\xi}\rangle\langle\boldsymbol{\xi}|\,.$$

It is important to note two things. First, we count every link only once; since $\triangle_{A,w}$ is symmetric, we have $(\boldsymbol{x}, \boldsymbol{y}) = (\boldsymbol{y}, \boldsymbol{x})$. Otherwise the second equality can not be derived. Second, this is not a diagonalization in the sense of a spectral decomposition because the link states do not represent a decomposition of unity.

This representation is possible for any symmetric graph Laplacian, in particular for the reference Laplacian:

$$K^*\triangle^* = \sum_{\boldsymbol{\xi}\in T(\Omega)} \left[-2\frac{A^*(\boldsymbol{\xi})\, K^*}{D^*(\boldsymbol{\xi})}\right] |\boldsymbol{\xi}\rangle\langle\boldsymbol{\xi}|\,.$$

When the difference between the Laplacian and the reference operator is called $\delta_{A,w} := \triangle_{A,w} - K^*\triangle^*$, we have

$$\delta_{A,w} := \sum_{\boldsymbol{\xi}\in T(\Omega)} \left[2\frac{A^*(\boldsymbol{\xi})\, K^*}{D^*(\boldsymbol{\xi})} - 2A(\boldsymbol{\xi})\, w(\boldsymbol{\xi})\right] |\boldsymbol{\xi}\rangle\langle\boldsymbol{\xi}|\,. \tag{A.10}$$

**Resolvent:** The resolvent of $K^*\triangle^*$ is defined by:

$$\mathcal{G}^*(s) := \left[s - K^*\triangle^*\right]^{-1}.$$

Similarly we may define the resolvent $\mathcal{G}_{A,w}$ of $\triangle_{A,w}$. The kernel of the resolvent operator is the Laplace transform of the propagator. In the effective medium approximation, we express the resolvent of the disordered system, $\mathcal{G}_{A,w}(s)$, in terms of the effective medium resolvent. After doing so, an average over the disorder is applied and the reference operator is chosen such that it coincides with the disorder averaged resolvent up to a certain order. Let us start with expressing the disordered resolvent in terms of $\delta_{A,w}$ and $\mathcal{G}^*$ (we drop the argument $s$ for clarity).

$$\begin{aligned}\mathcal{G}_{A,w} :=& \left[s - \triangle_{A,w}\right]^{-1} = \left[s - \triangle^* - \delta_{A,w}\right]^{-1} = \left[\mathbb{1} - \mathcal{G}^*\delta_{A,w}\right]^{-1}\mathcal{G}^* \\ =& \mathcal{G}^* + \mathcal{G}^*\delta_{A,w}\mathcal{G}^* + \mathcal{G}^*\delta_{A,w}\mathcal{G}^*\delta_{A,w}\mathcal{G}^* + \ldots.\end{aligned}$$

The Neumann series is split up into a sum of alternating products of $\delta_{A,w}$ and $\mathcal{G}^*$. Remember that $\delta_{A,w}$ is diagonal in the projectors $|\boldsymbol{\xi}\rangle\langle\boldsymbol{\xi}|$, we thus reorder the series in

[2] Note that when the state space $\Omega$ is the $d$-dimensional simple cubic lattice $a\mathbb{Z}^d$, the set of all possible transitions $T(\Omega)$ can also be identified with $a\mathbb{Z}^d$.

terms of powers of $\langle\boldsymbol{\xi}|\,\delta_{A,w}\,|\boldsymbol{\xi}\rangle$. This can be done by splitting up $\mathcal{G}^*$ into a part $\mathcal{G}^*_\parallel$ that is diagonal in the link projectors and a part $\mathcal{G}^*_\perp$ that is not. We have $\mathcal{G}^* = \mathcal{G}^*_\parallel + \mathcal{G}^*_\perp$, and

$$\mathcal{G}^*_\parallel(s) := \sum_{\boldsymbol{\xi}\in T(\Omega)} |\boldsymbol{\xi}\rangle\,\langle\boldsymbol{\xi}|\,\mathcal{G}^*(s)\,|\boldsymbol{\xi}\rangle\,\langle\boldsymbol{\xi}|\,.$$

At last we obtain:

$$\mathcal{G}_{A,w} - \mathcal{G}^* = \mathcal{G}^*\frac{\mathbb{1}}{\mathbb{1}-\delta_{A,w}\mathcal{G}^*_\parallel}\delta_{A,w}\mathcal{G}^* + \mathcal{G}^*\left\{\sum_{m=0}\left[(\delta_{A,w}\mathcal{G}^*)^m - (\delta_{A,w}\mathcal{G}^*_\parallel)^m\right]\right\}\delta_{A,w}\mathcal{G}^*$$

The expression on the right hand side is the deviation of the quenched disorder resolvent (or propagator) from the reference one $\mathcal{G}^*$. The first order of deviation is given by the first summand and only involves powers of one single $\langle\boldsymbol{\xi}|\,\delta_{A,w}\,|\boldsymbol{\xi}\rangle$.

**Application of the average:** Let us apply the average over disorder. The self-consistent condition, according to which $\triangle^*$ has to be chosen is

$$0 \stackrel{!}{=} \mathbb{E}\left[\mathcal{G}^*\frac{\mathbb{1}}{\mathbb{1}-\delta_{A,w}\mathcal{G}^*_\parallel}\delta_{A,w}\mathcal{G}^*\right]$$

and states that the first order of the deviation vanishes.

If the bond-properties are independent of each other, $\langle\boldsymbol{\xi}|\,\delta_{A,w}\,|\boldsymbol{\xi}\rangle$ is stochastically independent of $\langle\boldsymbol{\xi}'|\,\delta_{A,w}\,|\boldsymbol{\xi}'\rangle$. Then every individual summand of the last equation has to vanish. We can fix one link $\boldsymbol{\xi} = (\boldsymbol{x},\boldsymbol{y})$ and apply the condition bond-wise:

$$0 \stackrel{!}{=} \mathbb{E}\left[\frac{\langle\boldsymbol{\xi}|\,\delta_{A,w}\,|\boldsymbol{\xi}\rangle}{1-\langle\boldsymbol{\xi}|\,\delta_{A,w}\,|\boldsymbol{\xi}\rangle\,\langle\boldsymbol{\xi}|\,\mathcal{G}^*_\parallel\,|\boldsymbol{\xi}\rangle}\right].$$

The remaining matrix element of the resolvent is related to the frequency dependent resistance distance of the reference graph [141]:

$$\mathrm{d}^*_{\mathrm{res}}(\boldsymbol{x},\boldsymbol{y};s) := 2K^*\,\langle(\boldsymbol{x},\boldsymbol{y})|\,\mathcal{G}^*(sK^*)\,|(\boldsymbol{x},\boldsymbol{y})\rangle\,. \tag{A.11}$$

In the stationary limit, $s\to 0$, it is proportional to the total resistance $R^*$ between the two nodes:[3]

$$R^*(\boldsymbol{x},\boldsymbol{y}) := \langle\boldsymbol{x}|\,\mathcal{G}^*(0)\,|\boldsymbol{x}\rangle + \langle\boldsymbol{y}|\,\mathcal{G}^*(0)\,|\boldsymbol{y}\rangle - 2\,\langle\boldsymbol{x}|\,\mathcal{G}^*(0)\,|\boldsymbol{y}\rangle = \frac{\mathrm{d}^*_{\mathrm{res}}(\boldsymbol{x},\boldsymbol{y};0)}{K^*}. \tag{A.12}$$

The limit $s\to 0$ corresponds to measuring the effective diffusivity in the stationary state. Plugging Eq.s (A.11) and (A.10) into the link-wise condition, taking the limit $s\to 0$ and multiplying by $R^*/2$ bears:

$$0 \stackrel{!}{=} \mathbb{E}\left[\frac{\frac{A^*K^*}{D^*} - Aw}{\frac{K^*}{\mathrm{d}^*_{\mathrm{res}}} - \frac{A^*K^*}{D^*} + Aw}\right] = \mathbb{E}\left[\frac{A^*K^* - AwD^*}{\left(\frac{D^*}{\mathrm{d}^*_{\mathrm{res}}} - A^*\right)K^* + AwD^*}\right]. \tag{A.13}$$

[3] Taking the limit $s\to 0$ requires inverting the singular operator $\triangle^*$. This is only possible, when the inverse is properly interpreted, for instance in the sense of a Moore-Penrose inverse.

The last equation is exactly the same that we derived by means of the electrical analogy. The reader may keep in mind that all quantities here are link dependent, although the argument $(\boldsymbol{x}, \boldsymbol{y})$ has been omitted for clarity. Hence all quantities are also spatially dependent. This implicit equation has to be solved for $A^*(\boldsymbol{x}, \boldsymbol{y})$, $D^*(\boldsymbol{x}, \boldsymbol{y})$ and $K^*$. This will be discussed in the next section.

### A.4.3. Finding the effective medium quantities

The average in Eq.(A.13), has to be performed with respect to the joint probability of the adjacency matrix and the equilibrium rates. Let us consider first the adjacency matrix. Let's say, we have a link $\boldsymbol{\xi}$ that does not exist in the reference graph but *could exists* in the original graph, i.e. $A^* = 0$, but possibly $A = 1$. This would require that $\mathbb{E}[AwD^*/(AwD^* + (K^*D^*/\mathrm{d}^*_{\mathrm{res}})]$ vanishes, which is only possible when $A = 0$ almost surely. Hence, in this particular case, the effective medium requirement can not be met. Consequently, we have to require that $A^* = 1$, wherever $A$ could be unity. The reference transition graph *necessarily* has a link, wherever the original graph *possibly* has a link. The probability that there is a link along $(\boldsymbol{x}, \boldsymbol{y})$ in the original graph is denoted with $p(\boldsymbol{x}, \boldsymbol{y})$.

Next, consider Eq.(A.13), for a link that exists in the in the reference graph, i.e. $A^*(\boldsymbol{x}, \boldsymbol{y}) = 1$. The average over $A$ can be computed by using $p$. The remaining function is a continuous, bounded, and monotonously decaying rational function of $w$. We may thus apply the mean value theorem of integration, which ensures the existence of a constant $w'(\boldsymbol{x}, \boldsymbol{y})$, such that:

$$
\begin{aligned}
0 =& \mathbb{E}\left[\frac{A^*K^* - AwD^*}{\left(\frac{D^*}{\mathrm{d}^*_{\mathrm{res}}} - A^*\right)K^* + AwD^*}\right] = \quad \frac{1-p}{\frac{D^*}{\mathrm{d}^*_{\mathrm{res}}} - 1} + p\mathbb{E}\left[\frac{K^* - wD^*}{wD^* + K^*\left(\frac{D^*}{\mathrm{d}^*_{\mathrm{res}}} - 1\right)}\right] \\
=& \frac{1-p}{\frac{D^*}{\mathrm{d}^*_{\mathrm{res}}} - 1} + p\frac{K^* - w'D^*}{w'D^* + K^*\left(\frac{D^*}{\mathrm{d}^*_{\mathrm{res}}} - 1\right)}
\end{aligned}
$$

In terms of this constant, we can express the diffusivity as:

$$
K^* = w'(\boldsymbol{x}, \boldsymbol{y})\, D^*(\boldsymbol{x}, \boldsymbol{y})\, \frac{\mathrm{d}^*_{\mathrm{res}}(\boldsymbol{x}, \boldsymbol{y}) - p(\boldsymbol{x}, \boldsymbol{y})\, D^*(\boldsymbol{x}, \boldsymbol{y})}{\mathrm{d}^*_{\mathrm{res}}(\boldsymbol{x}, \boldsymbol{y}) - D^*(\boldsymbol{x}, \boldsymbol{y})}. \tag{A.14}
$$

Since $K^*$ must not depend on the link $(\boldsymbol{x}, \boldsymbol{y})$, $D^*$ must be chosen to compensate the joint effects of $w'$ and $\mathrm{d}^*_{\mathrm{res}}$. The resistance distance $\mathrm{d}^*_{\mathrm{res}}$ is itself a functional of $D^*$. Therefore Eq.(A.11) poses another self-consistency condition. This influence is only short-range, because the fraction in Eq.(A.14) approaches a constant. The reason is that $D^*(\boldsymbol{x}, \boldsymbol{y})$ constitutes the resistance on the link $(\boldsymbol{x}, \boldsymbol{y})$, but $\mathrm{d}^*_{\mathrm{res}}(\boldsymbol{x}, \boldsymbol{y})$ is the total resistance between the nodes $\boldsymbol{x}$ and $\boldsymbol{y}$ in the reference graph. Since there are in general many connections parallel to the direct link $(\boldsymbol{x}, \boldsymbol{y})$, it is safe to assume that $\mathrm{d}^*_{\mathrm{res}}$ is smaller than $D^*$. Furthermore, the resistive distance distance should grow slower with $|\boldsymbol{x} - \boldsymbol{y}|$ than $D^*$. Hence, for large distances, the terms proportional to $\mathrm{d}^*_{\mathrm{res}}$ can safely be neglected and the fraction approaches $p(\boldsymbol{x}, \boldsymbol{y})$.

Consequently, for $K^*$ to be finite, $D^*$ has to compensate for the distance dependence of $w'(\boldsymbol{x},\boldsymbol{y})\,p(\boldsymbol{x},\boldsymbol{y})$. When this is not the case – which indicates that we chose the wrong reference network – we put either $K^*=\infty$ or $K^*=0$. We put the diffusivity to zero, when $w'pD^*$ decays and we put it to infinity when $w'pD^*$ diverges. Eq.(A.14) and Eq.(A.11) determine the functional form of $D^*$.

Last but not least, note the following: When $f_{(\boldsymbol{x},\boldsymbol{y})}(A,w)$ denotes the joint pdf of $A(\boldsymbol{x},\boldsymbol{y})$ and $w(\boldsymbol{x},\boldsymbol{y})$, and it has the form

$$\begin{aligned} f_{(\boldsymbol{x},\boldsymbol{y})}(A,w)\,\mathrm{d}A\mathrm{d}w =& (1-p(\boldsymbol{x},\boldsymbol{y}))\,\delta(A)\,\delta(w)\,\mathrm{d}A\mathrm{d}w \\ &+ p(\boldsymbol{x},\boldsymbol{y})\,\delta(A-1)\,\bar{f}(wD(\boldsymbol{x},\boldsymbol{y}))\,D(\boldsymbol{x},\boldsymbol{y})\,\mathrm{d}A\mathrm{d}w, \end{aligned}$$

then the constant $w'(\boldsymbol{x},\boldsymbol{y})$ in Eq.(A.14) is $\bar{w}'/D(\boldsymbol{x},\boldsymbol{y})$ and the "correct" distance factor is $D^*(\boldsymbol{x},\boldsymbol{y})=D(\boldsymbol{x},\boldsymbol{y})\,/p(\boldsymbol{x},\boldsymbol{y})$. Hence, for $K^*$ to be finite, the distance factor $D^*$ has to be chosen as the "stochastic" distance dependence of the rates.

Since $A$, $w$ and $D^*$ together determine the value of the effective medium diffusivity, we denote the solution to equation (A.13) suggestively with

$$K^* = \mathbb{E}[AwD^*]_{\mathrm{EM}} \tag{A.15}$$

This also reflects the correct physical dimension of $[D^*]$/sec.

## A.4.4. The coefficient of normal diffusion

When does the reference model exhibit normal diffusion? The answer is easy to give, when one repeats the procedure applied for the binomial random walk. Consider the master equation:

$$\dot{\rho}(\boldsymbol{x};t) = K^*\triangle^*\rho(\boldsymbol{x};t)\,.$$

The time derivative of the msd is given by:

$$\begin{aligned} \frac{\mathrm{d}}{\mathrm{d}t}\left\langle \boldsymbol{X}^2(t)\right\rangle =& \frac{\mathrm{d}}{\mathrm{d}t}\sum_{\boldsymbol{x}\in a\mathbb{Z}^d}\boldsymbol{x}^2\rho(\boldsymbol{x};t) = \sum_{\boldsymbol{x}\in a\mathbb{Z}^d}\boldsymbol{x}^2\dot{\rho}(\boldsymbol{x},t) \\ =& K^*\sum_{\boldsymbol{\xi}\in T(\Omega)}\sum_{\boldsymbol{x}\in\Omega}\boldsymbol{x}^2\frac{A^*(\boldsymbol{x},\boldsymbol{x}+\boldsymbol{\xi})}{D^*(\boldsymbol{x},\boldsymbol{x}+\boldsymbol{\xi})}\left[\rho(\boldsymbol{x}+\boldsymbol{\xi};t)-\rho(\boldsymbol{x};t)\right] \\ =& K^*\sum_{\boldsymbol{x}\in\Omega}\rho(\boldsymbol{x};t)\sum_{\boldsymbol{\xi}\in T(\Omega)}\frac{A^*(\boldsymbol{x},\boldsymbol{x}+\boldsymbol{\xi})}{D^*(\boldsymbol{x},\boldsymbol{x}+\boldsymbol{\xi})}\left[(\boldsymbol{x}-\boldsymbol{\xi})^2-\boldsymbol{x}^2\right] \\ =& K^*\sum_{\boldsymbol{x}\in\Omega}\rho(\boldsymbol{x};t)\sum_{\boldsymbol{\xi}\in T(\Omega)}\frac{A^*(\boldsymbol{x},\boldsymbol{x}+\boldsymbol{\xi})}{D^*(\boldsymbol{x},\boldsymbol{x}+\boldsymbol{\xi})}\left[\boldsymbol{\xi}^2-2\boldsymbol{x}\boldsymbol{\xi}\right] \\ =& K^*\left[\sum_{\boldsymbol{x}\in\Omega}\rho(\boldsymbol{x};t)\right]\left\{\sum_{\boldsymbol{\xi}\in T(\Omega)}\frac{A^*(|\boldsymbol{\xi}|)}{D^*(|\boldsymbol{\xi}|)}\boldsymbol{\xi}^2\right\} = K^*C_{\triangle^*}. \end{aligned}$$

In the last line, we identified the connection factor. We assumed that the reference model is translationally invariant, i.e. $A^*(\boldsymbol{x},\boldsymbol{y})$ and $D^*(\boldsymbol{x},\boldsymbol{y})$ are only functions of the difference $\boldsymbol{x}-\boldsymbol{y}$. Additionally we assumed that it is isotropic, such that $A^*$ and $D^*$ are only functions of $|\boldsymbol{\xi}|$. That way, the $2\boldsymbol{x}\boldsymbol{\xi}$ term vanishes and the series over $\boldsymbol{\xi}$ becomes independent of $\boldsymbol{x}$. This ensures that the connection factor is well-defined.

## A.5. Projection in state space

The following derivation is closely related to the one by Haus and Kehr in [40, especially sec. 5 ], but essentially a reformulation of [152] in Laplace space. Again we start from Eq.(2.16). Suppose there is a projection operator $\mathcal{Q}$, and set $\overline{\rho} := \mathcal{Q}\rho$, as well as $\widetilde{\rho} := (\mathbb{1}-\mathcal{Q})\rho$. Writing Eq.(2.16) in Laplace domain and applying the projector and its orthogonal bears:

$$\begin{cases} s\overline{\rho}(s) - \overline{\rho_0} = \overline{\triangle_\lambda}\overline{\rho}(s) + \mathcal{Q}\triangle_\lambda(\mathbb{1}-\mathcal{Q})\widetilde{\rho}(s) \\ s\widetilde{\rho}(s) - \widetilde{\rho_0} = \widetilde{\triangle_\lambda}\widetilde{\rho}(s) + (\mathbb{1}-\mathcal{Q})\triangle_\lambda\mathcal{Q}\overline{\rho}(s) \end{cases}.$$

Here we put $\overline{\triangle_\lambda} := \mathcal{Q}\triangle_\lambda\mathcal{Q}$ and $\widetilde{\triangle_\lambda} := (\mathbb{1}-\mathcal{Q})\triangle_\lambda(\mathbb{1}-\mathcal{Q})$. Both equations are a simple reordering of the original one. When $(s-\widetilde{\triangle_\lambda})$ is not singular on the subspace generated by $\mathbb{1}-\mathcal{Q}$, there is an inverse in this subspace, which we denote by: $(\mathbb{1}-\mathcal{Q})/(s-\widetilde{\triangle_\lambda})$. (Note that this is the case when $\mathrm{Re}(s) > 0$, because $\triangle_\lambda$ has one eigenvalue zero and otherwise negative eigenvalues. Assuming that the equilibrium state lies in the range of $\mathcal{Q}$, it also holds for $\mathrm{Re}(s) = 0$, and hence in the whole relevant domain.) In this case, the second equation can be solved for $\widetilde{\rho}(s)$. The result is

$$\widetilde{\rho}(s) = \frac{\mathbb{1}-\mathcal{Q}}{s-\widetilde{\triangle_\lambda}}\{\widetilde{\rho_0} + (\mathbb{1}-\mathcal{Q})\triangle_\lambda\mathcal{Q}\overline{\rho}\}$$

and can be plugged into the other equation. The final result is

$$s\overline{\rho}(s) - \overline{\rho_0} = \left\{\overline{\triangle_\lambda} + \mathcal{Q}\triangle_\lambda\frac{\mathbb{1}-\mathcal{Q}}{s-\widetilde{\triangle_\lambda}}\triangle_\lambda\mathcal{Q}\right\}\overline{\rho} + \mathcal{Q}\triangle_\lambda\frac{\mathbb{1}-\mathcal{Q}}{s-\widetilde{\triangle_\lambda}}\widetilde{\rho_0}. \tag{A.16}$$

where we have a coarse grained Laplacian $\overline{\triangle_\lambda}$ and an additional frequency dependent part

$$\mathcal{M}_{\triangle_\lambda}(s) := \mathcal{Q}\triangle_\lambda\frac{\mathbb{1}-\mathcal{Q}}{s-\widetilde{\triangle_\lambda}}\triangle_\lambda\mathcal{Q} \tag{A.17}$$

that depends on $s$, hence acts as a convolution operator in the time domain. Furthermore a source term

$$\mathcal{I}_{\triangle_\lambda}[\widetilde{\rho_0};s] := \mathcal{Q}\frac{\mathbb{1}-\mathcal{Q}}{s-\widetilde{\triangle_\lambda}}\widetilde{\rho_0} \tag{A.18}$$

appears. Reverting the Laplace transform bears Eq.(2.36) of the main text.

# B. Bibliography


[1] Roberto Maiocchi. The case of brownian motion. *The British Journal for the History of Science*, 23(3):257–283, September 1990.

[2] Derek Abbott, Bruce R. Davis, Nicholas J. Phillips, and Kamran Eshraghian. Simple derivation of the thermal noise formula using window-limited fourier transforms and other conundrums. *IEEE Transactions on education*, 39(1):1–13, February 1996.

[3] Igor M. Sokolov. Models of anomalous diffusion in crowded environments. *Soft Matter*, 8:9043–9052, 2012.

[4] Adolf Fick. Ueber diffusion. *Annalen der Physik*, 170(1):59–86, 1855.

[5] Albert Einstein. Über die von der molekularkinetischen theorie der wärme geforderte bewegung von in ruhenden flüssigkeiten suspendierten teilchen. *Annalen der Physik*, 322(8):549–560, 1905.

[6] Marian von Smoluchowski. Zur kinetischen theorie der brownschen molekularbewegung und der suspensionen. *Annalen der Physik*, 326(14):756–780, 1906.

[7] Paul Langevin. Sur la théorie du mouvement brownien [on the theory of brownian motion]. *C. R. Acad. Sci. (Paris)*, 146:530–533, 1908. reviewed by D. S. Lemons & A. Gythiel: Paul Langevins 1908 paper "On the Theory of Brownian Motion" [...], Am. J. Phys. 65, 1079 (1997).

[8] Jean Baptiste Perrin. L'agitation moleculaire et le mouvement brownien. *Comptes Rendus. Academie des Sciences (Paris)*, 146:967–970, 1908.

[9] W. Schottky. Über spontane stromschwankungen in verschiedenen electrizitätsleitern. *Annalen der Physik*, 57:541–567, 1918.

[10] L. Bachelier. Théorie de la spéculation. *Annales Scientifiques de lÈcole Normale Supérieure*, 3(17):21–86, 1900.

[11] D. Brockmann, L. Hufnagel, and T. Geisel. The scaling laws of human travel. *Nature*, 439:462–465, January 2006.

[12] Marta C. González, César A. Hidalgo, and Albert-László Barabási. Understanding individual human mobility patterns. *Nature*, 453(7196):779–782, June 2008.

[13] Vitaly Belik, Theo Geisel, and Dirk Brockmann. Natural human mobility patterns and spatial spread of infectious diseases. *Physical Review X*, 1:011001, Aug 2011.

[14] L. Hufnagel, D. Brockmann, and T. Geisel. Forecast and control of epidemics in a globalized world. *Proceedings of the National Academy of Sciences of the United States of America*, 101(42):15124–15129, 2004.

[15] Frederic Bartumeus and Simon A. Levin. Fractal reorientation clocks: Linking animal behavior to statistical patterns of search. *Proceedings of the National Academy of Sciences*, 105(49):19072–19077, 2008.

[16] O. Bénichou, C. Loverdo, M. Moreau, and Voituriez. Intermittent search strategies. *Reviews of Modern Physics*, 83:81–129, Mar 2011.

[17] Rachel R. Bennett and Ramin Golestanian. Chlamydomonas with intrinsic noise. *Physical Review Letters*, 110:148102, Apr 2013.

[18] Edward A Codling, Michael J Plank, and Simon Benhamou. Random walk models in biology. *Journal of The Royal Society Interface*, 5(25):813–834, 2008.

[19] Monique de Jager, Franz J. Weissing, Peter M. J. Herman, Bart A. Nolet, and Johan van de Koppel. Lévy walks evolve through interaction between movement and environmental complexity. *Science*, 332(6037):1551–1553, 2011.

[20] Andrew M. Edwards, Richard A. Phillips, Nicholas W. Watkins, Mervyn P. Freeman, Eugene J. Murphy, Vsevolod Afanasyev, Sergey V. Buldyrev, M. G. E. da Luz, E. P. Raposo, H. Eugene Stanley, and Gandhimohan M. Viswanathan. Revisiting lévy flight search patterns of wandering albatrosses, bumblebees and deer. *Nature*, 449:1044–1049, October 2007.

[21] Tajie H. Harris, Edward J. Banigan, David A. Christian, Christoph Konradt, Elia D. Tait Wojno, Kazumi Norose, Emma H. Wilson, Beena John, Wolfgang Weninger, Andrew D. Luster, Andrea J. Liu, and Christopher A. Hunter. Generalized lévy walks and the role of chemokines in migration of effector cd8+ t cells. *Nature*, 486(7404):545–548, June 2012.

[22] Nicolas E. Huphries, Nuno Queiroz, Jennifer R. M. Dyer, Nicolas G. Payde, Michael K. Musyl, Kurt M. Schaefer, Daniel W. Fuller, Juerg M. Brunnschweiler, Thomas K. Doyle, Jonathan D. R. Houghton, Graeme C. Hays, Catherine S. Jones, Leslie R. Noble, Victoria J. Wearmouth, Emiliy J. Southall, and David W. Sims. Lévy and brownian movement patters of marine predators. *Nature*, 465(7301):1066–1069, 2010.

[23] Friedrich Lenz, Thomas C. Ings, Lars Chittka, Aleksei V. Chechkin, and Rainer Klages. Spatiotemporal dynamics of bumblebees foraging under predation risk. *Physical Review Letters*, 108:098103, Mar 2012.

[24] Michael A. Lomholt, Koren Tal, Ralf Metzler, and Joseph Klafter. Lévy strategies in intermittent search processes are advantageous. *Proceedings of the National Academy of Sciences*, 105(32):11055–11059, 2008.

[25] Alla Mashanova, Tom H. Oliver, and Vincent A. A. Jansen. Evidence for intermittency and a truncated power law from highly resolved aphid movement data. *Journal of The Royal Society Interface*, 7(42):199–208, 2010.

[26] F. Matthäus, M. Jagodič, and J. Dobnikar. E. coli superdiffusion and chemotaxis – search strategy, precision and motility. *Biophysical Journal*, 97:946–957, August 2009.

[27] Franziska Matthäus, Mario S. Mommer, Tine Curk, and Jure Dobnikar. On the origin and characteristics of noise-induced lévy walks of *E. Coli*. *PLoS ONE*, 6(4):e18623, 04 2011.

[28] Dan V. Nicolau Jr., Judith P. Armitage, and Philip K. Maini. Directional persistence and the optimality of run-and-tumble chemotaxis. *Computational Biology and Chemistry*, 33(4):269–274, 2009.

[29] H. G. Othmer, S. R. Dunbar, and W. Alt. Models of dispersal in biological systems. *Journal of Mathematical Biology*, 26(3):263–298, 1988.

[30] E. P. Raposo, S. V. Buldyrev, M. G. E. da Luz, G. M. Viswanathan, and H. E. Stanley. Lévy flights and random searches. *Journal of Physics A: Mathematical and Theoretical*, 42(43):434003, 2009.

[31] Andy M. Reynolds. Lévy flight patterns are predicted to be an emergent property of a bumblebees' foraging strategy. *Behavioral Ecology and Sociobiology*, 64(1):19–23, 2009.

[32] Andreé Visser and Uffe Hogsbro Thygesen. Random motility of plankton: diffusive and aggregative contributions. *Journal of Plankton Research*, 25(9):1157–1168, September 2003.

[33] Gandimohan M. Viswanathan. Fish in lévy-flight foraging. *Nature*, 465(7301):1018–1019, June 2010.

[34] Taeseok Daniel Yang, Jin-Sung Park, Youngwoon Choi, Wonshik Choi, Tae-Wook Ko, and Kyoung J. Lee. Zigzag turning preference of freely crawling cells. *PLoS ONE*, 6(6):e20255, 06 2011.

[35] Marcus G. Daniels, J. Doyne Farmer, László Gillemot, Giulia Iori, and Eric Smith. Quantitative model of price diffusion and market friction based on trading as a mechanistic random process. *Physical Review Letters*, 90:108102, Mar 2003.

[36] Hui Gu, Jin-Rong Liang, and Yun-Xiu Zhang. Time-changed geometric fractional brownian motion and option pricing with transaction costs. *Physica A: Statistical Mechanics and its Applications*, 391(15):3971 – 3977, 2012.

[37] I. Mastromatteo, B. Tóth, and J.-P. Bouchaud. Anomalous impact in reaction-diffusion financial models. *Phys. Rev. Lett.*, 113:268701, Dec 2014.

[38] Mark M. Meerschaert and Enrico Scalas. Coupled continuous time random walks in finance. *Physica A: Statistical Mechanics and its Applications*, 370(1):114 – 118, May 2006. Econophysics Colloquium Proceedings of the International Conference International Conference.

[39] Vasiliki Plerou, Parameswaran Gopikrishnan, , Luís A. Nunes Amaral, Xavier Gabaix, and H. Eugene Stanley. Economic fluctuations and anomalous diffusion. *Physical Review E*, 62:R3023–R3026, Sep 2000.

[40] J.W. Haus and K.W. Kehr. Diffusion in regular and disordered lattices. *Physics Reports*, 150(56):263 – 406, 1987.

[41] Sebastian Bustingorry. Diffusion and percolation in anisotropic random barrier models. *Physical Review E*, 69:031107, March 2004.

[42] Shlomo Havlin and R. Nossal. Topological properties of percolation clusters. *Journal of Physics A: Mathematical and General*, 17(8):L427, 1984.

[43] S Havlin, B Trus, G H Weiss, and D Ben-Avraham. The chemical distance distribution in percolation clusters. *Journal of Physics A: Mathematical and General*, 18(5):L247, 1985.

[44] Filip J.R. Meysman, Volodymyr S. Malyuga, Bernard P. Boudreau, and Jack J. Middelburg. A generalized stochastic approach to particle dispersal in soils and sediments. *Geochimica et Cosmochimica Acta*, 72(14):3460 – 3478, 2008.

[45] F. Camboni and I. M. Sokolov. Normal and anomalous diffusion in random potential landscapes. *Physical Review E*, 85:050104, May 2012.

[46] F. Camboni, A. Koher, and Igor M. Sokolov. Diffusion of small particles in a solid polymeric medium. *Physical Review E*, 88:022120, August 2013.

[47] Markus Bier, René van Roij, Marjolein Dijkstra, and Paul van der Schoot. Self-diffusion of particles in complex fluids: Temporary cages and permanent barriers. *Phys. Rev. Lett.*, 101:215901, Nov 2008.

[48] S. Condamin, O. Bénichou, V. Tejedor, R. Voituriez, and Joseph Klafter. First-passage times in complex scale-invariant media. *Nature Letters*, 450(7166):77–80, November 2007.

[49] Gernot Guigas, Claudia Kalla, and Matthias Weiss. Probing the nanoscale viscoelasticity of intracellular fluids in living cells. *Biophysical Journal*, 93:316–323, July 2007.

[50] Yasmine Meroz, Iddo Eliazar, and Joseph Klafter. Facilitated diffusion in a crowded environment: from kinetics to stochastics. *Journal of Physics A: Mathematical and Theoretical*, 42(43):434012, 2009.

[51] Jedrzej Szymanski and Matthias Weiss. Elucidating the origin of anomalous diffusion in crowded fluids. *Physical Review Letters*, 103:038102, Jul 2009.

[52] Sebastian Bustingorry. Universality of ac conduction in anisotropic disordered systems: An effective-medium-approximation study. *Physical Review B*, 71:132201, April 2005.

[53] Jean-Philippe Bouchaud and Antoine Georges. Anomalous diffusion in disordered media: Statistical mechanisms, models and physical applications. *Physics Reports*, 195(45):127 – 293, 1990.

[54] J. W. Haus, K. W. Kehr, and J. W. Lyklema. Diffusion in a disordered medium. *Physical Review B*, 25:2905–2907, Feb 1982.

[55] Jonathan Machta. Generalized diffusion coefficient in one-dimensional random walks with static disorder. *Physical Review B*, 24:5260–5269, Nov 1981.

[56] Domingos H. U. Marchetti and Roberto da Silva. Brownian motion limit of random walks in symmetric non-homogeneous media. *Brazilian Journal of Physics*, 29:492 – 509, 09 1999.

[57] Paul C. Bressloff and Jay M. Newby. Stochastic models of intracellular transport. *Reviews of Modern Physics*, 85:135–196, Jan 2013.

[58] I. Bronstein, Y. Israel, E. Kepten, S. Mai, Y. Shav-Tal, E. Barkai, and Y. Garini. Transient anomalous diffusion of telomeres in the nucleus of mammalian cells. *Physical Review Letters*, 103:018102, Jul 2009.

[59] L. Bruno, V. Levi, M. Brunstein, and M. A. Despósito. Transition to superdiffusive behavior in intracellular actin-based transport mediated by molecular motors. *Physical Review E*, 80:011912, Jul 2009.

[60] Krysztof Burnecki, Eldad Kepten, Joanna Janczura, Irena Bronshtein, Yuval Garini, and Aleksander Weron. Universal algorithm for identification of fractional brownian motion: A case of telomere subdiffusion. *Biophysical Journal*, 103:1839–1847, November 2012.

[61] Avi Caspi, Rony Granek, and Michael Elbaum. Diffusion and directed motion in cellular transport. *Physical Review E*, 66:011916, Jul 2002.

[62] Toni .J. Feder, Ingrid Brust-Mascher, James P. Slattery, Barbara Baird, and Watt W. Webb. Constrained diffusion or immobile fraction on cell surfaces: a new interpretation. *Biophysical Journal*, 70(6):2767–2773, June 1996.

[63] Ido Golding and Edward C. Cox. Physical nature of bacterial cytoplasm. *Physical Review Letters*, 96:098102, Mar 2006.

[64] Marcel Hellmann, Joseph Klafter, Dieter W Heermann, and Matthias Weiss. Challenges in determining anomalous diffusion in crowded fluids. *Journal of Physics: Condensed Matter*, 23(23):234113, 2011.

[65] Felix Höfling and Thomas Franosch. Anomalous transport in the crowded world of biological cells. *Reports on Progress in Physics*, 76(4):046602, 2013.

[66] Jae-Hyung Jeon, Vincent Tejedor, Stas Burov, Eli Barkai, Christine Selhuber-Unkel, Kirstine Berg-Sørensen, Lene Oddershede, and Ralf Metzler. In vivo anomalous diffusion and weak ergodicity breaking of lipid granules. *Physical Review Letters*, 106:048103, Jan 2011.

[67] Carlo Manzo, Juan A. Torreno-Pina, Pietro Massignan, Gerald J. Lapeyre, Maciej Lewenstein, and Maria F. Garcia Parajo. Weak ergodicity breaking of receptor motion in living cells stemming from random diffusivity. *Physical Review X*, 5:011021, Feb 2015.

[68] C. Metzner, C. Raupach, D. Paranhos Zitterbart, and B. Fabry. Simple model of cytoskeletal fluctuations. *Physical Review E*, 76:021925, Aug 2007.

[69] Georg Seisenberger, Martin U. Ried, Thomas Endreß, Hildegard Büning, Michael Hallek, and Christoph Bräuchle. Real-time single-molecule imaging of the infection pathway of an adeno-associated virus. *Science*, 294(5548):1929–1932, 2001.

[70] S. M. Ali Tabei, Stanislav Burov, Hee Y. Kim, Andrey Kuznetsov, Toan Huynh, Justin Jureller, Louis H. Philipson, Aaron R. Dinner, and Norbert F. Scherer. Intracellular transport of insulin granules is a subordinated random walk. *Proceedings of the National Academy of Sciences*, 110(13):4911–4916, 2013.

[71] Stephanie C. Weber, Andrew J. Spakowitz, and Julie A. Theriot. Bacterial chromosomal loci move subdiffusively through a viscoelastic cytoplasm. *Physical Review Letters*, 104:238102, Jun 2010.

[72] Aubrey V. Weigel, Blair Simon, Michael M. Tamkun, and Diego Krapf. Ergodic and nonergodic processes coexist in the plasma membrane as observed by single-molecule tracking. *Proceedings of the National Academy of Sciences*, 108(16):6438–6443, 2011.

[73] Matthias Weiss, Markus Elsner, Fredrik Kartberg, and Tommy Nilsson. Anomalous subdiffusion is a measure for cytoplasmic crowding in living cells. *Biophysical Journal*, 87(5):3518–3524, November 2004.

[74] R. Aldrin Denny, David R. Reichman, and Jean-Philippe Bouchaud. Trap models and slow dynamics in supercooled liquids. *Physical Review Letters*, 90:025503, Jan 2003.

[75] J. Helfferich, F. Ziebert, S. Frey, H. Meyer, J. Farago, A. Blumen, and J. Baschnagel. Continuous-time random-walk approach to supercooled liquids. i. different definitions of particle jumps and their consequences. *Phys. Rev. E*, 89:042603, April 2014.

[76] J. Helfferich, F. Ziebert, S. Frey, H. Meyer, J. Farago, A. Blumen, and J. Baschnagel. Continuous-time random-walk approach to supercooled liquids. ii. mean-square displacements in polymer melts. *Phys. Rev. E*, 89:042604, April 2014.

[77] Julian Helfferich. Renewal events in glass-forming liquids. *The European Physical Journal E*, 37(8):73, 2014.

[78] J. L. A. Dubbeldam, V. G. Rostiashvili, A. Milchev, and T. A. Vilgis. Driven translocation of a polymer: Fluctuations at work. *Physical Review E*, 87:032147, Mar 2013.

[79] Andrea Zoia, Alberto Rosso, and Satya N. Majumdar. Asymptotic behavior of self-affine processes in semi-infinite domains. *Physical Review Letters*, 102:120602, Mar 2009.

[80] Masao Doi and S. F. Edwards. *The Theory of Polymer Dynamics*. Oxford University Press, New York, 1986.

[81] Prince E. Rouse Jr. A theory of the linear viscoelastic properties of dilute solutions of coiling polymers. *The Journal of Chemical Physics*, 21(7):1272–1280, 1953.

[82] J.-C. Walter, A. Ferrantini, E. Carlon, and C. Vanderzande. Fractional brownian motion and the critical dynamics of zipping polymers. *Physical Review E*, 85:031120, Mar 2012.

[83] G. K. Batchelor. Diffusion in a field of homogeneous turbulence ii. the relative motion of particles. *Mathematical Proceedings of the Cambridge Philosophical Society*, 48:345–362, 4 1952.

[84] Lewis F. Richardson. Atmospheric diffusion shown on a distance-neighbor graph. *Proceedings of the Royal Society of London. A. Mathematical and Physical Sciences*, 110:709–737, April 1926.

[85] M. F. Shlesinger, B. J. West, and Joseph Klafter. Lévy dynamics of enhanced diffusion: Application to turbulence. *Physical Review Letters*, 58:1100–1103, Mar 1987.

[86] Michael F. Shlesinger, George M. Zaslavsky, and Uriel Frisch, editors. *Lévy Flights and Related Topics*. Springer, 1995.

[87] Ramin Golestanian. Anomalous diffusion of symmetric and asymmetric active colloids. *Physical Review Letters*, 102:188305, May 2009.

[88] T. Keyes and J.W. Lyklema. A theory of exciton dynamics with a percolation threshold. *Journal of Statistical Physics*, 27(3):487–498, 1982.

[89] Scott Kirkpatrick. Percolation and conduction. *Reviews of Modern Physics*, 45:574–588, Oct 1973.

[90] Harvey Scher and Elliott W. Montroll. Anomalous transit-time dispersion in amorphous solids. *Physical Review B*, 12:2455–2477, Sep 1975.

[91] O. Bénichou, C. Chevalier, J. Klafter, B. Meyer, and R. Voituriez. Geometry-controlled kinetics. *Nature Chemistry*, 2:472–477, April 2010.

[92] S. Eule, R. Friedrich, F. Jenko, and I. M. Sokolov. Continuous-time random walks with internal dynamics and subdiffusive reaction-diffusion equations. *Physical Review E*, 78:060102, Dec 2008.

[93] Michael A. Lomholt, Irwin M. Zaid, and Ralf Metzler. Subdiffusion and weak ergodicity breaking in the presence of a reactive boundary. *Physical Review Letters*, 98:200603, May 2007.

[94] V. A. Avetisov, A. Kh. Bikulov, and V. Al. Osipov. p-adic description of characteristic relaxation in complex systems. *Journal of Physics A: Mathematical and General*, 36(15):4239, 2003.

[95] V. A. Avetisov and A. Kh. Bikulov. Ultrametricity of fluctuation dynamic mobility of protein molecules. *Proceedings of the Steklov Institute of Mathematics*, 265(1):75–81, July 2009.

[96] V Bercu, M Martinelli, C A Massa, L A Pardi, and D Leporini. A study of the deep structure of the energy landscape of glassy polystyrene: the exponential distribution of the energy barriers revealed by high-field electron spin resonance spectroscopy. *Journal of Physics: Condensed Matter*, 16(43):L479, 2004.

[97] Jasna Brujić, Rodolfo I. Hermans Z., Kirstin A. Walther, and Julio M. Fernandez. Single-molecule force spectroscopy reveals signatures of glassy dynamics in the energy landscape of ubiquitin. *Nature Physics*, 2:282–286, April 2006.

[98] P. Didier, L. Guidoni, and F. Bardou. Infinite average lifetime of an unstable bright state in the green fluorescent protein. *Physical Review Letters*, 95:090602, Aug 2005.

[99] B. Esser. Random walk on ultrametric spaces: A time dependent bond construction. *Physica Status Solidi (b)*, 141(1):K11–K14, 1987.

[100] M. Lederman, R. Orbach, J. M. Hammann, M. Ocio, and E. Vincent. Dynamics in spin glasses. *Physical Review B*, 44:7403–7412, Oct 1991.

[101] L. Lundgren, P. Svedlindh, P. Nordblad, and O. Beckman. Dynamics of the relaxation-time spectrum in a cumn spin-glass. *Physical Review Letters*, 51:911–914, Sep 1983.

[102] G. Margolin and E. Barkai. Nonergodicity of blinking nanocrystals and other lévy-walk processes. *Physical Review Letters*, 94:080601, Mar 2005.

[103] Andrew T. Ogielski and D. L. Stein. Dynamics on ultrametric spaces. *Physical Review Letters*, 55:1634–1637, October 1985.

[104] J. Schnakenberg. Network theory of microscopic and macroscopic behavior of master equation systems. *Reviews of Modern Physics*, 48:571–585, Oct 1976.

[105] Douglas Magde, Elliot Elson, and W. W. Webb. Thermodynamic fluctuations in a reacting system–measurement by fluorescence correlation spectroscopy. *Phys. Rev. Lett.*, 29:705–708, Sep 1972.

[106] Akihiro Kusumi, Chieko Nakada, Ken Ritchie, Kotono Murase, Kenichi Suzuki, Hideji Murakoshi, Rinshi S. Kasai, Junko Kondo, and Takahiro Fujiwara. Paradigm shift of the plasma-membrane concept from the two-dimensional continuum fluid to the partitioned fluid: High-speed single-molecule tracking of membrane molecules. *Annual Review of Biophysics and Biomolecular Structure*, 34(1):351–378, 2005. PMID: 15869394.

[107] Daniel Campos and Vicenç Méndez. Superdiffusive-like motion of colloidal nanorods. *The Journal of Chemical Physics*, 130(13):134711, 2009.

[108] Donald L. Kramer and Robert L. McLaughlin. The behavioral ecology of intermittent locomotion. *American Zoologist*, 41:137–153, 2001.

[109] Jean-René Martin. A portrait of locomotor behaviour in drosophila determined by a video-tracking paradigm. *Behavioural Processes*, 67(2):207 – 219, 2004.

[110] M. C. Santos, D. Boyer, O. Miramontes, G. M. Viswanathan, E. P. Raposo, J. L. Mateos, and M. G. E. da Luz. Origin of power-law distributions in deterministic walks: The influence of landscape geometry. *Physical Review E*, 75:061114, Jun 2007.

[111] Hartmut H. K. Lentz, Thomas Selhorst, and Igor M. Sokolov. Unfolding accessibility provides a macroscopic approach to temporal networks. *Physical Review Letters*, 110:118701, Mar 2013.

[112] O. Bénichou, M. Coppey, M. Moreau, P-H. Suet, and R. Voituriez. Optimal search strategies for hidden targets. *Physical Review Letters*, 94:198101, May 2005.

[113] Gandimohan M. Viswanathan, Sergey V. Buldyrev, Shlomo Havlin, M. G. E. da Luz, E. P. Raposo, and H. Eugene Stanley. Optimizing the success of random searches. *Nature*, 401:911–914, October 1999.

[114] George D. Birkhoff. Proof of the ergodic theorem. *Proceedings of the National Academy of Sciences*, 17(12):656–660, 1931.

[115] Johann von Neumann. Proof of the quasi-ergodic hypothesis. *Proceedings of the National Academy of Sciences*, 18(1):70–82, 1932.

[116] A. I. Khinchin. *Mathematical Foundations of Statistical Mechanics*. Dover Publications, Inc., New York, 1949.

[117] Y. He, S. Burov, R. Metzler, and E. Barkai. Random time-scale invariant diffusion and transport coefficients. *Physical Review Letters*, 101:058101, Jul 2008.

[118] Ariel Lubelski, Igor M. Sokolov, and Joseph Klafter. Nonergodicity mimics inhomogeneity in single particle tracking. *Physical Review Letters*, 100:250602, Jun 2008.

[119] Yasmine Meroz, Igor M. Sokolov, and Joseph Klafter. Subdiffusion of mixed origins: When ergodicity and nonergodicity coexist. *Physical Review E*, 81:010101, Jan 2010.

[120] G. Bel and E. Barkai. Weak ergodicity breaking in the continuous-time random walk. *Physical Review Letters*, 94:240602, Jun 2005.

[121] A. Rebenshtok and E. Barkai. Distribution of time-averaged observables for weak ergodicity breaking. *Physical Review Letters*, 99:210601, Nov 2007.

[122] Igor M. Sokolov, E. Heinsalu, P. Hänggi, and I. Goychuk. Universal fluctuations in subdiffusive transport. *EPL (Europhysics Letters)*, 86(3):30009, 2009.

[123] Jing-Dong Bao, Peter Hänggi, and Yi-Zhong Zhuo. Non-markovian brownian dynamics and nonergodicity. *Physical Review E*, 72:061107, Dec 2005.

[124] S. Burov and E. Barkai. Weak subordination breaking for the quenched trap model. *Physical Review E*, 86:041137, Oct 2012.

[125] Andrey G. Cherstvy, Aleksei V. Chechkin, and Ralf Metzler. Anomalous diffusion and ergodicity breaking in heterogeneous diffusion processes. *New Journal of Physics*, 15(8):083039, 2013.

[126] Weihua Deng and Eli Barkai. Ergodic properties of fractional brownian-langevin motion. *Physical Review E*, 79:011112, Jan 2009.

[127] A. Fuliński. Anomalous diffusion and weak nonergodicity. *Physical Review E*, 83:061140, Jun 2011.

[128] Luciano C. Lapas, Rafael Morgado, Mendeli H. Vainstein, J. Miguel Rubí, and Fernando A. Oliveira. Khinchin theorem and anomalous diffusion. *Physical Review Letters*, 101:230602, Dec 2008.

[129] Marcin Magdziarz and Aleksander Weron. Anomalous diffusion: Testing ergodicity breaking in experimental data. *Physical Review E*, 84:051138, Nov 2011.

[130] Tomoshige Miyaguchi and Takuma Akimoto. Intrinsic randomness of transport coefficient in subdiffusion with static disorder. *Physical Review E*, 83:031926, Mar 2011.

[131] Tomoshige Miyaguchi and Takuma Akimoto. Anomalous diffusion in a quenched-trap model on fractal lattices. *Phys. Rev. E*, 91:010102, Jan 2015.

[132] 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. *Physical Review Letters*, 90:120601, Mar 2003.

[133] Ronald D. Vale. The molecular motor toolbox for intracellular transport. *Cell*, 112(4):467–480, February 2003.

[134] Roop Mallik and Steven P. Gross. Molecular motors: Strategies to get along. *Cell*, 14(22):R971–R982, November 2004.

[135] S. Condamin, V. Tejedor, R. Voituriez, O. Bénichou, and Joseph Klafter. Probing microscopic origins of confined subdiffusion by first-passage observables. *Proceedings of the National Academy of Sciences*, 105(15):5675–5680, 2008.

[136] Yasmine Meroz, Igor M. Sokolov, and Joseph Klafter. Test for determining a subdiffusive model in ergodic systems from single trajectories. *Physical Review Letters*, 110:090601, Feb 2013.

[137] Marcin Magdziarz, Aleksander Weron, Krzysztof Burnecki, and Joseph Klafter. Fractional brownian motion versus the continuous-time random walk: A simple test for subdiffusive dynamics. *Physical Review Letters*, 103:180602, Oct 2009.

[138] Barry D. Hughes. *Random Walks and Random Environments, Volume 1: Random Environments*. Oxford Science Publications, 1995.

[139] Barry D. Hughes. *Random Walks and Random Environments, Volume 1: Random Environments*. Oxford Science Publications, 1995.

[140] Takashi Kumagai. *Random Walks on Disordered Media and their Scaling Limits*. Springer, 2014.

[141] Ravindra B. Bapat. *Graphs and Matrices*. Universitext. London: Springer, second edition edition, 2014.

[142] Piet Van Mieghem. *Graph Spectra for Complex Networks*. Cambridge University Press, Cambridge, first paperback edition edition, 2012.

[143] Andries E. Brouwer and Willem H. Haemers. *Spectra of Graphs*. Springer New York, 2012.

[144] William Feller. *An Introduction to Probability Theory*. Wiley series in probability and mathematical statistics. John Wiley & Sons, Inc., 2nd edition, 1968.

[145] Daniel W. Stroock. *An Introduction to Markov Processes*. Graduate Texts in Mathematics. Springer, second edition edition, 2014.

[146] P. Hanggi, H. Thomas, H. Grabert, and P. Talkner. Note on time evolution of non-markov processes. *Journal of Statistical Physics*, 18(2):155–159, 1978.

[147] Shechao Feng and Pabitra N. Sen. Percolation on elastic networks: New exponent and threshold. *Phys. Rev. Lett.*, 52:216–219, Jan 1984.

[148] Stphane Roux. Generalized brownian motion and elasticity. *Journal of Statistical Physics*, 48(1-2):201–213, 1987.

[149] Alessandra Faggionato. Spectral analysis of 1d nearest-neighbor random walks and applications to subdiffusive trap and barrier models. *Electronic Journal of Probability*, 17:no. 15, 1–36, 2012.

[150] Charles Stone. Limit theorems for random walks, birth and death processes, and diffusion processes. *Illinois Journal of Mathematics*, 7(4):638–660, 12 1963.

[151] L. R. G. Fontes, M. Isopi, and C. M. Newman. Random walks with strongly inhomogeneous rates and singular diffusions: convergence, localization and aging in one dimension. *The Annals of Probability*, 30(2):579–604, April 2002.

[152] Robert Zwanzig. Memory effects in irreversible thermodynamics. *Phys. Rev.*, 124:983–992, Nov 1961.

[153] Gustav Kirchhoff. Ueber die auflösung der gleichungen, auf welche man bei der untersuchung der linearen vertheilung galvanischer ströme gefhrt wird. *Annalen der Physik*, 148(12):497–508, 1847.

[154] Philipp Maass, Bernd Rinn, and Walter Schirmacher. Hopping dynamics in random energy landscapes: An effective medium approach. *Philosophical Magazine Part B*, 79(11/12):1915–1922, 1999.

[155] Kiyoshi Kawazu and Harry Kesten. On birth and death processes in symmetric random environment. *Journal of Statistical Physics*, 37(5-6):561–576, 1984.

[156] D. A. G. Bruggeman. Berechnung verschiedener physikalischer konstanten von heterogenen substanzen. *Annalen der Physik*, 24:636–664, 1935.

[157] Sebastian Bustingorry, Manuel O. Cáceres, and Eduardo R. Reyes. Effective-medium approximation with asymmetric transition rates. *Physical Review B*, 65:165205, April 2002.

[158] Jonathan Machta. Random walks on site disordered lattices. *Journal of Physics A: Mathematical and General*, 18(9):L531, June 1985.

[159] Yasmine Meroz, Igor M. Sokolov, and Joseph Klafter. Unequal twins: Probability distributions do not determine everything. *Phys. Rev. Lett.*, 107:260601, Dec 2011.

[160] M. Silver, G. Schoenherr, and H. Baessler. Dispersive hopping transport from an exponential energy distribution of sites. *Phys. Rev. Lett.*, 48:352–355, Feb 1982.

[161] E. M. Bertin and J.-P. Bouchaud. Subdiffusion and localization in the one-dimensional trap model. *Phys. Rev. E*, 67:026128, Feb 2003.

[162] E. M. Bertin and J.-P. Bouchaud. Linear and nonlinear response in the aging regime of the one-dimensional trap model. *Physical Review E*, 67:065105, Jun 2003.

[163] S. Burov and Eli Barkai. Occupation time statistics in the quenched trap model. *Physical Review Letters*, 98:250601, Jun 2007.

[164] S. Burov and E. Barkai. Time transformation for random walks in the quenched trap model. *Physical Review Letters*, 106:140602, Apr 2011.

[165] Cécile Monthus. Anomalous diffusion, localization, aging, and subaging effects in trap models at very low temperature. *Phys. Rev. E*, 68:036114, Sep 2003.

[166] J.P. Bouchaud. Weak ergodicity breaking and anging in disordered systems. *J. Phys. I France*, 2(9):1705–1713, September 1992.

[167] Cécile Monthus and Jean-Philippe Bouchaud. Models of traps and glass phenomenology. *J. Phys. A: Math. Gen.*, 29:3847–3869, 1996.

[168] H. Kesten and F. Spitzer. A limit theorem related to a new class of self similar processes. *Zeitschrift fr Wahrscheinlichkeitstheorie und Verwandte Gebiete*, 50(1):5–25, 1979.

[169] Michael A. Lomholt, Ludvig Lizana, Ralf Metzler, and Tobias Ambjörnsson. Microscopic origin of the logarithmic time evolution of aging processes in complex systems. *Physical Review Letters*, 110:208301, May 2013.

[170] John H. Cushman, Daniel O'Malley, and Moongyu Park. Anomalous diffusion as modeled by a nonstationary extension of brownian motion. *Physical Review E*, 79:032101, Mar 2009.

[171] Eugene B. Postnikov and Igor M. Sokolov. Model of lateral diffusion in ultrathin layered films. *Physica A: Statistical Mechanics and its Applications*, 391(21):5095 – 5101, 2012.

[172] Miquel Montero and Jaume Masoliver. Nonindependent continuous-time random walks. *Physical Review E*, 76:061115, Dec 2007.

[173] Felix Thiel, Lutz Schimansky-Geier, and Igor M. Sokolov. Anomalous diffusion in run-and-tumble motion. *Physical Review E*, 86:021117, Aug 2012.

[174] Elliott W. Montroll and George H. Weiss. Random walks on lattices. ii. *Journal of Mathematical Physics*, 6(2):167–181, 1965.

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

[176] R. Friedrich, F. Jenko, A. Baule, and S. Eule. Anomalous diffusion of inertial, weakly damped particles. *Physical Review Letters*, 96:230601, Jun 2006.

[177] Thomas Neusius, Igor M. Sokolov, and Jeremy C. Smith. Subdiffusion in time-averaged, confined random walks. *Physical Review E*, 80:011109, Jul 2009.

[178] Mark M. Meerschaert, Erkan Nane, and Yimin Xiao. Correlated continuous time random walks. *Statistics & Probability Letters*, 79(9):1194 – 1202, 2009.

[179] Felix Thiel and Igor M. Sokolov. Scaled brownian motion as a mean-field model for continuous-time random walks. *Physical Review E*, 89:012115, Jan 2014.

[180] Michael J. Saxton. Anomalous subdiffusion in fluorescence photobleaching recovery: A monte carlo study. *Biophysical Journal*, 81:2226–2240, October 2001.

[181] Jianrong Wu and Keith M. Berland. Propagators and time-dependent diffusion coefficients for anomalous diffusion. *Biophysical Journal*, 95(4):2049–2052, August 2008.

[182] Gennady SamorodnitskyAND Murad S. Taqqu. *Stable Non-Gaussian Random Processes:Stochastic Models with Infinite Variance*. Chapman & Hall / CRC, London, 1994.

[183] Ken-Iti Sato. *Lévy Processes and Infinitely Divisible Distributions*. Cambridge University Press, Cambridge, 1999.

[184] D. Froemberg, M. Schmiedeberg, E. Barkai, and V. Zaburdaev. Asymptotic densities of ballistic lévy walks. *Physical Review E*, 91:022131, Feb 2015.

[185] Benoit B. Mandelbrot and John van Ness. Fractional brownian motions, fractional noises and applications. *SIAM Review*, 10(4):pp. 422–437, October 1968.

[186] Shlomo Havlin and Daniel Ben-Avraham. Diffusion in disordered media. *Advances in Physics*, 36(6):695–798, 1987.

[187] Felix Thiel, Franziska Flegel, and Igor M. Sokolov. Disentangling sources of anomalous diffusion. *Physical Review Letters*, 111:010601, Jul 2013.

[188] Jacques Janssen and Raimond Manca. *Applied Semi-Markov Processes*. Springer Science + Businesss Media, 2006.

[189] Aleksander Weron and Marcin Magdziarz. Generalization of the khinchin theorem to lévy flights. *Physical Review Letters*, 105:260603, Dec 2010.

[190] A. G. Cherstvy and R. Metzler. Nonergodicity, fluctuations, and criticality in heterogeneous diffusion processes. *Phys. Rev. E*, 90:012134, Jul 2014.

[191] D. Froemberg and E. Barkai. Time-averaged einstein relation and fluctuating diffusivities for the lévy walk. *Physical Review E*, 87:030104, Mar 2013.

[192] Eric Lutz. Power-law tail distributions and nonergodicity. *Phys. Rev. Lett.*, 93:190602, Nov 2004.

[193] Jae-Hyung Jeon and Ralf Metzler. Fractional brownian motion and motion governed by the fractional langevin equation in confined geometries. *Physical Review E*, 81:021103, Feb 2010.

[194] Marcin Magdziarz and Aleksander Weron. Ergodic properties of anomalous diffusion processes. *Annals of Physics*, 326(9):2431 – 2443, 2011.

[195] Felix Thiel and Igor M. Sokolov. Weak ergodicity breaking in an anomalous diffusion process of mixed origins. *Physical Review E*, 89:012136, Jan 2014.

[196] Joseph Klafter and Igor M. Sokolov. *First steps in random walks*. Oxford UP, 1st edition, 2011.

[197] Milton Abramowitz and Irene A. Stegun, editors. *Handbook of mathematical functions*. U.S. Government Printing Office, 10th edition, December 1972.

## B.1. Selbständigkeitserklärung

Ich erkläre, dass ich die Dissertation selbständig und nur unter Verwendung der von mir gemäß §7 Abs. 3 der Promotionsordnung der Mathematisch-Naturwissenschaftlichen Fakultät, veröffentlicht im Amtlichen Mitteilungsblatt der Humboldt-Universität zu Berlin Nr. 126/2014 am 18.11.2014, angegebenen Hilfsmittel angefertigt habe.

Berlin, den ..... Unterschrift .....