INAUGURAL–DISSERTATION

zur

Erlangung der Doktorwürde

der

Naturwissenschaftlich-Mathematischen Gesamtfakultät

der

Ruprecht–Karls–Universität
Heidelberg

Thema:

Stochastic model and intermediate asymptotics for charge transport in organic semiconductors

vorgelegt von

Diplom-Mathematiker Sven Stodtmann

aus

Neunkirchen -Saar-

Betreuer: . . . . . Prof. Dr. Dr. h.c. mult. Willi Jäger

Prof. Dr. Dr. h.c. Vincenzo Capasso

Tag der mündlichen Prüfung: . . . . . .

Stochastic model and intermediate asymptotics for charge transport in organic semiconductors

Ph.D. Thesis

Sven Stodtmann

Advisors
Prof. Dr. Dr. h.c. mult. Willi Jäger
Prof. Dr. Dr. h.c. Vincenzo Capasso

Acknowledgments

I would like to express my gratitude towards professor W. Jäger for his advice and encouragement throughout my research. I am especially thankful for the great deal of freedom I have been allowed in the choice of both, research topic and methods. The same goes for professor V. Capasso, whom I would like to thank for his advice and for always being able to point me to exactly the literature I needed. A big thank you also goes to the secretaries of professor Jäger, Ina Scheid and Gabriele Schocke, who always helped me find a slot in his very busy schedule.

I gratefully acknowledge funding by BASF SE. I want to thank everyone in Scientific Computing and the OLED team for the great working atmosphere, stimulating discussions and for introducing me to the topic. Particularly Alexander Badinski, who made this work possible, Christian Lennartz and Falk May, for teaching me a a lot about the quantum chemical fundamentals. I am also very glad that I have had the opportunity to work with experimentalists within BASF, U. Heinemeyer and M. Al-Helwi, whose input was very valuable for me.

I thank the Andrienko Group at the Max Planck Institute for Polymer Research in Mainz, R. Coehoorn and H. van Eersel from Phillips/TU Eindhoven, M. Mommer from IWR Heidelberg and R. Nitsche from sim4tec for fruitful discussions and collaborations.

I am also very grateful to my friends for providing the occasional distraction needed to clear my mind. Especially Kai and Lars for being my “intelligent walls” to bounce off random thoughts – often related to this work.

I am deeply indebted to my family, who have done so much for me.

Contents

List of figures and tables

List of symbols

BM	Brownian motion	p. 111
CTRW	continuous time random walk	p. 24
DOS	density of states	p. 20
IP	inverse power-law	p. 2
kMC	kinetic Monte Carlo	p. 79
RV	random variable	p. 104
SRW	simple random walk	p. 115
β	inverse temperature	p. 16
$\mathcal{B}$	bad traps	p. 41
$B_x(r)$	ball centered at x, radius r	p. 32
$\mathbf{C}$	class of functions	p. 74
$\mathrm{Conv}(A)$	convex hull	p. 124
$C_{d,\alpha}$	scaling constant	p. 35
Δ	(discrete) Laplacian	p. 116
∇_y	discrete difference	p. 115
${}_a\mathrm{D}_x^\alpha$	$\alpha \in \mathbb{R}$, fractional integral/differential	p. 74
D	diffusion constant	p. 19
$D(I, \mathbb{R}^k)$	Skorokhod space	p. 124
$D_{\uparrow\uparrow}$	strictly increasing functions	p. 126
$\mathbb{E}[X\|\mathcal{A}]$	conditional expectation	p. 107
$\mathbb{E}[X]$	expected value	p. 107
$\mathcal{E}$	tame sites	p. 41
$\mathcal{E}_0$	very tame sites	p. 43
$\tilde{\mathbf{E}}_x$	random energy	p. 32
$D^{\mathrm{Emp.}}(t)$	empirical diffusion coefficient	p. 86
$D_\alpha^{\mathrm{Emp.}}(t)$	empirical anomlalous diffusion coefficient	p. 86
$E_\alpha(x)$	Mittag-Leffler function	p. 113

$\Phi_X(\lambda)$	characteristic exponent	p. 109
$\phi_X(\lambda)$	characteristic function	p. 108
$\|\psi\rangle$	state vector	p. 12
$\Psi_X(\lambda)$	Laplace exponent	p. 109
$\psi_X(\lambda)$	moment generating function	p. 108
$Q_x(r)$	cube centered at x, side-length r	p. 32
ρ	coarse graining scale	p. 40
ρ	density	p. 19
r	spatial (macro-)scale	p. 40
$R(n)$	observation horizon	p. 36
$r^n(x)$	conditional spatial increment	p. 43
$r_{x\to y}^{\textbf{(MA)}}$	Miller-Abrahams rate	p. 15
$r_{x\to y}^{\textbf{(Marcus)}}$	Marcus rate	p. 15
$\hat{s}$	scale base	p. 40
$\hat{\sigma}$	effective disorder	p. 33
$s^n(x)$	conditional score	p. 43
$S^{(n)}$	scaled clock	p. 35
$S^{\hat{\sigma}}(k)$	clock process	p. 33
s_i^n	score	p. 43
$\boldsymbol{\tau}$	random environment	p. 33
τ_x	mean waiting time	p. 32
T	absolute temperature	p. 15
T^ε	shallow traps	p. 40
T_ε^M	deep traps	p. 40
T_M	very deep traps	p. 40
V_α	stable subordinator	p. 112
$X^{(n)}(t)$	scaled SRW	p. 35
$X^{(x)}(t)$	simple random walk	p. 33
$Y(t)$	charge transport process	p. 33

CHAPTER 1

Introduction

Organic electronics is a relatively young and very promising field of research. In general, any compound that contains carbon is considered organic – with the exception of some very simple compounds, such as carbon dioxide or carbonates. While it was known since mid 19th century that some organic materials conduct electricity [57], they were mostly considered insulators. Since the 1950's, there were more reports on conductive organic materials [49, 68, 71, 90]. The breakthrough of these materials was the 1977 discovery by Shirakawa, MacDiarmid, Heeger and co-workers [29, 104]. They reported a strong effect of chemical doping on the conductivity of certain polymers, making it possible to reach conductivities comparable to metallic conductors . Their work was awarded the Nobel prize in 2000 [84].

Organic materials have several advantages over their inorganic counterparts, including easier processing, mechanical flexibility and the possibility to design organic molecules for specific purposes. Among the numerous applications for these materials are light-emmitting diodes (OLEDs [113]), photovoltaic devices (OPV [112]) and field-effect transistors (OFET [52]).

As many new technologies, organic semiconductors pose challenges to science. In this thesis we closely look at one aspect of these materials which is vital for their function – namely charge transport. Despite the fact that microscale charge transport has been modeled as a discrete stochastic process already in the 1980's, to our knowledge it has not yet been taken advantage of the methods and results of modern probability theory, which is the aim of the present work.

The transition from micro- to macroscale is particularly important in organic semiconductors, since the properties of the molecules can easily be manipulated. We will rigorously investigate the question of this upscaling via a scaling limit, *i.e.* we will present a way to rescale the discrete space microscopic charge transport process in a way, that the rescaled process converges to a continuous space one. In order not to miss important effects, we have to be particularly careful in the choice of the scales here and not choose them too coarse.

This kind of rescaling has not been done until now, leaving open the question of

validity of continuous space device models. In particular, since the most widely used approach, based on a drift-diffusion equation, can not account for the dispersive effects in thin devices. Dispersion stems from a broad distribution of local transport parameters and manifests in seemingly time- or thickness dependent macroscopic transport properties. This effect is observed in many similar situations in different application areas such as transport in porous media[65], polymer translocation[47] or chromatin-binding proteins [118].

We will start our modeling of charge transport at the molecular level, building on inputs from quantum chemistry. We find that the materials are amorphous. This means, there is no long range order. Therefore, the charge carriers (more precisely, their wavefuctions) are strongly localized on (parts of) molecules. Hence, transport does not take place in delocalized states, as it is the case in conventional semiconductors, but is, at a fundamental level, discrete. Analyzing the structure of this model, we identify the main ingredient, which we believe to cause the dispersive effects. We propose a simplified model based on this intuition which is then mathematically treated.

The mathematical treatment makes use of a decomposition of the charge transport process into a spatial and a temporal component, both of which can be represented as a sum of random variables. We use a modified coarse graining scheme inspired by Ben Arous & Černý [15] to treat the strong correlations in the temporal component.

We show, that our simplified model admits a strongly dispersive regime. More particularly, we show that suitably rescaled trajectories of the charge carriers converge to the *fractional kinetics process* FK_α, a time change of Brownian motion by a stable subordinator. This allows us to connect charge transport to time-fractional diffusion, which governs the evolution of the density of FK_α. While we rigorously treat only the simplified model, we provide theoretical and simulation based arguments for the validity of our theory in the full, unsimplified model.

The simplified model we propose has been previously studied in mathematical physics, mainly in the context of spin-glasses. It is known as *Bouchaud's trap model* [23]. Typically, it is analyzed on the complete graph and with inverse power-law (IP) waiting time distribution. In contrast to that, our model uses the graph $\mathbb{Z}^d$ for $d \geq 3$, and waiting times will not be IP. The case of non-IP tails has been considered on the complete graph by Gayrard [42] and the case of $\mathbb{Z}^d$ but with IP tails by Ben Arous & Černý [15]. This last work has strongly inspired our main theorem and its proof. The difference between their and our main theorem is in the assumption on the distribution of the mean waiting times τ_x. While they assume

$$\mathbb{P}\left(\tau_x > u\right) = u^{-\alpha}(1 + L(u)), \tag{1.1}$$

for some $L(u)$ which goes to 0 as u goes to infinity, we wish to extend their result to the practically very relevant case of lognormal energies, *i.e.* we require instead (cf.

cor. A.6.3) as $n \to \infty$,

$$\mathbb{P}\left(\tau_x^n \geq g(n)u\right) = \left(1 + L(n,u)\right) \frac{1}{\sqrt{2\pi\alpha}} g(n)^{-\alpha/2} \hat{\sigma}(n)^{-1} u^{-\alpha}, \tag{1.2}$$

where $\tau_x^n = (\tau_x)^{\hat{\sigma}(n)}$ are rescaled mean waiting times, $\hat{\sigma}(n)$ and $g(n)$ deterministic sequences. Here $L(n,u) \to 0$ as $u \to \infty$ uniformly in n.

The results and new contributions of this thesis are the following:

- We generalize the main theorem of Ben Arous & Černý [15] to the practically very important case of lognormally distributed waiting times.

  While we may adopt the structure of the proof of ref. 15, we are obliged to use different scales (see tab. 5.1 and *e.g.* eq. (5.17)) to ensure that the theorem still holds and adapt the technical details (*e.g.* in eq. (5.27) and the proof of lem. 5.2.1).

- We thoroughly investigate the connection between charge transport in organic semiconductors and fractional diffusion. In particular,
  - we analyze the finite size effect in kinetic Monte Carlo simulations and suggest a way of simulating charge transport, that completely eliminates it.
  - This enables us to computationally study the dispersive nature of charge transport.
  - We show, that a fractional diffusion coefficient is an invariant of the microscopic charge transport process (at least for the interesting time period) – while the classical diffusion coefficient is not (see fig. 1.1).
  - Due to our generalization of the convergence result of ref. 15, we can rigorously connect a simplified charge transport model to fractional diffusion.

- We present our intuitions about generalizations of the model and show that a simple correction in existing diffusive models can already describe the missing dispersive effects (see fig. 1.2).

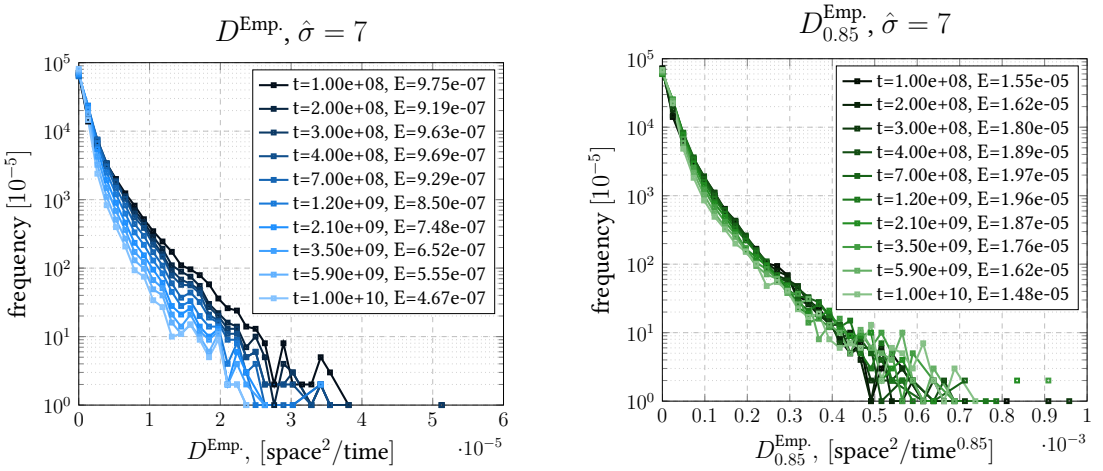

Figure 1.1: "Normal" diffusion coefficient $D(t)$ (left) vs. anomalous diffusion coefficient $D_{0.85}$ for different times at effective disorder $\hat{\sigma} = 7$. See fig. 7.9, p. 88.

1.1 Overview of the thesis

The **first part** of this thesis is devoted to the physics of charge transport in organic semiconductors and reviewing the approaches that have already been taken at the problem.

In **chap. 2: Charge transport in organic semiconductors**, we start with the physical description of charge transport in organic materials. In **sec. 2.1: Morphology and disorder**, we consider the microscale structure of the material. Upon this understanding, we then build the physical charge transport models in **sec. 2.2: Microscale charge transport**.

In **chap. 3: Overview of macroscopic transport models**, we present a brief overview of analytic approaches (**sec. 3.1: Bottom-up/analytical approaches**) and macroscopic charge transport models used in the literature. Many recent approaches are based on Monte Carlo simulation studies of the so-called Gaussian disorder model (GDM, see **sec. 3.2: Simulation based aproaches**, [13]), which is also the basis of our treatment. A very common feature in macroscopic device models is the assumption that transport is diffusive. For the sake of completeness, in **sec. 3.3: Experimental/heuristic approaches**, we review heuristic models, which are frequently used to interpret experimental results.

In the **second part** of the thesis, we introduce a simplified mathematical model and apply methods of modern probability theory to show convergence of this model to the fractional kinetics process.

In **chap. 4: Mathematical model and main result**, we introduce our simplified model and state the main result. Due to the discrete nature of the charge transport, we model the motion of the charge carriers as a spatially discrete, continuous time

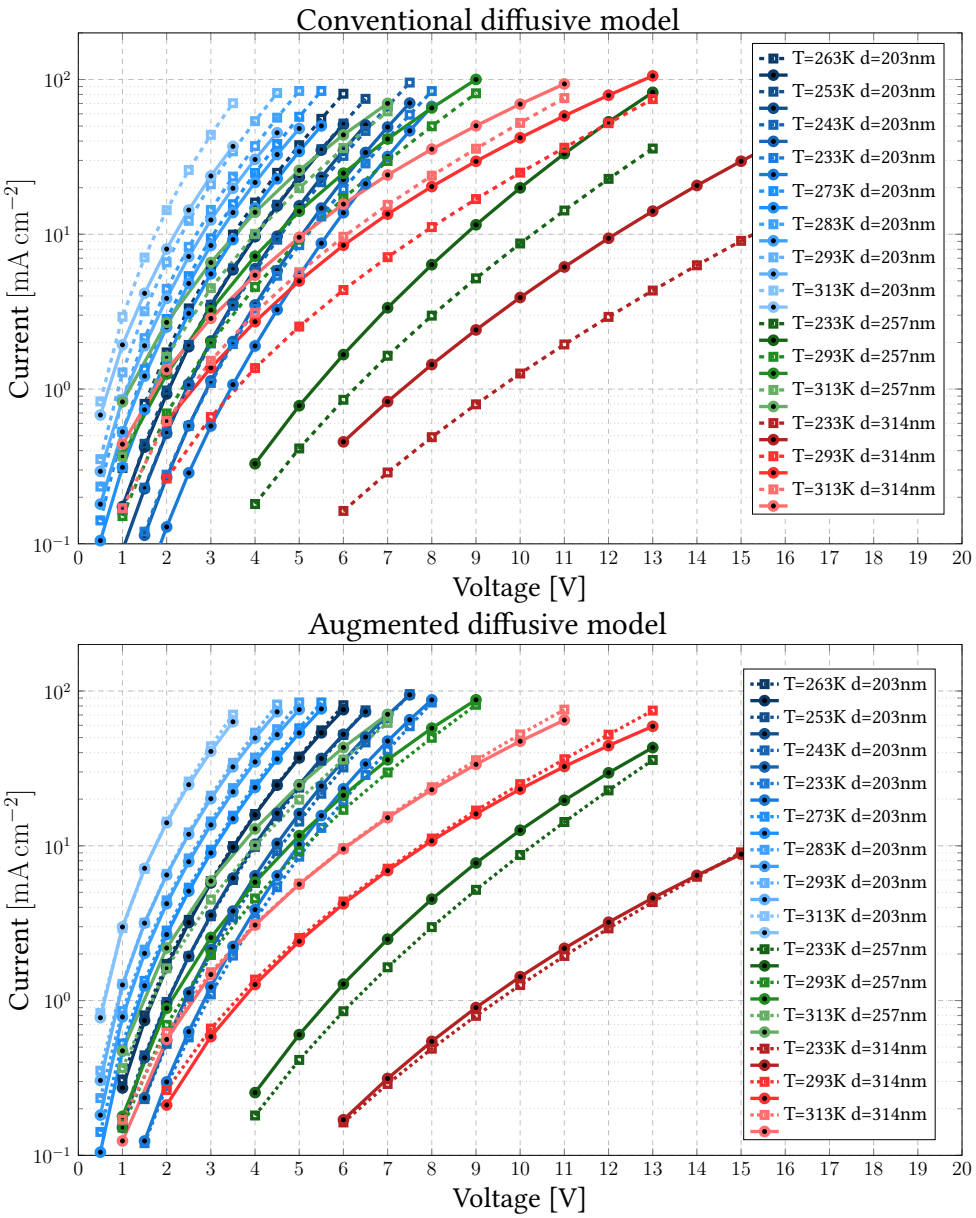

Figure 1.2: Solid lines: Diffusive (upper) and augmented diffusive model (lower), dashed lines: Experimental data for different temperatures and thicknesses. See figs. 7.7, p. 85 and 7.6, p. 84.

stochastic process. The transition rates are taken from the discussion in chap. 2 and are random due to their dependence of the site energies. We follow the GDM and model these energies as independent, zero-mean, normal random variables with variance σ^2. We show that the physically motivated charge transfer rates have a common

reversible structure, which we use to define a simplified model on the integer lattice $\mathbb{Z}^3$ in **sec. 4.1: Simplified mathematical model**. Furthermore, in **sec. 4.2: Main result and rough outline of the proof** we state our main result and present a rough sketch of its proof.

Chapter 5: Scaling limit contains the proof of the main theorem. Due to its length, the proof is partitioned into several sections, **sec. 5.1: The coarse graining procedure**, **sec. 5.2: Approximation of the clock process** and finally **sec. 5.3: Convergence of the process**.

In the **third part** of the thesis, we interpret our mathematical result and apply it to the physical problem.

In **chap. 6: The fractional kinetics process and fractional calculus**, we collect some of the properties of the fractional kinetics process we consider important in this context, especially its connection to time-fractional diffusion. For this to make sense, we introduce in **sec. 6.1: Fractional calculus and fractional differential equations** the relevant methods and results to treat the equation. The properties then are collected in **sec. 6.2: Properties of the fractional kinetics process**.

In **chap. 7: Simulations and experiments** we use the theoretical results of the previous sections to interpret both, computationial and experimental results. In **sec. 7.1: The finite-size effect**, we introduce special kinetic Monte Carlo algorithm, which we show is necessary to avoid computational artifacts.
In **sec. 7.2: Dispersion**, we investigate the change of transport parameters over time (or sample thickness) with different experimental and computational methods.
In **sec. 7.4: Simplifications**, we present evidence from literature, complemented with theoretical and computational arguments of our own to justify the simplifications made in sec. 4.1.
In **sec. 7.3: Comparison with the continuum limit**, we study the dispersive behavior on both the micro- and the macroscale. The microscopic model is studied using the algorithm introduced in sec. 7.1.
We also show how, based on the results of this thesis, a macroscopic device model can be augmented with a simple trick, to get a correction which accounts for dispersive effects. As a motivation for the relatively technical work, fig. 1.2 gives a pre-post comparison for a realistic unipolar diode. While the diffusive model fails to take into account the apparent density dependence of the transport, a model taking the dispersive behavior into account yields better results.

In **chap. 8: Conclusions and outlook**, we propose some generalizations of the main theorem to accommodate more realistic microscopic models (**sec. 8.1: More realistic models**). We outline how we expect they can be proven. However, in many cases no qualitatively different result is expected. Furthermore, in **sec. 8.2: Extensions**, we offer some intuitions about how we expect the model to behave in the cases of multiple interacting carriers and additional electric fields.

In **sec. A: Concepts of probability theory**, we provide the reader with the definitions needed to formulate the problem, as well as some results from the literature we need in the proof of our main theorem. Starting from **sec. A.1: Fundamentals** , we consider Lévy proceses (**sec. A.2: Lévy processes, subordinators and their inverses**), potential theory for the simple random walk on $\mathbb{Z}^d$ (**sec. A.3: Potential theory of the simple random walk**) and some limit theorems for sums of random variables (**sec. A.4: Limit theorems for sums of random variables**), some of which are simpler predecessors of our main result and thus provide us with the intuition it should be true – others are needed as intermediate results for our proof. Because we deal with the convergence of stochastic processes, we have to introduce Skorokhod spaces (**sec. A.5: Skorokhod spaces**), which are the natural spaces for the analysis of stochastic processes. Finally, in **sec. A.6: Powerlaws and the lognormal distribution**, we show that a triangular array of rescaled lognormal random variables converges to a sequence of random variables with inverse power law tail at infinity. **Appendix B** contains a flowchart of the proof, which gives an overview on the dependence of the various substeps needed to show our main result.
In **sec. C: Device simulation models**, we present two state of the art continuum device models and suggest an augmentation based on the ideas of this thesis.
In **sec. D: Some details on the kinetic Monte Carlo implementation**, the implementation of our kinetic Monte Carlo method is described and some details on particularly performance relevant aspects are discussed.

PART I

A Short Introduction to the Physics

CHAPTER 2

Charge transport in organic semiconductors

In theory, organic semiconductor based devices can often outperform their inorganic counterparts, because the material properties can be engineered on a molecular level. This means, that in order to profit from this abilities, it is crucial to understand the relationship between molecular properties and macroscopic behavior.

While not going into too much detail, this chapter introduces the physics of charge transport in organic semiconductors. We assume familiarity with some basic chemistry, such as molecular orbitals. A reference for basic chemistry is *e.g.* Tro [114].

There are two main types of organic semiconductors. One is based on polymers, large molecules with a carbon-chain backbone. The other type are the so-called small molecules. These materials have a much lower molecular weight compared to polymers and consist of only few organic functional groups. This is the class we will focus on in this work.

The individual molecules only interact weakly via the van-der-Waals (vdW) force unlike most inorganic semiconductors or metals, which bond covalently (*i.e.* by sharing electrons). Therefore, the materials only have a very weak long range order. They are *amorphic* – the positions of the individual molecules are not following a regular pattern. This lack of periodicity on the microscopic scale gives rise to the main difference in transport mechanism compared to their inorganic counterpart.

2.1 Morphology and disorder

In the following, we sometimes take the perspective of electron transport. The same arguments are true if we consider the absence of an electron – *i.e.* a positive charge, called a hole – instead of an electron.

Charge transport on the molecular scale is governed by the *Schrödinger equation* for

the *state vector* $|\psi\rangle$ in some separable Hilbert space[1]

$$i\hbar \frac{\partial}{\partial t} |\psi\rangle = \hat{\mathbf{H}} |\psi\rangle . \tag{2.1}$$

Here, $\hbar$ is the *reduced Planck constant* and $\hat{\mathbf{H}}$ the *Hamiltonian* operator of the system, which models the physics.

We will not go into further detail on how to solve this equation or why and how $\hat{\mathbf{H}}$ models to the physics. The interested reader is referred to the standard works in the field, such as Dirac [36] or Shankar [103]. For the application to chemistry, see Levine [58].

In principle this solves the problem. It is possible to have the *state vector* ψ describe the whole physical system and include all (relevant) dynamics in the Hamilton operator $\hat{\mathbf{H}}$. The problem is: if we do that, the equation for a full device is completely unsolvable in practice. Therefore, simplifying assumptions are necessary.

There are unlimited possibilities in the choice of simplifications. The main idea behind all of them is, to either get to a system less degrees of freedom or one with a simpler Hamiltonian operator. For example, if we assume that the molecules themselves do not move on the same timescale as the charges we are interested in, we can decouple the dynamics of the molecules and charges (this is essentially the Born-Oppenheimer approximation [22]). If this background (*i.e.* the positive charges modeling the nuclei of the atoms making up the molecules) now were perfectly periodic, the Bloch-Theorem (Ashcroft & Mermin [7, chap. 8]) would apply and we could effectively shrink the whole system to one periodic box. This approach is very successful in explaining properties of charge transport in crystalline inorganic semiconductors (such as Si, Ge, GaAs, see Sze [110]). However, as we already mentioned, in organic semiconductors, there is no long-range order, and this approach does not work.

Instead, we will first have to understand how neutral Molecules interact in a material. Even further, to do this we will have to first understand a single molecule. We follow the bottom-up approach chosen *e.g.* in the software package VOTCA (see Rühle *et al.* [95]). We will briefly comment on the way up here, a complete treatment of this procedure and the underlying theories is again out of the scope of this thesis.

2.1.1 Morphology

The molecules interact mainly through their electron shells, repulsing each other (so they do not get too close) and inducing dipoles (see fig. 2.1) – this is the van-der-Waals force, the main adhesive force in the materials. Therefore, the main effort on this stage is to understand the behavior of the electron shells of the molecules'. This is done by solving the Schrödinger equation for a single molecule approximatively *e.g.* using

[1] this is not completely true, but it gives the right idea. For the full story, see Gadella & Gòmez [40].

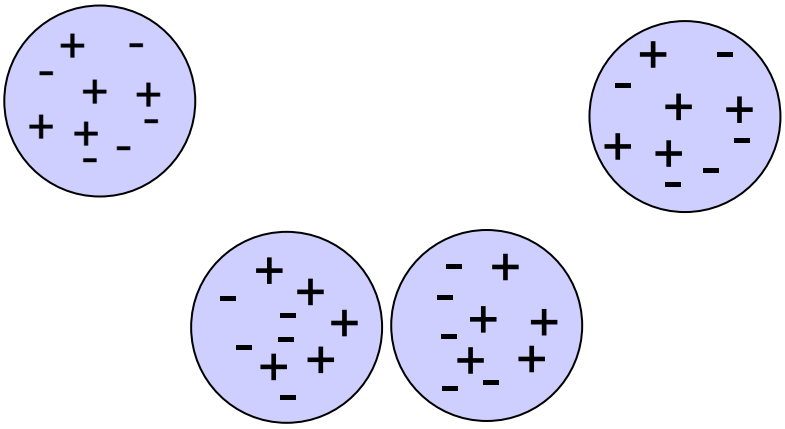

Figure 2.1: Illustration of the van-der-Waals force between two molecules. Internally, the charges reassemble and a dipole is induced.

(hybrid) density functional theory (Hohenberg & Kohn [45] and Kohn & Sham [53]) or semi-empirical methods (Ridley & Zerner [92]).

Once this is solved, external forces are added to investigate the molecules response to deformation. This procedure is guided by heuristics – the idea is to capture the response of a molecule to the main deformation processes it is subjected to in a solid. These responses are parametrized and summarized in what is called the *force field* of that molecule.

To get an idea how the materials are assembled on molecular scale, this force field is now used in classical simulations with a large number of molecules, represented by the centers of mass of their constituting atoms and some potentials taking into account the interactions parametrized in the forcefield, as well as standard terms for electrostatic interaction.

With this approach, it is possible to calculate with reasonable computational effort the arrangement of about 4×10^4 molecules. Because the arrangement of the molecules is not completely static, one should more precisely speak of snapshots of the morphology, since they change over time. We assume that the timescale for this change is much slower than the one of charge transport. Another important point is that the dynamics should be run until the statistical properties of the sample do not change significantly between snapshots.

We will later show that, in order to capture some properties of the transport, larger simulation domains are necessary (cf. sec. 7.1). In order to overcome this problem, one can use the statistical properties of the box and find an algorithm that generates a larger morphology with the same properties. See Baumeier *et al.* [14] for a mathematical and Kordt *et al.* [54] for a more physics based approach to this problem.

2.1.2 Energetic disorder

After computing a sample morphology, we are not done yet. In order to understand how charges are transported, we need to revisit each individual molecule, this time with a static background given by the other molecules. We already mentioned earlier, that because of the absence of long-range order and periodicity, the wavefunctions of the electrons will be localized to a narrow spatial area – typically (some part of) a molecule.

Therefore, electron transport will happen by the transition of electrons from one place of localization to another. This is referred to as *hopping transport*, or in terms of chemistry, a red-ox reaction. The rates, at which these reactions take place are investigated in sec. 2.2. We anticipate that the rates will depend on the energy needed to charge (positively or negatively) the particular localization site. Another influence is the spatial proximity to the target, measured by the overlap of the wavefunctions (*i.e.* corresponding solutions to the Schrödinger equation eq. (2.1)). This is called the *transfer integral*.

Even if all molecules are of the same type (which need not be the case in actual devices), the energy of an additional (or for the removal of an) electron on that molecule will differ from molecule to molecule. There are several reasons for this, the main ones are the static dipole background and the non-constant polarizability of the surrounding.

The static dipole background stems from the molecules themselves, whose charge distribution will most likely not be spherically symmetric, therefore creating a dipole field. It has been shown by Dunlap *et al.* [37] that this leads to normally distributed energies at the molecules. The non-constant polarizability directly follows from the inhomogeneous distribution of the molecules and their random arrangement - therefore the surrounding electron clouds (which are polarizable) will vary for different positions in space. The influence of this is usually smaller than that of the static dipoles. For a more complete account of the different types of disorder and their calculation see the Ph.D. thesis of May [67].

Supported by the findings of ref. 37 and following the GDM (sec. 3.2), the energies are modeled as normally distributed. The variance depends on the particular material and can in principle, using the procedure we just outlined, be computed from just the knowledge of the chemical formula of the molecule. However, since this involves many computations and approximations, the results are not necessarily very accurate.

2.2 Microscale charge transport

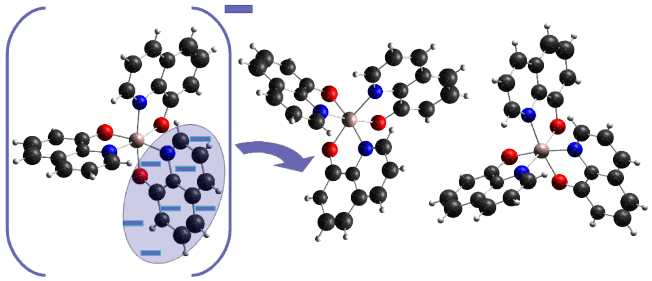

Figure 2.2: Comic of charge transport in Alq_3, a blue-emitting organic semiconductor

In the previous section, we have seen that transport in organic semiconductors happens via hopping of the charge carriers between localized states, this is illustrated in fig. 2.2. The rates of these transitions depend on how energetically favorable it is for the charge carrier to make that particular move. Therefore, the time spent on a molecule before going to a neighbor is essentially a function of to the free energy E_x of the carrier on that particular molecule, labeled by its center of mass at x here and in the following. We will refer to E_x as the *site energy at* x.

We now look a bit more closely at two widely used rates:

i) The *Miller-Abrahams* (MA) rate (see Miller & Abrahams [73]),

$$r_{x\to y}^{(\mathrm{MA})} = \nu_0 \exp\left(\gamma_{\mathrm{decay}} d_{xy}\right) \min\left\{\exp\left(-\frac{E_x - E_y}{k_B T}\right), 1\right\} . \tag{2.2}$$

This rate is also known as the Metropolis rate. ν_0 is the *attempt-to-escape frequency*, d_{xy} the *distance* of origin an destination molecule γ_{decay} is related to the inverse of the *wave function decay length*. T is the *absolute temperature* and k_B the *Boltzmann constant*.

ii) The *Marcus* rate (see Marcus [64]),

$$r_{x\to y}^{(\mathrm{Marcus})} = \frac{2\pi |J_{xy}|^2}{\hbar\sqrt{4\pi\lambda k_B T}} \exp\left(\frac{(E_x - E_y + \lambda)^2}{4\lambda k_B T}\right) , \tag{2.3}$$

J_{xy} is the transfer integral, computed from the wavefunction overlap of initial and target state, λ the reorganization energy. The orbital overlap is symmetric, so $J_{xy} = J_{yx}$. To understand λ and the rates involved, one can consider molecules as harmonic potential wells, see fig. 2.3.

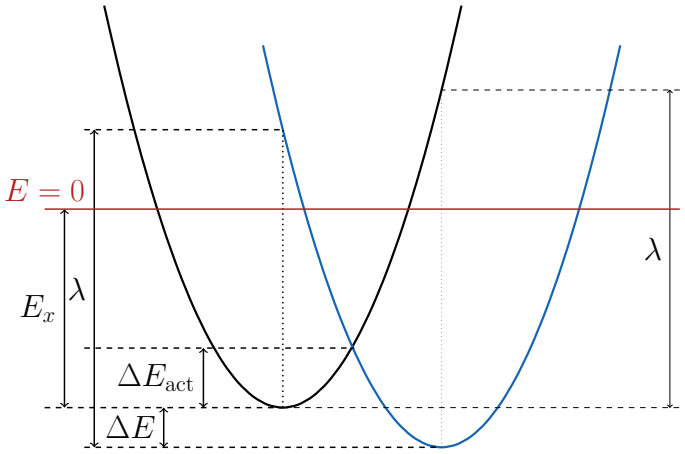

Figure 2.3: The Marcus and Miller-Abrahams rates can be understood from considering the molecules as harmonic oscillators (*i.e.* with a parabolic potential). The energy barrier assumed in the MA rate is ΔE, the one in the Marcus rate ΔE_{act}.

More important for us than the explicit formula of the above rates is the fact that both dynamics are reversible with respect to the same measure, which at $x \in \mathbb{Z}^3$ has the weight $e^{-\beta E_x}$, with $\beta = 1/k_B T$ the *inverse temperature*. Indeed:

$$\begin{aligned} r_{x\to y}^{\text{(MA)}} &= \nu_0 \exp\left(\gamma_{\text{decay}} d_{xy}\right) \exp\left(\frac{|E_x - E_y|}{2k_B T}\right) \exp\left(-\frac{E_x - E_y}{2k_B T}\right) \\ &= \underbrace{\nu_0 \exp\left(\gamma_{\text{decay}} d_{xy}\right) \exp\left(\frac{|E_x - E_y|}{2k_B T}\right) \exp\left(\frac{E_x + E_y}{2k_B T}\right)}_{\text{symmetric}} \exp\left(-\frac{E_x}{k_B T}\right) \end{aligned} \tag{2.4}$$

and

$$\begin{aligned} r_{x\to y}^{\text{(Marcus)}} &= \frac{2\pi |J_{xy}|^2}{\hbar\sqrt{4\pi\lambda k_B T}} \exp\left(-\frac{\lambda}{4k_B T}\right) \exp\left(-\frac{(E_x - E_y)^2}{4k_b T}\right) \exp\left(-\frac{E_x - E_y}{2k_B T}\right) \\ &= \frac{2\pi |J_{xy}|^2}{\hbar\sqrt{4\pi\lambda k_B T}} \exp\left(-\frac{\lambda}{4k_B T}\right) \times \\ &\qquad \underbrace{\times \exp\left(-\frac{(E_x - E_y)^2}{4k_B T}\right) \exp\left(\frac{E_x + E_y}{2k_B T}\right)}_{\text{symmetric}} \exp\left(-\frac{E_x}{k_B T}\right). \end{aligned} \tag{2.5}$$

In sec. 4.1, we will use this structure to justify using a simpler model for our analysis.

After the discussion of this section, we can describe charge transport at the microscopic level as a hopping process, which hops from a site at x to y at rate $r_{x\to y}$. We

can calculate the current density, *i.e.* the number of charge carriers passing through a plane in a given time interval either directly by simulating the above stochastic process – this is called the Monte-Carlo approach (see sec. 7) or by solving the master equation (Ambegaokar *et al.* [2]), setting $p(x,t) := \mathbb{P}\left(x \text{ is occupied at time t}\right)$

$$\frac{\partial\, p(x,t)}{\partial t} = \sum_y r_{y\to x}(1-p(x,t))p(y,t) - \sum_y r_{x\to y}(1-p(y,t))p(x,t), \tag{2.6}$$

for all *sites* x in the simulation area.

Both methods are computationally prohibitive for full devices. It is however possible to simulate a small part of a device – because the devices are often very thin, it is even possible in some cases to simulate the full longitudinal dimension. However, since the lateral dimensions are about $1\mathrm{mm}$, there are of the order of 10^6 sites along each of these dimensions, full devices are not feasible.

CHAPTER 3

Overview of macroscopic transport models

We have discussed models for microscale transport in the previous chapter. In this chapter we give an overview of models which have been used to describe charge transport on the device scale.

Most of the literature on macroscopic transport models for organic electronics falls into one of three categories – bottom-up, top-down and computational. The bottom-up approach is concerned with deriving simplified transport relations and analytic expressions for the parameters in these equations. We review in sec. 3.1 some approaches based on Green's functions, percolation theory and the effective medium approximation, as well as more heuristic approaches based on the concept of transport energy.

The top-down approach is mainly concerned with reproducing experimental data. Usually it is based on the assumption that transport on the scale it is observed, is described by the drift diffusion equation, that is for the *current* j,

$$j = \mu\rho F + D\nabla\rho. \tag{3.1}$$

Here, μ is the so-called *mobility*, D the *diffusion constant*, ρ the *charge carrier density* and F the *electric field*. Diffusion constant and mobility are assumed to be connected by a (generalized) Einstein relation

$$D = f(\rho)k_BT\mu, \tag{3.2}$$

in the normal case, $f(\rho) \equiv 1$.

Most top-down approaches take into account in some heuristic way the nature of microscale charge transport described in chap. 2. This usually includes assumptions about the distributions of site energies, field- and density dependence of the mobility. It is not possible to cover all mobility models, so we restrict to the most widely used ones.

Finally, somewhere in between these two approaches are the computation focused ones. These usually build upon the same microscopic models as the bottom-up models,

couple them with assumptions about the transport coming from experiments and try to link the free parameters (such as the mobility) to the microscale physics by simulation. The most influential line of work here is the family of Gaussian disorder models.

Digression: The density of states

A concept, which is widely used in solid state physics is the density of states (DOS) [7]. It usually refers to the density of quantum mechanically possible states over *e.g.* an energy or wavevector coordinate. However, the full density of states usually exists as possible states anywhere in physical space (as long as the model considered holds for this point in space).

In organic semiconductors, the same name is used to describe the *probability distribution* of the site energies. However, one has to always keep in mind that in an actual device, at any point in physical space, only one of these energies will be realized. Even more, will be fixed for all time – *i.e.* when a charge carrier visits the place again, the energy necessarily will be the same as on any visit before or thereafter.

It is not problematic to talk about the DOS of an organic semiconductor as long as one keeps this in mind. However, some of the theoretical approaches confuse this at some point, either assigning the full density of states to every molecule/site, or allowing one particular point in space to have different energies upon different visits. Both could of course be understood as an approximation technique, but it is rarely stated as such and – at least in the first case – not a good approximation since it neglects the percolative nature of charge transport.

Remark. We actually show (prop. 5.2.7) that the second technique *can be* a reasonable approximation, as long as we keep the memory of the energy resp. waiting time while the charge carrier is still nearby. Once it has been sufficiently far away, the waiting times become essentially independent again.

3.1 Bottom-up/analytical approaches

It seems that there is not much recent literature on bottom-up approaches, most of the work having been done in the 1980s and 1990s. However, the materials used in organic semiconductors have changed. In particular, some materials – which are in the focus of this work – exhibit a much stronger disorder than the first organic semiconductors had. The large disorder is particularly pronounced in blue emitting, phosphorescent molecules, which are very important for lighting applications.

Especially for those materials, there is a wide consensus that for materials with very large disorder, hopping motion is the correct microscopic description. Some earlier models, borrowing from inorganic semiconductors, assume a so-called “mobility edge”, *i.e.* the existence of delocalized states, to which charge carriers can be excited and then

transported efficiently.

While not applicable to the microscopic problem as we have posed it, these theoretical considerations are still interesting, because they can lead to a fractional differential equation as macroscopic description. In particular, most literature on fractional diffusion equations contains references to the continuous time random walk (CTRW) model we introduce in sec. 3.1.5.

3.1.1 Effective medium theory

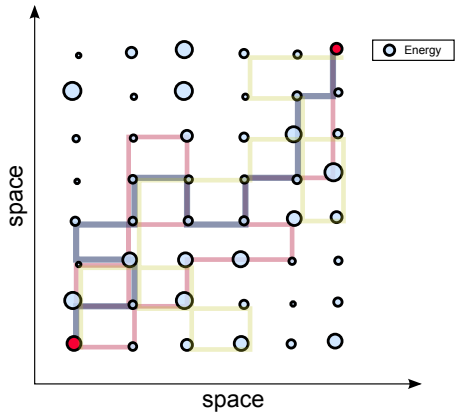

Figure 3.1: The effective medium approximation uses techniques from theoretical physics to sum over all possible paths a charge carrier can take (three paths shown as illustration).

A major line of theoretical investigation was carried out by Movaghar, Grünewald and co-workers [44, 79–82]. Their approximation techniques are referred to as effective medium approximation. The starting point of their theory are the Green's functions of the master equation (eq. (2.6)). They use perturbation theory for the Green's functions, which are expanded into the contributions of different possible paths the charge carrier can take, this approach is visualized in fig. 3.1. Via renormalization techniques developed for tight binding problems in solid state physics, they arrive at the conclusion that hopping transport admits a dispersive regime which can formally explained by both hopping and multiple trapping models.

The arguments and results are often complex and stated on the level of Laplace transforms. This makes it hard to understand, as Baranovskii *et al.* [9] put it, the results are reached "by very sophisticated and nontransparent analytical calculations".

Therefore, it is not always easy to tell the exact mathematical assumptions that enter the approximations borrowed from other branches of theoretical physics and sometimes argued not as approximations, but rather as analogies. Nevertheless, they develop a

powerful formal machinery which reaches tractable expressions keeping both, the diffusive and the percolative nature of the problem. Powerful enough to both predict a dispersive regime and the correct Temperature dependence of the equilibrium transport parameters of the system.

However, maybe due to the lack of transparency in the statement of the results criticized in ref. 9, most of their work is only rarely used in contemporary approaches to the problem. Another drawback is, that it is hard to "modularize" these methods and add different organic materials or physical effects, such as charge carrier interaction or boundary conditions.

Note. We hope that the relatively simple to state results of this work do not suffer the same fate despite the rather technical route we have to go to rigorously prove them.

3.1.2 Percolation theory

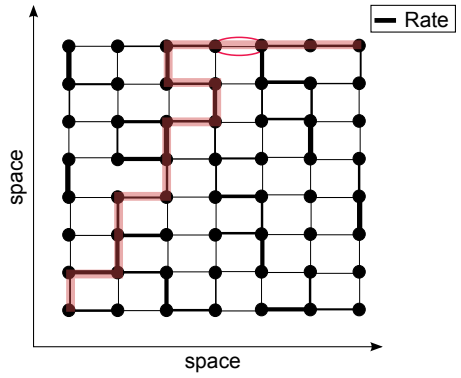

Figure 3.2: Percolation theory applied to charge transport: most current is assumed to flow along a percolating cluster of high conductances. The smallest conductance needed to get a percolating cluster is the critical conductance (circled red).

Percolation theory is based on the idea that most current flows through a few highly conducting paths (in the language of this work, paths that do not contain deep traps). These paths are found by removing all connections and successively adding them back starting with the best conducting ones until a percolation cluster is formed – see fig. 3.2.

Ambegaokar *et al.* [2] were among the first to apply this idea to a charge transport problem, they assumed a uniform distribution of energies over an interval. Vissenberg & Matters [117] applied it to an exponential distribution of the site energies.

Recently, Cottaar *et al.* [34] and Cottaar *et al.* [35] picked up this idea and applied it to a Gaussian distribution of energies, comparing in addition Miller-Abrahams (eq. (2.2)) and Marcus (eq. (2.3)) rate expressions. They augment the pure percolation theory by

adding some paths, which are "nearly critical" – the concept is called *fat percolation* and was introduced by Schrøder & Dyre [102].

3.1.3 Multiple trapping models

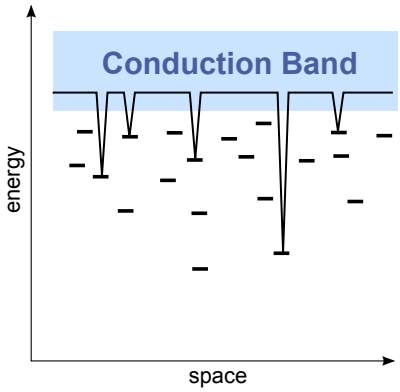

Figure 3.3: Illustration of charge transport in the multiple trapping and release model.

The multiple trapping and release (MTR) models assume the existence of some "extended states", *i.e.* delocalized quantum states, in which charges can be transported efficiently [4, 93, 94]. This transport is assumed to be interrupted, as charge carriers get trapped in localized, low energy states, usually these states are assumed to have an exponential energy distribution. The carriers are assumed to stay trapped for an exponentially distributed time with mean e^{E/k_BT} until they are released.

In order to reproduce time-dependent signals [111], or to take into account dispersion effects (see sec. 7.2), sometimes trapping models are coupled to drift-diffusion mobility models such as the EGDM or ECDM . Another reason to do this is the actual presence of traps due to impurity or degradation effects [100].

3.1.4 Transport energy based approaches

The concept of transport energy can be viewed as an attempt to make the arguments of percolation theory local in some sense, while arguing more along the lines of physical plausibility than abstract mathematics. However, one still needs to be careful when interpreting the transport energy. While physical plausibility is at the core of its derivation, its quantitative value differs from the actual energy charge carriers have when transported [86].

The concept was developed by Monroe [76] for hopping among exponentially distributed energies. The work shows, that even though no extended states are assumed to

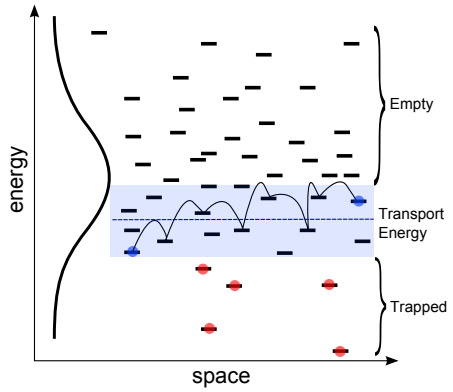

Figure 3.4: Illustration of charge transport in the transport energy model

exist, a multiple trapping model (see sec. 3.1.3) adequately describes charge transport, when the mobility edge is replaced by a certain transport energy.

The idea is that hops with higher energies are unlikely, since these energies are not populated, while hops at lower energies are unlikely, because the waiting times at sites with low energy are exponentially longer – we will see this theme again more rigorously in our proof, when we introduce the coarse graining scheme in sec. 5.1 – the idea is visualized in fig. 3.4.

Baranovskii *et al.* [9] use a similar concept to simplify the transport problem for a Gaussian energy distribution – they correctly acknowledge, that their approach only works if the relaxation time of the charge carriers is smaller than the time needed to exit the device. This approach can also account for the temperature dependence of mobility. It was extended to include density- and field dependence effects [5, 10]. Oelerich *et al.* [86] investigates how the transport energy can be extracted from computer simulations.

3.1.5 Continuous time random walk

The continuous time random walk (CTRW) is a stochastic model for transport in disordered media introduced by Montroll & Weiss [77] as a random walk on a lattice with random waiting times, drawn from a waiting time distribution independently in every step.

Remark. In a mathematical sense, the original CTRW [77] is the "completely annealed" [15] version of the model we consider in chap. 4. Completely annealed means in this context that the randomness of the energies is not fixed and attached to the sites as we do in our model (this also-called the quenched case). In the annealed model, the spatial correlations in the time increments are completely neglected, which makes the scaling limit considerably easier (though the result is the same).
This gives another way to interpret the main result of this work. One can also read it as

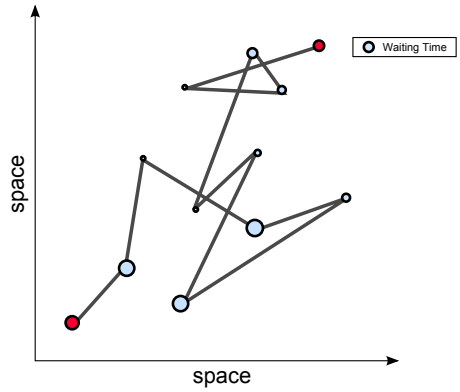

Figure 3.5: Illustration of charge motion in the CTRW model.

"the spatial correlations in the waiting times vanish faster than the dispersive effects" (*i.e.* the incomplete sampling of the statistics of the waiting times).

The idea was extended by Scher & Lax [98] and Scher & Montroll [99] to include arbitrary distributions of spatial increments and applied successfully to charge transport in inorganic amorphous semiconductors. Figure 3.5 illustrates the continuous space version of this model.

Noolandi [85] shows the formal equivalence of CTRW and MTR (see sec. 3.1.3) models. In the light of our results, it is worth mentioning that under certain conditions, the CTRW has exactly the same continuum limit as the process we consider here [70].

3.1.6 Fractional diffusion

The idea of fractional diffusion has been around for a long time and was rediscovered on different occasions. However, this work is the first to rigorously connect fractional diffusion with the full 3-dimensional microscale process.

Fractional diffusion, as already mentioned, emerges naturally in CTRW models (see *e.g.* Metzler & Klafter [72]) and has been applied to charge transport in organic semiconductors by Sibatov & Uchaikin [105]. They show how it can explain many transport phenomena and in a the review paper [106], they give an overview of fractional approaches to charge transport problems.

In a later paper [107], they suggest an integral operator with exponentially truncated Lévy kernels instead of fractional derivatives. This approach covers the long time regime which is diffusive too while have the correct intermediate scale behavior (dispersive). However, the type of transition (exponential truncation) is completely heuristic and no connection between molecule properties and transport parameters is made.

3.2 Simulation based aproaches

A very straightforward way to transition from micro- to macroscale is simulation. While devices are only a few hundred nm (*i.e.* molecules) across in transport direction, their size is several mm in the other directions, which makes direct simulation computationally impossible.

What can be done, is either periodic boundary conditions in all directions or the full size in transport direction and periodic boundary conditions in the others. However, as we will see, with a large disorder, finite size effects can be really severe, even in those cases. This problem is addressed *e.g.* by Lukyanov & Andrienko [61]. In sec. 7.1 we present a modification of the kMC algorithm, which can be used to investigate these effects in simplified situations.

Remark. The main result of this thesis can, loosely speaking, be seen as kind of the inverse to the above problem. Models found as long time limits of the system are free of finite size effects, but real world, thin film devices, are not.

3.2.1 The Gaussian disorder model

Discrete stochastic models for charge transport with rates depending on an energy coordinate, which has Gaussian distribution, are in the literature often referred to as the *Gaussian disorder models*. It is a widely used theoretical approach (see secs. 3.1.1, 3.1.4). Besides analytical calculations, one can use the kinetic Monte Carlo technique with this model to extract diffusion constants or charge carrier mobilities from the microscopic model [101] – a good summary of this line of work is the review by Bässler [13].

This, of course relies on the assumption that the corresponding macroscopic model is a diffusive one. Typically, drift-diffusion-Poisson systems [116] are used to describe charge transport on the device level.

The three dimensional case was studied computationally in ref. 13, where the results of kinetic Monte Carlo computations are compared to experimental ToF data and found in good agreement. A slight drawback (see ref. 41) is that the Poole-Frenkel behavior (*i.e.* $\log \mu \sim \sqrt{F}$) is only found for a narrow range of electric fields. This contradicts experimental data, which is described well using this kind of relationship (see sec. 3.3.2).

While the model is stochastic, to our knowledge it has not yet been treated with the methods of probability theory anywhere prior to this thesis.

3.2.2 Extensions to the GDM

Gartstein & Conwell [41] extend the GDM to the case of spatially correlated disorder (*correlated disorder model* (CDM)). Their simulations with this model exhibit the Poole-Frenkel type field dependence over a wider range, resolving this controversy.

Pasveer *et al.* [88] add charge carrier interaction effects to the computational model. To efficiently solve the model for high carrier densities, an iterative solution of the master equation is used rather than kinetic Monte Carlo. The same work gives the mobility as a relatively simple function approximating the computation results – this approximation is known as the extended GDM mobility function (EGDM). Coehoorn *et al.* [33] compares the new model to various other mobility functions, giving a good review of those approaches.

The same is done for the CDM by Bouhassoune *et al.* [24], extending the CDM to multiple carriers as well. Again, an approximate mobility function in compact form is given, which is known as the extended correlated disorder model (ECDM) mobility function.

In both cases a strong dependence of the mobility on the charge carrier density is found. The physical origin of this effect is the gradual filling of the deep traps as the charge carrier density is increased. Once most of the deep states are occupied, all but a few (those in the traps) carriers can move at a much higher speed because the probability of getting trapped approaches zero.

The resulting mobility models are considered the state-of-the-art for macroscopic modeling and are incorporated into commercial device modeling software (see Knapp *et al.* [51], Nitsche *et al.* [83]).

It has on many occasions been demonstrated (see *e.g.* refs. [50, 100, 115]) that drift-diffusion-Poisson equations with mobility/diffusion coefficients extracted from kMC data are capable of describing realistic devices to some extent.

3.3 Experimental/heuristic approaches

We want to mention here two widely used transport models, which are not directly connected to the microscopic transport. While the effects they describe are observed in experiments, the mechanics behind the emergence of these effects are different as the models were developed for inorganic semiconductors.

3.3.1 Space charge limited current

The concept of space charge limited current (SCLC) was first considered by Child [30] for vacuum. Mott & Gurney [78] show, how to extend the idea to insulators resp. semiconductors with a constant mobility. Leading to the Current-Voltage relation

$$j_{\mathrm{SCLC}} = \frac{9}{8}\varepsilon_r\varepsilon_0\mu\frac{V^2}{d^3}. \tag{3.3}$$

The idea behind this approximation is that one would expect ohmic behavior at very high voltages (*i.e.* $j \sim V$), when the field is strong enough to ionize the molecules.

However, at small fields, charges are still bound to the molecules and build up a field acting against the external potential. This charge is basically integrated up over the device and subtracted from the external potential. Finally the field is assumed to be constant throughout the device, and we arrive at the above expression for the current.

This is obviously a coarse approximation, nevertheless it is still widely used because it is a step in the right direction coming from the ohmic law.

3.3.2 Poole-Frenkel

The Poole-Frenkel effect (Frenkel [39]) describes the influence of an electric field on charge carrier mobility. The physical reasoning behind this is, that in the presence of a field, the typical energetic barrier a charge carrier has to overcome is smaller. The relationship between field and mobility is assumed to be exponential in the square root of the field

$$\mu \propto \exp\left(\gamma\sqrt{F}\right). \tag{3.4}$$

Here, γ is some constant that depends mainly on the dielectric properties of the material.

The original argument was that thermal excitation into the conduction band is aided by being pulled by the field. In our microscopic model, the field would directly enter the rate, and this can be seen as a try to capture the average additional nonlinear effect of this (in addition to the linear response – drift motion – considered anyway).

The general idea to include an exponential dependence on the squre root of the electric field has been quite fruitful in describing charge transport in organic electronics – see for example the Gaussian disorder models sec. 3.2.

PART II

Mathematical Model and Scaling Limit

CHAPTER 4

Mathematical model and main result

In this part, we use the information about the microscopic physical process of charge transport (sec. 2.2) to formulate a simplified mathematical model. Within this model, we will then show, that on certain scales (for time, space and energy) the interaction-free charge transport process admits a continuum limit whose evolution is governed by a fractional diffusion equation. We now introduce the model and discuss the simplifying assumptions we make. Then we state the principal theorem and give an outline of the proof. The rigorous proof is given in chap. 5.

We recall the Marcus (eq. (2.3)) and Miller-Abrahams rates (eq. (2.2)) are both reversible w.r.t. the measure with weights $e^{-\beta E_x}$. In order to make the model more easily treatable, we will completely drop the symmetric part in the mathematical model. This can be made plausible in several ways:

- Within the organic electronics community, the choice of the correct rate is still subject to discussion. One thing both rate-expressions have in common is the reversible measure.

- Handwavingly, one can say that the symmetric part can not trap the random walk, since the probability to go into an area where all outgoing rates are low is small. This is supported by the simulation study of Cottaar *et al.* [35] and Massé *et al.* [66], who find no qualitative impact of the particular choice of rate function. We go even further and show that completely dropping the symmetric part doesn't change the essential properties of charge transport (see fig. 7.10).

- For a similar model, it has been shown that the scaling limit is the same, regardless of the symmetric part of the rates by Barlow & Černý [11].

Furthermore, we reduce the real locations and connectivity of the molecules to a lattice. This approach has usually been taken already for Monte Carlo studies (*e.g.* [13,

24, 32]). The effect of replacing a realistic morphology by a simple cubic lattice was investigated in refs. 35, 66 and found to not make a qualitative difference.

We will comment on on the treatment of the model without some of these simplifications in chap. 8.

Before starting, we introduce some general notation and definitions.

- We say that for $x \to m \in \bar{\mathbb{R}}$, that

$$f(x) = \mathcal{O}\left(g(x)\right) :\Leftrightarrow \limsup_{x \to m} g(x)^{-1} f(x) \leq \infty.$$

- Similarly, $f(x) = o(g(x))$ if $\lim_{x \to m} g(x)^{-1} f(x) = 0$.
- $f(x) \gg g(x) :\Leftrightarrow$ there exists some $\delta > 0$ such that $g(x) = \mathcal{O}\left(f(x) e^{-\delta x}\right)$.
- It is often convenient to use $a \wedge b := \min\{a, b\}$ and $a \vee b := \max\{a, b\}$.
- By $B_x(r)$ we denote a ball of radius r around x:

$$B_r(x) = \left\{y \,:\, \mathrm{dist}_2\left(x, y\right) < r\right\}.$$

- By $Q_r(x)$ we denote the cube with side-length r centered at x

$$Q_r(x) = \left\{y \,:\, \mathrm{dist}_\infty\left(x, y\right) < r\right\}.$$

Definition 4.0.1. The *gamma function* is defined for all $x \in \mathbb{C}$ as

$$\Gamma(x) = \int_0^\infty t^{x-1} e^{-t} \mathrm{d}t.$$

Remark. The gamma function can be understood as the continuous extension of the factorial, since $x! = \Gamma(x+1)$ for all $x \in \mathbb{N}$.

4.1 Simplified mathematical model

The basis of our simplified model for charge transport is the graph $\mathbb{Z}^3$ with nearest neighbor edges. One can picture each vertex as a molecule, or a part of the molecule, which can hold a charge carrier. We consider *i.i.d.* normal random variables $\tilde{\mathbf{E}}_x \sim \mathcal{N}(0, \sigma)$ defined on the vertices of the graph, which model the energy landscape.
For a more unified treatment, we introduce the family $\mathbf{E}_x \sim \mathcal{N}(0, 1)$ on $\mathbb{Z}^d$ and work with $\tilde{\mathbf{E}}_x \sim \sigma \mathbf{E}_x$ in the following. This way, the mean waiting times τ_x have the simple

form $e^{-\hat{\sigma}\mathbf{E}_x}$, where we used the *effective disorder* $\hat{\sigma} = \beta\sigma$, giving us the opportunity to rescale them with a sequence $\hat{\sigma}(n)$ later on.

The charge transport process is a combination of two processes, namely a simple random walk (SRW) $\left(X_t^{(x)}\right)_{t\in\mathbb{N}}$ on $\mathbb{Z}^3$ starting in $x \in \mathbb{Z}^3$ (see def. A.3.1), which jumps to every neighbor with probability $1/6$ and the so-called clock process, which records the physical time that has passed during the first k steps.

Definition 4.1.1. The *clock process* $(S(k))_{k\in\mathbb{N}}$ for the *effective disorder* $\hat{\sigma}$ is defined as

$$S^{\hat{\sigma}}(k) := \sum_{i=1}^{k} e_i \exp\left(-\hat{\sigma}\mathbf{E}_{X_i}\right) = \sum_{i=1}^{k} e_i \left(\tau_x\right)^{\hat{\sigma}}. \tag{4.1}$$

The e_i are *i.i.d.* exponential random variables with mean 1.

This sum has a complicated correlation structure, because we are summing the independent e_i and $\mathbf{E}_x$ along X_i, which is random itself and may backtrack, *i.e.* revisit sites with the same mean waiting time.

In the following we will write for realization of the random environment $\boldsymbol{\tau} := (\tau_x)_{x\in\mathbb{Z}^d}$ and we will denote the probability measure on the space of environments with $\mathbf{P}(\cdot)$, which is due to the independence of the energies just a $\mathbb{Z}^3$-fold lognormal product measure with parameters 0 and 1.

Since all the increments of the clock process are strictly positive, there exists a right inverse $S^- : \mathbb{R} \to \mathbb{N}$. The charge transport process $Y(t) : \mathbb{R} \to \mathbb{Z}^3$ is then defined as

$$Y(t) := X\left(S^-(t)\right). \tag{4.2}$$

The goal of this work is to find a continuous description for this discrete process in order to model charge transport in organic semiconductor devices. As discussed in chap. 3, the focus in continuous modeling has been on diffusive transport. This, however, does not always give a satisfactory description for very thin film devices when the disorder (measured by the parameter $\hat{\sigma}$) is large compared to the size of the device. In these devices the mobility seems to depend on the thickness (see *e.g.* fig. 7.3, p. 82).

The physical process behind this behavior is the relaxation of the charge carriers. As they move through the device they "discover" molecules with successively more favorable, *i.e.* lower, energies $\mathbf{E}_x$. In terms of waiting times, this means they discover longer and longer waiting times as they move. If the device is very thin, they do not have the time to fully discover the statistics of the energy landscape before exiting the device again. The thickness dependence of the transport properties stems from this fact - the thicker the device is, the more chance the carriers have to encounter a particularly low energy and therefore long waiting time.

It is a well known fact in statistics, that a broad lognormal distribution is in the "bulk" very close to a power law distribution [89]. This means, judging from the waiting

times encountered along a trajectory until exiting the device, it may not be possible to distinguish whether they are drawn from a lognormal or a powerlaw distribution. In chap. 5, we will show that there exists a rescaling of the process, such that it converges to a continuous process that is not a diffusion. The next chapter gives an overview of how we will achieve this.

4.2 Main result and rough outline of the proof

First, we introduce the limiting process.

Definition 4.2.1. The *fractional kinetics* process with index α in dimension d, $\mathrm{FK}_{d,\alpha}$ is defined as

$$\mathrm{FK}_{d,\alpha}(t) := \mathrm{BM}_d\left(V_\alpha^{-1}(t)\right),$$

where $\mathrm{BM}_d(t)$ is a d-dimensional Brownian motion and $V_\alpha(t)$ an independent α-stable subordinator.

For more details on Brownian motion and inverse subordinators, see sec. A.2, the properties of $\mathrm{FK}_{d,\alpha}$ are investigated in chap. 6. The scaling we use is motivated by a result for the distribution of triangular arrays of rescaled lognormal variables.

Note (see cor. A.6.3)**.** For $g(n) := \exp\left(\alpha\hat{\sigma}(n)^2\right)$, $\mathbf{E}_x \sim \mathcal{N}(0,1)$ and any sequence $\hat{\sigma}(n)$, which goes to ∞ as $n \to \infty$, it holds that,

$$\mathbb{P}\left(\exp\left(-\hat{\sigma}(n)\mathbf{E}_x\right) \geq g(n)u\right) = \left(1 + L(n,u)\right)\frac{1}{\sqrt{2\pi}\alpha} g(n)^{-\alpha/2}\hat{\sigma}(n)^{-1}u^{-\alpha},$$

where $L(n,u) \xrightarrow{u\to\infty} 0$ uniformly in n.

With this result in mind, we define

$$Y^{(n)}(t) := \frac{\sqrt{d}}{C_{d,\alpha}r(n)} = X^{(n)}\left(\left(S^{(n)}\right)^{-1}(t)\right), \tag{4.3}$$

where

$$X^{(n)}(t) := \frac{\sqrt{d}}{C_{d,\alpha}r(n)} X\left(\left\lfloor C_{d,\alpha}^2 r(n)^2 t\right\rfloor\right), \tag{4.4}$$

$$S^{(n)}(t) := \frac{1}{g(n)}\tilde{S}^{\hat{\sigma}(n)}\left(\left\lfloor C_{d,\alpha}^2 r(n)^2 t\right\rfloor\right), \tag{4.5}$$

and

$$C_{d,\alpha} = \sqrt{\frac{(2\pi\alpha)^{1/2}}{\Gamma(1+\alpha)\Gamma(1-\alpha)G_d(0)^\alpha}}, \tag{4.6}$$
$$\hat{\sigma}(n) = \hat{\sigma}_0\sqrt{n}, \quad r(n) = \hat{\sigma}(n)^{-1/2}e^{\alpha^2\hat{\sigma}(n)^2/4}, \quad g(n) = e^{\alpha\hat{\sigma}(n)^2}.$$

Here, $G_d(0)$ is the full-space Green's function of the simple random walk on the integer lattice $\mathbb{Z}^d$ (see sec. A.3).

Now we can formulate the main theorem.

Theorem 4.2.2 (Convergence to fractional kinetics)**.** *For* $\mathbf{P}$*-almost every* $\boldsymbol{\tau}$*,* $\alpha < 1$*,*

$$Y^{(n)}(t) \to \mathrm{FK}_{d,\alpha}(\cdot), \tag{4.7}$$

weakly in distribution on $D([0,T],\mathbb{R}^d)$*, equipped with the uniform topology.*

$D([0,T],\mathbb{R}^d)$ is a Skorokhod space (see sec. A.5). For more on the fractional kinetics process $\mathrm{FK}_{d,\alpha}$ and its properties, see chap. 6.

Note that in addition to time and space, we also rescale temperature resp. energy – or from a mathematical viewpoint – the distribution of waiting times, in order to get an interesting limit.

The additional rescaling is needed to make the intermediate behavior "survive" in the limit. This intermediate behavior can be characterized as the regime where the clock process is dominated by very few, very large contributions, whose expected depth depends on the time. The time dependence is caused by the fact that over time, the process can "discover" lower energies.

This is important, because in realistic, small devices, the charge carriers do not have the opportunity to completely explore the statistics of the waiting time landscape. The scaling limit shows, that in this situation, modeling the charge transport as a diffusion is not optimal. Instead, for either small devices, low temperatures, or large energetic disorder, a fractional time derivative is a better approach, as it can account for relaxation effects.

This becomes clear from looking at the definition the fractional derivative (def. 6.1.2). It is an integral over time, and thus introduces temporal memory into the equation. This nicely corresponds to the physical intuition of charge carriers slowing down over time, because they encounter successively lower energies.

However, we would like to point out, that on the microscopic scale, the picture of gradual relaxation is wrong. Due to the discrete nature of transport, every charge carrier will wildly jump up and down in energy as it traverses the device. The quantity that does indeed gradually change over time is the *probability* of finding a particularly low energy resp. long waiting time. This statistical behavior is, what is actually modeled by the fractional derivative.

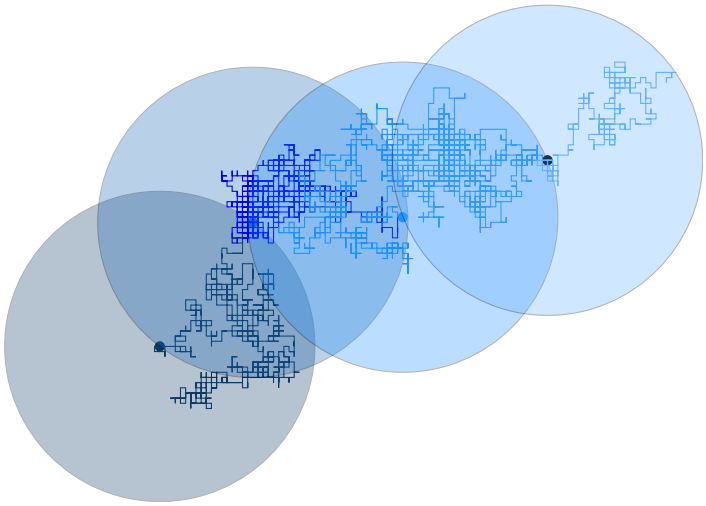

Figure 4.1: Illustration of the coarse graining procedure.

Before we start with the technical details, we give the heuristics that will guide the proof. In this part, we will bold references, in order to facilitate mapping between the heuristic and the rigorous part. For an actual map of the proof, see fig. B.1, p. 129.

Our approach for showing convergence of Y is to show convergence of the tuple $(S^{(n)}, X^{(n)})$ on a suitable product space. We pick the time horizon such that we observe the process only before it exits a ball with radius $R(n) = mr(n)$ for some $m \in \mathbb{N}$. This way we will, for every n, only have to deal with a finite subset of the trapping landscape $\boldsymbol{\tau}$. Our approach for showing weak convergence of the trajectories of the processes is proving convergence of the finite dimensional distributions (**lem. 5.3.3**) and tightness of the sequence $(S^{(n)}, X^{(n)})$ (**lem. 5.3.4**).

If we rescale $X^{(n)}(t) := 1/f(n)X\left(t/f(n)^2\right)$ for some f, which goes to infinity as $N \to \infty$, the limit is Brownian motion started at x (cf. [48, thm. 4.20]). The situation for the clock process $S^{(n)}$ is not as straightforward, but due to the representation of S as a sum of random variables, we are in a good shape. Assuming *i.i.d.* summands, S would converge to a stable subordinator (thm. A.4.9). This scaling also motivates our scaling for the dependent case.

To show convergence of the finite dimensional distributions, we approximate the processes by sums of suitable random variables, *i.e.* $X^n \approx \sum r^n$ resp. $S^n \approx \sum s^n$ (see **prop. 5.2.7** for the approximation of S^n; the approximation of X^n is considerably easier – cf. **eq. (5.36)**). The way we define these approximating random variables is based on the coarse-graining method introduced in ref. 15. We will describe this method now – the full formal description can be found in **sec. 5.1**.

- We group the waiting times between successive exits of the spatial process X of balls with a radius $\rho(n)$. We will refer to the trajectory between two successive

exits as a *segment*. This is illustrated in fig. 4.1

- We show, that with sufficiently high probability, these coarse grained segments have good properties. Sufficiently high in our case means $1 - o\left(h^{-2}\right)$ as $n \to \infty$, with $h(n) = r(n)/\rho(n)$. This is the content of **prop. 5.1.4**:
  - The process visits either no or exactly one deep trap on each segment. Deep trap in this context means a location $x \in \mathbb{Z}^d$ with average waiting time τ_x of order $g(n)$ (see **def. 5.1.1**).
  - This trap is not near the end of the segment nor is it near another trap (measured by the "proximity scale" $\nu(n)$).
  - The probability to actually hit a deep trap is exactly of the order $h(n)^{-2}$.
  - The trap may be visited multiple times, but once the process is at a distance $\nu(n)$, it does not return.
  - Furthermore, every segment ends at least $\nu(n)$ away from a trap.
- We then show (**prop. 5.2.7**) that the RVs $(s^n(x))_{x\in\mathbb{Z}}$, which are essentially defined as the time spent in the set of deep traps on a segment started at x, are good approximations of the total time the charge carrier spends on that segment (*i.e.* they can be used to approximate the increments of S^n).
- Similarly, the RVs $(r^n(x))_{x\in\mathbb{Z}}$, which are defined as having the same distribution as the endpoint of a segment started at x, approximate the increments of X^n (**eq. (5.36)**).
- By getting a fine control on the depth of the first (and, with probability $1-o\left(h^{-2}\right)$, only) trap in **lem. 5.2.1**, we can show that the (conditional) Laplace transform of the $s^n(x)$, rescaled with $g(n)$ converges to $1 + F_d(\lambda)/h(n)^2$ (**lem. 5.2.3**) – F_d is a function that converges to a constant multiple of λ^α.
- As an analogue for the spatial process, we show that the (conditional) Laplace transform of the $r^n(x)$ rescaled by $r(n)$ converges to $1-\xi^2/2dh(n)^2$ (**lem. 5.3.1**).
- In **lem. 5.3.2**, we extend this to the convergence of the joint Laplace transform of the tuple $(r^n/r(n), s^n/g(n))$ to $1 + F_d(\lambda)/h^2 - |\xi|/2dh^2$.

Using the coarse graining procedure, we have now shown that X^n and S^n can be well approximated by $\sum r^n$ and $\sum s^n$ respectively. Further we have computed the (rescaled, conditional) Laplace transforms of the r^n and s^n. In **lem. 5.3.3** we use this knowledge, to conclude the convergence of the finite dimensional distributions of (X^n, S^n) to those of $(\mathrm{BM}_d, V_\alpha)$.

The basic idea in this is, that the number of segments (*i.e.* steps with length ρ in a random direction) before reaching the boundary of a disk of size of $R = mr$ is by the law of large numbers proportional to $h^2 = r^2/\rho^2$ for large n. Using the Markov property of the tuple (X^n, S^n), the joint Laplace transform of the increments can be transformed into a product. Then, we approximate the increments with the independent families r^n resp. s^n which turns the Laplace transform into a power. Thus, the joint Laplace transform of the tuple (X^n, S^n) looks like

$$\left(1 + F_d(\lambda)/h^2 - |\xi|^2/(dh^2)\right)^{h^2} \xrightarrow{h\to\infty} e^{F_d(\lambda) - |\xi|^2/d}.$$

Because $F_d(\lambda) \xrightarrow{n\to\infty} \lambda^\alpha$, this converges to the joint Laplace transform of a stable subordinator and an independent d-dimensional Brownian motion.

What is left to show for the convergence of (X^n, S^n), is tightness of this sequence. This is done in **lem. 5.3.4** and allows us to conclude that the tuple also converges in distribution on the product space of the space- and time-trajectory spaces, equipped with the product topology. The final step is to show that convergence of S^n also implies convergence of its inverse, and that the composition of X^n and the inverse of S^n converges to the composition of the individual limits. Both of these are achieved using the *continuous mapping approach*, introduced in sec. A.5, which finishes the proof.

We need to modify the original method used by Ben Arous & Černý [15], to accommodate the changed scales and in particular the fact that we work with a sequence of random environments instead of a fixed one.

CHAPTER 5

Scaling limit

While we are mainly interested in the process on $\mathbb{Z}^3$, we will in the following use the more general graph $\mathbb{Z}^d$, $d \geq 3$. We expect the result in the case $d = 1$ to be different (see Fontes *et al.* [38] for a similar case) and similar for $d = 2$ with logarithmic correction in the scaling to account for the recurrence of the SRW on $\mathbb{Z}^2$ (see Ben Arous & Černý [15]). But since our physical motivation is in 3 dimensions and the correction terms would invoke additional technical difficulties (we will already need an extra correction term for the rescaling of the random environment, which would interfere), we do not consider this case here.

Note. In the following we will frequently use $C, C_1, C_2, \ldots$ to denote positive real constants. The actual value of these constants may change from line to line. They are generally assumed to be independent of all parameters that are important in the immediate context unless explicitly stated otherwise.

5.1 The coarse graining procedure

The difficult object in this limit is the clock process $S^n(t)$. Due to the correlation introduced by the spatial structure, we can not directly use the rich theory available for *i.i.d.* sums of random variables. We need to partition the sum into blocks in a way that the correlation between the blocks becomes negligible. As already mentioned, for the coarse graining, we follow Ben Arous & Černý [15]. We recall that the main difference is, that they assume

$$\mathbb{P}\left(\tau_x > u\right) = u^{-\alpha}(1 + L(u)), \tag{5.1}$$

for some $L(u)$ which goes to 0 as u goes to infinity. Whereas in this work, we have lognormal energies and therefore (cf. cor. A.6.3) as $n \to \infty$,

$$\mathbb{P}\left(\tau_x^n \geq g(n)u\right) = (1 + L(n, u)) \frac{1}{\sqrt{2\pi\alpha}} g(n)^{-\alpha/2} \hat{\sigma}(n)^{-1} u^{-\alpha}, \tag{5.2}$$

where $\tau_x^n = (\tau_x)^{\sigma(n)}$ are rescaled mean waiting times. Here $L(n,u) \to 0$ as $u \to \infty$ uniformly in n.

In this section we introduce the adapted coarse graining procedure and prove its properties.

5.1.1 Notation and definitions

Tab. 5.1 contains the definitions of the various scales that will be used. To simplify notation, we introduce the additional quantity $\hat{s}(n) := e^{\alpha^2 \hat{\sigma}(n)^2}$. The scale exponents are

$$\gamma = 1/4 - 1/12d^{-1}, \tag{5.3}$$
$$\kappa = 1/4d^{-1}, \tag{5.4}$$
$$\mu = 1/4 - 1/6d^{-1}. \tag{5.5}$$

Name	Symbol	Choice	using $\hat{s}$
spatial scale	r	$\hat{\sigma}^{1/2} e^{\alpha^2 \hat{\sigma}^2/4}$	$\hat{\sigma}^{1/2} \hat{s}^{1/4}$
trap energy scale	g	$e^{\alpha \hat{\sigma}^2}$	$\hat{s}^{1/\alpha}$
coarse graining scale	ρ	$e^{\gamma \alpha^2 \hat{\sigma}^2}$	$\hat{s}^{\gamma}$
proximity scale	ν	$e^{\kappa \alpha^2 \hat{\sigma}^2}$	$\hat{s}^{\kappa}$
intermediate scale	i	$e^{\mu \alpha^2 \hat{\sigma}^2}$	$\hat{s}^{\mu}$
observation scale	R	$mr(n)$	$m\hat{\sigma}^{1/2} \hat{s}^{1/4}$
coarse graining ratio	h	$r(n)/\rho(n)$	$\hat{\sigma}^{1/2} \hat{s}^{1/4-\gamma}$
"expansion" scale	h^{-2}	$(\rho(n)/r(n))^2$	$\hat{\sigma}^{-1} \hat{s}^{2\gamma - 1/2}$

Table 5.1: Overview of the scales involved.

Definition 5.1.1. The set of *deep traps* is defined as

$$T_\varepsilon^M(n) := \{x \in B_0(R(n)) \,:\, \varepsilon g(n) \le \tau_x^n \le M g(n)\}\,.$$

Similarly, we define the *very deep traps* T_M and *shallow traps* T^ε, setting the upper bound to ∞ resp. the lower bound to 0.

We will later show that the time spent in the deep traps is responsible for the main increments of the clock process. It is important to note that we can easily control the probability for some x to be a deep trap. Setting $p_\varepsilon^M = (2\pi)^{-1/2}\,\alpha^{-1}\left(\varepsilon^{-\alpha} - M^{-\alpha}\right)$, we have

$$\begin{aligned}\mathbb{P}\left(x \in T_\varepsilon^M\right) &= \mathbb{P}\left(\tau_x^n > \varepsilon g(n)\right) - \mathbb{P}\left(\tau_x^n > M g(n)\right) \\ &\overset{\text{cor. A.6.3}}{=} (1+o(1)) p_\varepsilon^M \hat{\sigma}(n)^{-1} g(n)^{-\alpha/2}.\end{aligned} \tag{5.6}$$

We want to exclude conglomerates of traps – which we can do as long as we show that the probability of encountering several deep traps in a small spatial region is sufficiently small. We will call these traps "bad" in the following.

Definition 5.1.2. The set of *bad traps* is defined as

$$\mathcal{B}(n) := \left\{x \in T_\varepsilon^M(n) \,:\, \left(B_x(\nu(n)) \setminus \{x\}\right) \cap T_\varepsilon^M \neq \emptyset\right\}. \tag{bad traps}$$

Another set we will need is the set of sites, which are not themselves near a trap (measured by the proximity scale ν).

Definition 5.1.3. We define the *tame sites* as

$$\mathcal{E}(n) := \left\{x \in B_0(R(n)) \,:\, \operatorname{dist}\left(x, T_\varepsilon^M\right) > \nu(n)\right\}. \tag{tame sites}$$

To formally treat the increments of the clock process over the segments, we have to introduce some stopping times. We set $j_1^n := 0$,

$$j_i^n := \min\left\{k > j_{i-1}^n \,:\, \operatorname{dist}\left(X(k), X(j_{i-1}^n\right) > \rho(n)\right\}. \tag{partition of the trajectory}$$

The j_i^n partition the trajectory into parts confined to a ball of radius $\rho(n)$ centered at the last exit point of such a ball. We are now interested in the statistics of the waiting times typically encountered on such a segment. To this end, we define additional random times,

$$\lambda_{i,1}^n := \min\left\{k \geq j_i^n \,:\, Y(k) \in T_\varepsilon^M\right\}, \tag{1st deep trap}$$

the time of the first visit to the set of deep traps on a segment, we denote this trap by $x_i^n := X(\lambda_{i,1}^n)$. Then we define some additional times which formalize the dependence structure, namely

$$\lambda_{i,2}^n := \min\left\{k \geq \lambda_{i,1}^n \,:\, \operatorname{dist}\left(X(k), x_i^n\right) > \nu(u)\right\}, \tag{escape from 1st trap}$$

i.e. the time, when we are no longer within a proximity scale of this trap. The last important stopping time is when the random walk visits the second trap on a segment or the same trap again after escaping from it.

$$\lambda_{i,3}^n := \min\left(\left\{k > \lambda_{i,1}^n \,:\, X(k) \in T_\varepsilon^M \setminus \{x_i^n\}\right\} \cup \left\{k \geq \lambda_{i,2}^n \,:\, X(k) = x_i^n\right\}\right)$$
(hitting 2nd trap or returning to 1st after escape)

We will approximate the increment of the clock process over one segment by the so-called *score* of the segment,

$$\tilde{s}_i^n := \sum_{k=\lambda_{i,1}^n}^{\lambda_{i,2}^n} e_k \tau_k^n \mathbf{1}_{\{X(k)=x_i^n\}}, \tag{5.7}$$

where e_k is a sequence of *i.i.d.* exponential RVs with mean 1. This is the contribution of the visits to the first deep trap encountered on the segment. We later work with something very similar to this to approximate the whole time increase over the segment. It is not a good approximation for all segments, we will introduce some conditions for a segment to be good and show that with sufficiently high probability, all segments are good. We call a segment *good* if the following conditions are fulfilled:

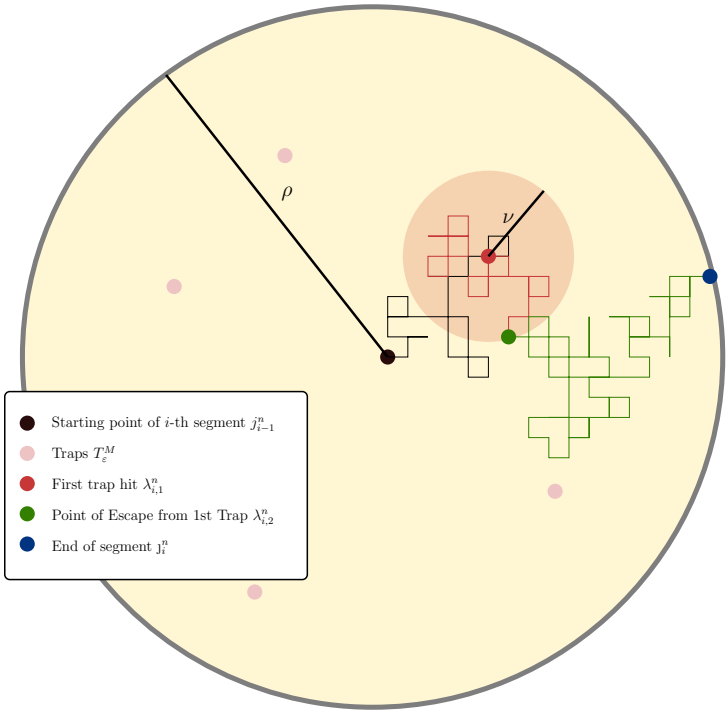

Figure 5.2: Illustration of a good segment

GS1 The distance to the observation horizon (*i.e.* the boundary of a ball with radius $R(n)$ around the origin) is at least one coarse graining scale:

$$\text{dist}\left(x_i^n, \partial B_0(R(n))\right) > \rho(n).$$

GS2 The segment starts and ends on a tame site:

$$X(j_i^n), X(j_{i+1}^n) \in \mathcal{E}(n).$$

GS3 The first trap is not near the boundary (measured again by the proximity scale):

$$\operatorname{dist}\left(x_i^n, \partial B_{X\left(j_i^n\right)}\left(\rho(n)\right)\right) > \nu(n).$$

GS4 The trap is not a bad trap: $x_i^n \notin \mathcal{B}(n)$.

GS5a We want the process to find a trap on the segment, to escape from it and then leave to the next segment. Neither another trap, nor the same trap should be visited again:

$$\lambda_{i,1}^n < \lambda_{i,2}^n < j_{i+1}^n \leq \lambda_{i,3}^n.$$

GS5b The other possibility where the approximation is still good is when no trap is encountered:

$$\lambda_{i,1}^n \geq j_{i+1}^n.$$

See fig. 5.2 for an illustration of the concept of a good segment. Obviously, when GS1,GS2 and GS5b hold, $\tilde{s}_i^n = 0$. In any other case, if one of GS1–GS5 fails, we set it to infinity. The scores we are going to work with are

$$s_i^n = \begin{cases} 0 & \text{, if GS1,GS2 and GS5b hold,} \\ \tilde{s}_i^n & \text{, if GS1–GS5a hold,} \\ \infty & \text{, otherwise.} \end{cases} \tag{5.8}$$

Finally we can define the family $\left(s^n(x)\right)_{x\in\mathbb{Z}^d}$, as the family with the same law as the law of s_i^n conditional on $X(j_i^n) = x$.

For X, we introduce the family $\left(r^n(x)\right)_{x\in\mathbb{Z}^d}$, defined by their law being the same as the law of $X^n(j_{i+1}^n) - X^n(j_i^n)$ conditional on $X(j_i^n) = x$.

In order to avoid explicitly dealing with GS1, we will show the properties only for starting points sufficiently far from the boundary and later argue that the error near the boundary is negligible, simply by choosing m large. To this end, we introduce

$$\mathcal{E}_0 := \left\{x \in \mathcal{E} \,|\, \operatorname{dist}\left(x, \partial B_0(R)\right) > \rho\right\}. \tag{5.9}$$

Halfway of the proof for the scaling limit is the following.

Proposition 5.1.4 (Coarse Graining)**.** *Conditional on the waiting time landscape $\boldsymbol{\tau}$, the probability that $s^n(x) = \infty$ is $o\left(h^{-2}\right)$ uniformly for all $x \in \mathcal{E}_0$ for $\mathbf{P}$-a.e. $\boldsymbol{\tau}$.*

To prove this we will check each of the conditions separately. The proposition follows from the assertions of the next section. Before going into the proof, we show a corollary.

Corollary 5.1.5 (First bad segment appears late)**.** *Let*

$$J(n) := \min\{i \,:\, s_i^n = \infty\}.$$

For all $\delta, T \geq 0$, there exists an m independent of ε and M , such that $\mathbb{P}$-a.s. for large n,

$$\mathbb{P}\left(h(n)^{-2} J(n) \geq T \middle| \boldsymbol{\tau}\right) \geq 1 - \delta.$$

Proof. We have

$$\mathbb{P}\left(0 \notin \mathcal{E}_0\right) = \sum_{x \in B_0(\nu)} \mathbb{P}\left(x \in T_\varepsilon^M\right) \leq C\nu^d \hat{\sigma}^{-1} g^{-\alpha/2} = \mathcal{O}\left(\hat{s}^{-1/2}\right),$$

which is summable. Thus we can apply the Borel-Cantelli lemma to conclude that for large n, $0 \in \mathcal{E}_0$. We apply prop. 5.1.4, which yields $\mathbb{P}\left(s_0^n = \infty\right) = o\left(h^{-2}\right)$. If $s_0^n < \infty$, we know that (GS2) is fulfilled and thus the first segment ends on a site in $\mathcal{E}$.

Even more, since $\operatorname{dist}(\mathcal{E} \setminus \mathcal{E}_0, 0) \geq mr - \rho \gg \rho$, the first segment ends in $\mathcal{E}_0$. We can repeat this argument inductively for all parts of the trajectory until we reach the boundary of $B_0(R)$.

By the law of large numbers and the fact that the expected number of steps to reach $\partial B_0(R)$ is $m^2h^2(1 + o(1))$,we can chose a large m which guarantees that, with probability larger than $1 - \delta/2$, Th^2, segments of the trajectory stay in $B_0(R - \rho)$. By using the discussion above, the probability that at least one of this segments has a score of infinity is $Th^2 o\left(h^{-2}\right) = o(1)$. □

5.1.2 Proof of proposition 5.1.4

We frequently need to locally control the density of the traps, which is achieved in the next lemma. In the following, to keep the equations more readable, we will often omit the dependence of the quantities on n explicitly.

Lemma 5.1.6 (Control of the trap-density)**.** *Let A_n be a sequence of subsets of $\mathbb{Z}^d$, such that $|A_n| \gg Ce^{\alpha^2 \hat{\sigma}(n)^2/2} = \hat{s}^{1/2}$. Then*

$$|A_n \cap T_\varepsilon^M(n)| \in |A_n| p_\varepsilon^M \hat{\sigma}(n)^{-1} g(n)^{-\alpha/2} (1 - \delta, 1 + \delta).$$

for large n.

Proof. We show the above by proving that the bounds are only violated finitely often. We apply the exponential Chebyshev inequality, that is, for $\lambda > 0$,

$$\mathbb{P}\left(|A \cap T_\varepsilon^M| \geq (1 + \delta) p_\varepsilon^M |A| \hat{\sigma}^{-1} g^{-\alpha/2}\right) \leq \mathbb{E}\left[e^{\lambda |A \cap T_\varepsilon^M|}\right] e^{-\lambda(1+\delta) p_\varepsilon^M |A| \hat{\sigma}^{-1} g^{-\alpha/2}}. \tag{5.10}$$

Similarly,

$$\begin{aligned}\mathbb{P}\left(|A\cap T_\varepsilon^M|\le(1-\delta)p_\varepsilon^M|A|\hat\sigma^{-1}g^{-\alpha/2}\right)&\\ =\mathbb{P}\left(-|A\cap T_\varepsilon^M|\ge-(1\pm\delta)p_\varepsilon^M|A|\hat\sigma^{-1}g^{-\alpha/2}\right)&\\ \le\mathbb{E}\left[e^{-\lambda|A\cap T_\varepsilon^M|}\right]e^{\lambda(1-\delta)p_\varepsilon^M|A|\hat\sigma^{-1}g^{-\alpha/2}}.&\end{aligned}\tag{5.11}$$

We compute the expectation using cor. 5.6,

$$\mathbb{E}\left[e^{\pm\lambda|A\cap T_\varepsilon^M|}\right]=\left(1+\left(e^{\pm\lambda}-1\right)\left(1+o(1)\right)p_\varepsilon^M\hat\sigma^{-1}g^{-\alpha/2}\right)^{|A|}.\tag{5.12}$$

Together with $\log(1+x)\le x$, and inserting into eq. (5.10) this yields the bound

$$\mathbb{P}\left(|A\cap T_\varepsilon^M|\ge(1+\delta)p_\varepsilon^M|A|\hat\sigma^{-1}g^{-\alpha/2}\right)\le e^{|A|p_\varepsilon^M\hat\sigma^{-1}g^{-\alpha/2}\left(-\lambda K+(1+o(1))\left(e^\lambda-1\right)\right)}.$$

The term $\left(-\lambda K+(1+o(1))\left(e^\lambda-1\right)\right)$ is 0 for $\lambda=0$ and

$$\frac{\partial}{\partial\lambda}\left(-\lambda(1+\delta)+(1+o(1))\left(e^\lambda-1\right)\right)=-(1+\delta)+(1+o(1))\,e^\lambda,$$

which is negative for a small λ and large n. By Taylor expansion, we can see that the expression itself will be negative for small λ and large n. By the growth condition on the A_n, $|A|p_\varepsilon^M\hat\sigma^{-1}g^{-\alpha/2}<p_\varepsilon^Me^{\delta n}$ for large n. Altogether this gives us

$$\mathbb{P}\left(|A\cap T_\varepsilon^M|\le Kp_\varepsilon^M|A|\hat\sigma^{-1}g^{-\alpha/2}\right)\le e^{-Ce^{\delta n}}.\tag{5.13}$$

Therefore, the probability for violating the upper bound is summable and by the Borel-Cantelli lemma, it is only violated for finitely many n.

We use the same strategy for the lower bound. Equation (5.11) together with eq. (5.12) gives

$$\mathbb{P}\left(|A\cap T_\varepsilon^M|\ge(1-\delta)p_\varepsilon^M|A|\hat\sigma^{-1}g^{-\alpha/2}\right)\le e^{|A|p_\varepsilon^M\hat\sigma^{-1}g^{-\alpha/2}\left(\lambda(1-\delta)+(1+o(1))\left(e^{-\lambda}-1\right)\right)}.$$

The expression $\left(\lambda(1-\delta)+(1+o(1))\left(e^{-\lambda}-1\right)\right)$ is also negative for small λ. This can be seen by expanding the exponential function in a series. The constant term cancels with the -1, the linear term almost cancels with the first part, but for large n, the expression will remain negative. □

Corollary 5.1.7 (Uniform control of the trap density)**.** *Inspecting the proof, one can see that the bound holds uniformly for collections* $\left(A_n^{(m)}\right)_{m\in I_n}$ *as long as* $|I_n|\ll e^{e^{\delta n}}$.

Proof. This can simply be achieved by taking the union over I_n before applying the Borel-Cantelli lemma in eq. (5.13). □

Now we can show the first property of the segments.

Lemma 5.1.8 (GS2). *The probability that the SRW on $\mathbb{Z}^d$ started at $x \in \mathcal{E}_0$ exits $B_x(\rho)$ at a site that is not in $\mathcal{E}$ is $\mathbf{P}$-a.s. $o\left(h^{-2}\right)$ as $n \to \infty$.*

Proof. Let $A_x = B_x(\rho+\nu) \setminus B_x(\rho-\nu)$. Since $|A_x| \sim \rho^{d-1}\nu = \hat{s}^{(d-1)\gamma+\kappa} \gg \hat{s}^{1/2}$, we can use the upper bound from lem. 5.1.6. For some $\delta > 0$ and sufficiently large n

$$|A_x \cap T_\varepsilon^M| \leq (1+\delta)\rho^{d-1}\nu\hat{\sigma}^{-1}g^{-\alpha/2}. \tag{5.14}$$

Therefore, there can be at most $\mathcal{O}\left(\rho^{d-1}\nu\hat{\sigma}^{-1}g^{-\alpha/2}\nu^{d-1}\right)$ points on $\partial B_x(\rho)$, that are not in $\mathcal{E}$. Namely in the case where all traps are exactly on $\partial B_x(\rho)$ and at least 2ν apart - every trap contributes $\mathcal{O}\left(\nu^{d-1}\right)$ points. By prop. A.3.19, the probability for a site on $\partial B_x(\rho)$ to be hit is $\mathcal{O}\left(\rho^{1-d}\right)$. Hence the probability to exit near a trap is

$$\mathcal{O}\left(\rho^{1-d}\nu\hat{\sigma}^{-1}g^{-\alpha/2}\nu^{d-1}\rho^{d-1}\right) = \mathcal{O}\left(\hat{\sigma}^{-1}\hat{s}^{d\kappa-1/2}\right) = o\left(\hat{\sigma}^{-1}\hat{s}^{2\gamma-1/2}\right) = o\left(h^{-2}\right).$$

□

With a similar argument, we can show that traps near the boundary of $B_x(\rho)$ are not important.

Lemma 5.1.9 (GS3). *The probability that a SRW on $\mathbb{Z}^d$ started in $x \in B_0(R)$ hits the set $T_\varepsilon^M \cap (B_x(\rho) \setminus B_x(\rho-\nu))$ before exiting $B_x(\rho)$ is $o\left(h^{-2}\right)$.*

Proof. Since $B_x(\rho) \setminus B_x(\rho-\nu) \subset A_x$, we know from eq. (5.14), that there are at most $\mathcal{O}\left(\rho^{d-1}\nu\hat{\sigma}^{-1}g^{-\alpha/2}\right)$ traps near the boundary. By lem. A.3.18, the probability to hit a certain site y before exiting $B_x(\rho)$ is $\mathcal{O}\left(|x-y|^{2-d}\right)$. For the traps near the boundary, this is $\mathcal{O}\left(\rho^{2-d}\right)$, because $|x-y| > \rho - \nu$.

The probability, that a segment of a trajectory contains a trap near the boundary is therefore

$$\mathcal{O}\left(\rho^{d-1}\nu\hat{\sigma}^{-1}g^{-\alpha/2}\rho^{2-d}\right) = \mathcal{O}\left(\hat{\sigma}^{-1}\hat{s}^{\gamma+\kappa-1/2}\right) = o\left(h^{-2}\right).$$

□

To show the other properties of the segments, we first need more technical results.

Lemma 5.1.10 (Cumulative probability to hit a trap). *For all segments starting in $\mathcal{E}_0$, for any $\delta > 0$, the following holds $\mathbf{P}$-a.s. for large n*

$$\sum_{y \in T_\varepsilon^M} \mathbb{P}\left(\mathsf{Hit}_x^{\{y\}} < \mathsf{Hit}_x^{\partial B_x(\rho)}\right) \in h^{-2} p_\varepsilon^M \left(1-\delta, 1+\delta\right).$$

Note. This lemma (more precisely, this type of lemma, lem. 5.2.1 later improves on this lemma, using the same techniques) is the only time where we do not only have $o\left(h^{-2}\right)$ but where a quantity is exactly h^{-2}. Therefore this lemma is crucial in determining the right scales.

Proof. We partition the sum into three parts

$$\Sigma_1 = \sum_{B_x(\rho-2i)\setminus Q_x(i)} \mathbb{P}\left(\mathsf{Hit}_x^{\{y\}} < \mathsf{Hit}_x^{\partial B_x(\rho)}\right), \qquad \text{(main part)}$$

$$\Sigma_2 = \sum_{Q_x(i)} \mathbb{P}\left(\mathsf{Hit}_x^{\{y\}} < \mathsf{Hit}_x^{\partial B_x(\rho)}\right), \qquad \text{(inner part)}$$

$$\Sigma_3 = \sum_{B_x(\rho)\setminus B_x(\rho-2i)} \mathbb{P}\left(\mathsf{Hit}_x^{\{y\}} < \mathsf{Hit}_x^{\partial B_x(\rho)}\right). \qquad \text{(outer part)}$$

To treat the main contribution (Σ_1) we use a covering of the main part with cubes of side-length i:

$$\Sigma_1 \leq \sum_{z\in H^+} \sum_{y\in Q_z(i)\cap T_\varepsilon^M} \mathbb{P}\left(\mathsf{Hit}_x^{\{y\}} < \mathsf{Hit}_x^{\partial B_x(\rho)}\right),$$
$$H^+ = \left\{z \in i\mathbb{Z}^d \setminus \{x\} \,:\, Q_z(i) \cap B_x(\rho - 2i) \neq \emptyset\right\},$$

for the other inequality, we use $H^- = \left\{z \in i\mathbb{Z}^d \setminus \{x\} \,:\, Q_z(i) \subset B_x(\rho - 2i)\right\}$.

By cor. 5.1.7, we can bound the number of traps in $Q_x(i)$ uniformly for all $x \in B_0(R - \sqrt{2}i)$. This is possible because $|Q_x(i)| = \mathcal{O}\left(i^d\right) = \hat{s}^{d/4-1/6} \overset{d\geq 3}{\gg} \hat{s}^{1/2}$ and $|B_0(R - \sqrt{2}i)| \sim \mathcal{O}\left(r^2\right) = o\left(e^{e^{\delta n}}\right)$. We get

$$|Q_x(i) \cap T_\varepsilon^M| \in i^d p_\varepsilon^M \hat{\sigma}^{-1} g^{-\alpha/2} \left(1-\delta, 1+\delta\right). \tag{5.15}$$

By lem. A.3.18, the probability to hit a certain trap at y before exiting $B_x(\rho)$ is

$$a_d \left(|y|^{2-d} - \rho^{2-d} + \mathcal{O}\left(|y|^{2-d}\right)\right)\left(1 + \mathcal{O}\left(\rho - |y|\right)^{2-d}\right),$$

with $a_d = \frac{d}{2}\Gamma\left(\frac{d}{2} - 1\right)\pi^{d/2}$. For $y \in Q_z(i)$, y, z bounded away from 0, we use a Taylor series

$$|y|^{2-d} = |z|^{2-d} + (2-d)|z|^{1-d}(|y| - |z|) + \mathcal{O}\left(|z|^{-d}\right) = |z|^{2-d} + \mathcal{O}\left(i|z|^{1-d}\right). \tag{5.16}$$

Now we can exchange $|y|^{2-d}$ for $|z|^{2-d}$ in the summation

$$\Sigma_1 \leq a_d \sum_{z\in H^+} \sum_{y\in Q_z(i)\cap T_\varepsilon^M} \left(|z|^{2-d} - \rho^{2-d} + \mathcal{O}\left(i|z|^{1-d}\right)\right)\left(1 + \mathcal{O}\left(\rho - |y|\right)^{2-d}\right).$$

Using the upper bound for the number of traps in each $Q_z(i)$ (eq. (5.15)),

$$\Sigma_1 \leq a_d(1+\delta)p_\varepsilon^M \hat{\sigma}^{-1} g^{-\alpha/2} \sum_{z\in H^+} \left(|z|^{2-d} - \rho^{2-d} + \mathcal{O}\left(i|z|^{1-d}\right)\right)\left(1+\mathcal{O}\left(\rho-|y|\right)^{2-d}\right).$$

Since $|y| < \rho - i$, we have $(\rho-|y|)^{2-d} = \mathcal{O}\left(i^{2-d}\right) = o(1)$.

The summand is smooth away from zero, therefore $\int_{B_x(\rho-2i)\setminus Q_x(i)}$ can be approximated by $\sum_{H^+} + \mathcal{O}\left(i|z|^{1-d}\right)$ (resp. $\sum_{H^-} + \mathcal{O}\left(i|z|^{1-d}\right)$ for the lower bound), thus

$$\begin{aligned}\Sigma_1 &\leq a_d(1+\delta)p_\varepsilon^M \hat{\sigma}^{-1} g^{-\alpha/2} \int_{B_x(\rho-2i)\setminus Q_x(i)} \left(|z|^{2-d} - \rho^{2-d} + \mathcal{O}\left(i|z|^{1-d}\right)\right) \mathrm{d}z \\ &\qquad + \mathcal{O}\left(i|z|^{1-d}\right) \qquad (5.17)\\ &\leq a_d(1+\delta)p_\varepsilon^M \hat{\sigma}^{-1} g^{-\alpha/2}\big(1+o(1)\big)\rho^2\omega_d\left(\frac{1}{2}-\frac{1}{d}\right) \leq (1+2\delta)p_\varepsilon^M h^{-2},\end{aligned}$$

because $a_d\omega_d = 2d/(d-2) = (1/2-1/d)^{-1}$ and $\hat{\sigma}^{-1}g^{-\alpha/2}\rho^2 = h^{-2}$.

Note. This step determines the exact relation between the scales r, ρ and g.

This finishes the upper bound for Σ_1. For the lower bound we use the lower sum, eq. (5.16) and approximate it by an integral again to arrive at

$$\Sigma_1 \geq (1-2\delta)p_\varepsilon^M h^{-2}.$$

Now, all that is left is to show that the inner and outer parts Σ_2, resp. Σ_3 are $o\left(h^{-2}\right)$.

For the inner part Σ_2, we proceed in a similar way and use a covering of $Q_x(i)$. However, we need a more elaborate covering this time to control the contributions of the sites close to x.
We let $k_{\max} = \min\{k \in \mathbb{N} : \hat{s}(k)\nu \geq i\}$, hence $k_{\max} \approx (\mu-\kappa)n$. The covering we use is

$$Q_x(i) \subset \left(\bigcup_{k=0}^{k_{\max}} \bigcup_{y\in\{-1,0,1\}^d\setminus\{0,0,0\}} Q_{\hat{s}(k)\nu y}(\hat{s}(k)\nu)\right) \cup Q_x(\nu).$$

We need to bound the number of traps in one of these cubes. Since $\mathbb{P}\left(x \notin \mathcal{E}_0\right) = o\left(h^{-2}\right)$ by lem. 5.1.8 and $Q_x(\nu) \subset B_x(\nu)$, we can assume $|Q_x(\nu) \cap T_\varepsilon^M| = 0$. For the other cubes, we have

$$\begin{aligned}|Q_x(\hat{s}(k)\nu) \cap T_\varepsilon^M| &\leq n \max\left\{1, \hat{s}\left(k\right)^d \nu^d \hat{\sigma}^{-1} g^{-\alpha/2}\right\} \\ &= n \max\left\{1, \hat{\sigma}^{-1}\hat{s}\left(dk + (d\kappa - 1/2)n\right)\right\}.\end{aligned} \qquad (5.18)$$

To show this, we need to slightly modify the proof of lem. 5.1.6, but the ingredients are the same. The exponential Chebyshev inequality yields, together with $\log(1+x) \leq x$,

$$\begin{aligned}
&\mathbb{P}\left(|Q_x(\hat{s}(k)\nu) \cap T_\varepsilon^M| \geq n \max\left\{1, \hat{s}\,(k)^d\, \nu^d \hat{\sigma}^{-1} g^{-\alpha/2}\right\}\right) \\
&\quad \leq e^{-\lambda n \max\left\{1, \hat{\sigma}^{-1}\hat{s}(dk+(d\kappa-1/2))\right\}} e^{(e^\lambda-1)(1+o(1))p_\varepsilon^M \hat{\sigma}^{-1}\hat{s}(dk+(d\kappa-1/2))} \\
&\quad \leq C_1 e^{-C_2\lambda n}.
\end{aligned}$$

The last inequality holds because the argument of the first exponential grows faster than the one of the second by an additional factor of n, the other terms have the same growth behavior with different constants.

There is enough room to get the bound uniformly for all $x \in B_0(R)$ and $k \in \{1, \ldots, k_{\max}\}$, because $C_1 n r^d e^{-C_2\lambda n}$ is still summable for sufficiently large λ and therefore the inequality is only violated finitely often by the Borel-Cantelli lemma.

Using eq. (5.18) together with lem. A.3.18, we get

$$\Sigma_2 \leq Cn \sum_{k=0}^{k_{\max}} \max\left\{1, \hat{s}\,(k)^d\, \nu^d \hat{\sigma}^{-1} g^{-\alpha/2}\right\} \left(\hat{s}\,(k)\,\nu\right)^{2-d}.$$

If $1 > \hat{s}\,(k)^d\, \nu^d \hat{\sigma}^{-1} g^{-\alpha/2}$, the summand is decreasing in k, thus we can replace that case simply by $\hat{s}(0)^{2-d}\nu^{2-d}$. This yields the bound

$$\Sigma_2 \leq C_1 n^2 \max\left\{C_2 \nu^{d-2}, \hat{s}\,(k)^2\, \nu^2 \hat{\sigma}^{-1} g^{-\alpha/2}\right\}.$$

Both of the possibilities are $o\left(h^{-2}\right)$, thus we are done. Indeed,

$$\nu^{2-d} = \hat{s}^{(2-d)\kappa} = o(\hat{s}^{2\gamma-1/2}) \;\Leftrightarrow\; (2-d)\kappa < 2\gamma - 1/2 \;\Leftrightarrow\; d \geq 8/3,$$

and

$$\hat{s}\,(k)^2\, \nu^2 g^{-\alpha/2} \leq \hat{s}^{2(\mu-\kappa)+2\kappa-1/2} = \hat{s}^{2\mu-1/2} \ll \hat{s}^{2\gamma-1/2},$$

because $\mu < \gamma$. The $\hat{\sigma}^{-1}$ term does not play a role here because the other terms are exponential in $\hat{\sigma}$ and $\hat{\sigma} \sim n$.

The outer sum Σ_3 can be bounded easily. Since

$$|B_x(\rho) \setminus B_x(\rho - 2i)| \sim \rho^{d-1} i = \hat{s}^{(d-1)\gamma+\mu} \gg \hat{s}^{1/2},$$

we can apply lem. 5.1.6. Using that together with the fact that by lem. A.3.18, for $x \approx \rho$, all summands are $\mathbb{P}\left(\mathsf{Hit}_x^{\{y\}} < \mathsf{Hit}_x^{\partial B_x(\rho)}\right) = \mathcal{O}\left(\rho^{2-d}\right)$, we get

$$\Sigma_3 \leq \rho i p_\varepsilon^M \hat{\sigma}^{-1} g^{-\alpha/2} = o\left(h^{-2}\right).$$

Indeed, $\rho i g^{-\alpha/2} = \hat{s}^{\gamma+\mu-1/2} \ll \hat{s}^{2\gamma-1/2}$, because $\mu < \gamma$. □

Now we can turn to GS4.

Lemma 5.1.11 (GS4)**.** *The probability for the SRW started at $x \in \mathcal{E}$ to hit a bad trap before exiting $B_x(\rho)$ is $o\left(h^{-2}\right)$.*

Proof. This time the quantity we are interested in is

$$\sum_{y \in \mathcal{B}(n)} \mathbb{P}\left(\mathsf{Hit}_x^{\{y\}} < \mathsf{Hit}_x^{\partial B_x(\rho)}\right). \tag{5.19}$$

We split up the sum again, keeping the notation the same as in the proof of lem. 5.1.10. This already eliminates Σ_2 and Σ_3 because $\mathcal{B} \subset T_\varepsilon^M$. Now we bound the main part

$$\Sigma_1 \leq \sum_{z \in H^+} \sum_{y \in Q_z(i) \cap \mathcal{B}(n)} \mathbb{P}\left(\mathsf{Hit}_x^{\{y\}} < \mathsf{Hit}_x^{\partial B_x(\rho)}\right).$$

We will show that

$$|Q_x(i) \cap \mathcal{B}(n)| \leq n^2 \nu^d r^d \hat{\sigma}^{-2} g^{-\alpha} =: \phi(n), \tag{5.20}$$

uniformly for all $x \in B_0(R)$. We again use a covering for better control. If there exists $x \in B_0(R)$ such that the Inequality 5.20 is violated, there also exists a cube $Q_z(2i)$ centered on $\mathcal{G} := i\mathbb{Z}^d \cap B_0(R)$ which violates it. Therefore it suffices to show that there exists no cube with side-length $2i$ centered on $\mathcal{G}$ violating (5.20) for sufficiently large n. We have, because the $(Q_x(2i))_{x \in \mathcal{G}}$ do not intersect,

$$\sum_{x \in \mathcal{G}} \mathbb{P}\left(|Q_x(2i) \cap \mathcal{B}| \geq \phi(n)\right) \leq |\mathcal{G}| \mathbb{P}\left(|Q_0(2i) \cap \mathcal{B}| \geq \phi(n)\right).$$

Now we estimate the probability for $x \in \mathcal{B}$ using the union bound

$$\mathbb{P}\left(x \in \mathcal{B}\right) \leq \mathbb{P}\left(x \in T_\varepsilon^M\right) \sum_{y \in B_x(\nu)} \mathbb{P}\left(y \in T_\varepsilon^M\right) = \mathcal{O}\left(\nu^d \hat{\sigma}^{-2} g^{-\alpha}\right).$$

This and the Markov inequality yield

$$\begin{aligned} |\mathcal{G}| \mathbb{P}\left(|Q_0(2i) \cap \mathcal{B}| \geq \phi(n)\right) &\leq |\mathcal{G}| \phi(n)^{-1} \mathbb{E}\left[\sum_{x \in Q_0(2i)} \mathbf{1}_{\{x \in \mathcal{B}\}}\right] \\ &= \mathcal{O}\left((r/i)^d\right) \mathcal{O}\left(n^{-2} \nu^{-d} r^{-d} \hat{\sigma}^2 g^\alpha\right) \mathcal{O}\left((2i)^d \nu^d \hat{\sigma}^{-2} g^{-\alpha}\right) \\ &= \mathcal{O}\left(n^{-2}\right). \end{aligned}$$

Since n^{-2} is summable, eq.5.20 follows from Borel-Cantelli. Therefore we have

$$|Q_x(i) \cap \mathcal{B}(n)| = \mathcal{O}\left(\nu^d r^d g^{-\alpha}\right) = \mathcal{O}\left(\hat{s}^{d\kappa + d/4 - 1}\right) \overset{d \geq 3}{=} o\left(\hat{s}^{d\kappa - 1/2}\right) = o\left(i^d g^{-\alpha/2}\right), \tag{5.21}$$

which is sufficient for showing $\Sigma_1 = o\left(h^{-2}\right)$, because in the proof of lem. 5.1.10, Σ_1 was of the exact order h^{-2}. However, in that case, we were summing equal contributions over a set of size $i^d g^{-\alpha/2}$. Since by eq. (5.21) the set we are summing over is of negligible size compared to that, the quantity we are looking for is $o\left(h^{-2}\right)$. □

Now we turn to proving (GS5). To show this, we first show that the number of distinct traps visited on one segment is typically equal to 1.

Lemma 5.1.12 (GS5a) 1st part)**.** *The probability for the SRW on $\mathbb{Z}^d$ started in $x \in \mathcal{E}_0$ to hit two different traps before exiting $B_x(\rho)$ is $\mathbf{P}$-a.s. $o\left(h^{-2}\right)$.*

Proof. We denote by $X_a^b := \bigcup_{a \le t \le b} X_t$ the set of sites visited by the SRW between steps a and b. We get an upper bound on the probability of hitting multiple traps on a segment by considering the cases where at least one good trap is hit and at least one bad trap is hit separately

$$\begin{aligned}
&\mathbb{P}\left(|X_{j_i^n}^{j_{i+1}^n} \cap T_\varepsilon^M| \ge 2\right) \\
&\quad \le \underbrace{\mathbb{P}\left(|X_{j_i^n}^{j_{i+1}^n} \cap T_\varepsilon^M| \ge 2 \middle| |X_{j_i^n}^{j_{i+1}^n} \cap T_\varepsilon^M \setminus \mathcal{B}| \ge 1\right)}_{\text{(I)}} \underbrace{\mathbb{P}\left(X_{j_i^n}^{j_{i+1}^n} \cap T_\varepsilon^M \setminus \mathcal{B}| \ge 1\right)}_{\text{(II)}} \\
&\qquad + \underbrace{\mathbb{P}\left(|X_{j_i^n}^{j_{i+1}^n} \cap T_\varepsilon^M| \ge 2 \middle| |X_{j_i^n}^{j_{i+1}^n} \cap \mathcal{B}| \ge 1\right)}_{\text{(III)}} \underbrace{\mathbb{P}\left(X_{j_i^n}^{j_{i+1}^n} \cap \mathcal{B}| \ge 1\right)}_{\text{(IV)}}.
\end{aligned}$$

By lem. 5.1.11, (IV) $= o\left(h^{-2}\right)$ – since (III) is bounded, this is also true for the product. For (II) lem. 5.1.10 and the union bound,

$$\mathbb{P}\left(X_{j_i^n}^{j_{i+1}^n} \cap T_\varepsilon^M \setminus \mathcal{B}| \ge 1\right) \le \sum_{x \in T_\varepsilon^M} \mathbf{1}_{\left\{x \in X_{j_i^n}^{j_{i+1}^n}\right\}} = \mathcal{O}\left(h^{-2}\right).$$

This bound is not enough, because we need $o\left(h^{-2}\right)$. This can be achieved by noting that independent of the location of the first trap x_i^n, $B_x(\rho) \subset B_{x_i^n}(2\rho)$. Using this and the strong Markov property,

$$\mathbb{P}\left(|X_{j_i^n}^{j_{i+1}^n} \cap T_\varepsilon^M| \ge 2 \middle| |X_{j_i^n}^{j_{i+1}^n} \cap T_\varepsilon^M \setminus \mathcal{B}| \ge 1\right) \le \sum_{z \in T_\varepsilon^M} \mathbb{P}\left(\mathsf{Hit}_{x_i^n}^{\{z\}} \le \mathsf{Hit}_{x_i^n}^{\partial B_{x_i^n}(2\rho)}\right).$$

This means that for (I), we can use lem. 5.1.10. Therefore (I) $= \mathcal{O}\left(h^{-2}\right)$. Hence, the product (I) × (II) $= \mathcal{O}\left(h^{-4}\right) = o\left(h^{-2}\right)$. The reason we can apply lem. 5.1.10 here is, that instead of $x \in \mathcal{E}$ we now use $x \notin \mathcal{B}$ which guarantees there is no trap in $B_{x_i^n}(v)$. Furthermore, doubling the radius will only change the prefactor, not the asymptotic behavior in n. □

The only thing that is left now, is to bound the probability of returning to the same trap after escaping it (*i.e.* after being more than ν away).

Lemma 5.1.13 (GS5a) 2nd part)**.** *The probability that the SRW started at $x \in \mathcal{E}_0$ hits a trap $y \in B_x(\rho) \cap T_\varepsilon^M$, exits $B_y(\nu)$ and then hits y again is $\mathbf{P}$-a.s. $o\left(h^{-2}\right)$.*

Proof. Because of lem. 5.1.9, we can assume (making an $o\left(h^{-2}\right)$ error) that y is at least ν away from $\partial B_x(\rho)$, thus, $B_y(\nu) \subset B_x(\rho)$.

We group the visits to y before exiting $B_y(\nu)$ and those after. Using the Green's function (cf. eq. A.3.14), we get

$$\begin{aligned} G_{B_x(\rho)}(y,y) &= \mathbb{E}\left[\sum_{i=0}^{\mathsf{Hit}_y^{\partial B_x(\rho)}} \mathbf{1}_{\left\{X_i^{(y)}=y\right\}}\right] \\ &= \mathbb{E}\left[\sum_{i=0}^{\mathsf{Hit}_y^{\partial B_y(\nu)}} \mathbf{1}_{\left\{X_i^{(y)}=y\right\}}\right] + \mathbb{E}\left[\sum_{i=\mathsf{Hit}_y^{\partial B_y(\nu)}}^{\mathsf{Hit}_y^{\partial B_x(\rho)}} \mathbf{1}_{\left\{X_i^{(y)}=y\right\}}\right]. \end{aligned}$$

The first term is $G_{B_y(\nu)}(y,y)$. The second term only counts the visits to y after exiting $B_y(\nu)$. The probability we are looking for is the probability to visit y at all after leaving $B_y(\nu)$. We denote this probability by $P_{\text{return}}(x,y)$. Only those trajectories that return at all do contribute to the second term, thus we can factor out $P_{\text{return}}(x,y)$, since all other trajectories do not contribute. Those who return may however visit y several times before exiting $B_x(\rho)$ again. This quantity is exactly $G_{B_x(\rho)}(y,y)$. Hence

$$\begin{aligned} G_{B_x(\rho)}(y,y) &= G_{B_y(\nu)}(y,y) + P_{\text{return}} G_{B_x(\rho)}(y,y) \\ \Rightarrow P_{\text{return}}(x,y) &= 1 - \frac{G_{B_y(\nu)}(y,y)}{G_{B_x(\rho)}(y,y)}. \end{aligned}$$

We have $G_{B_y(\nu)}(y,y) = G_{B_0(\nu)}(0,0)$ and $G_{B_x(\rho)}(y,y) \leq G_{B_0(2\rho)}(0,0)$. This yields

$$P_{\text{return}}(x,y) \leq 1 - \frac{G_{B_0(\nu)}(0,0)}{G_{B_0(2\rho)}(0,0)} = \mathcal{O}\left(\nu^{2-d}\right) = o\left(h^{-2}\right).$$

Indeed, by lem. A.3.18

$$\frac{G_{B_0(\nu)}(0,0)}{G_{B_0(2\rho)}(0,0)} = \frac{G(0) + \mathcal{O}\left(\nu^{2-d}\right)}{G(0) + \mathcal{O}\left(\rho^{2-d}\right)} = 1 + \mathcal{O}\left(\nu^{2-d}\right),$$

and $\nu^{2-d} = \hat{s}^{(2-d)\kappa} \ll \hat{s}^{2\gamma-1/2}$. □

This concludes the proof of prop. 5.1.4.

5.2 Approximation of the clock process

We have already shown that for fixed ε and M, $\mathbb{P}\left(0 \in T_\varepsilon^M\right)$ converges to an inverse power law. In the following, we will get a finer control on the depth of the first trap on a segment. To achieve that, we partition T_ε^M into smaller parts. The partition will depend on the parameter n, therefore we will need to make the deviation from a true power law explicit and control it. We recall that by cor. A.6.3, we have

$$\mathbb{P}\left(\tau_x^n \geq g(n)u\right) = (1 + L(n,u)) \frac{1}{\sqrt{2\pi\alpha}} g(n)^{-\alpha/2} \hat{\sigma}(n)^{-1} u^{-\alpha},$$

where $L(n,u) \to 0$ as $u \to \infty$, uniformly in n. We set

$$\theta(n) = \frac{1}{n} + \left(\max_{x \geq \varepsilon} |L(n, g(n)x)|\right)^{1/2}.$$

For every fixed ε,

$$\theta(n) \geq \frac{1}{n}, \quad \text{and } \lim_{n\to\infty} \theta(n) = 0.$$

Furthermore, due to the construction, $\max_{x\geq\varepsilon} |L(n, g(n)x)| = o(\theta(n))$.
Let $(z_n(k))_{k\in\{1,\dots R_n\}}$ be a strictly increasing, equidistant sequence with $z_n(0) = \varepsilon$ and $z_n(R_n) = M$, such that

$$\theta(n) < z_n(k+1) - z_n(k) < 2\theta(n). \tag{5.22}$$

Finally, we set

$$p_k^n := \frac{1}{\sqrt{2\pi\alpha}} \left(z_n(k)^{-\alpha} - z_n(k+1)^{-\alpha}\right).$$

Lemma 5.2.1 (Fine control on trap density)**.** *For any $\delta > 0$, almost every $\boldsymbol{\tau}$ and n large,*

i) $\mathbb{P}\left(0 \in T_{z_n(k)}^{z_n(k+1)}\right) \in p_k^n \hat{\sigma}^{-1} g^{-\alpha/2} (1-\delta, 1+\delta)$,

ii) $\forall x \in B_0(R)$, $k \in \{0, \dots, R_n - 1\}$,

$$\left|Q_x(i) \cap T_{z_n(k)}^{z_n(k+1)}\right| \in p_k^n i^d \hat{\sigma}^{-1} g^{-\alpha/2} (1-\delta, 1+\delta).$$

Proof. With the notation introduced above, we have

$$\begin{aligned}
&\mathbb{P}\left(0 \in T_{z(k)}^{z(k+1)}\right) \\
&\quad = \frac{1}{\sqrt{2\pi}\alpha}\hat{\sigma}^{-1}g^{-\alpha/2}\left(\frac{1+L(n,gz(k))}{z_n(k)^\alpha} - \frac{1+L(n,gz(k+1))}{z_n(k+1)^\alpha}\right) \\
&\quad = \hat{\sigma}^{-1}g^{-\alpha/2}\left(p_k^n + \underbrace{\frac{1}{\sqrt{2\pi}\alpha}\left(\frac{L(n,gz_n(k))}{z_n(k)^\alpha} - \frac{L(n,gz_n(k+1))}{z_n(k+1)^\alpha}\right)}_{\text{(I)}}\right).
\end{aligned}$$

As $n \to \infty$, (I) is negligible compared to p_k^n, which behaves as $z_n(k)^{-\alpha} - z_n(k)^{-\alpha}$. Indeed, omitting the dependences on n,

$$\begin{aligned}
\frac{L(k)}{z(k)^\alpha} - \frac{L(k+1)}{z(k+1)^\alpha} &= \frac{L(k)}{z(k)^\alpha} - \frac{L(k)}{z(k+1)^\alpha} + \frac{L(k)}{z(k+1)^\alpha} - \frac{L(k+1)}{z(k+1)^\alpha} \\
&= \underbrace{\underbrace{L(k)}_{=o(1)}\left(\frac{1}{z(k)^\alpha} - \frac{1}{z(k+1)^\alpha}\right)}_{o(z(k)^{-\alpha}-z(k+1)^{-\alpha})} + \underbrace{\frac{L(k)-L(k+1)}{z(k+1)^\alpha}}_{\text{(II)}}.
\end{aligned}$$

For (II), we show convergence to zero for

$$\begin{aligned}
\frac{\text{(II)}}{z(k)^{-\alpha} - z(k+1)^{-\alpha}} &= \frac{z(k+1)^{-\alpha}\left(L(k) - L(k+1)\right)}{z(k)^{-\alpha} - z(k+1)^{-\alpha}} \\
&= \frac{L(k) - L(k+1)}{\left(\frac{z(k)}{z(k+1)}\right)^{-\alpha} - 1}.
\end{aligned}$$

We expand $\left(\frac{z(k)}{z(k+1)}\right)^{-\alpha} = \left(\frac{z(k+1)}{z(k)}\right)^{\alpha}$ into a Taylor series in θ – due to eq. (5.22) for some $c \in (1,2)$,

$$\frac{L(k) - L(k+1)}{1 + \frac{\alpha c\theta}{\varepsilon} + \mathcal{O}\left(\theta^2\right) - 1} = \frac{L(k) - L(k+1)}{\varepsilon^{-1}\alpha c\theta + \mathcal{O}\left(\theta^2\right)} = \frac{o(\theta)}{C\theta + \mathcal{O}\left(\theta^2\right)} = o(1).$$

Thus, for every $\delta > 0$ we can make the modulus of (I) smaller than δp_k^n by choosing n large.

Since $|Q_x(i)| \sim i^d = \hat{s}^{1/4d-1/6} \gg \hat{s}^{1/2}$, a version of lem. 5.1.6 applies, where all occurrences of $\mathbb{P}\left(x \in T_\varepsilon^M\right)$ and its estimates are replaced with the equivalent versions for $\mathbb{P}\left(x \in T_{z_n(k)}^{z_n(k+1)}\right)$, which have exactly the same asymptotic behavior. This proves the second part of the lemma. □

Now we can get a control on the depth of the first deep trap that is hit on a segment.

Lemma 5.2.2 (Control on depth of first trap)**.** *For any $\delta > 0$, $x \in B_0(R)$, $\mathbf{P}$-a.s. there exists n_0 such that for $n > n_0$ and $k \in \{0, \dots, R_n - 1\}$,*

$$\mathcal{P}_x(n,k) := \mathbb{P}\left(\mathsf{Hit}_x^{T_{z_n(k)}^{z_n(k+1)}} < \mathsf{Hit}_x^{\partial B_x(\rho)}\right) \in p_k^n h(n)^{-2}(1-\delta, 1+\delta). \tag{5.23}$$

Proof. We mimic the proof of lem. 5.1.10, this time summing over $T_{z_n(k)}^{z_n(k+1)}$ instead of T_ε^M,

$$\sum_{y \in T_{z_n(k)}^{z_n(k+1)}} \mathbb{P}\left(\mathsf{Hit}_x^{\{y\}} < \mathsf{Hit}_x^{\partial B_x(\rho)}\right).$$

We then apply the procedure used to prove lem. 5.1.10, using lem. 5.2.1 to replace the estimate on $\mathbb{P}\left(x \in T_\varepsilon^M\right)$ to get

$$\sum_{y \in T_{z_n(k)}^{z_n(k+1)}} \mathbb{P}\left(\mathsf{Hit}_x^{\{y\}} < \mathsf{Hit}_x^{\partial B_x(\rho)}\right) \in p_k^n h^{-2}(1-\delta, 1+\delta).$$

This yields the upper bound for $\mathcal{P}_x(n,k)$.

For the lower bound, we first subtract all the doubly counted probabilities (actually more, because we subtract twice the triple hits etc., but it still yields a lower bound)

$$\begin{aligned}
&\mathcal{P}_x(n,k) \\
&\quad \geq \sum_{y \in T_{z_n(k)}^{z_n(k+1)}} \mathbb{P}\left(\mathsf{Hit}_x^{\{y\}} < \mathsf{Hit}_x^{\partial B_x(\rho)}\right) - \sum_{y,y' \in T_{z_n(k)}^{z_n(k+1)}} \mathbb{P}\left(\mathsf{Hit}_x^{y} \wedge \mathsf{Hit}_x^{y'} < \mathsf{Hit}_x^{\partial B_x(\rho)}\right),
\end{aligned}$$

and note that

$$\begin{aligned}
\sum_{y,y' \in T_{z_n(k)}^{z_n(k+1)}} \mathbb{P}\left(\mathsf{Hit}_x^{y} \wedge \mathsf{Hit}_x^{y'} < \mathsf{Hit}_x^{\partial B_x(\rho)}\right) &= \mathbb{P}\left(\left|X_{j_i^n}^{j_{i+1}^n} \cap T_{z_n(k)}^{z_n(k+1)}\right| \geq 2\right) \\
&\leq \mathbb{P}\left(\left|X_{j_i^n}^{j_{i+1}^n} \cap T_\varepsilon^M\right| \geq 2\right),
\end{aligned}$$

which is $o\left(h^{-2}\right)$ by lem. 5.1.12. Thus for any $\delta > 0$ there exists n_0, such that for $n \geq n_0$,

$$\begin{aligned}
\mathcal{P}_x(n,k) &\geq \sum_{y \in T_{z_n(k)}^{z_n(k+1)}} \mathbb{P}\left(\mathsf{Hit}_x^{\{y\}} < \mathsf{Hit}_x^{\partial B_x(\rho)}\right) - \mathbb{P}\left(\left|X_{j_i^n}^{j_{i+1}^n} \cap T_\varepsilon^M\right| \geq 2\right) \\
&\geq (1-\delta) p_i^n h^{-2},
\end{aligned}$$

which finishes the proof. □

Now we show an important property of the family of conditional scores $(s^n(x))_{x\in\mathbb{Z}^d}$.

Lemma 5.2.3 (Laplace transform of scores)**.** *The following holds* $\mathbf{P}$*-a.s. as* $n \to \infty$

$$\mathbb{E}\left[e^{-\lambda s^n(s)/g(n)}\middle| s^n(x) < \infty, \boldsymbol{\tau}\right] = 1 - \frac{F_d(\lambda)}{h(n)^2} + o\left(h^{-2}\right)$$

where

$$F_d(\lambda) = p_\varepsilon^M - \frac{1}{\sqrt{2\pi}} \int_\varepsilon^M \frac{z^{-(\alpha+1)}}{1+\lambda G(0) z} \mathrm{d}z.$$

Proof. We start by computing the conditional Laplace transform

$$\mathbb{E}\left[e^{-\lambda s^n(x)/g(n)}\middle| \tau_y, 0 < s^n(x) < \infty\right],$$

where τ_y is the (unscaled) mean waiting time at the trap that is hit on this segment. This makes sense because for $0 < s^n(x) < \infty$, there is exactly one trap contributing to $s^n(x)$. This trap is possibly visited multiple times before exiting $B_x(\nu)$. In fact, after each visit, the random walk escapes $B_y(\nu)$ with probability $\mathbb{P}\left(\mathsf{Hit}_y^{\partial B_y(\nu)} < \mathsf{Hit}_y^{\{y\}}\right)$ and conditional on $s^n < \infty$, never returns. Therefore, the number of hits follows a geometric distribution.

The expected number of hits is exactly the Green's function $G_{B_y(\nu)}(y,y)$ (cf. eq. A.9). On every visit, the time spent in the trap is exponentially distributed with mean τ_y^n, which is deterministic in this calculation, due to the conditioning. With the above discussion, $s^n(x)$ is distributed as $\sum_{i=0}^M Z_i$, where M is a geometric RV with parameter $G_{B_y(\nu)}(y,y)^{-1}$ and Z_i is an independent collection of *i.i.d.* exponential variables with parameter $1/\tau_y^n$. Thus, we have

$$\mathbb{E}\left[e^{-\lambda s^n(x)/g(n)}\middle| \tau_y, 0 < s^n(x) < \infty\right] = \frac{1}{G_{B_y(\nu)}(y,y)\lambda g(n)^{-1}\tau_y^n + 1}. \tag{5.24}$$

Indeed, we let $G := G_{B_y(\nu)}(y,y)$

$$\begin{aligned}
&\mathbb{E}\left[e^{-\lambda s^n(x)/g(n)}\middle| \tau_y, 0 < s^n(x) < \infty\right] = \mathbb{E}\left[e^{-\lambda \sum_{i=1}^M Z_i/g(n)}\right] \\
&= \int \mathrm{d}\mathbb{P}\left(M\right) \int \mathrm{d}\mathbb{P}\left(Z_i\right) e^{\lambda \sum_{i=1}^M g^{-1} Z_i)} \\
&= \sum_{y=1}^\infty \left(\frac{1}{G}\left(1-\frac{1}{G}\right)^{y-1} \int_0^\infty \frac{1}{\tau_y^n} e^{-x/\tau_y^n} \mathrm{d}z_i e^{-\lambda \sum_{i=1}^y g^{-1} z_i} \right) \\
&= \sum_{y=1}^\infty \left(\left(\frac{1}{\lambda g^{-1}\tau_y^n + 1}\right)^y \frac{1}{1+G}\left(1-\frac{1}{G}\right)^{y-1} \right) \\
&= \frac{1}{G\lambda g^{-1}\tau_y^n + G} \sum_{y=0}^\infty \left(\frac{G-1}{G\lambda g^{-1}\tau_y^n + G}\right)^y.
\end{aligned}$$

The last sum is a geometric series and can be evaluated, leading to

$$\frac{1}{G\lambda g^{-1}\tau_y^n + G}\frac{1}{1-\frac{G-1}{G\lambda g^{-1}\tau_y^n+G}} = \frac{1}{1+G\lambda g^{-1}\tau_y^n}.$$

By prop. A.3.17, $G_{B_y(\nu)}(y,y) = G(0) + \mathcal{O}\left(\nu^{2-d}\right)$, and since $G(0)$ does not depend on n, this means that also $G_{B_y(\nu)}(y,y) = G(0)\left(1+o(1)\right)$.

By prop. 5.1.4, $\mathbb{P}\left(s^n(x)=\infty\right) = o\left(h^{-2}\right)$ and the argument of the expectation is bounded between 0 and 1, thus we can remove the conditioning and get

$$\mathbb{E}\left[e^{-\lambda s^n(x)/g(n)}\middle|\boldsymbol{\tau}, s^n(x)<\infty\right] = \mathbb{E}\left[e^{-\lambda s^n(x)/g(n)}\middle|\boldsymbol{\tau}\right]\left(1+o\left(h^{-2}\right)\right). \tag{5.25}$$

The expectation conditional only on $\boldsymbol{\tau}$ can be estimated from below using lem. 5.2.2 and eq. (5.24),

$$\mathbb{E}\left[e^{-\lambda s^n(x)/g(n)}\middle|\boldsymbol{\tau}\right] \geq \left(1-(1+\delta)p_\varepsilon^M h^{-2}\right) + h^{-2}\sum_{k=1}^{R_n}\frac{p_k^n(1-\delta)}{1+\lambda z_n(k)G(0)\left(1+o(1)\right)},$$

where we used that

$$\begin{aligned}\mathbb{E}\left[e^{-\lambda s^n(x)/g(n)}\middle|\boldsymbol{\tau}\right] &= \underbrace{e^0\mathbb{P}\left(s^n(x)=0\right)}_{\text{(I)}} + \sum_{k=1}^{R_n}\mathbb{P}\left(\left\{\tau_y\in T_{z_n(k)}^{z_n(k+1)}\right\}\cap\left\{s^n(x)<\infty\right\}\right)\times\\ &\quad\times\mathbb{E}\left[e^{-\lambda s^n(x)/g(n)}\middle|\tau_y, 0<s^n(x)<\infty\right]\\ &\geq \mathbb{P}\left(\mathsf{Hit}_x^{\partial B_x(\rho)} < \mathsf{Hit}_x^{T_\varepsilon^M}\right) + h^{-2}\sum_{k=1}^{R_n}\frac{p_k^n(1-\delta)}{1+\lambda\tau_y^n g(n)G(0)\left(1+o(1)\right)}.\end{aligned}$$

Furthermore, conditional on hitting $T_{z_n(k)}^{z_n(k+1)}$, $\tau_y^n g(n)$ is bounded from above by $z_n(k)$ and from below by $z_n(k+1)$.

The probability to exit $B_x(\rho)$ before hitting a trap is bounded by lem.5.1.10 from above, resp. below, by $(1-(1\mp\delta)h^{-2}p_\varepsilon^M)$.

Since the probability to hit T_ε^M is at most $(1+\delta)\hat{\sigma}^{-1}p_\varepsilon^M h^{-2}$, (I) can be estimated from below by $\left(1-(1+\delta)\hat{\sigma}^{-1}p_\varepsilon^M h^{-2}\right)$.

In a completely analogous way, we get the upper bound

$$\begin{aligned}&\mathbb{E}\left[e^{-\lambda s^n(x)/g(n)}\middle|\boldsymbol{\tau}\right]\\ &\quad\leq\left(1-(1-\delta)p_\varepsilon^M h^{-2}\right) + h^{-2}\sum_{k=1}^{R_n}\frac{p_k^n(1+\delta)}{1+\lambda z_n(k+1)G(0)\left(1+o(1)\right)}.\end{aligned}$$

The sum can in both cases be approximated with an integral, yielding

$$1 - h^{-2}\left(p_\varepsilon^M - \frac{1}{\sqrt{2\pi}}\int_\varepsilon^M \frac{1}{1+\lambda G(d)z} z^{-(\alpha+1)}\mathrm{d}z + o\left(h^{-2}\right)\right) \pm \delta C h^{-2} p_\varepsilon^M.$$

We let $\delta \to 0$ and choose n large. This completes the proof. □

We already know that the probability for a bad segment is $o\left(h^{-2}\right)$. If a segment is not bad, there are still two cases possible – either a trap is being hit or not. The next lemma gives the probability that a segment actually contains a trap.

Lemma 5.2.4 (Probability for nonzero score)**.** *As $n \to \infty$, $\mathbf{P}$-a.s., we have, uniformly in $x \in \mathcal{E}_0$,*

$$\mathbb{P}\left(s^n(x) \neq 0 | \boldsymbol{\tau}\right) = \left(1 + o(1)\right) h(n)^{-2} p_\varepsilon^M.$$

Proof. Just as we bounded the probability to hit $T_{z_n(k)}^{z_n(k+1)}$ before exiting $B_x(\rho)$ in lem. 5.2.2, we can use the same kind of reasoning (even easier because the bounds do not depend on n) to show that

$$\mathbb{P}\left(\mathsf{Hit}_x^{T_\varepsilon^M} < \mathsf{Hit}_x^{\partial B_x(\rho)}\right) \in h^{-2} p_\varepsilon^M (1-\delta, 1+\delta).$$

Obviously, if a trap is hit, $s^n(x) = 0$ is no longer possible and we get a lower bound that is of order h^{-2}.

To get the upper bound, we remind ourselves that the two cases where $s^n \neq 0$ are either when the segment is good and T_ε^M is hit before exiting $B_x(\rho)$, or the properties are violated in a way such that $s^n(x) = \infty$. But we have already shown that $\mathbb{P}\left(s^n(x) = \infty\right) = o\left(h^{-2}\right)$. We get an upper bound $(1+\delta)h^{-2}p_\varepsilon^M + o\left(h^{-2}\right)$ which asymptotically behaves as $h^{-2}p_\varepsilon^M$, too. □

Now we want to show that the scores approximate the clock process. It follows from cor. 5.1.5 that the probability for the "good" event $\mathcal{G}_n := \left\{s_j^n < \infty \,\forall\, j \leq Th^2\right\}$ can be made arbitrarily large by choosing m large. Therefore, it suffices to approximate the clock process conditional on $\mathcal{G}_n$. The main work here is to show that the time spent in the *shallow traps* T^ε and *deep traps* T_M can be neglected.

Lemma 5.2.5 (Time spent in shallow traps negligible)**.** *For any $\delta > 0$ there exists an $\varepsilon > 0$ such that for almost every $\boldsymbol{\tau}$ and large n*

$$\mathbb{P}\left(\left\{\sum_{i=0}^{j^n_{Th(n)^2}} e_i \tau_x^n \mathbf{1}_{\{X_i \in T^\varepsilon\}} \geq g(n)\delta\right\} \cap \mathcal{G}_n\right) \leq \delta,$$

where $(e_i)_{i\in\mathbb{N}}$ is an i.i.d. family of exponential RVs with mean 1.

Proof. Given that the event $\mathcal{G}_n$ occurs, the first $Th(n)^2$ segments are in $B_0(R(n))$. Therefore, conditionally on $\mathcal{G}_n$, we get an upper bound by summing up to $\zeta_n := \mathsf{Hit}_0^{\partial B_0(R(n))}$.

We want to apply the Markov inequality – to do this, we compute $\mathbb{E}\left[\sum_{i=1}^{\zeta_n} e_i \tau_x^n\right]$. We split up the expectation into different parts, starting with the time spent in traps with $\tau_x^n \leq 1$:

$$\begin{aligned}\mathbb{E}\left[\sum_{i=1}^{\zeta_n} e_i \tau_x^n \mathbf{1}_{\left\{\tau_{X_i}\leq 1\right\}}\right] &= \sum_{x\in B_0(R(n))} \tau_x^n G_{B_0(R(n))}(0,x)\, \mathbf{1}_{\{\tau_x^n\leq 1\}} \\ &\leq \sum_{x\in B_0(R(n))} G_{B_0(R(n))}(0,x) = \mathbb{E}\left[\zeta_n\right] \\ &= \mathcal{O}\left(R^2\right) = \mathcal{O}\left(\hat{\sigma}\hat{s}^{1/2}\right) \overset{\alpha\leq 1}{=} o(\hat{s}^{1/\alpha}) = o(g). \end{aligned} \tag{5.26}$$

Where we have used

$$\begin{aligned}\sum_{x\in B_0(R)} G_{B_0(R)}(0,x) &= \sum_{x\in B_0(R)} \mathbb{E}\left[\sum_{j=1}^{\zeta_n} \mathbf{1}_{\left\{X_j^{(0)}=x\right\}}\right] \\ &= \mathbb{E}\left[\sum_{j=1}^{\zeta_n} \sum_{x\in B_0(R)} \mathbf{1}_{\left\{X_j^{(0)}=x\right\}}\right] = \mathbb{E}\left[\zeta_n\right].\end{aligned}$$

The remaining part of T^ε is divided into disjoint sets $T_{\tilde{\varepsilon}_k}^{\tilde{\varepsilon}_{k-1}}$, where $\tilde{\varepsilon}_k := g(-k) = \varepsilon e^{-\alpha\hat{\sigma}_0^2 k}$ and k ranges from 1 to $k_{\max} := \min\left\{k : e^{-\alpha\hat{\sigma}_0^2} \leq g(-k)\varepsilon g(n) < 1\right\}$. Obviously, by this, $k_{\max} \sim n$. We have

$$P_{n,k} := \mathbb{P}\left(0 \in T_{\tilde{\varepsilon}_k}^{\tilde{\varepsilon}_{k-1}}\right) \leq \mathbb{P}\left(\tau_x^n \geq g e^{-k\alpha\hat{\sigma}_0^2}\varepsilon\right) \leq C\left(1+o(1)\right)\hat{\sigma}^{-1} g^{-\alpha/2}\varepsilon^{-\alpha} e^{k\alpha^2\hat{\sigma}_0^2}. \tag{5.27}$$

In the following, we will show that the probability to have a large expected waiting time in the sets $T_{\tilde{\varepsilon}_k}^{\tilde{\varepsilon}_{k-1}}$ is small. To do this, we fix a k, and consider for some K',

$$\mathcal{P} := \mathbb{P}\left(\mathbb{E}\left[\sum_{j=1}^{\zeta_n-1} e_j \tau_{X_j}^n \mathbf{1}_{\left\{X_j\in T_{\tilde{\varepsilon}_k}^{\tilde{\varepsilon}_{k-1}}\right\}} \middle| \boldsymbol{\tau}\right] \geq K'\varepsilon^{1-\alpha} e^{k(\alpha-1)\alpha\hat{\sigma}_0^2} g(n)\right).$$

The conditional expectation can be bounded from above using $\tau_x^n \leq e^{(-k+1)\alpha\hat{\sigma}_0^2}\varepsilon g(n)$,

$$\begin{aligned}\mathbb{E}\left[\sum_{j=1}^{\zeta_n-1} e_j \tau_{X_j}^n \mathbf{1}_{\left\{X_j\in T_{\tilde{\varepsilon}_k}^{\tilde{\varepsilon}_{k-1}}\right\}} \middle| \boldsymbol{\tau}\right] &= \sum_{x\in B_0(R(n))} G_{B_0(R(n))}(0,x)\tau_x^n \mathbf{1}_{\left\{x\in T_{\tilde{\varepsilon}_k}^{\tilde{\varepsilon}_{k-1}}\right\}} \\ &\leq \sum_{x\in B_0(R(n))} G_{B_0(R(n))}(0,x) e^{(-k+1)\alpha\hat{\sigma}_0^2}\varepsilon g(n)\, \mathbf{1}_{\left\{x\in T_{\tilde{\varepsilon}_k}^{\tilde{\varepsilon}_{k-1}}\right\}},\end{aligned}$$

therefore, $\mathcal{P}$ is bounded from above by

$$\mathbb{P}\left(\sum_{x\in B_0(R(n))} G_{B_0(R(n))}(0,x)e^{(-k+1)\alpha\hat{\sigma}_0^2}\varepsilon g(n)\,\mathbf{1}_{\left\{x\in T^{\tilde{\varepsilon}_{k-1}}_{\tilde{\varepsilon}_k}\right\}} \geq K'\varepsilon^{1-\alpha}e^{k(\alpha-1)\alpha\hat{\sigma}_0^2}g(n)\right).$$

We multiply both sides by $g(n)^{-1}e^{(k-1)\alpha\hat{\sigma}_0^2}\varepsilon^{-1}$. The terms disappear on the left side and the right becomes $K'\varepsilon^{-\alpha}e^{(k\alpha-1)\alpha\hat{\sigma}_0^2}$. We apply the exponential Chebyshev inequality, this time allowing λ to be chosen different for each n,

$$\mathcal{P} \leq e^{-\lambda_n K'\varepsilon^{-\alpha}e^{(k\alpha-1)\alpha\hat{\sigma}_0^2}} \underbrace{\prod_{x\in B_0(R(n))} \left((1-P_{n,k}) + P_{n,k}e^{\lambda_n G_{B_0(R(n))}(0,x)}\right)}_{\text{(I)}}. \tag{5.28}$$

The term (I) can be bounded using eq. (5.27) and $\log(x+1)\leq x$

$$\begin{aligned}\text{(I)} &\leq \prod_{x\in B_0(R(n))} \left(1 + C\left(1+o(1)\right)\hat{\sigma}^{-1}g^{-\alpha/2}\varepsilon^{-\alpha}e^{k\alpha^2\hat{\sigma}_0^2}\left(e^{\lambda_n G_{B_0(R(n))}(0,x)}-1\right)\right)\\ &\leq \exp\left(\sum_{x\in B_0(R(n))} C\left(1+o(1)\right)\hat{\sigma}^{-1}g^{-\alpha/2}\varepsilon^{-\alpha}e^{k\alpha^2\hat{\sigma}_0^2}\left(e^{\lambda_n G_{B_0(R(n))}(0,x)}-1\right)\right).\end{aligned} \tag{5.29}$$

To estimate the Green's function in the exponential, we divide the sum in a part close to the origin, $B_0(n^{2/(d-2)})$, which we will control by the smallness of the set, and $B_0(R(n))\setminus B_0(n^{2/(d-2)})$, where we can bound the Green's function.

We start with the inner sum. We set

$$\lambda_n := \alpha^2\hat{\sigma}_0^2 n/(4G_{B_0(R(n)))}(0,0)).$$

Since $G_{B_0(R(n)))}(0,0)$ converges to $G(0)$, this means that $\lambda_n \sim n$. Furthermore, we note that $G_{B_0(R(n)))}(0,x)\leq G_{B_0(R(n)))}(0,0)$, arriving at

$$\begin{aligned}(\text{I})_{\text{inner}} &\leq \exp\left(\sum_{x\in B_0(n^{2/(d-2)})} C\hat{\sigma}^{-1}g^{-\alpha/2}\varepsilon^{-\alpha}e^{k\alpha^2\hat{\sigma}_0^2}\left(e^{\alpha^2\hat{\sigma}_0^2 n/4}-1\right)\right)\\ &\leq \exp\left(Cn^{2d/(d-2)}\hat{\sigma}^{-1}\varepsilon^{-\alpha}e^{\alpha^2\hat{\sigma}_0^2(-n/2+k)}\left(e^{\alpha^2\hat{\sigma}_0^2 n/4}-1\right)\right) = e^{o(1)}.\end{aligned}$$

For $x\in B_0(R(n))\setminus B_0\left(n^{2/(2-d)}\right)$, thm. A.3.13 yields $G_{B_0(R(n))}(0,x)\leq cn^{-2}$. Therefore $\lambda_n G_{B_0(R(n))}(0,x) = \mathcal{O}\left(n^{-1}\right)$ and we can use for large n that $e^x - 1 \leq 2x$ as $x\to 0$.

$$e^{\lambda_n G_{B_0(R(n))}(0,x)} - 1 \leq cnG_{B_0(R(n))}(0,x).$$

By plugging this into eq. (5.29), we get for the terms away from zero,

$$(\mathrm{I})_{\text{outer}} \leq \exp\left(\sum_{x\in B_0(R(n))\setminus B_0(n^{2/(d-2)})} C\hat{\sigma}^{-1}\varepsilon^{-\alpha}g^{-\alpha/2}e^{k\alpha^2\hat{\sigma}_0^2}nG_{B_0(R(n))}(0,x)\right).$$

Since the Green's function is the only expression depending on x and

$$\sum_{x\in B_0(mr(n)^2)} G_{B_0(R(n))}(0,x) \sim r(n)^2 \sim \hat{\sigma}g^{\alpha/2},$$

we can cancel $\hat{\sigma}^{-1}g^{-\alpha/2}$ against the sum of the Green's functions to arrive at

$$(\mathrm{I})_{\text{outer}} \leq e^{Cn\varepsilon^{-\alpha}e^{\alpha^2\hat{\sigma}_0^2k}}.$$

Inserted into eq. (5.28), this yields

$$\mathcal{P} \leq e^{-C_1nK'\varepsilon^{-\alpha}e^{C_2k}}\underbrace{e^{C_3n\varepsilon^{-\alpha}e^{C_2k}}}_{=(\mathrm{I})_{\text{outer}}}\underbrace{e^{o(1)}}_{(\mathrm{I})_{=\text{inner}}}.$$

By choosing $K' > C_3C_1^{-1}$, this exponentially decreases in n for every k and ε. Taking the union over all k we get an additional factor n. The expression remains summable and by the Borel-Cantelli lemma, we have for sufficiently large n, for almost every $\boldsymbol{\tau}$, uniformly in k,

$$\mathbb{E}\left[\sum_{j=1}^{\zeta_n-1} e_j\tau_{X_j}^n\,\mathbf{1}_{\left\{X_j\in T_{\tilde{\varepsilon}_k}^{\tilde{\varepsilon}_{k-1}}\right\}}\middle|\boldsymbol{\tau}\right] < K'\varepsilon^{1-\alpha}e^{k(\alpha-1)\alpha\beta^2}g(n).$$

This, together with eq. (5.26) yields

$$\mathbb{E}\left[\sum_{i=0}^{\zeta_n} e_i\tau_{X_i}^n\,\mathbf{1}_{\{X_i\in T^\varepsilon\}}\right] \leq K\varepsilon^{1-\alpha}g(n).$$

Therefore, by the Markov,

$$\mathbb{P}\left(\left\{\sum_{i=0}^{j_{Th(n)^2}^n} e_i\tau_x^n\,\mathbf{1}_{\{X_i\in T^\varepsilon\}} \geq g(n)\delta\right\}\cap\mathcal{G}_n\right) \leq \frac{K\varepsilon^{1-\alpha}g(n)}{g(n)\delta}.$$

We choose $\varepsilon \geq K^{-1}\delta^{2/(\alpha-1)}$. This finishes the proof. □

Now we have to deal with the very deep traps. Obviously, if such a trap is ever hit, the contribution to the clock process is very large. So our strategy is, to show that T_M is never hit.

Lemma 5.2.6 (Very deep traps not hit)**.** *For every $\delta > 0$ there exist $M > \varepsilon > 0$ such that for large n and almost every $\boldsymbol{\tau}$*

$$\mathbb{P}\left(\left\{\mathsf{Hit}_x^{T_M} < j^n_{Th(n)^2}\right\} \cap \mathcal{G}_n \middle| \boldsymbol{\tau}\right) \leq \delta.$$

Proof. As in the proof of lem. 5.2.5, we can assume that the complement of $\mathcal{G}_n$ is arbitrarily small and replace $j^n_{Th(n)^2}$ by ζ_n. We again use the exponential Chebyshev inequality to prepare the application of the Borel-Cantelli lemma

$$\mathbb{P}\left(\mathbb{P}\left(\left\{\mathsf{Hit}_x^{T_M} < \zeta_n\right\} \cap \mathcal{G}_n \middle| \boldsymbol{\tau}\right) \geq \delta\right) \leq e^{-\lambda_n \delta} \mathbb{E}\left[e^{\lambda_n \mathbb{P}\left(\mathsf{Hit}_x^{T_M} < \zeta_n \middle| \boldsymbol{\tau}\right)}\right]. \tag{5.30}$$

Using $\mathbb{P}(x \in T_M) \leq C\hat{\sigma}^{-1} g^{-\alpha/2} M^{-\alpha}$ and $\log(1+x) \leq x$,

$$\begin{aligned}
\log \mathbb{E}\left[e^{\lambda_n \mathbb{P}\left(\mathsf{Hit}_x^{T_M} < \zeta_n \middle| \boldsymbol{\tau}\right)}\right] &\leq \log \mathbb{E}\left[e^{\lambda_n \sum_{y \in B_0(R(n))} \mathbb{P}\left(\mathsf{Hit}_x^{y} < \zeta_n \middle| \boldsymbol{\tau}\right) \mathbf{1}_{\{y \in T_M\}}}\right] \\
&\leq \sum_{y \in B_0(R(n))} \log\left(1 + C\hat{\sigma}^{-1} g^{-\alpha/2} M^{-\alpha}\left(e^{\lambda_n \mathbb{P}\left(\mathsf{Hit}_x^{y} \leq \zeta_n\right)} - 1\right)\right) \\
&\leq \sum_{y \in B_0(R(n))} C\hat{\sigma}^{-1} g^{-\alpha/2} M^{-\alpha}\left(e^{\lambda_n \mathbb{P}\left(\mathsf{Hit}_x^{y} \leq \zeta_n\right)} - 1\right).
\end{aligned} \tag{5.31}$$

Again, just as for the proof of lem. 5.2.5, we control the contributions by splitting them up in an inner, spatially small part near the starting point and an outer part, where the Green's function is small. We start with the inner part, that is $B_0(n^{2/(d-2)})$. We choose $\lambda_n = \alpha^2 \hat{\sigma}(n)^2/4$, leading to

$$\begin{aligned}
(5.31)_{\text{inner}} &\leq \sum_{y \in B_0(n^{2/(d-2)})} C\hat{\sigma}^{-1} g^{-\alpha/2} M^{-\alpha}\left(e^{\alpha^2\hat{\sigma}^2/4\, \mathbb{P}\left(\mathsf{Hit}_x^{y} \leq \zeta_n\right)} - 1\right) \\
&\leq C n^{2d/(d-2)} \hat{\sigma}^{-1} e^{-\alpha^2\hat{\sigma}^2/2} M^{-\alpha}\left(e^{\alpha^2\hat{\sigma}^2/4} - 1\right).
\end{aligned}$$

This tends to 0 as $n \to \infty$ exponentially, thus the contribution of $B_0(n^{2/(d-2)})$ to the expected value in eq. (5.30) is $e^{o\left(e^{-Cn}\right)}$.

For the part far from the origin, we use lem. A.3.18 for the argument of the exponential where $y \notin B_0(n^{2/(d-2)})$ and therefore $|y|^{2-d} \leq \left(n^{2/(d-2)}\right)^{2-d}$,

$$\begin{aligned}
\frac{\alpha^2\hat{\sigma}^2}{4} \mathbb{P}\left(\mathsf{Hit}_x^{y} \leq \zeta_n\right) &\leq Cn\left(\left(\left(n^{2/(d-2)}\right)^{2-d} - R(n)^{2-d}\right) + \mathcal{O}\left(\left(n^{2/(d-2)}\right)^{1-d}\right)\right) \\
&= \mathcal{O}\left(n^{-2}\right).
\end{aligned}$$

Thus, we can use $e^x - 1 \leq 2x$ for large n and get

$$\begin{aligned}(5.31)_{\text{outer}} &\leq \sum_{y \in B_0(R(n)) \setminus B_0(n^{2/(d-2)})} C_1 \hat{\sigma}^{-1} g^{-\alpha/2} M^{-\alpha} \left(e^{-C_2 |y|^{-2}} - 1 \right) \\ &\leq C \underbrace{\hat{\sigma}^{-1} g^{-\alpha/2}}_{=r^{-2}} M^{-\alpha} n \underbrace{\sum_{y \in B_0(R(n)) \setminus B_0(n^{2/(d-2)})} \underbrace{|y|^{2-d}}_{\leq 1}}_{\sim r^2} \leq C M^{-\alpha} n.\end{aligned}$$

Put into eq. (5.30), this yields

$$\mathbb{P}\left(\mathbb{P}\left(\left\{ \mathsf{Hit}_x^{T_M} < \zeta_n \right\} \cap \mathcal{G}_n \middle| \boldsymbol{\tau} \right) \geq \delta \right) \leq C_1 e^{-\delta n + o(e^{-C_2 n}) + C_3 M^{-\alpha} n},$$

which is summable when M is chosen sufficiently large. We can now apply Borel-Cantelli and the proof is complete. □

Now we can quantify the approximation error of the scores for approximating the clock process.

Proposition 5.2.7 (Scores approximate clock process)**.** *Let $\delta > 0, T > 0$. Then there exist ε, M, m such that $\mathbf{P}$-a.s. for large n,*

$$\mathbb{P}\left(\frac{1}{g(n)} \max \left\{ \left| S(j_k^n) - \sum_{j=0}^{k-1} s_j^n \right| \,:\, k \in \left\{ 1, \dots h^2 T \right\} \right\} \geq \delta \right) < \delta.$$

Proof. As mentioned, we can make the probability for the complement of $\mathcal{G}_n$ smaller than $\delta/2$ by chosing m sufficiently large. Conditional on $\mathcal{G}_n$, the difference is positive and increasing in k, since the score only counts contributions to the clock process when it is in deep traps. Therefore it can be bounded with the difference at $k = Th^2$. By lem. 5.2.5 and lem. 5.2.6 this is done and the proof is finished. □

5.3 Convergence of the process

We have shown that the coarse graining scheme has good properties and we have used these properties, to show that the scores approximate the clock process and that their Laplace transform converges to $F_d(\lambda)$ within an $o\left(h^{-2}\right)$ error. In this chapter, we show how this translates into convergence of the rescaled process Y^n.

First, we will show a result for the Laplace transform of the spatial process X^n similar to lem. 5.2.3 for the clock.

Lemma 5.3.1 (Laplace transform of spatial increments)**.** *For all $x \in \mathcal{E}(n)$, $\xi \in \mathbb{R}^d$, as $n \to \infty$,*

$$\mathbb{E}\left[e^{-\xi \cdot r^n(x)/r(n)} \right] = 1 + \frac{|\xi|^2}{2dh(n)^2} + o\left(h^{-2}\right).$$

Proof. We recall that $r^n(x)$ has the same law as the spatial increment of a coarse graining segment, which was defined by partitioning the trajectory into successive exits of spheres with radius ρ. Therefore, $|r^n(x)| = \rho\,(1 + o(1))$. We expand the expectation as a Taylor series w.r.t $r^n(x)$, where, due to the symmetry of the distribution only even powers play a role

$$\mathbb{E}\left[e^{-\xi\cdot r^n(x)/r(n)}\right] = 1 + \mathbb{E}\left[\frac{1}{2}h(n)^{-2}\left(\xi\cdot\frac{r^n(x)}{\rho(n)}\right)^2\right] + o\left(h^{-2}\right).$$

By prop. A.3.19 and (asymptotic) rotational symmetry, the distribution of $r^n(x)/\rho(n)$ converges as $n\to\infty$ to the uniform distribution on the sphere of radius one. Therefore

$$\mathbb{E}\left[\frac{1}{2}\left(\xi\cdot\frac{r^n(x)}{\rho(n)}\right)^2\right] = -\frac{|\xi|^2}{2d}.$$

□

Now we are ready to prove convergence of the joint conditional Laplace transforms.

Lemma 5.3.2 (Joint Laplace transform). *For* $\mathbf{P}$*-almost every* $\boldsymbol{\tau}$*, and all* $\lambda < 0$, $\xi\in\mathbb{R}^d$*, as* $n\to\infty$*,*

$$\mathbb{E}\left[\exp\left(-\frac{\lambda s^n(x)}{g(n)} - \frac{\xi\cdot r^n(x)}{r(n)}\right)\middle| s^n(x)<\infty, \boldsymbol{\tau}\right] = 1 - \frac{F_d(\lambda)}{h(n)^2} + \frac{|\xi|^2}{2dh(n)^2} + o\left(h^{-2}\right).$$

Proof. By prop. 5.1.4, the probability that $s^n(x)=\infty$ is $o\left(h^{-2}\right)$ and the contribution of those is bounded – thus we can remove the conditioning at an error of $o\left(h^{-2}\right)$. We decompose the expectation upon whether a trap is hit on the segment or not

$$\begin{aligned}\mathbb{E}\left[\exp\left(-\frac{\lambda s^n(x)}{g(n)} - \frac{\xi\cdot r^n(x)}{r(n)}\right)\middle|\boldsymbol{\tau}\right] &= \underbrace{\mathbb{E}\left[\exp\left(-\frac{\xi\cdot r^n(x)}{r(n)}\right)\mathbf{1}_{\{s^n(x)=0\}}\middle|\boldsymbol{\tau}\right]}_{\text{(I)}} \\ &+ \underbrace{\mathbb{E}\left[\exp\left(-\frac{\lambda s^n(x)}{g(n)} - \frac{\xi\cdot r^n(x)}{r(n)}\right)\mathbf{1}_{\{s^n(x)\neq 0\}}\middle|\boldsymbol{\tau}\right]}_{\text{(II)}}.\end{aligned} \tag{5.32}$$

Since $|r^n(x)| = \rho\,(1+o(1))$,

$$e^{-\frac{\xi\cdot r^n(x)}{r(n)}} = e^{|\xi|\mathcal{O}\left(h(n)^{-1}\right)} = (1+o(1)).$$

By using this,

$$\text{(II)} = \mathbb{E}\left[\exp\left(-\frac{\lambda s^n(x)}{g(n)}\right)\mathbf{1}_{\{s^n(x)\neq 0\}}\middle|\boldsymbol{\tau}\right](1+o(1)). \tag{5.33}$$

When no trap is hit on a segment we can use lem. 5.3.1 to get

$$\begin{aligned}(\mathrm{I}) &= \mathbb{E}\left[\exp\left(-\frac{\xi\cdot r^n(x)}{r(n)}\right)\middle|\boldsymbol{\tau}\right] - \mathbb{E}\left[\exp\left(-\frac{\xi\cdot r^n(x)}{r(n)}\right)\mathbf{1}_{\{s^n(x)\neq 0\}}\middle|\boldsymbol{\tau}\right]\\ &= 1+\frac{|\xi|^2}{2dh(n)^2}+o(h(n)^{-2})-(1+o(1))\,\mathbb{P}\left(s^n\neq 0\right).\end{aligned}$$

For eq. (5.33), we apply lem. 5.2.3

$$\begin{aligned}&\mathbb{E}\left[\exp\left(-\frac{\lambda s^n(x)}{g(n)}\right)\mathbf{1}_{\{s^n(x)\neq 0\}}\middle|\boldsymbol{\tau}\right]\\ &\qquad=\mathbb{E}\left[\exp\left(-\frac{\lambda s^n(x)}{g(n)}\right)\middle|\boldsymbol{\tau}\right]d-\mathbb{E}\left[\exp\left(-\frac{\lambda s^n(x)}{g(n)}\right)\mathbf{1}_{\{s^n(x)=0\}}\middle|\boldsymbol{\tau}\right]\\ &\qquad=1-h(n)^{-2}F_d(\lambda)+o(h(n)^{-2})-\mathbb{P}\left(s^n(x)=0\right)\\ &\qquad=-h(n)^{-2}F_d(\lambda)+o(h(n)^{-2})+\mathbb{P}\left(s^n\neq 0\right).\end{aligned}$$

We put these results into eq. (5.32) and arrive at

$$\begin{aligned}&\mathbb{E}\left[\exp\left(-\frac{\lambda s^n(x)}{g(n)}-\frac{\xi\cdot r^n(x)}{r(n)}\right)\middle|\boldsymbol{\tau}\right]\\ &\quad=1+\frac{|\xi|^2}{2dh(n)^2}-\frac{F_d(\lambda)}{h(n)^2}+\underbrace{\mathbb{P}\left(s^n(x)\neq 0\right)(1+o(1)-1)}_{=o(h^{-2})\text{ by lem. 5.2.4}}+o(h(n)^{-2}),\end{aligned}$$

which finishes the proof. □

The following lemma is where it all comes together. We will prove the convergence of the processes themselves. However, we only show the convergence of the finite dimensional marginal distributions. Then we extend this, first to joint weak convergence, and finally convergence of the original process $Y^n = X^n(S^{-1}(t))$.

Lemma 5.3.3 (Convergence of finite dimensional marginals)**.** *The finite dimensional distributions of* (S^n, Y^n) *converge to those of* $(\mathrm{V}_\alpha, \mathrm{BM}_d)$.

Proof. We recall the definitions of X^n (eq. (4.4)) and S^n (eq. (4.5))

$$\begin{aligned}X^n(t) &= \frac{\sqrt{d}}{C_{d,\alpha}r(n)}X\left(\left\lfloor C_{d,\alpha}^2 r(n)^2 t\right\rfloor\right),\\ S^n(t) &= \frac{1}{g(n)}\tilde{S}^{\hat{\sigma}(n)}\left(\left\lfloor C_{d,\alpha}^2 r(n)^2 t\right\rfloor\right).\end{aligned}$$

We want to approximate the processes by their increments r_k^n and s_k^n in a first step and then show that, due to the Markovian structure, we can use the families $(r^n(x))_{x\in\mathbb{Z}^d}$ resp. $(s^n(x))_{x\in\mathbb{Z}^d}$ instead, for which we have a control on the Laplace transforms.

By lem. A.3.20, $\mathbb{E}\left[j_1^n\right] = \rho(n)^2\left(1+o(1)\right)$ and $\mathbb{E}\left[\left(j_1^n/\rho(n)^2\right)^2\right]$ is bounded. Hence, by the law of large numbers for triangular arrays (see thm. A.4.3), $j^n_{C^2_{d,\alpha}h(n)^2u}$ strongly converges to $C^2_{d,\alpha}r(n)^2u$ and for any $\delta' > 0$, $u \leq T$ and large n, almost surely,

$$C^2_{d,\alpha}r(n)^2u \in \left[j^n_{(1-\delta')C^2_{d,\alpha}h(n)^2u}, j^n_{(1+\delta')C^2_{d,\alpha}h(n)^2u)}\right]. \tag{5.34}$$

Due to the monotonicity of $S^n(\cdot)$, this is also true if we apply S^n to both sides of the inclusion (5.34).

From prop. 5.2.7, it follows that

$$\mathbb{P}\left(\left|\frac{1}{g(n)}S^n\left(j_{(1\pm\delta')C^2_{d,\alpha}h(n)^2u}\right) - \sum_{k=0}^{j_{(1\pm\delta')C^2_{d,\alpha}h(n)^2u}} s^n_k\right| \geq \delta \middle| \boldsymbol{\tau}\right) \leq \delta, \tag{5.35}$$

and it is easy to see that also for large n,

$$\mathbb{P}\left(\left|\frac{\sqrt{d}}{C_{d,\alpha}r(n)}X\left(j_{(1\pm\delta')C^2_{d,\alpha}h(n)^2u}\right) - \frac{\sqrt{d}}{C_{d,\alpha}r(n)}\sum_{k=0}^{j_{(1\pm\delta')C^2_{d,\alpha}h(n)^2u}} r^n_k\right| \geq \delta\right) \leq \delta. \tag{5.36}$$

Indeed, we can just pull out the constant, which has $r(n)$ in the denominator, and note that the difference we see now is at most $\rho(n)$, due to the definition of the r^n_k. Therefore the quantity in question behaves as $\rho(n)/r(n) = h(n)^{-1}$ which goes to 0 as $n \to \infty$.

We chose a partition of $[0,T]$: let $0 = u_0 < u_1, \ldots, < u_q = T$ and consider the joint Laplace transform of the increments of both processes on this partition

$$\mathbb{E}\left[\exp\left(-\frac{1}{g(n)}\sum_{i=0}^{q}\lambda_i\left(S\left(C^2_{d,\alpha}r(n)u_i\right) - S\left(C^2_{d,\alpha}r(n)u_{i-1}\right)\right)\right.\right.$$
$$\left.\left.-\frac{\sqrt{d}}{C_{d,\alpha}r(n)}\sum_{i=0}^{q}\xi_i\cdot\left(X\left(C^2_{d,\alpha}r(n)u_i\right) - X\left(C^2_{d,\alpha}r(n)u_{i-1}\right)\right)\right)\middle|\boldsymbol{\tau}\right].$$

By the eqs. (5.35) and (5.36), it is sufficient to show convergence of

$$\mathbb{E}\left[\exp\left(-\sum_{i=0}^{q}\sum_{k\in S_{\pm\delta'}(n,i)}\frac{\lambda_i s^n_k}{g(n)} - \sum_{i=0}^{q}\sum_{k\in S_{\pm\delta'}(n,i)}\frac{\sqrt{d}\xi_i\cdot r^k_n}{C_{d,\alpha}r(n)}\right)\middle|\boldsymbol{\tau}\right], \tag{5.37}$$

where

$$S_{\pm\delta'}(n,i) := \left\{\left\lfloor(1\pm\delta')\,C^2_{d,\alpha}h(n)^2u_{i-1}\right\rfloor, \ldots, \left\lfloor(1\pm\delta')\,C^2_{d,\alpha}h(n)^2u_i\right\rfloor - 1\right\}.$$

Now we consider each possible trajectory separately. Let $c_0(I) := \left\{(x_l)_{l\in I} \subset \mathbb{Z}^d\right\}$ and set $I_n := \left\{1, \dots, \left\lfloor C_{d,\alpha}^2 h(n)^2 u_q \right\rfloor\right\}$. Then, the expextation in eq. (5.37) is equal to

$$\sum_{(x_l)\in c_0(I_n)} \mathbb{P}\left(X(j_l^n) = x_l \,\forall l \in I_n\right) \times$$
$$\times \mathbb{E}\left[\exp\left(-\sum_{i=0}^{q}\sum_{k\in S_{\pm\delta'}(n,i)} \frac{\lambda_i s_k^n}{g(n)} - \sum_{i=0}^{q}\sum_{k\in S_{\pm\delta'}(n,i)} \frac{\sqrt{d}\xi_i \cdot r_n^k}{C_{d,\alpha} r(n)}\right)\middle| \boldsymbol{\tau},\, X(j_l^n) = x_l \,\forall l \in I_n\right].$$

Because of cor. 5.1.5, the sequences where one of the x_l is not in $\mathcal{E}_0$ contribute at most $o(h(n)^{-2})$ and can thus be neglected.

For the other sequences, all the $X(j_k^n)$ are in $\mathcal{E}_0$. In this case, lem. 5.3.2 can be applied. In order to do that, we first have to replace the r_n^k and s_n^k by suitable members of the families $(r^n(x))_{x\in\mathbb{Z}^d}$ and $(s^n(x))_{x\in\mathbb{Z}^d}$.

Let $\omega := \left\lfloor C_{d,\alpha}^2 h(n) u_q \right\rfloor - 1$. Due to the Markovian structure of the process, conditionally on $\boldsymbol{\tau}$ and the event that $X(j_\omega^n) = x_\omega$, the distribution of (s_ω^n, r_ω^n) is independent of the rest of the process and, by definition, its distribution is the same as the distribution of $(s^n(x_\omega), r^n(x_\omega))$. Therefore,

$$\mathbb{E}\left[\exp\left(-\sum_{i=0}^{q}\sum_{k\in S_{\pm\delta'}(n,i)}\left(\frac{\lambda_i s_k^n}{g(n)} - \frac{\sqrt{d}\xi_i \cdot r_n^k}{C_{d,\alpha} r(n)}\right)\right)\middle| \boldsymbol{\tau},\, X(j_l^n) = x_l \,\forall l \in I_n\right]$$
$$= \mathbb{E}\left[\exp\left(-\sum_{i=0}^{q}\sum_{\substack{k\in S_{\pm\delta'}(n,i)\\ k\le\omega-1}}\left(\frac{\lambda_i s_k^n}{g(n)} - \frac{\sqrt{d}\xi_i \cdot r_n^k}{C_{d,\alpha} r(n)}\right)\right)\middle| \boldsymbol{\tau},\, X(j_l^n) = x_l \,\forall l \in I_n\right] \times$$
$$\times \mathbb{E}\left[\exp\left(-\frac{\lambda_q s^n(x_\omega)}{g(n)} - \frac{\sqrt{d}\xi_q \cdot r^n(x_\omega)}{C_{d,\alpha} r(n)}\right)\middle| \boldsymbol{\tau}\right].$$

By lem. 5.3.2, the last expectation – with only $s^n(x_\omega)$ and $r^n(x_\omega)$ appearing – is, uniformly in x_ω, bounded from above, resp. below, by

$$1 + (1 \pm \delta')h(n)^{-2}\left(\frac{|\xi_q|^2}{2d}\frac{d}{C_{d,\alpha}^2} - F_d(\lambda_q)\right).$$

We iterate backwards, *i.e.* use the same procedure for $\omega - 1, \omega - 2, \dots, 2, 1$. This gives us an upper resp. lower bound of the joint Laplace transform of the score and spatial

increment at all u_i, as

$$\mathbb{P}\left(X(j_l^n) = x_\forall l \in I_n\right)\mathbb{P}\left(X(j_l^n) \in \mathcal{E}_0 \forall l \in I_n\right) \times \\ \times \prod_{i=1}^{q} \underbrace{\prod_{k \in S_{\pm\delta}(n,i)} 1 + (1 \pm \delta')h(n)^{-2}\left(\frac{|\xi_q|^2}{2d}\frac{d}{C_{d,\alpha}^2} - F_d\left(\lambda_q\right)\right)}_{\text{(I)}}. \tag{5.38}$$

Because the families $(s^n(x))_{x\in\mathbb{Z}^d}$ and $(r^n(x))_{x\in\mathbb{Z}^d}$ are *i.i.d.* and using the fact that $|S_{\pm\delta}(n,i)| = C_d^2 h(n)^2\left(u_i - u_{i-1}\right)$, we get,

$$\begin{aligned}\text{(I)} &= \left(1 + (1 \pm \delta')h(n)^{-2}\left(\frac{|\xi_q|^2}{2d}\frac{d}{C_{d,\alpha}^2} - F_d\left(\lambda_q\right)\right)\right)^{C_{d,\alpha}^{-2}h(n)^2(u_i - u_{i-1})} \\ &= \left(1 + \frac{(1 \pm \delta')\left(|\xi_q|^2/(2d) - C_{d,\alpha^2}F_d(\lambda_q)\right)}{C_{d,\alpha}^2 h(n)^2}\right)^{C_{d,\alpha}^2 h(n)^2(u_i - u_{i-1})}.\end{aligned} \tag{5.39}$$

As $C_{d,\alpha}^2 h(n)^2 \to \infty$, this converges to an exponential function,

$$\text{(I)} \to \exp\left((1 \pm \delta')\left(\frac{|\xi_q|^2}{2d} - C_{d,\alpha}^2 F_d(\lambda_q)\right)\right). \tag{5.40}$$

Putting this back into eq. (5.38), we get the lower, resp. upper, bound

$$\mathbb{P}\left(X(j_l^n) = x_l \forall l \in I_n\right)\mathbb{P}\left(X(j_l^n) \in \mathcal{E}_0 \forall l \in I_n\right) \times \\ \times \prod_{i=1}^{q} \exp\left(\left((1 \mp \delta')\frac{|\xi_i|^2}{2} - (1 \pm \delta')C_{d,\alpha}^2 F_d\left(\lambda_i\right)\right)\left(u_i - u_{i-1}\right)\right)\left(1 + o(1)\right). \tag{5.41}$$

We now let $\delta' \to 0$ and see that upper and lower bound coincide. Further, we note that

$$\lim_{\varepsilon\downarrow 0,\, M\uparrow\infty} F_d(\lambda) = \frac{1}{\sqrt{2\pi}\alpha}\left(G_d(0)\lambda\right)^\alpha \Gamma(1+\alpha)\Gamma(1-\alpha). \tag{5.42}$$

Indeed,

$$\begin{aligned}&\frac{1}{\sqrt{2\pi}\alpha}\left(\varepsilon^{-\alpha} - M^{-\alpha}\right) - \frac{1}{\sqrt{2\pi}}\int_\varepsilon^M \frac{1}{1 + \lambda G_d(0) z} z^{-(\alpha+1)}\mathrm{d}z \\ &= \frac{1}{\sqrt{2\pi}}\left(\int_\varepsilon^M x^{-1-\alpha}\mathrm{d}x - \int_\varepsilon^M \frac{1}{1 + \lambda G_d(0) z} z^{-1-\alpha}\mathrm{d}z\right) \\ &= -\frac{1}{\sqrt{2\pi}}\int_\varepsilon^M \frac{\lambda G_d(0)}{1 + \lambda G_d(0) z} z^{-\alpha}\mathrm{d}z.\end{aligned}$$

Changing variables $z \to (\lambda G_d(0))^{-1} z$ and letting $\varepsilon \downarrow 0$ and $M \uparrow \infty$, this equals

$$\frac{(\lambda G_d(0))^\alpha}{\sqrt{2\pi}} \int_0^\infty \frac{z^{-\alpha}}{1+z} \mathrm{d}z.$$

This integral can be evaluated and simplified using [1, eqs. 6.2.1, 6.2.2],

$$\frac{(\lambda G_d(0))^\alpha}{\sqrt{2\pi}} \Gamma(1-\alpha)\Gamma(\alpha) = \frac{(\lambda G_d(0))^\alpha}{\sqrt{2\pi}\alpha} \Gamma(1-\alpha)\Gamma(1+\alpha).$$

Therefore,

$$C_{d,\alpha}^2 F_d(\lambda) \xrightarrow{\varepsilon \downarrow 0,\, M \uparrow \infty} = \lambda^\alpha.$$

Inserting this into eq. (5.41), summing over all possible sequences x_l and using that by cor. 5.1.5,

$$\mathbb{P}\left(X(j_l^n) \in \mathcal{E}_0 \forall l \in I_n\right) = 1 + o(h^{-2}),$$

we end up with a formula for the limit of the joint Laplace transform of spatial and temporal process – except for an event with a probability that can be made arbitrarily small (cf. eqs. 5.36 resp. 5.35) – which reads

$$(1+o(1)) \prod_{i=1}^q \exp\left(\left(\frac{|\xi_i|^2}{2d} - \lambda_i^\alpha\right)(u_i - u_{i-1})\right),$$

which is the Laplace transform of the increments of a Brownian motion and an independent α-stable subordinator at finitely many points (see prop. A.2.10 and prop. A.2.9). Since the Laplace transform uniquely determines the distribution, this concludes the proof. □

To extend the convergence to almost sure convergence of the whole trajectory in the $D([0,T], M_1) \times D^d([0,T], U)$ product-topology (see sec. A.5 for the definition of these spaces and the topologies), we will show tightness of the individual components.

Lemma 5.3.4 (Tightness.). *The sequences S^n and X^n are tight in $D([0,T], M_1)$ resp. $D^d([0,T], U)$.*

Proof. We will use the characterization from lem. A.5.6. We start with the clock process S^n. We will show that it fulfills the assumptions of the lemma, *i.e.*

i) $\forall \varepsilon > 0 \exists c : \mathbb{P}\left(\|S^n(\cdot)\| > c\right) < \varepsilon$,

ii) $\forall \varepsilon > 0, \eta > 0, \exists \delta > 0 : \mathbb{P}\left(\omega'(S^n(\cdot), \delta) \geq \eta\right) \leq \varepsilon$.

Because $S^n(\cdot)$ is increasing, it is sufficient to check condition i) for $S^n(T)$. From lem. 5.3.3, we know that $S^n(T) \to V_\alpha(T)$ in distribution, which fulfills i).

To check ii), we treat the three possible values of $\omega'(S^n(\cdot), \delta)$ separately, starting with $\omega_w(S^n(\cdot), \delta)$

$$\sup_{0\le t\le T} \sup_{(0\vee t-\delta)\le t_1\le t_2\le t_3\le (t+\delta\wedge T)} \left\{ \left\| S^n(t_2) - [S^n(t_1), S^n(t_3)] \right\|_\infty \right\}.$$

Due to the monotonicity of $S^n(\cdot)$, $S^n(t_2)$ is always contained in $[S^n(t_1), S^n(t_3)]$ and thus the expression is always zero. Next, we consider

$$\bar{v}(x, 0, \delta) = \sup_{0\le t_1\le t_2\le \delta} \left\{ \left\| S^n(t_1) - S^n(t_2) \right\|_\infty \right\}, \text{ and}$$
$$\bar{v}(x, T, \delta) = \sup_{T-\delta\le t_1\le t_2\le T} \left\{ \left\| S^n(t_1) - S^n(t_2) \right\|_\infty \right\}.$$

Again we use lem. 5.3.3 to replace the distributions of $S^n(\cdot)$ with those of $V_\alpha(\cdot)$. The distribution of a single process trivially fulfills the conditions - since the distribution of the $S^n(\cdot)$ is close to this in the limit, this is sufficient to show it for all $n \ge 1$ by adjusting the constant c if necessary.

For the spatial process $X^n(\cdot)$, tightness immediately follows from prop. A.4.7.

Tightness in both components implies tightness in the product topology on the product space $D([0,T], M_1) \times D([0,T], U)$ because both spaces are polish (Whitt [119, thm. 11.6.7]). □

Together with the convergence of the finite dimensional distributions, tightness of $(S^n(\cdot), X^n(\cdot))$ yields weak convergence by [48, thm. 4.15]. To finish the proof of our main theorem, we recall that $Y^n(\cdot) = X^n(S^n)^{-1}(\cdot)$.

By prop. A.5.8, the inversion map from $D_{u,\uparrow}([0,T], M_2)$, the subset of increasing functions to $D_{u,\uparrow}([0,T], U)$ is continuous at strictly increasing functions. Since the M_2 topology is coarser than the M_1 topology, this implies the same statement for the M_1 topology.

Since the limiting process $V_\alpha(\cdot)$ has almost surely strictly increasing paths, we can conclude, that $(S^n)^{-1}(\cdot) \to V_\alpha^{-1}(\cdot)$ weakly in distribution on $D_{u,\uparrow}([0,T], U)$ – since we applied a continuous map to both sides. In contrast to V_α, which has jumps the inverse subordinator V_α^{-1} is almost surely continuous – the jumps translate to constant parts. The same is true for the paths of Brownian motion.

Therefore we can use that the composition as a mapping from $D^d([0,T], U) \times D([0,T], U)$ to $D^d([0,T], U)$, $(f, g) \mapsto f \circ g$ is continuous at $C^d \times C$. Indeed, in the topology of uniform convergence, continuity of the composition is implied by showing that

$$(f_n \to f) \wedge (g_n \to g) \implies f_n \circ g_n \to f \circ g.$$

This is true, because all the f_n and f are continuous and f_n uniformly converges to f.

Therefore weak convergence of $X^n(\cdot)$ to $\mathrm{FK}_{d,\alpha}$ follows.

PART III

Applications: Modeling, Simulation and Experiments

CHAPTER 6

The fractional kinetics process and fractional calculus

In this chapter, we will look into the limiting process and investigate its properties, including the equation governing the evolution of its probability density. To state some of the results, we first need to develop the toolbox for it – fractional calculus.

6.1 Fractional calculus and fractional differential equations

The name fractional calculus is a bit of a a misnomer, since it usually refers to the generalization of derivative and integral to any order $\alpha \in \mathbb{R}$ and not only fractions. Even more, we will see that from a mathematical-systematical viewpoint the name fractional differential equation is misleading too. A more appropriate name would be "integro-differential equations with convolution-type integral operator with a weakly singular kernel of powerlaw type" – but since the term fractional is widely used and more catchy than this, we will use it here too.

Both the name and the general idea stem from mathematical curiosity that goes back to the time of Leibnitz, who was asked by l'Hôpital in 1695 about the meaning of $D^n f$ for the fraction $n = 1/2$. We will not further pursue the historical path, but directly introduce modern fractional calculus as it applies to our situation. For a more complete account on the history and the subject itself, the reader is referred to Oldham & Spanier [87].

We start by introducing the fractional integrals, then we will define fractional derivatives as a conjunction of integer order derivatives and a fractional integral. Because the two operations don't commute, there will be different types of fractional derivatives. Finally we consider fractional differential equations, which govern the evolution of the density of the fractional kinetics process.

We will be using D^n for both differential $n > 0$ and integral $n < 0$. We first define

the fractional integral and the class of functions, we want to use.

Definition 6.1.1. We define the class $\mathbf{C}$ as the class of functions that are continuous on $[0, \infty)$ and integrable on every compact interval contained in $[0, \infty)$.

Whenever needed, we extend these functios to $-\infty$ by setting them equal to 0 for $x < 0$ (such functions are called *causal*).

Definition 6.1.2. For $f \in \mathbf{C}$, we define the *fractional integral of order* $\alpha > 0$, ${}_a\mathrm{D}_x^{-\alpha}$, as

$$({}_a\mathrm{D}_x^{-\alpha} f)(x) := \frac{1}{\Gamma(\alpha)} \int_a^x f(\xi)\,(x-\xi)^{\alpha-1}\,\mathrm{d}\xi.$$

For $\alpha = 0$, we set ${}_a\mathrm{D}_x^0 = \mathrm{I}$.

The choice of the lower bound is somewhat arbitrary. In Miller & Ross [74], the version with lower boundary $c \in \mathbb{R}$ as parameter is called *Riemann version*, the version with $-\infty$ as lower boundary is called *Liouville version*. The version we will mainly use is the *Riemann-Liouville* version, where we set $a = 0$. In that case, we just write D_x^α.

Remark. Another possibility is to fix the upper bound at ∞ and put x as a lower bound – this is called the *Weyl fractional integral*.

Remark. We can change variables $\zeta = (x - \xi)$, leading to

$$\mathrm{D}_x^\alpha f(x) = \frac{1}{\Gamma(\alpha)} \int_{-\infty}^x f(\xi)\,(x-\xi)^{\alpha-1}\,\mathrm{d}\xi = \frac{1}{\Gamma(\alpha)} \int_0^\infty f(x-\zeta)\zeta^{\alpha-1}\mathrm{d}\zeta.$$

Note that because $f(\xi) = 0$ for $x < 0$ we can just extend the integral to $-\infty$ without changing its value. This is the formula used *e.g.* in Meerschaert & Scheffler [70], which we later cite for some results.

This formula can be understood e.g. by recalling the Cauchy formula (see *e.g.* Miller & Ross [74, pp. 24–25]) for repeated integration, which states, that for $n \in \mathbb{N}$,

$$\mathrm{D}_x^{-n} f(x) = \int_a^x \frac{(x-\xi)^{n-1}}{(n-1)!} f(\xi)\mathrm{d}\xi.$$

In other words, the n-fold iterated integral operator D_x^{-n} can be understood as a convolution operator with kernel $K(x) = x^{n-1}/(n-1)!$ – our definition of the fractional integral for arbitrary order is just the extension of this kernel to non-integer values of n.

Another way to arrive at the formula which will be useful for us is via Laplace transforms. It is well known that the Laplace transform of the n-fold iterated integral of f is given by $s^{-n}\mathcal{L}[f](s)$. Again, it is straightforward to replace n by a non-integer. Inverting this transform, we would arrive at the formula in def. 6.1.2.

There are several other ways to find the above formula (see Miller & Ross [74, chap. II]). The way we define fractional integration now is via conjunction of integer-order differentiation and a fractional integration.

Definition 6.1.3. We define the *Riemann-Liouville fractional derivative of order* $\alpha \geq 0$ as

$$\left(\mathrm{D}_x^\alpha f\right)(x) = \left(\mathrm{D}_x^{\lceil \alpha \rceil} \left(\mathrm{D}_x^{\alpha - \lceil \alpha \rceil}\right) f\right)(x).$$

Unfortunately, fractional integration and differentiation do not commute (for a counterexample, see [74, pp. 104]). This also means, that the choice of the order of differentiation and integration in def. 6.1.3 is not arbitrary. Indeed, if we change the order, we end up with the so-called *Caputo* fractional derivative.

Just as the integral, the fractional derivative generalizes the Laplace transform of integer order derivatives.

Proposition 6.1.4. *For the Laplace transform of a fractionally differentiated function, the following holds*

$$\mathscr{L}\left[\mathrm{D}_x^\alpha f\right](s) = s^\alpha \mathscr{L}\left[f\right](s) - \sum_{k=0}^{\lfloor \alpha \rfloor} s^k \, \mathrm{D}_x^{\alpha-1-k} f(0). \tag{6.1}$$

Proof. See Oldham & Spanier [87, pp. 134]. □

Remark. In the important case (actually the only case we will consider) of $0 < \alpha < 1$, the formula above simplifies to

$$\mathscr{L}\left[\mathrm{D}_x^\alpha f\right](s) = s^\alpha \mathscr{L}\left[f\right](s) - \mathrm{D}_x^{\alpha-1} f(0). \tag{6.2}$$

The additional term is completely analogous to the additional terms that appear in the transforms of regular derivatives. Intuitively, it tells us that we need initial conditions for these objects if we want to solve differential equations – of integer order for integer derivatives, and of fractional order for fractional derivatives.

Remark. If fractional order initial conditions are a problem, one can switch to the Caputo formulation, which only needs integer order initial conditions.

The (partial-,integro-) differential equation, we want to solve is the so-called *time-fractionial diffusion equation*

$$\mathrm{D}_t^\alpha f = \Delta f + \delta(0) \frac{t^{-\alpha}}{\Gamma(1-\alpha)}, \tag{6.3}$$

because it – as we will soon find out – governs the evolution of the density of the fractional kinetics process (thm. 6.2.3). Therefore we want to study its solutions now.

Theorem 6.1.5. *The time-fractional diffusion equation (6.3) possesses a solution, which can be expressed as*

$$u(x,t) = \frac{t}{\alpha} \int_0^\infty k(x,\xi) g_\alpha \left(\frac{t}{\xi^{1/\alpha}} \right) \mathrm{d}\xi, \tag{6.4}$$

where g_α is the density of a stable subordinator and $k(x,y)$ is the convolution kernel of the heat semigroup.

Proof. See Baeumer & Meerschaert [8, thm. 3.1] and replace the general Lévy process with Brownian motion. □

In the next chapter, we connect the fractional diffusion equation to the fractional kinetics process $\mathrm{FK}_{d,\alpha}$. This will also give us another representation for a solution eq. (6.3).

We now present another, computationally more accessible representation of the solution. While it is less easily interpretable and generalizable than the representation of thm. 6.1.5, a convergent series expansion and asymptotics are available and can easily be implemented. It is centered on the following representation of the fundamental solution.

Theorem 6.1.6. *The fundamental solution $G_\alpha(x,t)$ of the time fractional diffusion equation (eq. (6.3)) for $0<\alpha<1$ is given by*

$$G_\alpha(x,t) = t^{-\alpha/2} K_\alpha\left(\frac{|x|}{t^{\alpha/2}}\right), \tag{6.5}$$

where

$$K_\alpha(x) = \frac{1}{2}\sum_{n=0}^{\infty} \frac{(-1)^n x^n}{n!\Gamma\left(-\alpha n/2 + (1-\alpha/2)\right)}. \tag{6.6}$$

Proof. See Mainardi & Pagnini [62, eqs. (2.6) and (3.11)]. □

To be able to still efficiently compute $K_\alpha(x)$ for large values of x, we need the following asymptotic representation.

Proposition 6.1.7. *Asymptotically, as $x\to\infty$,*

$$K_\alpha(x) \sim A x^a e^{-b c^c}, \tag{6.7}$$

where

$$A = \sqrt{2\pi(2-\alpha)2^{\alpha/(2-\alpha)}\alpha^{(2-2\alpha)/(2-\alpha)}},$$
$$a = \frac{2\alpha-2}{2(2-\alpha)}, \quad b = (2-\alpha)2^{-2/(2-\alpha)}\alpha^{\alpha/(2-\alpha)}, \quad c = \frac{2}{2-\alpha}.$$

Proof. See Braaksma [26] as in Mainardi & Pagnini [62, eq. (3.12)]. □

6.2 Properties of the fractional kinetics process

We start off with self similarity and regularity properties.

Proposition 6.2.1. *For the fractional kinetics process* $\mathrm{FK}_{d,\alpha}$ *the following holds*

i) $\mathrm{FK}_{d,\alpha}$ *neither has stationary nor stable increments.*

ii) $\mathrm{FK}_{d,\alpha}$ *is self-similar:* $\mathrm{FK}_{d,\alpha}(t) = r^{-\alpha/2}\mathrm{FK}_{d,\alpha}(rt)$.

iii) $\mathrm{FK}_{d,\alpha}$ *is* γ*-Hölder continuous for any* $\gamma < \alpha/2$.

Proof. i) see Meerschaert & Scheffler [70, cor. 4.3, thm. 4.3].

ii) We use the scaling properties of V_α (see A.2.12 i)) and BM_d (see A.2.9 i))

$$\mathrm{FK}_{d,\alpha}(\lambda t) = \mathrm{BM}_d\left(V_\alpha^{-1}(\lambda t)\right) = \mathrm{BM}_d\left(\lambda^{-\alpha}V_\alpha^{-1}(t)\right) = \lambda^{-\alpha/2}\mathrm{BM}_d\left(V_\alpha^{-1}(t)\right).$$

iii) Just as the self-similarity, this immediately follows from the properties of Brownian motion and the inverse subordinator prop. A.2.9 ii) and prop. A.2.12 iii) respectively.

□

The exponent for the self-similar rescaling is less than $1/2$ (which would be the case for Brownian motion). Such processes are called *subdiffusive*, because they spread in space slower than a diffusive process would.

Now, we connect the fractional kinetics process to the fractional diffusion equation.

Definition 6.2.2. We say that a stochastic process X is a *stochastic solution* to a (partial-/integro-) differential equation, if its density $\rho_X(t, x)$ solves the PDE.

Theorem 6.2.3. *The fractional kinetics process is a stochastic solution to the fractional diffusion equation (6.3).*

Proof. See Meerschaert *et al.* [69, main theorem]. □

CHAPTER 7

Simulations and experiments

In this chapter, we connect the mathematical results to practice. To do this, we will frequently make use of kinetic Monte Carlo (kMC) [43] simulations. To be confident that what we see are not numerical artifacts, we study in sec. 7.1 the influence of a finite simulation domain. We show that gives rise to numerical problems which lead to errors similar to the physical effect we want to investigate. Therefore we suggest and implement an algorithm which does not assume *a priori* finite simulation domains

In a next step we analyze lab experiments, which show that the postulated dispersive effect indeed appears in practice. If the same material is measured with varying thickness, the transport properties (the diffusion constant resp. mobility) seem to change. This work clearly shows that this dependence is not an artifact of boundary effects or a measurement error, but an inherent property of bulk transport in strongly disordered materials.

To make sure our simplified model qualitatively captures the dispersive behavior,we check in sec. 7.4 the simplifying assumptions made in sec. 4.1.

To get a better grasp of the validity of our scaling limit, we compare the evolution of the density of simulations with the evolution of solutions to the limit equation for different timescales and disorder stengths in sec. 7.3.

7.1 The finite-size effect

Within the (possibly correlated) GDM, the computational results can depend on the size of the computation domain. This is called a finite size effect. The finite size effect strikes, when the realization of the trapping landscape is too small to contain traps below a certain depth, which leads to an overestimation of the mobility.

A more unlikely, but possible case would be, if very deep traps are indeed present but then are hit too often due to the limited size. The second error can be easily avoided, by using different realizations of the disordered energy landscape. It is very unlikely that atypically deep traps occur independently in several morphologies. However, for every fixed size of a morphology, there exists an effective disorder $\hat{\sigma}$ such that the

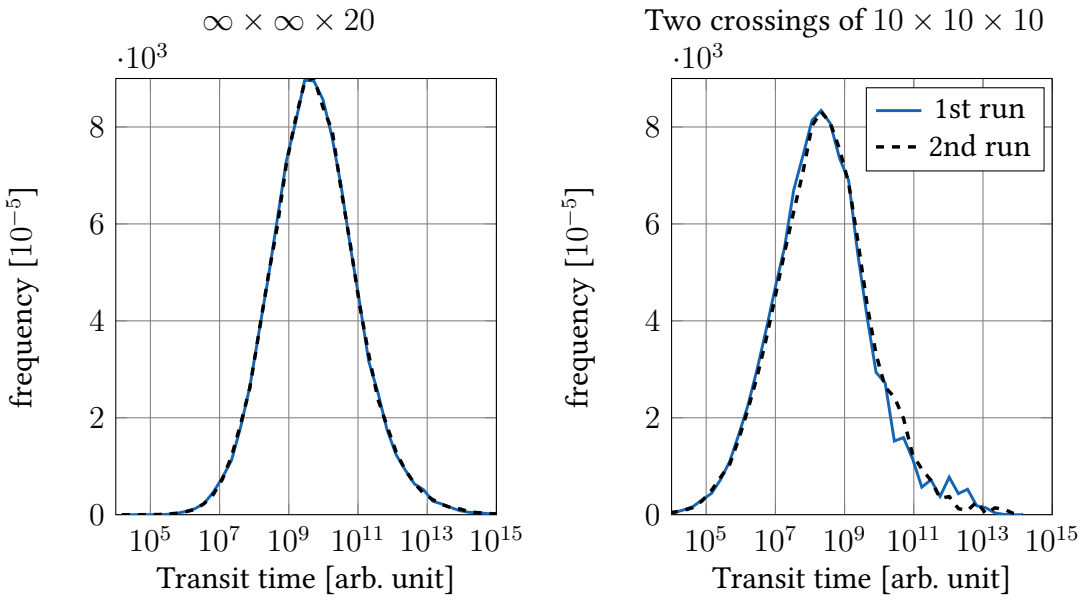

Figure 7.1: The finite size effect: The right tail of the empirical transit time distribution depends on the realisation and is incorrectly reproduced.

morphology is too small to properly account for the broadness of the rate resp. waiting time distribution.

We are not the first to point out this effect. In fact, Lukyanov & Andrienko [61, fig. 1] very nicely show that the size of the morphology needed to accurately reproduce charge transport grows exponentially in the effective disorder parameter $\hat{\sigma}$. In the same paper [61], it is suggested to simulate at higher temperatures, thus lower $\hat{\sigma}$, and extrapolate. This approach works perfectly fine for the equilibrium ($t \to \infty$) mobility resp. diffusion.

However, since we are interested in the short time behavior, we have to be very careful to separate numerical finite-size effects from the effects of the actual smallness of the devices we are considering. Using the extrapolation approach suggested in ref. 61, we solve the numerical problem, but we also completely get rid of the physical effect.

Therefore, we choose a different approach. We have developed an algorithm which does not use precomputed morphologies of a fixed size, but rather generates the environment along the trajectories of the charge carriers. This way we can simulate virtually infinitely large devices.

In a next step, we can restrict the size of the device in one dimension and study the effects of the smallness in this direction without needing to worry about interfering numerical finite-size effects.

Figure 7.1 shows the difference of the diffusive (no external potential applied) transit times of a $20\mathrm{nm}$ slab. On the left we use our algorithm, 20 layers in x direction, infinite

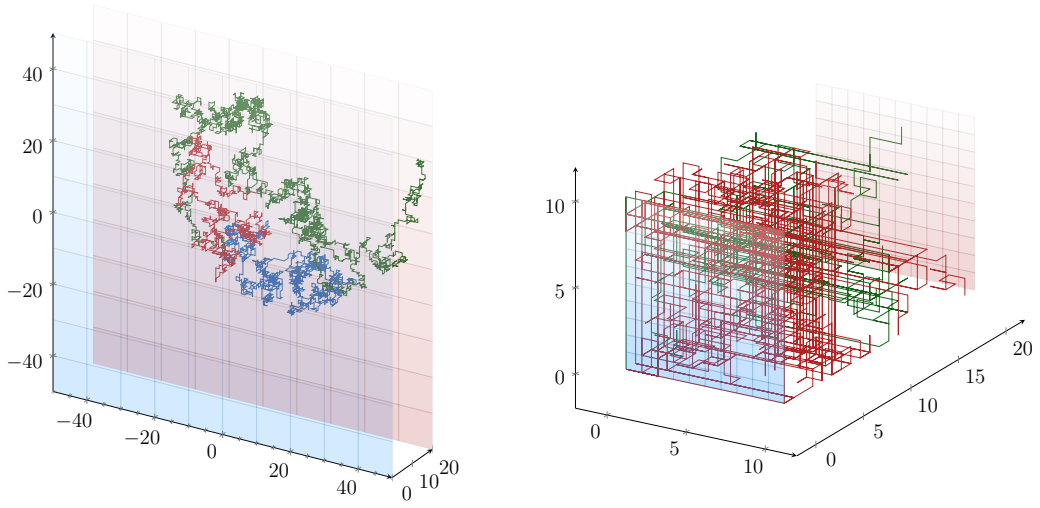

Figure 7.2: Comparison of the trajectories generated in the kMC simulations of a Slab.

in the other two. On the right, we use a $10 \times 10 \times 10$ box, periodic in y and z direction, which is traversed two times in x direction, see fig. 7.2. We used $\hat{\sigma} = 7$, which is a typical value for relevant devices at room temperature. In all kMC simulations we use 10^5 trajectories as sample size.

Obviously, the tails of the distribution are not correctly reproduced in the finite simulation, while, as we will see later (sec. 7.2), for $\hat{\sigma} = 7$, normal diffusion (*i.e.* equilibrium transport) is far from being attained in a 20nm device. Due to the strong influence of the far right tail on the expected value, the relatively small differences in tail lead to fluctuations of the expected value by orders of magnitude (note that the x-axis scaling is logarithmic). Therefore it is important to properly account for these effects when doing Monte Carlo simulations.

7.2 Dispersion

In this section we will introduce some data on the phenomenon that motivated our research – namely dispersion. By dispersion, we mean transport properties which are changing over time, specifically the diffusion constant resp. mobility in our case.

Direct experimental measurement – Impedance spectroscopy

Besides measuring IV-curves, another way to extract charge transport parameters is by impedance spectroscopy. We do not go into further detail how this is done here, neither will we model it. However, we want to include this because it is a very straightforward way to see that the effect we are describing is important.

Figure 7.3 shows field dependent mobilities extracted for three different device thicknesses. It is obvious that the extracted mobility parameter – thus by the Einstein relation (eq. (3.2)) also the diffusion constant – varies with the thickness of the sample.

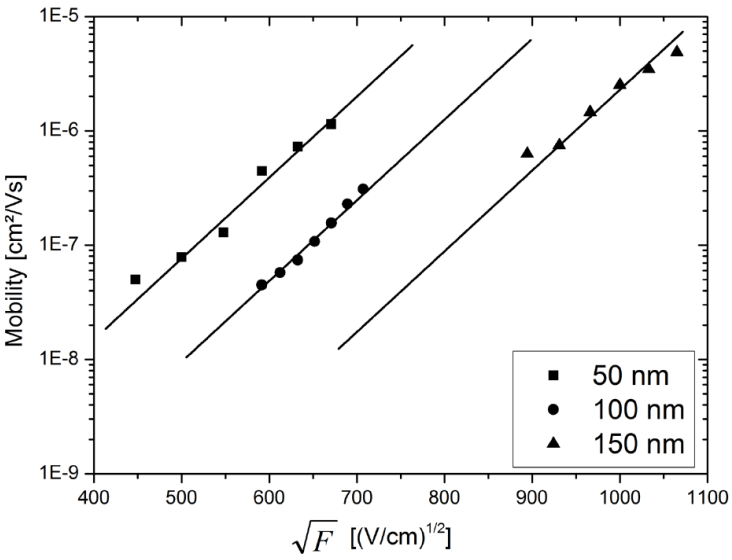

Figure 7.3: Mobility extracted from impedance spectroscopy (courtesy of M. Al-Helwi).

Indirect experimental measurement – IV curves

This investigation is a bit less rigorous compared to the rest of the thesis. Mostly, because here, it is hard to determine rigorously where effects come from and which parameters influence what kind of behavior exactly. We are going to model a full device, trying to reproduce actual IV curve (*i.e.* current density vs. voltage) measurements. We still want to include this discussion, because it shows that the ideas and results of this thesis can be applied to current cutting edge engineering problems.

The data consists of 9 IV curves for the same material, measured at different temperatures and with different thicknesses of the intrinsic (*i.e.* undoped) middle layer (see fig. 7.4 for a schematic of the diode). We apply a state of the art model (ECDM) and estimate several parameters (including the ECDM mobility parameters, doping concentrations for the injection layers and energy barriers) in order to describe the data.The model is presented in chap. C; For the numerical procedure used to solve the model equations, see Stodtmann *et al.* [109].

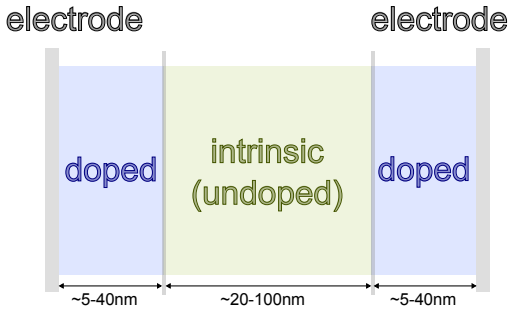

Figure 7.4: Schematic of a unipolar organic device

The initial fit shows a systematic failure to describe the correct variation of the IV curves with the varied thickness of the intrinsic layer – see fig. 7.5. If we fit each thickness independently, we get a good agreement for the temperature and voltage dependence, but of course the fit of the other thicknesses becomes even worse. This is shown in fig. 7.7.

To remedy this problem, we augment the model, by adding an additional state for charge carriers, which we call traps in the following. The parameters for the traps are energetic depth and concentration. In each step of the drift diffusion computation, the charge carriers are split up into free and trapped species according to the Fermi-Dirac distribution. The trapped carriers only contribute to the field, while the free carriers also enter into the expression for the current density. The result of this is shown in fig. 7.6. Adding the traps makes it possible to reproduce the correct variation of the current density with the thickness. The model is described in more detail in sec. C.2.

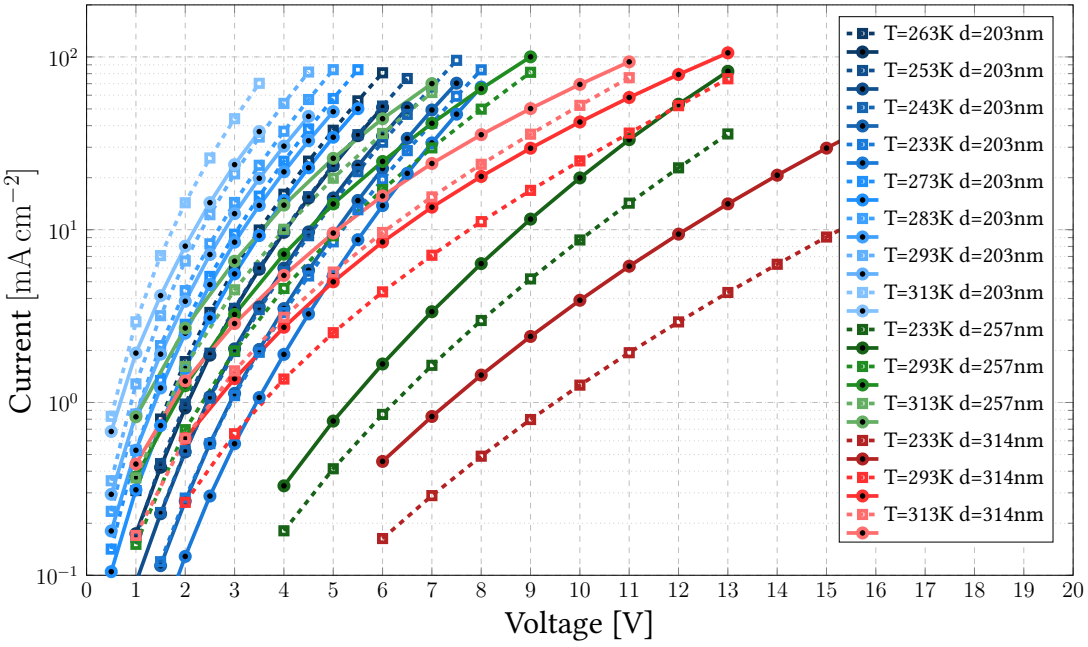

Figure 7.5: Best fit for the IV curves using the ECDM model. Dashed/empty square are the experimental curves, solid/dots the model response.

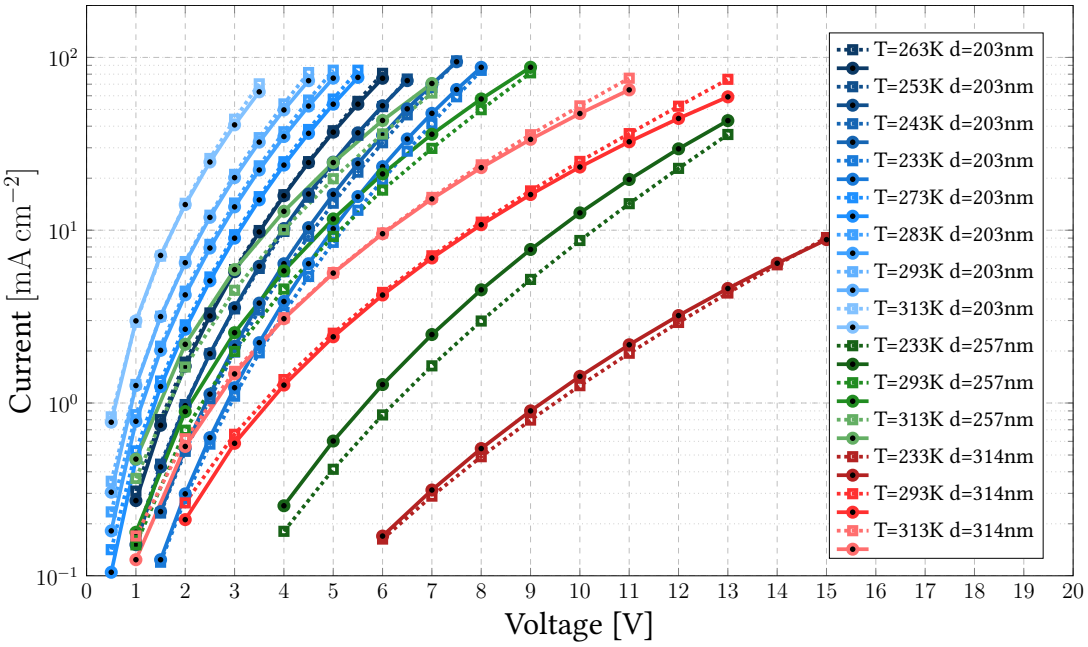

Figure 7.6: Best fit for the IV curves using the augmented ECDM model.

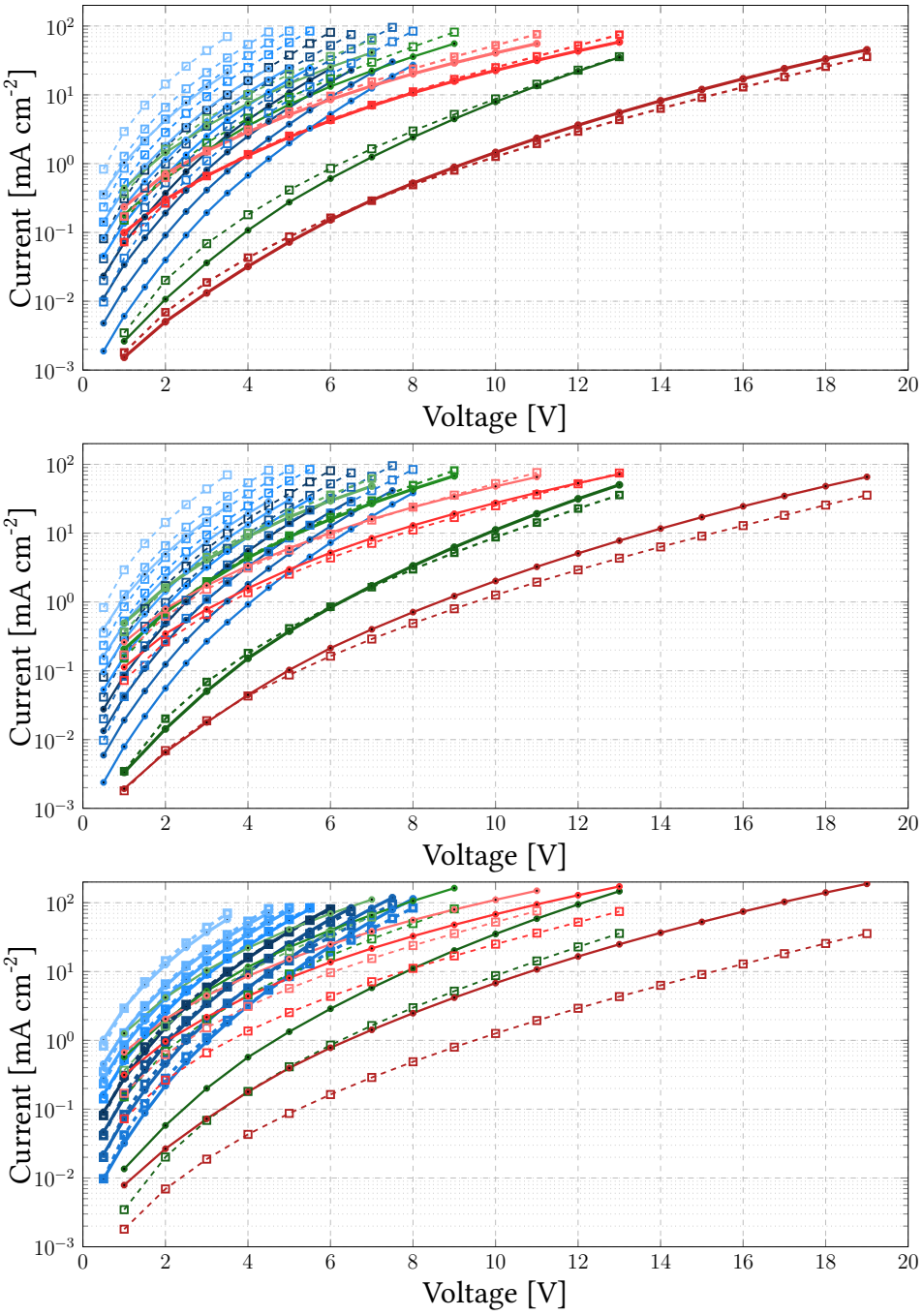

Figure 7.7: Best fit for the IV curves using the ECDM model. Here, the parameters are extracted for one thickness at a time, thick (left) to thin (right).

7.3 Comparison with the continuum limit

We test the validity of our proposed continuum model by comparing an invariant of the limit equation with the microscopic motion. We consider instead of the empirical diffusion coefficient

$$D^{\text{Emp.}}(t) := \frac{X(t)^2 - X(0)}{t}, \tag{7.1}$$

an *empirical anomalous diffusion coefficient*

$$D_\alpha^{\text{Emp.}}(t) := \frac{X(t)^2 - X(0)}{t^\alpha}. \tag{7.2}$$

As we have shown in sec. 7.1, we have to be careful about finite size effect. Therefore, we can not resort to other charge transport simulation packages, but have to use our own code which uses an infinitely large morphology. We will investigate the distribution of the empirical diffusion coefficient, again using a histogram for 10^5 realizations – the results are shown in fig. 7.9.

Motivated by the scaling limit, we use logarithmically spaced points in time. We find that for large values of the effective disorder $\hat{\sigma}$, the distributions change relatively uniformly between these times if plotted on a logarithmic scale too. This is exactly what one would expect from the scaling limit, since both, time and energy, are exponentially scaled.

We also see that for medium values of $\hat{\sigma}$ this behavior saturates eventually and diffusion will be a good description of the process at long times. Even more, for very low or no disorder, the distribution only changes slightly resp. not at. This is what one would expect, since the model then degenerates to the simple random walk on $\mathbb{Z}^3$, which rapidly converges to Brownian motion.

In fig. 7.8, we compare the evolution of the histograms of the particle positions for kMC simulations of the simple model for various effective disorders $\hat{\sigma}$ with the evolution of the solution to the fractional diffusion equation (eq. (6.3)) for different dispersion exponent α. Note that the case $\alpha = 1$ corresponds to normal diffusion and, again, the case $\hat{\sigma} = 0$ corresponds to the simple random walk.
One can clearly see that increasing $\hat{\sigma}$ corresponds to decreasing α and thus a more anomalous behavior. What we mainly are interested in in these images is the qualitative similarities, as we would not expect perfect agreement. In the beginning there are still residuals of the discreteness of the model, and eventually a diffusive limit will be attained. However, it is apparent that for large disorder at intermediate times, the density is much closer to a fractional than a normal diffusive evolution.

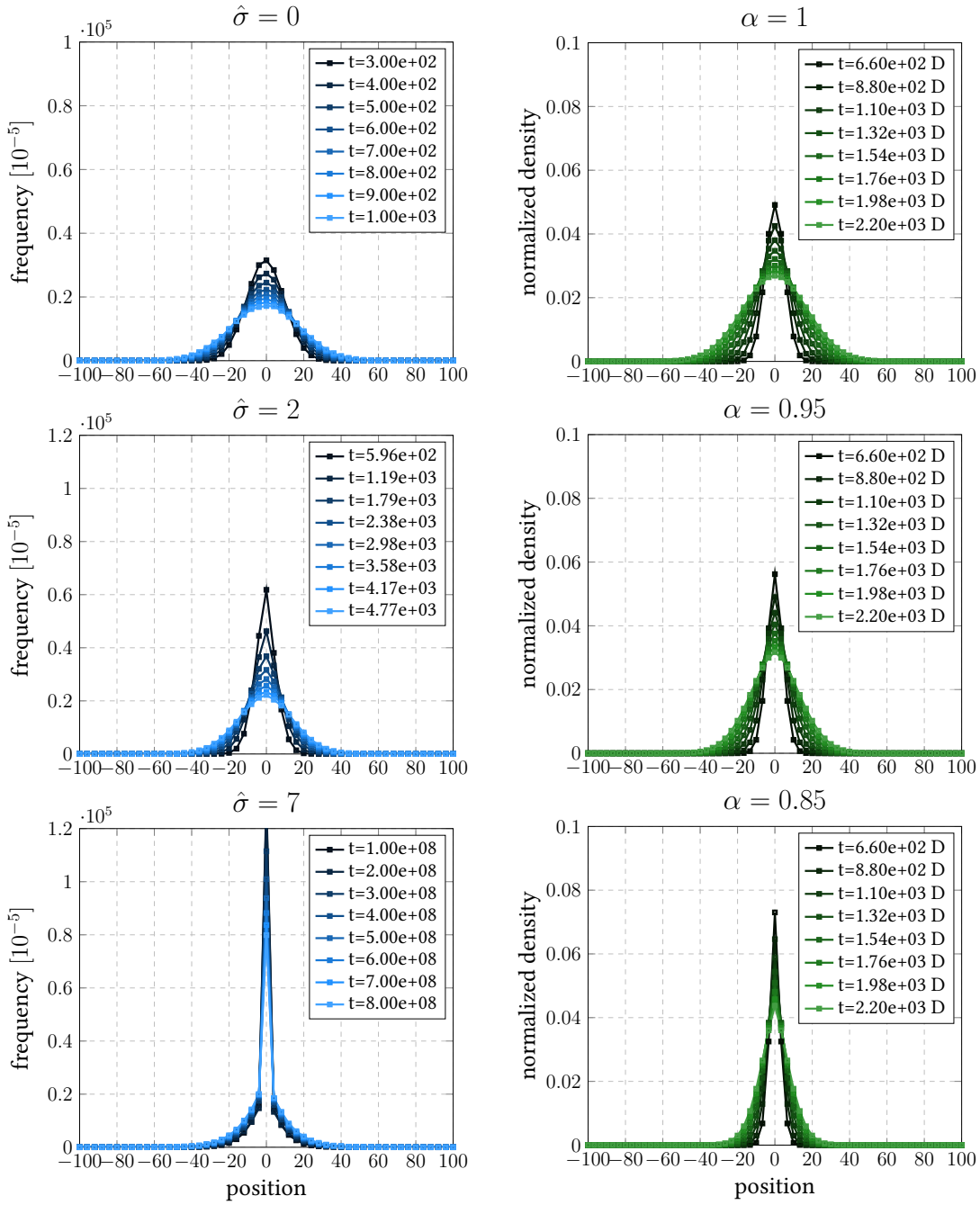

Figure 7.8: (left) Evolution of the probability density of a 1-dimensional projection (arbitrary due to rotational symmetry) of the simplified process for different effective disorder strengths $\hat{\sigma}$. (right) Solutions of the fractional diffusion equation for different dispersion exponents α. Time is measured in multiples of the (generalized) diffusion constant D.

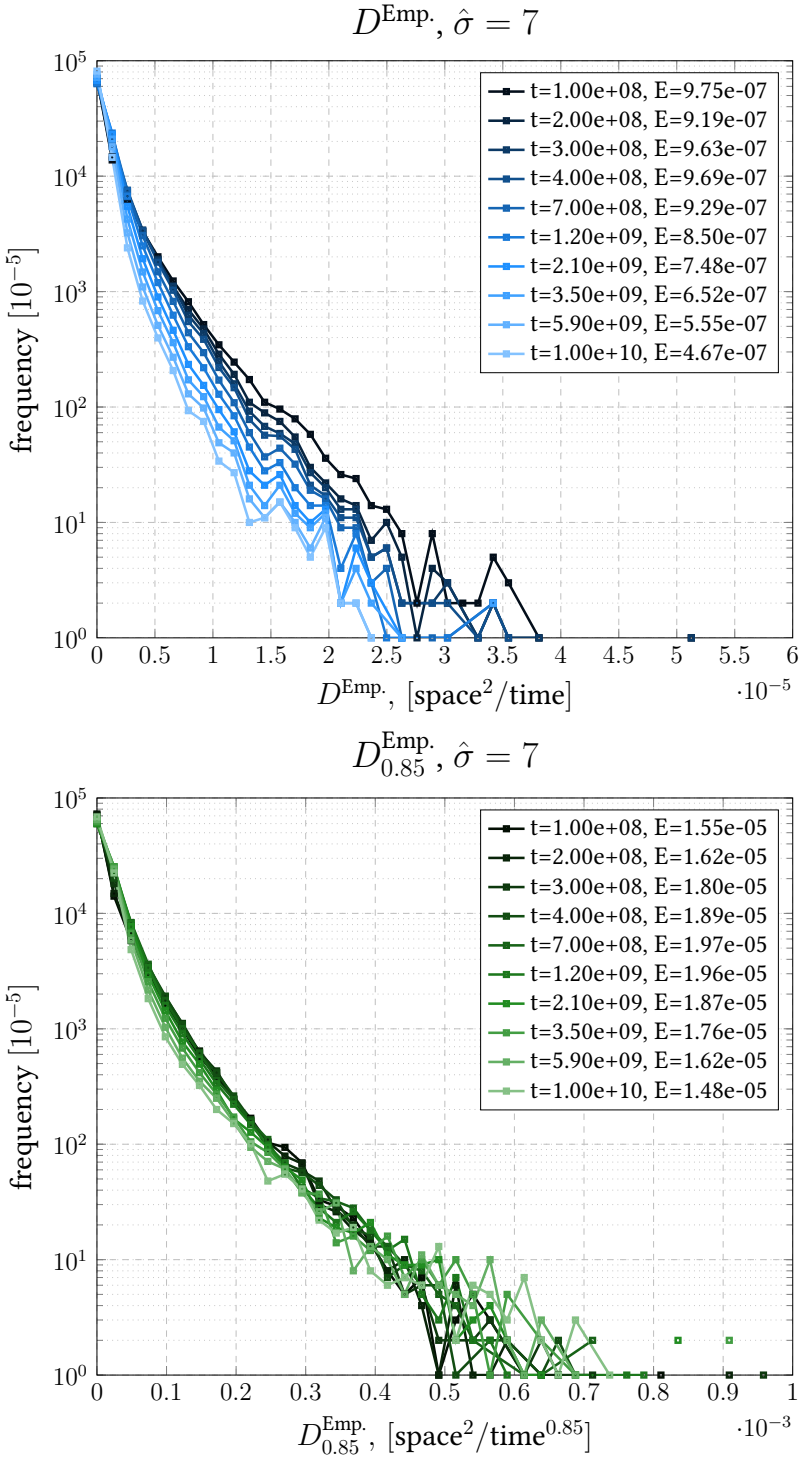

Figure 7.9: Distribution of "normal" diffusion coefficient $D(t)$ (upper) vs. anomalous diffusion coefficient $D_{0.85}$ (lower) for different times at effective disorder $\hat{\sigma} = 7$.

7.4 Simplifications

We recall the simplifications we have made in our model of bulk charge transport for organic materials:

- complex amorphous morphology replaced by simple cubic lattice,
- long range transfers neglected (only nearest neighbor),
- Marcus/Miller-Abrahams reaction rates replaced by simple holding times τ_x.

These simplifications have in some cases recently been treated in the literature. Where we feel, that additional justification is needed, we perform our own kinetic Monte Carlo simulations to study the effect of these simplifications.

Simplified geometry

The impact of using different geometries was studied by Cottaar *et al.* [34] and Massé *et al.* [66]. Both works compare simple cubic lattice with face centered cubic lattice geometry as well as Miller-Abrahams and Marcus rates. In all cases, they only observe a change of the transport behavior by a constant factor. This is expected anyway because of the different parametrization of the models. Therefore, we can conclude that for qualitative aspects, it is not necessary to distinguish.

In particular, if we still allow for a constant factor as fit parameter when comparing to experimental data and concentrate on whether or not certain effects, their direction and relative magnitude are included in the model, it is justifiable to go to the simplest case, which is done in this thesis.

Simplified rate expressions

As already pointed out above, refs. 34, 66 compare Marcus and Miller-Abrahams rate expressions and do not find a qualitative difference. Inspired by the fact that both of these rate expressions have a common invariant measure (cf. eqs. 2.4, 2.5), we introduced a simplified model in sec. 4.1, which is the simplest rate with this reversible measure.

In fig. 7.10, we compare the evolution of the charge carrier density for the Miller-Abrahams against our simplified rate expressions. Furthermore we investigate the distribution of the normal and fractional empirical diffusion constant. For the Miller-Abrahams rates, the distance d_{xy} are assumed to be normally distributed, this corresponds to a lognormal distribution of the overlap parameter $e^{\gamma_{\text{decay}} d_{xy}} \sim J_{xy}$ (cf. May [67, fig. 2.9]).

As expected, the results for the full rate and the simplified rate are qualitatively similar. They differ in a constant in time and space scaling. While we already accounted for the

additional constants in the rate expressions (*i.e.* rescaled to prefactor 1), the additional stochastic terms in the full model are not taken account for. Furthermore we sampled for a more restricted time range in the full model, because the memory requirements for longer trajectories grow much faster than for the simplified model. Still the basic features – the sharp peaked density and the fact that a fractional diffusion coefficient better describes the motion – persist in the general case too.

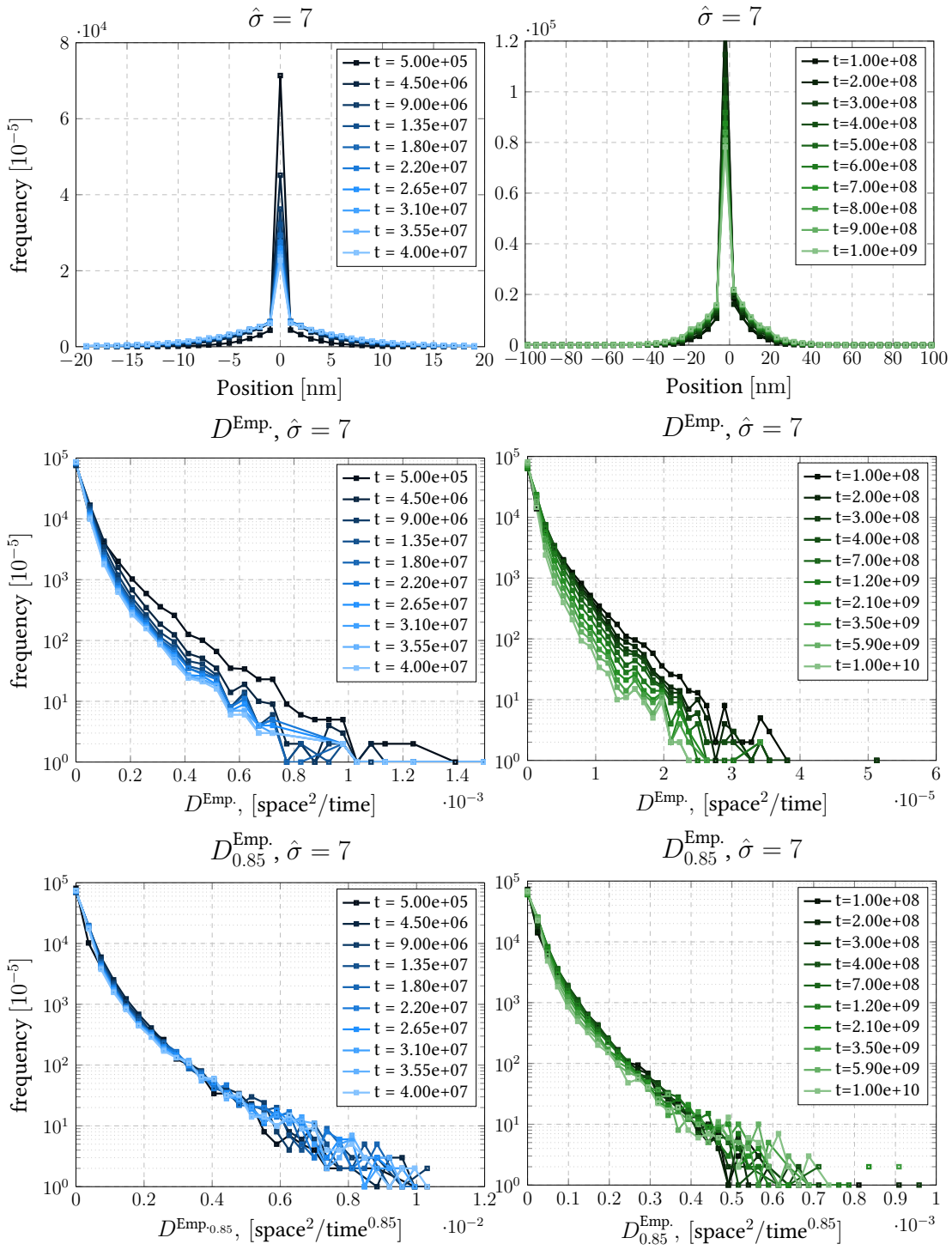

Figure 7.10: Comparison of transit times for the full Miller Abrahams rate expression (left) against the simplified model proposed in sec. 4.1 (right).

CHAPTER 8

Conclusions and outlook

Starting out from a physically motivated microscale model, we have used mathematical and intuitive arguments to deduce a simplified stochastic microscale model. We identified this model as a variant of the *Bouchaud trap model* with lognormally distributed mean waiting times. Generalizing a previous result for the BTM by Ben Arous & Černý [15], we show that, on carefully chosen scales, the microscopic model admits a scaling limit. The density of the stochastic process which occurs as the limit, can be computed using a time-fractional diffusion equation. This is particularly appealing, since it gives a explanation for dispersive transport effects, which is intuitively easy to understand.

We have chosen the term "intermediate asymptotics" to point out that due to our choice of scales, the limit is most suited for the description of the process on scales which, while they are large compared to molecules, are still macroscopically small. In particular, both mathematical and experimental arguments suggest that on large scales, the process will behave diffusively. However, we show in simulations, that for a $20\mathrm{nm}$ thick device with effective disorder $\sigma/k_bT = 7$, the fractional diffusion coefficient $D_{0.85} = x^2/t^{0.85}$ (x is the distance traveled, t time) yields a more reliable characterization of the charge transport than the conventional diffusion coefficient $D = x^2/t$.

To be able to study the dispersive effects postulated by the scaling limit numerically, we develop a Monte Carlo algorithm, which does not suffer from the finite size effect. This is important because dispersion can be considered a physical finite size effect. Due to the broad distribution of a local transport parameter (here the waiting time) the small macroscopic dimensions of the device do not allow a complete sampling of the statistics of the microscopic properties. Via the scaling limit, we can extract the effect of this incomplete exploration on a typical charge carrier. As suggested by the results of our simulations, the limit is valid for devices with large energetic disorder or at low temperatures at the practically relevant time- and lengthscales.

Another advantage of the scaling limit is, that in principle, full device simulations similar to drift-diffusion models, which were very successful in the past, are possible. The only modification needed is replacing the time derivative in those equations

by a fractional one. A minor drawback here is, that we only rigorously considered a highly simplified situation in this work. Still, we are confident that this approach can be generalized to those situations, such as the full physical rate expressions or realistic morphologies. In the following, we outline how we expect that this could be achieved. Furthermore we explore extensions that include the effects of charge carrier interaction and electrostatic potential.

8.1 More realistic models

In this section, we look at models closer to the microscopic physical model (cf. sec. 2.2) and how they can be treated. We expect the result to be the same in almost all cases.The arguments are similar to those, we used to justify our simplified model in sec. 4.1. Having seen the details of the proof, we can be more precise now.

Realistic rates

In order to treat realistic rates (eqs. (2.2) and (2.3)), we have to take into account the symmetric part too, which depends on the energy at both, the origin and the target of the jump. Furthermore through the wavefunction overlap, an additional (symmetric) randomness would need to be introduced . A case, which has a simple dependence of both origin and target and yet the same reversible measure is the one with transition rates

$$\tau_{xy} = \tau_x^{a-1}\tau_y^a, \tag{8.1}$$

for some $a \in [0, 1]$. It is known in the literature as Bouchauds *asymmetric* trap model. This case is treated by Barlow & Černý [11], again with IP tails for the τ_x, which still are the reversible measure, but can no longer be interpreted straightforwardly as waiting times.

Note. This is the model treated here for $a = 0$, and almost looks like the Metropolis/Miller-Abrahams rate for $a = 1/2$ (except for the cutoff when the rate exceeds 1). This case has the rates $e^{-\beta/2(\mathbf{E}_x - \mathbf{E}_y)}$. They are sometimes referred to as Boltzmann factor.

Remark. This treatment only deviates from Miller-Abrahams rates, when the transitions are exceptionally fast. Since even in the case where fast transitions are not cut off, sub-diffusive behavior emerges, we can expect this to hold as well in the case when we cut off the transition rates at a maximum.

In fact, the proof is very similar to the proof in chap. 5. The only difference being the spatial object, which is not the simple random walk, but a more complicated process – one, which jumps from x to y with rate $\tau_x^a \tau_y^a$ (instead of $1/6$). This process is referred to

as the variable speed random walk (VSRW) for the random conductance model (RCM) or a random walk among random conductances (RWRC) and is harder to deal with than the simple random walk.

For the particular case of IP-distributed mean waiting times which additionally are almost surely bounded away from zero, Green's function estimates are given by Barlow & Deuschel [12] (here, the estimates themselves are random, which complicates the coarse graining procedure). The proof for this particular case by Barlow & Černý [11] makes use of those Green's function estimates and introduces a more abstract coarse graining procedure which only needs those random bounds as input.

The limit is up to a constant the same as in the model we considered here. Therefore, we would expect that using the full rate expressions does not change the result in our case as well.

To really prove this, one has to check that the triangular array methods we have used here are compatible with the methods used in ref. 11. Additionally, Green's function estimates in the spirit of ref. 12 would be required for non-IP waiting times which furthermore are not bounded away from zero. A result in this direction is the invariance principle for the RWRC by Andres *et al.* [3]. Their result only assumes finite moments of the conductances of order $q \in (1, \infty)$ and finite moments of the inverses of order $p \in (1, \infty)$ for $1/q + 1/p < 2/d$.

The Marcus and Miller-Abrahams rates also include transfer integrals and additional symmetric terms which would be covered by the invariance principle too, as they are all assumed to have distributions for which all moments exist.

Realistic morphology

In this work, we assumed a simple cubic lattice with only nearest neighbor transitions. This choice is mainly due to the readily available sharp Green's function bounds in this case. For (even simple) random walks on more general graphs like face centered cubic lattices, next-nearest-neighbor hopping or even morphology snapshots from molecular dynamics computations, these estimates are not available. At least, if bounds are available, they are not as sharp as the ones used in the proof.

Nevertheless, we expect no qualitative difference as long as the number of neighbors stays finite and the longest possible jump is still small compared to the dimensions of the device – or in the terms of the proof, the observation horizon. Due to the exponential decay of the wavefunctions, these assumption is not violated in realistic systems.

The important effect of the morphology is the influence on the transition rates resp. site energies, which we already have decoupled from it by explicitly modeling the energies as random variables and not computing it from the geometry.

Correlated energies

This case is particularly hard to treat mathematically, but we want to mention it for the sake of completeness. It is sometimes assumed that the random energies are spatially correlated [41]. This makes sense *e.g.* if one assumes that the major part of the randomness is caused by a random dipolar background [37].

If the correlation is very strong and the scale is comparable to the scale we want to observe charge transport on, we probably do not have much of a chance to get the same scaling limit. We expect, that in this case, the correlation will influence the limit.

If the correlation is very short ranged (*i.e.* small compared to the observation scale), we can probably just include it in the proof by introducing a correlation scale χ exponentially smaller than the proximity scale ν and use balls with radius χ around traps instead of traps themselves. The arguments should be similar in this case.

Very long ranged correlations can probably be considered as macroscopic electrostatic field in the limit equation, neglecting the influence on the neighbor transitions on the microscale. The consequences of electrostatic fields are discussed later.

8.2 Extensions

Naturally, once we have understood the behavior of a single carrier in the absence of an electric field, we want to go beyond this and study the influence of charge carrier interactions and electrostatic potential. We will sketch here, how this could be done in our framework.

Low density

For small densities, we conjecture that the exact same limiting equation holds. This could be proven by choosing a density which goes to ∞ as $n \to \infty$ in such a way that the probability of two carriers meeting before exiting $B_0(R)$ is $o(h^{-2})$.

If this is fulfilled, we can introduce the Pauli principle – or site-exclusion as it is often called in mathematics – as an additional interaction mechanism for the charge carriers. Due to our choice of the density, we can condition on the event that no two charge carriers ever meet while keeping the error introduced by the conditioning negligible compared to h^{-2}. Therefore the proof could be essentially carried out as before, only that now, one would consider the joint distribution functions and Laplace transforms of multiple trajectories, which, due to our diluteness assumption can be treated as independent at an $o\left(h^{-2}\right)$ error.

High density

We again consider site exclusion as interaction mechanism, but this time choose the density in a way, such that the number of deep traps is small compared to the density. If we do that, we can assume that a typical charge carrier does not spend any time in the set of deep traps, therefore no jumps in the clock process will typically occur and the charge transport is diffusive, even on small spatial scales and for large disorder.

Larger scales

For the a CTRW-model with the same energetic landscape as we use (which leads to essentially the same model without the correlation in the waiting times, see sec. 3.1.5), Kotulski [55] shows, that the density is asymptotically Gaussian. On longer scales, the effect of dependence plays even less a role than in this work. Here, the limit already does not qualitatively deviate from the limit assuming independence. Therefore, we expect that the result from the CTRW case carries over and that the limit on a diffusive scale exists and is Brownian motion – this is not too surprising, since all moments of the waiting time (and also the spatial increment) distribution exist finitely.

Digression: Possible continuous models with extended range of validity

Considering our intuition for low density, high density and large scales, a model, which interpolates between fractional behavior at low densities and short times and normal diffusion at high densities or on large scales, would be a good candidate for a continuum model with an extended range of validity. We will now briefly introduce two such models.

The first model has been introduced in Mantegna & Stanley [63] and has been applied to the problem of charge transport in organic semiconductors recently by Sibatov & Uchaikin [107]. They suggest exponentially truncating the (Lévy) integral kernel in the fractional derivative, therefore making the operator essentially local on large scales. While it does not come naturally with a density dependence, it captures the relaxation behavior over time well. The corresponding probability distributions of the time steps are called *tempered stable*.

The second model has been proposed by Mommer & Lebiedz [75] and introduces an adaptive model with multiple species with different diffusion constants. Over time, the faster species are converted into the slower ones. With appropriate rates, it is shown that this can mimic fractional behavior on short timescales while “essentially returning to normal diffusion” on large ones [75].

By adding density constraints for the slow species, it is also possible to naturally include the effect of density dependence in these models. However, neither of the two models presented here has been rigorously connected to the microscale. Therefore it is not

clear which one is the better description or how the macroscopic model parameters depend on the molecular properties of the material.

We will later show (see sec. 7.2) that a very simplistic version of the second model already yields a correction of the drift diffusion equations which can help to better describe lab experiments. However, in doing this we sacrifice the possibility to microscopically interpret the model parameters.

While it is not in the scope of this thesis, we believe that a further investigation into this type of models (*i.e.* structured population dynamics/diffusion models) could re-establish this connection and make them valuable methods for modeling charge transport in organic devices, or more generally, (initially) dispersive transport.

Electrostatic potential

Typically, one is not only interested in the pure diffusion of charge carriers, but also migration due to an applied electrostatic potential difference – and thus an electric field.

Including the field can in principle be done at the microscopic scale - by adding the electrostatic potential to the random energies $\mathbf{E}_x$. This would influence both the Green's functions and the mean waiting times.

For small fields, the qualitative behavior is not likely to change until the effect of the field becomes such, that either the paths become almost straight lines or the potential gain in between two sites is of the order of the typical (deep) trap depth.

In the first case, the number of different sites visited will become smaller, thus making the device thinner effectively. This enhances the deviation from long time equilibrium (*i.e.* diffusive behavior) further.
In the second case, the field cancels the effect of being in a deep trap, in this case, fractional diffusion is not expected, since it is the signature of the broad trap depth distribution.

If the applied potential difference is small, a natural approach to incorporate the effect of the field would be using a *fractional drift-diffusion equation* [59], denoting with ψ the *electrostatic potential*

$$\frac{\partial^\alpha \rho}{\partial t^\alpha} = \mu \nabla \psi \cdot \nabla \rho + D \Delta \rho. \tag{8.2}$$

Coulomb interaction

We have already discussed adding the Pauli principle as interaction mechanism. One could go even further and consider Coulomb interaction. The problem with handling this type of interaction is that it is relatively long range.

Obviously, if we add an assumption on the diluteness of the system which ensures that the interactions are negligible, the result will not differ from the low density limit described above. While medium or higher densities will be very hard to handle analytically, the Coulomb interaction can be taken into account on the macroscopic scale simply by solving a Poisson equation for the electrostatic potential, where the boundary values are given by the applied voltage and the source density by the charge carrier densities. Thus, if we know how to handle electric fields, the mean field effect of Coulomb interaction can be easily added to the model.

PART IV

APPENDIX

APPENDIX A

Concepts of probability theory

In this appendix, we introduce the mathematical concepts used in the work. These are mainly drawn from probability theory and adjacent areas like potential theory or Skorokhod spaces. Even in those cases, our approach will be a probabilistic one.

We start from the very basic concepts in sec. A.1, but in later sections we will also introduce specialized results that are needed as ingredients for the proof of the principal theorem. In sec. A.2, we introduce Lévy Processes, with a particular focus on subordinators. These are then used to define the fractional kinetics process which appears as scaling limit of the charge transport process (cf. thm. 4.2.2). Section A.3 treats the potential theory of the simple random walk. The proof of the main result will heavily make use of the estimates derived in this section. Some limit theorems for sums of random variables we will use for intermediate steps are formulated in sec. A.4. For some steps in the proof, we will have to work in so-called Skorokhod spaces. We give some results and definitions about these spaces in sec. A.5. Finally, sec. A.6 discusses the similarities of distributions with inverse polynomial tail at infinity and the lognormal distribution. This section motivates the scaling used to derive our result.

A.1 Fundamentals

This work is aimed to be as self contained as possible, but it is not meant to be educational. A good introduction to measure and elementary probability theory can be found e.g. in Breiman [27] and Billingsley [19].

The general arena where probability theory is set is a complete filtrated probability space. In order to define this space, we need some other definitions first.

Definition A.1.1. A *σ-algebra* $\mathcal{A}$ on a space Ω is a family of subsets of Ω, which fulfills

i) $\Omega \in \mathcal{A}$,

ii) $A \in \mathcal{A} \implies A^{\text{Comp}} \in \mathcal{A}$,

iii) $A_i \in \mathcal{A} \forall i \in \mathbb{N} \implies \bigcup_{i\in\mathbb{N}} A_i \in \mathcal{A}$.

Definition A.1.2. A *filtration* $\mathcal{F} := (\mathcal{F}_i)_{i\in I}$, is a family of σ-algebras, which fulfills $\mathcal{F}_s \subset \mathcal{F}_t$ for all $s \leq t \in I$.

Remark. The index set I of the filtration is typically either $\mathbb{N}$ for discrete time or $\mathbb{R}^+$ for continuous time.

Definition A.1.3. A tuple $(\Omega, \mathcal{A})$ is called *measurable space*, a function $\mu : \mathcal{A} \to \mathbb{R}$ is called a *measure*, if

i) For all $A \in \mathcal{A}$, $\mu(A) \geq 0$,

ii) $\mu(\emptyset) = 0$,

iii) $A_i \in \mathcal{A}, \forall i \in \mathbb{N}\,,\ A_i \cap A_j = \emptyset\, \forall i \neq j \implies \mu\left(\bigcup_{i\in\mathbb{N}} A_i\right) = \sum_{i\in\mathbb{N}} \mu\left(A_i\right)$.

Definition A.1.4. A measure $\mathbb{P} : \mathcal{A} \to [0, 1]$ on $(\sigma, \mathcal{A})$ fulfilling $\mathbb{P}\left(\Omega\right) = 1$ is called a *probability measure*. The triple $(\Omega, \mathcal{A}, \mathbb{P})$ is called a *probability space*.

Definition A.1.5. A triple $(\Omega, \mathcal{F}, \mathbb{P})$ is called a *filtrated probability space*. If additionally, all $\mathbb{P}\left(\cdot\right)$ nullsets are contained in $\mathcal{F}_0$, it is called *complete*.

Now that we have set up the arena, we will introduce the objects we are going to study, namely random variables, stochastic processes and stopping times.

Definition A.1.6. Let $(\Omega, \mathcal{A})$, $(\Omega', \mathcal{A}')$ be measurable spaces. A map $f : \Omega \to \Omega'$ such that for all $A \in \mathcal{A}'$, $f^{-1}(A) \in \mathcal{A}$ is called *measurable*.

Definition A.1.7. A measurable map from a probability space to a measurable space is called *random variable* (RV).

Note. We will not go into the topics of versions, distinguishability and modifications here. If we say a process has some property, what we usually mean is, that it has a modification which has the property.

Furthermore, since all processes we are interested in are right continuous, this already implies that they are indistinguishable, which means that the probability for a sample path of the modification to be different from the original process is 0.

Therefore it would make some of the upcoming statements more technical without adding any further information. So, if in doubt, all properties hold only for modifications of the processes and we always chose the most regular modification of the processes in question.

See *e.g.* Karatzas & Shreve [48, sec. 1.1] for a discussion.

Definition A.1.8. Let $X : (\Omega, \mathcal{F}, \mathbb{P}) \to (\Omega', \mathcal{F}')$. The measure $\mathbb{P}_X(\cdot) := \mathbb{P}\left(X^{-1}(\cdot)\right)$ is called the *distribution* or the *law* of X.

Remark. It is easy to check, that this turns $(\Omega', \mathcal{F}', \mathbb{P}_X)$ into a probability space.

Definition A.1.9. For a random variable X, the *σ-algebra generated by X*, denoted $\sigma(X_i)$, is the smallest σ-algebra w.r.t. which X_i is measurable.

Definition A.1.10. A family $(A_i)_{i\in I} \subset \mathcal{A}$ is called *independent* if for all $J \subset I$,

$$\mathbb{P}\left(\bigcap_{i\in J} A_i\right) = \prod_{i\in J} \mathbb{P}(A_i).$$

A family of random variables is called independent, if the σ-algebras it generates are independent.

An easy to prove, but immensely powerful result we will use very often is the following.

Lemma A.1.11 (Borel-Cantelli Lemma)**.** *Let $(\Omega, \mathcal{A}, \mathbb{P})$ be a probability space, $A_i \in \mathcal{A}$ for all $i \in \mathbb{N}$. We define*

$$A := \limsup A_n = \bigcap_{n\in\mathbb{N}} \bigcup_{i=n}^{\infty} A_i.$$

Then

i) $\sum_{n\in\mathbb{N}} \mathbb{P}(A_n) \leq \infty \implies \mathbb{P}(A) = 0$,

ii) $\sum_{n\in\mathbb{N}} \mathbb{P}(A_n) = \infty \implies \mathbb{P}(A) = 0$, *if additionally the family $(A_n)_{n\in\mathbb{N}}$ is independent.*

Proof. see Breiman [27, chap. 3, sec. 3]. □

Remark. The event A can be described as the event that infinitely many of the A_n occur. This means, if $\mathbb{P}(A) = 0$, only finitely many of the A_n can occur and therefore it is possible to choose some n_0, such that for $n \geq n_0$ A_n does not occur with probability 1.

Definition A.1.12. Collections of random variables $X = (X_i)_{i\in I}$ are called stochastic process. If X_i is $\mathcal{F}_i$ measurable, for all $i \in I$, the process is called *adapted* to $\mathcal{F}$.

Note. For some applications, *e.g.* to conclude the measurability of stopped processes (see A.1.14 for the definition of stopping times) to be adapted, a stronger property - *progressive measurability* is in principle needed. However, all processes we consider are càdlàg (see def. A.1.20), and càdlàg, adapted processes are progressively measurable. Therefore we can skip this and rest assured that all stopped processes we consider are adapted.

We will often refer to stochastic processes just by X and to filtrations by $\mathcal{F}$, omitting the index in cases where no confusion can arise.

Definition A.1.13. The filtration $(\sigma(X_i))_{i\in I}$ is called *natural* for the process X.

Definition A.1.14. A random variable $\tau : \Omega \to \mathbb{N} \cup \{\infty\}$ is called a *random time*. A random time is called *stopping time* w.r.t. a filtration $\mathcal{F}$, if $\{\tau = t\} \in \mathcal{F}_t \, \forall t \in I$. A process $(X_{t\wedge\tau})_{t\in I}$ is called a *stopped process*.

An example for a stopping times, we will use throughout the work is the hitting times of a set A for a stochastic process $(X_i)_{i\in I}$:

$$\mathsf{Hit}_x^A := \inf\left\{i \in I \, : \, X_i^{(x)} \in A\right\}$$

is a stopping time for the natural Filtration associated to $(X_i^{(x)})_{i\in\mathbb{N}}$. An important way to create stopping times is the following.

Lemma A.1.15. *Let τ_1, τ_2 be stopping times for a filtration $\mathcal{F}$, then $\tau_1 \vee \tau_2, \tau_1 \wedge \tau_2, \tau_1 + \tau_2$, and $\alpha\tau_1$ ($\alpha \geq 1$) are stopping times for $\mathcal{F}$ as well.*

Proof. see Sato [97, prop. 4.0.8]. □

We need some more concepts to formulate our results.

Definition A.1.16. The *sigma algebra at a stopping time* τ is defined as

$$\mathcal{F}_\tau := \{A \in \mathcal{F}_\infty \, : \, A \cap \{\tau \leq n\} \in \mathcal{F}_n \forall n \geq 0\}.$$

This definition makes sense, because $\mathcal{F}_\tau$ is indeed a σ-algebra (see Sato [97, prop. 4.0.8]).

Definition A.1.17. A filtration $\mathcal{F}$ is called *right continuous, i.e.* $\mathcal{F}_t = \bigcap_{s>t} \mathcal{F}_s$.

This property is of course not applicable for countable I, because in that case, only a constant Filtration can have the property. However, in that case we also never need it.

Definition A.1.18. We define *left* and *right limit* $f(t-)$ resp. $f(t+)$ as

$$f(t-) := \lim_{s\uparrow t} f(s), \text{resp. } f(t+) := \lim_{s\downarrow t} f(s).$$

$\uparrow$ and $\downarrow$ denote the one-sided limits.

Definition A.1.19. A process is called right/left continuous if $X(t\pm) = X(t)$ for all $t \in I$.

Definition A.1.20. A stochastic process X is called *càdlàg* if the *sample paths* – the maps $i \mapsto X_i$ for fixed $\omega \in \Omega$ – are almost surely right continuous and have left limits.

Càdlàg processes with some additional properties, so-called Lévy processes are treated in sec. A.2, the spaces of càdlàgfunctions as such, so-called Skorokhod spaces in sec. A.5.

We now want to introduce the expected value and conditional expected value, culminating in the definition of martingales, a very important class of stochastic processes.

Definition A.1.21. We define the *expected value* of a random variable X as

$$\mathbb{E}[X] = \int_\Omega X(\omega) \mathrm{d}\mathbb{P}(\omega).$$

If $\mathbb{E}[X]$ exists and is finite, we say that X is *integrable* and write $X \in L^1$.

Definition A.1.22. The *conditional expectation* of a random variable X w.r.t a σ-algebra $\mathcal{A}$, $\mathbb{E}[X|\mathcal{A}]$ is defined as as the $\mathcal{A}$-measurable random variable, which fulfills

$$\mathbb{E}\left[\mathbf{1}_{\{A\}} \mathbb{E}[X|\mathcal{A}]\right] = \mathbb{E}\left[\mathbf{1}_{\{A\}}\right], \quad \text{for all } A \in \mathcal{A}.$$

Note. It is justified to speak of "the" conditional expectation, as it can be shown that it is almost surely unique (see e.g. Billingsley [19, p. 445]).

Definition A.1.23. A real-valued, $\mathcal{F}$-adapted process X is called a *martingale* w.r.t $\mathcal{F}$, if $X \in L^1(\mathbb{P}(\cdot))$ and for $s \leq t$,

$$\mathbb{E}[X_t|\mathcal{F}_s] = X_s \text{ almost surely.}$$

Since we will not use them, we skip for better readability the treatment of sub- and supermartingales. However, many results for martingales hold in these cases, too. An important property of martingales is, that their expectation at any bounded stopping time is the same as in the beginning.

Theorem A.1.24 (Optional sampling)**.** *Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a filtrated probability space, M_n a $\mathcal{F}$-adapted L^1-martingale, τ an a.s. finite stopping time. Then,*

$$\int_{\{\tau > n\}} |M_n| \mathrm{d}\mathbb{P} \overset{n\to\infty}{\longrightarrow} 0 \implies \mathbb{E}[M_\tau] = \mathbb{E}[M_0].$$

Proof. See Breiman [27, thm. 5.10]. □

Another property that gives rise to a rich structure for stochastic processes is the so-called Markov property – or intuitively – memorylessness of the process.

Definition A.1.25. Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a filtrated probability space, $(\Omega', \mathcal{F}')$ a measurable space. A Ω'-valued stochastic process adapted to $\mathcal{F}$ is called *Markov*, if for each $A \in \mathcal{F}'$, and $s < t$

$$\mathbb{P}(X_t \in A|\mathcal{F}_s) = \mathbb{P}(X_t \in A|X_s).$$

A subtly stronger property, which extends the above property from constant points in time to random ones, is the strong Markov property.

Definition A.1.26. In the situation of def. A.1.25, X is said to have the *strong Markov property*, if for each stopping time τ, conditionally on $\{\tau < \infty\}$, for all $t > 0$,

$$\mathbb{P}\left(X_{\tau+t} \in A | \mathcal{F}_\tau\right) = \mathbb{P}\left(X_{\tau+t} \in A | X_\tau\right).$$

Remark. Since every constant is also a stopping time, the strong Markov property implies the Markov property.

Definition A.1.27. For a random variable X, the *characteristic function* ϕ_X is defined as

$$\phi_X(\lambda) = \mathbb{E}\left[e^{i\lambda X}\right].$$

Definition A.1.28. For a random variable X, the *moment generating function* ψ_X is defined as

$$\psi_X(\lambda) = \mathbb{E}\left[e^{\lambda X}\right].$$

Both uniquely determine the function. Even more, we will see, that convergence of either of them implies convergence of the random variables.

The moment generating and characteristic function are the values of the *Laplace transform* along the real and imaginary axes.

Definition A.1.29. The Laplace transform of a real valued function f is defined for $s \in \mathbb{C}$ as

$$\mathcal{L}[f](s) = \int_0^\infty f(x) e^{-sx} \mathrm{d}x.$$

Note. To see that this is indeed related to the moment generating and characteristic functions, assume the random variable X has a density f.

Theorem A.1.30. *Both, characteristic and moment generating function uniquely determine the distribution of X.*

Proof. See Ash & Doléans-Dade [6, cor. 7.1.4] for the characteristic function and Billingsley [19, p. 147] for the moment generating function. □

Proposition A.1.31. *Let $X_1, \ldots, X_n$ be indepentent random variables, then*

$$\phi_{\sum_{i=1}^n X_i} = \prod_{i=1}^n \phi_{X_i}.$$

The same is true for the moment generating function.

Proof. Straightforward computation yields

$$\mathbb{E}\left[e^{\eta\lambda\sum_{i=1}^{n}X_i}\right] = \mathbb{E}\left[\prod_{i=1}^{n} e^{\eta\lambda X_i}\right] \overset{\text{indep.}}{=} \prod_{i=1}^{n}\mathbb{E}\left[e^{\eta\lambda X_i}\right].$$

Setting $\eta = i$, yields the result for characteristic functions $\eta = 1$, the result for moment generating functions. □

Remark. As it can be seen from this proof, we can essentially just switch between moment generating and characteristic function – by exchanging adding a factor of i in the exponent. However, when doing so, we must make sure that the moment generating function exists, which fails in many cases. So, it is safer to work with the characteristic function.

We now define, what we mean by positive definite functions and state a theorem that establishes a one-to-one correspondence between these functions and characteristic functions of real valued random variables. This will enable us to define the characteristic exponent.

Definition A.1.32. A function on $\mathbb{R}$ is called *positive definite*, if $f(0) = 0$ and $f(x) > 0$ for all $x \neq 0$.

Theorem A.1.33 (Bochner)**.** *Every characteristic function of a real valued random variable is positive definite. Conversely, for every positive definite, normalized (i.e. $\int_{\mathbb{R}} f = 1$) function, there exists a random variable, which has f as its characteristic function.*

Proof. See Loève [60, pp. 220–222]. □

Now that we know that $\phi_X(\lambda)$ is positive away from 0, we can define the characteristic exponent, which uses the logarithm of $\phi_X(\lambda)$.

Definition A.1.34. The *characteristic exponent* of a random variable is defined as $\Phi_X(\lambda) := -\log\left(\phi_X(\lambda)\right)$.

Similarly, the *Laplace exponent* is defined as the logarithm of the moment generating function $\Psi_X(\lambda) = -\log\left(\psi_X(\lambda)\right)$.

We now introduce some modes of convergence for random variables we will use. For all of the following definitions, we let $(X_n)_{n\in N}$ be a sequence of random variables and X a random variable on the same space.

Definition A.1.35. We say that X_n converges to X *almost surely*, if

$$\mathbb{P}\left(\lim_{x\to\infty} X_n = X\right) = 1.$$

Definition A.1.36. We say that X_n converges to X *in probability*, if

$$\lim_{n\to\infty} \mathbb{P}\left(|X_n - X| \geq \varepsilon\right).$$

Definition A.1.37. We say that X_n converges to X *in distribution*, if for all $A \in \mathcal{A}$,

$$\lim_{n\to\infty} \mathbb{P}_{X_n}(A) = \mathbb{P}_X(A).$$

There is a connection between convergence in distribution and convergence of the characteristic and moment generating functions.

Theorem A.1.38 (Continuity theorem)**.** *Let $(X_n)_{n\in\mathbb{N}}$ be a sequence of random variables.*

- *$\phi_{X_n} \to \phi_X$ pointwise $\implies$ $X_n \to X$ weakly in distribution.*
- *$\psi_{X_n} \to \psi_X$ pointwise $\implies$ $X_n \to X$ weakly in distribution.*

Proof. For (i), see Billingsley [19, thm. 26.3] – for (ii), see Billingsley [18, ex. 5.5]. □

A.2 Lévy processes, subordinators and their inverses

Lévy processes are a very important class of processes, closely related to scaling limits of sums of random variables. We have this structure in both, the spatial process X and the clock process S. Thus, it is not surprising that Lévy processes indeed play an important role in the limit.

We start with some definitions.

Definition A.2.1. A stochastic process X is called a *Lévy process*, if it

i) is càdlàg,

ii) has independent increments: For any $0 = t_0 < t_1 \ldots < t_k \leq \infty$, the family $(X_{t_i} - X_{t_{i-1}})_{i\in I}$ is independent,

iii) has stationary increments: for $s > 0$ the law of $X_{t+s} - X_t$ is independent of t.

Due to the restrictions posed, the fixed-time distributions of Lévy processes belong to the class of infinitely divisible distributions (see Sato [97, sec. 2.7]).

Definition A.2.2. A probability measure μ on $\mathbb{R}^d$ is called *infinitely divisible*, if for each $n \in \mathbb{N}$ there exists a probability measure μ^n on $\mathbb{R}^d$ such that for $(X_i)_{i=1}^n$ which are *i.i.d.* distributed with law μ_n, $\sum_{i=1}^n X_i$ is distributed with law μ.

Definition A.2.3. The Lévy process starting at 0 with continuous paths and $X_1 - X_0 \sim \mathcal{N}(0,1)$ is called *Brownian motion* (BM).

In the following, we mainly focus on the subclass of subordinators.

Definition A.2.4. A Lévy process with almost surely increasing paths is called a *subordinator*.

This means, that the distribution of the increments of this process must be supported on $\mathbb{R}^+$ only. The following theorem gives a complete characterization of infinitely divisible distributions with support on $\mathbb{R}^+$. A similar theorem for $\mathbb{R}^k$-valued distributions can be found in Reed & Simon [91, thm. XIII.55].

Theorem A.2.5 (Lévy-Kintchine for $\mathbb{R}^+$)**.** *A $\mathbb{R}^+$-valued random variable X is infinitely divisible, iff its Laplace exponent has the form*

$$\Psi_X(\lambda) = c\lambda - \int_0^\infty \left(1 - e^{-\lambda x}\right) \mu(\mathrm{d}x), \tag{A.1}$$

where $c \geq 0$ and μ is a measure on $(0,\infty)$ fulfilling

$$\int_0^\infty x \wedge 1 \mu(\mathrm{d}x) \leq \infty. \tag{A.2}$$

The measure μ is called the Lévy measure of X.

Proof. See Bertoin [17, thm. 1.2]. □

Remark. Here, we did not use the characteristic function for the characterization, but the moment generating function (see defs. A.1.28, A.1.27). This is possible, because for distributions supported only on $\mathbb{R}^+$ it always exists.

We have shown in prop. A.1.31, that $S(t) := \sum_{i=1}^t X_i$ for *i.i.d.* X_i has the so-called semigroup property $\phi_{S_{t+s}}(\lambda) = \phi_{S_t}(\lambda)\phi_{S_s}(\lambda)$. By analogy, due to the independent, stationary increments, the characteristic resp. moment generating functions of Lévy processes at time t are the same as the ones of their increments over one unit of time multiplied by e^t. Therefore, it makes sense to speak of the characteristic exponent of the process, when we refer to $\Phi(\lambda) = -t^{-1}\log\left(\phi_{X(t)}\right)$ resp. the Laplace exponent $\Psi(\lambda) = -t^{-1}\log\left(\psi_{X(t)}\right)$.

We now introduce the class of stable random variables.

Definition A.2.6. A random variable X is called *stable*, if there exists a number $n \geq 2$ *i.i.d.* random variables $(Y_i)_{i=1}^n$ and constants $a_n \in \mathbb{R}$ and $b_n \in \mathbb{R}^+$, such that

$$\sum_{i=1}^n Y_i \sim a_n + b_n X.$$

Remark. Stable distributions are a subclass of infinitely divisible distributions.

We have a characterization in terms of characteristic functions for these distributions.

Proposition A.2.7. *A random variable X is stable, iff*

$$\Phi_X(\lambda) = \begin{cases} \sigma^\alpha |\lambda|^\alpha \left(1 - i\beta \operatorname{sgn}(\lambda) \tan\left(\frac{\pi\alpha}{2}\right)\right) + ic\lambda & \text{, for } \alpha \in (0,2] \setminus \{1\}, \\ \sigma^\alpha |\lambda|^\alpha \left(1 - i\beta \frac{2}{\pi} \operatorname{sgn}(\lambda) \log(|\lambda|)\right) + ic\lambda & \text{, for } \alpha = 1, \end{cases} \tag{A.3}$$

where $\beta \in [-1,1]$, $\sigma^2 \in \mathbb{R}^+$, $c \in \mathbb{R}$ are uniquely determined constants.

Proof. See Samoradnitsky & Taqqu [96]. □

Remark. A distribution with characteristic exponent of the form (A.3) called α-stable.

We now define the inverse of a subordinator.

Definition A.2.8. The *inverse (or hitting time process)* of a subordinator V is defined as

$$V^{-1}(s) := \inf\{t > 0 \,:\, V(t) > s\}.$$

Obviously, this process is non-decreasing. However, we will later see that this process itself is no longer a Lévy process (cf. prop. A.2.12). If a subordinator is also stable with exponent α we write V_α.

We now compute the Laplace exponents and investigate the properties of three processes we will need later – Brownian motion, α-stable subordinators and their inverses. Note that only the first two of them are actually Lévy processes, the inverse of a stable subordinator itself is not Lévy as we will see.

Proposition A.2.9 (Properties of Brownian motion)**.** *For the Brownian motion in d dimensions BM_d, the following holds*

i) The Laplace exponent of d-dimensional Brownian motion is $\frac{1}{2}\lambda^2$.

ii) BM_d is γ-Hölder continuous for $\gamma < 1/2$.

iii) BM_d is self-similar: $\mathrm{BM}_d(t) = c^{-1/2}\mathrm{BM}_d(ct)$ in law.

Proof. For i), note all moments of X_t exist, therefore so does the moment generating function, which can easily be computed. In fact we can compute the Laplace transform. Let $X \sim \mathcal{N}(0,1)$, then

$$\mathbb{E}\left[e^{-\lambda X}\right] = \int_{-\infty}^{\infty} e^{-\lambda x} \frac{1}{\sqrt{2\pi}} e^{-x^2/2} \mathrm{d}x.$$

We can complete the square by adding $\frac{1}{2}\lambda$ in the exponential. We end up with

$$\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} e^{-\frac{1}{\sqrt{2}}(x+\lambda)^2} e^{\frac{1}{2}\lambda}\mathrm{d}x.$$

The last term does not depend on x and can be taken out of the integral, then we substitute $x \rightsquigarrow x-\lambda$ – the limits don't change and the integral is $\sqrt{2\pi}$, which cancels with the prefactor, leading to

$$\mathbb{E}\left[e^{-\lambda X}\right] = e^{-\frac{1}{2}\lambda^2}. \tag{A.4}$$

For ii), see Karatzas & Shreve [48, thm. 2.8], for iii), lem. 9.4, (i) in the same book. □

Proposition A.2.10 (Properties of V_α)**.** *For the subordinator V_α, the following holds:*

i) its Laplace exponent is $k\lambda^\alpha$.

ii) V_α is not continuous.

iii) V_α is self-similar: $V_\alpha(t) = r^{-1/\alpha}V_\alpha(rt)$ in law.

Proof. For i), we work with the Laplace transform. Therefore can use the results of both prop. A.2.7 and thm. A.2.5, to conclude that the Laplace exponent has the form

$$\Psi(\lambda) = k\lambda^\alpha. \tag{A.5}$$

ii) It follows from the Lévy-Itô-decomposition (see *e.g.* Sato [97, thms. 19.2, 19.3 and discussion]) that Brownian motion is the only continuous Lévy process.

For iii), we notice that the Laplace exponent of $r^{-1/\alpha}V_\alpha(rt)$ is $k\lambda^\alpha$. Since it uniquely determines the distribution, this suffices. □

Remark. The Laplace exponent in prop. A.2.10 i) corresponds to the Lévy measure $\nu(\mathrm{d}x) = \dfrac{k\alpha}{\Gamma(1-\alpha)}x^{-1-\alpha}\mathrm{d}x$. We will use the symbol V_α to refer to the process with this Laplace exponent and $k=1$.

As we already mentioned, inverse subordinators are not Lévy processes, therefore, it is not a good idea to consider their Laplace exponent. As we will see, their Laplace transforms are not of exponential type. We will actually need to define a class of functions first before we can describe it.

Definition A.2.11. The *Mittag-Leffler function* $E_\alpha(x)$ is defined as

$$E_\alpha(x) = \sum_{k=0}^{\infty}\frac{x^k}{\Gamma(\alpha k+1)}.$$

Remark. Note that for $\alpha = 1$ this is just the exponential function. Other notable cases include

- $\alpha = 1/2$, whence $E_{1/2}(x) = e^{x^2}(1 - \operatorname{erf}(x))$ and $\operatorname{erf}(x)$ is the Gaussian error function (recall that $\operatorname{erf}(x) := 2\pi^{-1/2} \int_0^x e^{-\xi^2} \mathrm{d}\xi$),
- $\alpha = 2$, whence $E_2(x) = \cosh(\sqrt{x})$.

Proposition A.2.12. *The inverse $V_\alpha^{-1}(\cdot)$ of an α-stable subordinator has the following properties*

i) It does not have stationary increments, therefore it is also not a Lévy process.

ii) Its Laplace transform is given in terms of Mittag-Leffler functions

$$\mathbb{E}\left[-\lambda V_\alpha^{-1}(t)\right] = E_\alpha\left(\lambda t^\alpha\right).$$

iii) V_α^{-1} is a.s. Hölder continuous for all $\gamma < \alpha$.

iv) V_α^{-1} is self-similar: $V_\alpha^{-1}(t) = r^{-\alpha} V_\alpha^{-1}(rt)$.

Proof. For i), see Meerschaert & Scheffler [70, cor. 3.2]. The second statement, is shown in Bingham [20, prop. 1(a)]. For iii), see Bertoin [16, thm. III.17]. For iv), see Meerschaert & Scheffler [70, prop. 3.1]. □

Having studied the constituent processes, we recall the definition of the fractional kinetics process.

Note. (def. 4.2.1) The *fractional kinetics* process with index α in dimension d, $\mathrm{FK}_{d,\alpha}$ is defined as

$$\mathrm{FK}_{d,\alpha}(t) := \mathrm{BM}_d\left(V_\alpha^{-1}(t)\right),$$

where $\mathrm{BM}_d(t)$ is a d-dimensional Brownian motion and $V_\alpha(t)$ an independent α-stable subordinator.

We will explore the process and its properties later in chapter 6.2.

A.3 Potential theory of the simple random walk

In this section, we will derive some needed results on the potential theory of the simple random walk on $\mathbb{Z}^d$. For most of this, we will follow the book of Lawler [56].

Definition A.3.1. Let $(Z_i)_{i\in\mathbb{N}}$ be a family of independent, uniformly distributed variables on $\left\{x \in \mathbb{Z}^d \,:\, |x| = 1\right\}$. We define the *simple random walk (SRW) on $\mathbb{Z}^d$ started in x*, as

$$X^{(x)}(t) = x + \sum_{i=1}^{t} Z_i.$$

We will proceed along the lines of Lawler [56]. First, we introduce some notation.

Definition A.3.2. We say that x and y have the same *parity* and write $x \leftrightarrow y$ if $\frac{\|x\|_1 + \|y\|_1}{2} \in \mathbb{Z}$.

Definition A.3.3. We define the *n-step kernel of the SRW*, $p_n(x,y) := \mathbb{P}\left(X^{(x)}(n) = y\right)$.

Definition A.3.4. The *kernel of Brownian motion at time n* is defined as

$$\bar{p}_n(x) = 2\left(\frac{d}{2\pi n}\right)^{d/2} e^{-\frac{d|x|^2}{2n}}.$$

Definition A.3.5. The error for the approximation of the n-step kernel of the simple random walk with the kernel of a diffusion at time n is

$$E(n,x) = \left(p(n,x) - \bar{p}(n,x)\right)\mathbf{1}_{\{x\leftrightarrow n\}}.$$

Definition A.3.6. The discrete difference ∇_y and second difference ∇_y^2 are given for $f : \mathbb{R}^d \to \mathbb{R}$ by

$$\begin{aligned}\nabla_y f(x) &= f(x+y) - f(x),\\ \nabla_y^2 f(x) &= f(x+y) + f(x-y) - 2f(x).\end{aligned}$$

Theorem A.3.7 (Local central limit theorm)**.** *For $E(n,x)$ defined as above the following holds:*

$$\begin{aligned}|E(n,x)| &= \mathcal{O}\left(n^{-(d+2)/2}\right),\\ |E(n,x)| &= \mathcal{O}\left(|x|^{-2}n^{-(d)/2}\right).\end{aligned}$$

Proof. See Lawler [56, thm. 1.2.1]. □

Proposition A.3.8 (Strong Markov Property)**.** *The SRW has the strong Markov property.*

Proof. See Lawler [56, thm. 1.3.2]. □

Now we can get started on the potential theory of the simple random walk. The central object of the study will be the discrete Laplacian, $\Delta := 1/(2d) \sum_{j=1}^{d} \nabla_{e_j}^2 f(x)$, where $e_1, \dots, e_d$ are the d-dimensional unit vectors. This operator can be related to the simple random walk via

$$\Delta f(x) = \mathbb{E}\left[f(X_1^{(x)}) - f(x)\right].$$

Definition A.3.9. A function is called (sub-, super-) harmonic on A if for all $x \in A$, $\Delta f(x) = 0$ (resp. $\geq 0, \leq 0$).

Harmonic functions have a close relationship with martingales.

Proposition A.3.10. *Let f be a harmonic function on A, $\tau := \mathsf{Hit}_x^{\partial A}$ the exit time for the SRW started in x, then $f(X_{\max\{\tau,n\}})$ is a martingale w.r.t the natural filtration of the SRW.*

Proof. See Lawler [56, prop. 1.4.1]. □

An easy way to validate the almost sure finiteness of a exit times is to check boundedness of the set in question as the following lemma shows.

Lemma A.3.11. *For any finite $A \subset \mathbb{Z}^d$ and for all $x \in A$,*

$$\mathbb{P}\left(\mathsf{Hit}_x^{\partial A} \geq n\right) \leq C\rho^n,$$

for some $C > 0$, $0 < \rho < 1$.

Proof. See Lawler [56, lem. 1.4.4]. □

For the proof of the main result of the thesis, we mainly need estimates for the Green's function.

Definition A.3.12. The n-step Green's function is defined as

$$G_n(x, y) = \mathbb{E}\left[\sum_{j=0}^{n} \mathbf{1}_{\left\{X_j^{(x)}\right\}} = y\right] = \sum_{j=0}^{n} p_j(x, y). \tag{A.6}$$

For $d \geq 3$ the limit $n \to \infty$ of this expression exists finitely (Lawler [56, p. 29]) and we write

$$G(x,y) := \lim_{n\to\infty} G_n(x,y). \tag{A.7}$$

The Green's function is essentially translation invariant, in the sense that if we shift the starting point and the endpoint by the same amount, the value of the Green's function doesn't change. Therefore we introduce the notation $G(x) := G(0,x)$, whence $G(y,z) = G(y-z)$. As in classical PDE theory, the Green's function solves a discrete analogue of Poisson's equation (here solved on the full space $\mathbb{Z}^d$ with δ RHS)

$$\Delta G(x) = \mathbb{E}\left[\sum_{j=1}^{\infty} \mathbf{1}_{\{X_j=x\}}\right] - \mathbb{E}\left[\sum_{j=0}^{\infty} \mathbf{1}_{\{X_j=x\}}\right] = \mathbb{E}\left[-\mathbf{1}_{\{X_0=x\}}\right] = -\delta_0(x).$$

As already suggested by this equation, many results from elliptic PDE theory carry over to the discrete Laplacian, like the maximum principle and, to some degree, existence and uniqueness of solutions. For a discussion of this, see Lawler [56, sec. 1.4-1.7].

Now we will state the first of a series of estimates on the decay of the Green's function.

Theorem A.3.13 (Decay of the full space Green's function)**.** *For $d \geq 3$,*

$$G(x) \sim a_d |x|^{2-d}, \quad \text{as } |x| \to \infty,$$

where

$$a_d = \frac{d}{2}\Gamma\left(\frac{d}{2}-1\right)\pi^{-d/2} = \frac{2}{(d-2)\omega_d}, \tag{A.8}$$

ω_d being the volume of the unit ball in $\mathbb{R}^d$. More precisely, for any $\alpha < d$,

$$\lim_{|x|\to\infty} |x|^{\alpha}\left(G(x) - a_d|x|^{2-d}\right) = 0.$$

Proof. See Lawler [56, thm. 1.5.4]. □

What we will mainly need for the proof of the principal theorem, are estimates for the Green's function of a random walk, killed on exiting some set A.

Definition A.3.14. For $A \subset \mathbb{Z}^d$, the *Green's function on A*, $G_A(x,y)$ is the expected number of visits to y before leaving A of a SRW started at x:

$$G_A(x,y) := \mathbb{E}\left[\sum_{j=1}^{\mathsf{Hit}_x^{\partial A}} \mathbf{1}_{\left\{X_j^{(x)}=y\right\}}\right] = \sum_{j=1}^{\infty} \mathbb{P}\left(X_j^{(x)} = y, \tau > j\right). \tag{A.9}$$

Another important concept is the harmonic measure, which basically gives the distribution of the different exit points for a set A.

Definition A.3.15. We define the *harmonic measure* for $y \in \partial A$ as

$$H_{\partial A}(x, y) := \mathbb{P}\left(X^{(x)}_{\mathsf{Hit}^{\partial A}_x} = y\right),$$

the measure induced by the hitting distribution of the SRW started at x.

The harmonic measure can be used to relate the Green's function killed on exit of a set to the full space Green's function.

Proposition A.3.16. *Let $A \subset \mathbb{Z}^d$, $d \geq 3$ be finite, $x, z \in A$. Then,*

$$G_A(x, z) = G(z - x) - \sum_{x \in \partial A} H_{\partial A}(x, y) G(z - y).$$

Proof. The proof is a straightforward application of definitions:

$$\begin{aligned} G_A(x, z) &= \mathbb{E}\left[\sum_{j=0}^{\mathsf{Hit}^{\partial A}_x - 1} \mathbf{1}_{\left\{X^{(x)}_j = z\right\}}\right] = \mathbb{E}\left[\sum_{j=0}^{\infty} \mathbf{1}_{\left\{X^{(x)}_j = z\right\}} - \sum_{\mathsf{Hit}^{\partial A}_x}^{\infty} \mathbf{1}_{\left\{X^{(x)}_j = z\right\}}\right] \\ &= G(0, z - x) - \sum_{y \in \partial A} H_{\partial A}(x, y) G(0, z - y). \end{aligned}$$

□

Finally, we can state the estimates for the Green's functions of the simple random walk killed on the exit of a ball around its starting point.

Proposition A.3.17 (Killed GF and return probability)**.** *Let $x \in B_0(r), \tau := \mathsf{Hit}^{\partial B_0(r) \cup \{0\}}_x$, then*

$$\mathbb{P}\left(X^{(x)}_\tau = 0\right) = \frac{a_d}{G(0)}\left(|x|^{2-d} - r^{2-d}\right) + \mathcal{O}\left(|x|^{1-d}\right). \tag{A.10}$$

Furthermore, for the Green's function,

$$G_{B_0(r)}(0, 0) = G(0) + \mathcal{O}\left(r^{2-d}\right), \text{ and} \tag{A.11}$$

$$G_{B_0(r)}(x, 0) = a_d\left(|x|^{2-d} - r^{2-d}\right) + \mathcal{O}\left(|x|^{1-d}\right). \tag{A.12}$$

Proof. Since $G(x)$ is harmonic on $\mathbb{Z}^d \setminus \{0\}$, $M^{(x)}_j := G\left(X^{(x)}_{\max\{j,\tau\}}\right)$ is a martingale by prop. A.3.10. By the optional sampling theorem, $G(x) = \mathbb{E}\left[M^{(x)}_0\right] = \mathbb{E}\left[M^{(x)}_\tau\right]$. This expectation can be computed

$$\mathbb{E}\left[M^{(x)}_\tau\right] = G(0)\mathbb{P}\left(X^{(x)}_\tau = 0\right) + \mathbb{E}\left[G(X^{(x)}_\tau)\middle| X^{(x)}_\tau \in \partial B_0(r)\right]\mathbb{P}\left(X^{(x)}_\tau \in \partial B_0(r)\right). \tag{A.13}$$

For all $y \in \partial B_0(r)$, $|y| \in [r, r+1]$, hence by thm. A.3.13,

$$G(y) = a_d r^{2-d} + o\left(r^{1-d}\right).$$

Since this is true for all terms of the conditional expectation, we also have

$$\mathbb{E}\left[G(X_\tau^{(x)}) \middle| X_\tau^{(x)} \in \partial B_0(r)\right] = a_d r^{2-d} + \mathcal{O}\left(r^{1-d}\right).$$

We plug this into eq. (A.13) and rearrange to get the first result. For the second we note that

$$G_{B_0(r)}(x,0) = \mathbb{P}\left(X_\tau^{(x)} = 0\right) G_{B_0(r)}(0,0). \tag{A.14}$$

Indeed, if a trajectory hits the boundary before returning to 0, this trajectory will not contribute to $G(x,0)$. Therefore, what we see here is actually a decomposition on those trajectories who hit the boundary first, and those who go to 0 at least once before leaving, just that all trajectories of the second summand don't contribute at all. For those who hit 0 before the boundary, the successive visits have the same distribution as the visits of trajectories already starting at 0.

Furthermore,

$$\begin{aligned} G_{B_0(r)}(0,0) &= \mathbb{E}\left[\sum_{i=0}^{\mathsf{Hit}_0^{\partial B_0(r)}} \mathbf{1}_{\left\{X_i^{(0)}=0\right\}}\right] \\ &= \mathbb{E}\left[\sum_{i=0}^{\infty} \mathbf{1}_{\left\{X_i^{(0)}=0\right\}}\right] - \mathbb{E}\left[\sum_{i=\mathsf{Hit}_0^{\partial B_0(r)}}^{\infty} \mathbf{1}_{\left\{X_i^{(0)}=0\right\}}\right]. \end{aligned}$$

In the last expectation we only sum hits to 0 after hitting the boundary, therefore the term is the same as $\sum_{x\in\partial B_0(r)} G(x,0)$. Using thm. A.3.13, we see that this term is $\mathcal{O}\left(r^{2-d}\right)$, thus

$$G_{B_0(r)}(0,0) = G(0) + \mathcal{O}\left(r^{2-d}\right). \tag{A.15}$$

This together with eq. (A.14) and the first part of the proposition yields

$$G_{B_0(r)}(x,0) = \left(\frac{a_d}{G(0)}\left(|x|^{2-d} - r^{2-d}\right) + \mathcal{O}\left(|x|^{1-d}\right)\right)\left(G(0) + \mathcal{O}\left(r^{2-d}\right)\right),$$

and thus the second part is proven too. □

From this proposition, we deduce now the probability for a simple random walk to hit a certain site x before exiting a ball around the origin. In contrast to the previous, the following is not taken from Lawler [56] but from Ben Arous & Černý [15].

Lemma A.3.18 (Probability to hit before exiting a ball)**.** *Denote by $P_r(0,x)$ the probability that $X^{(0)}$ hits x before exiting $B_0(r)$. Then*

$$P_r(0,x) \geq \frac{a_d}{G(0)} \left(|x|^{2-d} - r^{2-d}\right) + \mathcal{O}\left(|x|^{1-d}\right), \tag{A.16}$$

$$P_r(0,x) \leq a_d \left(|x|^{2-d} - r^{2-d}\right) + \mathcal{O}\left(|x|^{1-d}\right). \tag{A.17}$$

Another upper bound is given by

$$P_r(0,x) \leq \left(\frac{a_d}{G(0)} \left(|x|^{2-d} - r^{2-d}\right) + \mathcal{O}\left(|x|^{1-d}\right)\right) \left(1 + \mathcal{O}\left((r-|x|)^{2-d}\right)\right). \tag{A.18}$$

Proof. The same way as we discussed in the proof of prop. A.3.17, we can decompose $G_{B_0(r)}(0,x)$ into a part with zero contribution, and the part where x is hit before exiting $B_0(r)$, whence the contribution is the same as if the walk had started at x

$$G_{B_0(r)}(0,x) = P_r(0,x) G_{B_0(r)}(x,x). \tag{A.19}$$

Since furthermore $1 \leq G_{B_0(r)}(x,x) \leq G(0,0)$ - the first inequality is because the walk starts at x, the second one because $G(0,0) = G(x,x) \geq G_A(x,x)$ for any $A \subset \mathbb{Z}^d$ - the first two bounds on $P_r(0,x)$ immediately follow from eq. (A.19) and the second part of prop. A.3.17.

For the second upper bound, we again use eq. (A.19) and

$$G_{B_0(r)}(x,x)^{-1} \leq G_{B_x(r-|x|)}(x,x)^{-1} = G_{B_0(r-|x|)}(0,0)^{-1} = G(0)^{-1} + \mathcal{O}\left((r-|x|)^{2-d}\right),$$

where we again used eq. (A.15). □

We need two more statements about the SRW for our main theorem, which we state now.

Proposition A.3.19 (Harmonic measure on spheres)**.** *For the harmonic measure on a sphere, it holds that*

$$H_{\partial B_0(r)}(0,y) \asymp r^{1-d}.$$

Proof. See Lawler [56, Lem 1.7.4]. □

Lemma A.3.20 (Exit time for SRW)**.** *The exit time of SRW for a ball with radius r behaves as r^2 for large r.*

Proof. See Ciesielski & Taylor [31, p. 445] for the proof of the same statement for Brownian motion. Together with the local limit theorem A.3.7, the result follows for large r. Sketch: □

A.4 Limit theorems for sums of random variables

In this section, we discuss limit theorems for sums of independent variables. The theory here is not strong enough to treat the charge transport process itself, but the results are needed in intermediate steps. Particularly, in addition to the usual limit theorems for sequences of random variables we treat triangular arrays. We immediately state the results for triangular arrays, since they contain their counterparts for sequences as a special case. A good introduction to generalized limit theorems and techniques can be found in Bovier [25].

Definition A.4.1. A *triangular array* is a sequence of finite sequences, where each row is only as long as its own index, we write $\left(a_i^{(n)}\right)_{n\in\mathbb{N}} = \left(\left(a_i^{(n)}\right)_{i=1,\dots,n}\right)_{n\in\mathbb{N}}$.

We start with the strong law of large numbers. For that we need the notion of stochastic domination.

Definition A.4.2. We say say that a family $(X_i)_{i\in I}$ of random variables is *stochastically dominated* by a random variable X, if

$$\mathbb{P}\left(|X_i| \geq x\right) \leq \mathbb{P}\left(|X| \geq x\right) \text{ for all } i \in I.$$

Theorem A.4.3 (Strong LLN for triangular arrays)**.** *Let* $\left(\left(X_i^{(n)}\right)_{i=1,\dots,n}\right)_{n\in\mathbb{N}}$ *be a triangular array of rowwise independent random variables such that* $\mathbb{E}\left[X_i^{(n)}\right] = 0$ *for all* i, n*. Furthermore, suppose that the* $X_i^{(n)}$ *are uniformly dominated by some random variable* $X \in L^{2p}$*, for some* $p \in [1, 2]$*. Then*

$$n^{-1/p}\sum_{i=1}^{n} X_i^{(n)} \to 0 \text{ a.s. as } n \to \infty.$$

Proof. This is the main result of Hu *et al.* [46]. In fact they prove an even stronger result, namely complete convergence, which via Borel-Cantelli implies almost sure convergence. □

We now prove a variant of the central limit theorem – Donsker's Theorem (also-called a functional central limit theorem). Because we will use a very similar technique for the spatial evolution of the charge transfer, we will cover its proof in more detail.

Donsker's theorem is a limit theorem for the measure on the space of trajectories of processes – the theory of these spaces is covered in sec. A.5. Since the measures do

not neccessary live in finite dimensional spaces, boundedness of a sequence will not suffice to extract a convergent subsequence. To show compactness in these spaces of measures, we introduce the concept of tightness and relate it to compactness.

Definition A.4.4. A collection of probability measures $(\mu_i)_{i\in I}$ on a topological space is called *(uniformly) tight* if for any $\varepsilon > 0$, there exists a compact set K_ε such that for all $i \in I$ $\mu_i(K_\varepsilon) > 1 - \varepsilon$.

Theorem A.4.5 (Prohorov)**.** *A family of probability measures on a complete, separable space is compact, iff it is tight.*

Proof. See Billingsley [18, thm. 5.1, 5.2] □

Lemma A.4.6 (FDD and tight implies weak)**.** *If a sequence $(X_n)_{n\in\mathbb{N}}$ of RVs is tight and the finite dimensional distributions converge to some limit X, it converges weakly to X.*

Proof. See Karatzas & Shreve [48, thm. 4.15]. □

Remark. A single measure on a complete, separable space is always tight (see. Billingsley [18, thm. 1.3]). All spaces we will consider are polish, *i.e.* completely metrizable, spaces. For definitions and properties of (infinite dimensional) spaces (separability, completeness, metrizability, etc.) the reader is referred to Yoshida [120]. They are important for the underlying structure, but not needed explicitly once we know that all spaces considered in this work have these properties.

We now show a result, which is sufficiently general to be used for the coarse grained spatial process later, as well as to prove a functional central limit theorem. The estimate is used to show uniform tightness of the process.

Proposition A.4.7 (Tightness for L^2 zero mean RV)**.** *For $S(t) = \sum_{i=1}^{t} \xi_i$, where $(\xi_i)_{i\in\mathbb{N}}$ is an i.i.d. sequence with mean zero and variance $\sigma < \infty$, we have that for any $T > 0$,*

$$\lim_{\delta\downarrow 0} \limsup_{n\to\infty} \frac{1}{\delta}\mathbb{P}\left(\max_{\substack{1\le j\le \lfloor n\delta\rfloor+1 \\ 0\le k\le \lfloor nT\rfloor+1}} |S(j+k) - S(k)| > \varepsilon\sigma\sqrt{n} \right) = 0.$$

Proof. See Karatzas & Shreve [48, lem. 4.19]. □

Theorem A.4.8 (Donsker's invariance principle)**.** *Let ξ_i be a collection of $\mathbb{R}^k$-valued mean zero random variables with finite variance $\sigma > 0$. We define*

$$S(t) := \sum_{i=1}^{\lfloor t\rfloor} \xi_i, \quad X(t) := S(t) + (t - \lfloor t\rfloor)\xi_{t+1}.$$

Then, $\tilde{X}^n(t) := \dfrac{1}{\sqrt{n}\sigma} X(nt)$ converges in distribution on $C([0,\infty), \mathbb{R}^k)$ (with the uniform topology) to BM_d.

Proof. $S^n(t)$ fulfills the requirements of prop. A.4.7, so we can show that

$$\mathbb{P}\left(\max_{\substack{|s-t|\leq\delta\\0\leq s,t\leq T}}\left|\tilde{X}^n(s)-\tilde{X}^n(t)\right|>\varepsilon\right)=\mathbb{P}\left(\max_{\substack{|s-t|\leq n\delta\\0\leq s,t\leq nT}}|X(s)-X(t)|>\varepsilon\right)$$
$$\leq\mathbb{P}\left(\max_{\substack{1\leq j\leq\lfloor n\delta\rfloor+1\\0\leq k\leq\lfloor nT\rfloor+1}}|S^n(j+k)-S^n(j)|>\varepsilon\right)\to 0,$$

where we used that by the piecewise linearity of X,

$$\max_{\substack{|s-t|\leq n\delta\\0\leq s,t\leq nT}}|X(s)-X(t)|\leq\max_{\substack{|s-t|\leq\lfloor\delta n\rfloor+1\\0\leq s,t\leq\lfloor nT\rfloor+1}}|X(s)-X(t)|\leq\max_{\substack{1\leq j\leq\lfloor\delta n\rfloor+1\\0\leq j\leq\lfloor nT\rfloor+1}}|S(j+k)-S(k)|\,.$$

Hence, $\tilde{X}^n$ is tight in the uniform topology. Furthermore, one can show that the finite dimensional distributions converge to those of Brownian motion (see Karatzas & Shreve [48, thm. 4.17]). This, together with lem. A.4.6 finishes the proof. □

To conclude this section, we state a generalized triangular array version of the functional CLT for convergence to a stable subordinator. It will not be explicitly used, but it motivates our approach to treat the lognormal waiting time landscape.

Theorem A.4.9. *Let $X_i^{(n)}$ be a triangular array of random variables, i.i.d. for fixed n and with support in $\mathbb{R}^+$. Assume there exists $\alpha\in(0,1)$ and sequences c_n, a_n such that*

$$a_n\mathbb{P}\left(X_1^n>c_nx\right)\to x^{-\alpha},$$

furthermore, let them satisfy

$$\lim_{\varepsilon\downarrow 0}\limsup_{n\uparrow\infty}c_n^{-1}a_n\mathbb{E}\left[\mathbf{1}_{\left\{X_1^n\leq c_n\varepsilon\right\}}X_1^n\right]=0.$$

Then,

$$S_n(t):=c_n^{-1}\sum_{i=1}^{\lfloor ta_n\rfloor}X_i^n\to V_\alpha(t),$$

in distribution on $D(\mathbb{R}^+,J_1)$.

Proof. See Bovier [25, thm. 4.1.7]. □

A.5 Skorokhod spaces

We give the basic definitions and properties of Skorokhod spaces needed for the proof of the main theorem. A more detailed discussion can be found in Whitt [119], Billingsley [19] or the original article by Skorokhod [108].

Definition A.5.1. The *Skorokhod spaces* $D(I, \mathbb{R}^k)$ are defined as the spaces of all right-continuous $\mathbb{R}^k$-valued functions with left limits (càdlàg) defined on $I \subset \mathbb{R}^+$.

Remark. To understand why theorems about convergence of random variables and Skorokhod spaces are interesting for the treatment of stochastic processes, note that a stochastic process on $\mathbb{R}^d$ observed up to a time T is a random variable on $D([0,T], \mathbb{R}^d)$.

In the case of $d = 1$, we omit the second argument – in fact, where no confusion can arise, we frequently use second argument to denote the topology.

These spaces are important because they are sufficiently regular to still allow us to do analysis on them, while they are rich enough to contain the trajectories of all stochastic processes of interest for us.

As already pointed out, the space $D(I)$ can be equipped with various topologies, leading to different convergence and compactness properties. The finest topology we will use is the uniform topology, generated by the max-norm $\|\cdot\|_\infty$. However, while the space of continuous functions $C^0(I) \subset D(I)$ is complete with this norm, the space $D(I)$ itself is not.

To deal with this problem, Skorokhod introduced in his seminal paper [108] several other metrics which generate coarser topologies. We will only discuss one of these metrics here, namely the M_1 metric, which we will use later for our limit.

Definition A.5.2. The *completed graph* of a function $x(t)$ is defined as

$$\Gamma_x := \left\{(z,t) \in \mathbb{R}^d \times I \,:\, z \in \mathrm{Conv}\left\{x(t-), x(t)\right\}\right\},$$

where $\mathrm{Conv}(A) := \{\sum_{i=1}^\infty \alpha_i a_i \,:\, \sum_{i=1}^\infty \alpha_i = 1,\, a_i \in A\}$ is the *convex hull of the set* A.

One can define an order on Γ_x:

$$(z_1,t_1) \leq (z_2,t_2) \;:\Leftrightarrow\; (t_1 < t_2) \vee ((t_1 = t_2) \wedge (|x(t_1-) - z| \leq |x(t_2-) - z_2|))\,.$$

Definition A.5.3. A map $\Gamma_x \to [0,1]$ is called a *parametric representation* of Γ_x iff it is continuous and nondecreasing.

We let $\Pi(x)$ be the set of parametric representations for Γ_x.

Definition A.5.4. The M_1 metric is defined as

$$\mathrm{dist}_{M_1}(x_1,x_2) := \inf_{\substack{(u_1,r_1)\in\Pi(x_1)\\(u_2,r_2)\in\Pi(x_2)}} \{|u_1 - u_2| \vee |r_1 - r_2|\}\,.$$

Remark. See ref. 119, thm. 12.3.1 for the fact, that the above is indeed a metric

We can extend this to $D(I, \mathbb{R}^d)$ in two ways, either by replacing $|\cdot|$ by $\|\cdot\|_\infty$, leading to the strong M_1 topology. The other possibility is to identify $D(I, \mathbb{R}^k) = D(I, \mathbb{R})^k$ and use the product topology induced by the metric

$$\operatorname{dist}_{\text{Prod}}(x, y) := \sum_{i=1}^{d} \operatorname{dist}\left(x^{(i)}, y^{(i)}\right).$$

Since convergence in this topology is implied by convergence in the strong topology, we call it the weak (M_1) topology. However, in the following, will either work in one dimension when using Skorokhod spaces or d dimensional ones equipped with the uniform topology. In both cases, this distinction is not necessary (obvious for one dimension, see Whitt [119, p. 83] for the uniform topology).

To show convergence of measures on Skorokhod spaces we will, similar to treatment in Donsker's theorem (thm. A.4.8), appeal to tightness of the sequence. Since this is hard to check explicitly, we will now give an alternative characterization of tightness for the M_1 topology.

Definition A.5.5. The *standard segment* is defined as as

$$[a, b] := \{\alpha a + (1 - \alpha)b \,:\, 0 \leq \alpha \leq 1\}.$$

We introduce several *oscillation functions*

$$\bar{v}(x, t, \delta) := \sup_{(0 \vee t-\delta) \leq t_1 \leq t_2 \leq (t+\delta \wedge T)} \{\|x(t_1) - x(t_2)\|_\infty\}, \tag{A.20}$$

$$\omega_s(x, t, \delta) := \sup_{(0 \vee t-\delta) \leq t_1 \leq t_2 \leq t_3 \leq (t+\delta \wedge T)} \{\|x(t_2) - [x(t_1), x(t_3)]\|_\infty\}, \tag{A.21}$$

where $\|x - A\| := \inf_{y \in A} \|x - y\|$. Finally, we let

$$\omega_s(x, \delta) := \sup_{0 \leq t \leq T} \omega_w(x, t, \delta), \tag{A.22}$$

$$\omega_s'(x, \delta) := \max\left(\omega_w(x, \delta), \bar{v}(x, 0, \delta), \bar{v}(x, T, \delta)\right). \tag{A.23}$$

Now we have what we need to give a characterization of tightness in the M_1 topology.

Lemma A.5.6 (Characterization of tightness)**.** *A sequence μ_n of probability measures on $D([0, T])$ is tight in the strong M_1 topology iff*

i) For each $\varepsilon > 0$ there exists c such that

$$\mu_n\left(\{x \in D([0, T]) \,:\, \|x\| \geq c\}\right) \leq \varepsilon.$$

ii) For each $\varepsilon > 0$ and $\eta > 0$, there exists a $\delta > 0$ such that

$$\mu_n\left(\{x \in D([0,T]) \,:\, \omega'(x,\delta) \geq \eta\}\right) \leq \varepsilon.$$

Proof. See Whitt [119, thm. 12.12.3].

Another approach to conclude convergence of complicated processes is, to write it as a continuous function of simpler processes. This is the so-called continuous mapping approach, which relies on the following theorem.

Theorem A.5.7 (Continuous mapping)**.** *Let (S,m) and (S',m') be metric spaces, $X_n \to X$ weakly in (S,m) and $g : (S,m) \to (S',m')$ continuous, then also $g(X_n) \to g(X)$ weakly in (S',m').*

Proof. See Whitt [119, thm. 3.4.1]. □

Later we will use one particular function, namely the one mapping strictly increasing trajectories onto their inverses. To make the continuous mapping theorem applicable, we now show continuity of this mapping with respect to the topologies we need.

Proposition A.5.8 (Inverse is continuous)**.** *The inversion map from the subspace of strictly increasing functions $D_{\uparrow\uparrow}(M_1) \subset D(M_1)$ to $D(U)$ is continuous.*

Proof. This immediately follows from Whitt [119, cor. 13.6.4], which states that the inverse map from $D(M_2)$ to $D(U)$ is continuous and the fact that the M_1 topology is finer than the M_2 topology (see Skorokhod [108, eq. 2.9]). □

A.6 Powerlaws and the lognormal distribution

We start with a classical bound for the probability of a standard normal variable to exceed a certain level.

Lemma A.6.1. *For $X \sim \mathcal{N}(0,1)$, it holds that*

$$\frac{u}{1+u^2}\frac{e^{-u^2/2}}{\sqrt{2\pi}} \leq \mathbb{P}\left(X \geq u\right) \leq \frac{1}{u}\frac{e^{-u^2/2}}{\sqrt{2\pi}}.$$

Proof. We note that because for $\xi \in (x,\infty)$, $1 < \xi/x$,

$$\mathbb{P}\left(X \geq u\right) = \int_u^\infty \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}\xi^2} \mathrm{d}\xi < \int_0^\infty \frac{1}{\sqrt{2\pi}} \frac{\xi}{u} e^{-\frac{1}{2}\xi^2} \mathrm{d}\xi.$$

Substituting $\eta = \xi^2/2$, we get

$$\mathbb{P}\left(X \geq u\right) < \int_{u^2/2}^\infty \frac{1}{\sqrt{2\pi}u} e^{-\eta} \mathrm{d}\eta = -\frac{1}{\sqrt{2\pi}u}\left[e^{-\eta}\right]_{\eta=u^2/2}^{\eta=\infty} = \frac{1}{\sqrt{2\pi}}\frac{1}{u} e^{-\frac{1}{2}u^2}.$$

For the other inequality, we start with

$$\left(1+\frac{1}{u^2}\right)\mathbb{P}\left(X\geq u\right)=\int_u^\infty\left(1+\frac{1}{u^2}\right)\frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}\xi^2}\mathrm{d}\xi$$
$$>\int_u^\infty\left(1+\frac{1}{\xi^2}\right)\frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}\xi^2}\mathrm{d}\xi=-\left[\frac{e^{-\frac{1}{2}\xi^2}}{\sqrt{2\pi}\xi}\right]_{\xi=u}^{\xi=\infty}=\frac{e^{-\frac{1}{2}u^2}}{\sqrt{2\pi}u}.$$

Rearranging this leads to the desired result. □

The following lemma motivates introducing an energy scale in addition to the spatial and temporal scales.

Lemma A.6.2. *There exist sequences $\hat{\sigma}(n)$, $c(n)$ and $g(n)$, such that for $\mathbf{E}_x\sim\mathcal{N}(0,1)$, as $n\to\infty$,*

$$c(n)\mathbb{P}\left(\exp\left(-\hat{\sigma}(n)\mathbf{E}_x\right)\geq g(n)u\right)=u^{-\alpha}\left(1+L(n,u)\right),$$

where $L(n,u)\overset{u\to\infty}{\longrightarrow}0$ uniformly in n.

Proof. Since the $\mathbf{E}_x$ are normally distributed, we can use the bounds from lem. A.6.1 for the tail probabilities to get

$$\underbrace{\left(\frac{u^2}{1+u^2}\right)}_{1+o(1)}\frac{1}{u}\exp\left(-\frac{u^2}{2}\right)\leq\sqrt{2\pi}\mathbb{P}\left(\mathbf{E}_x\geq u\right)\leq\frac{1}{u}\exp\left(-\frac{u^2}{2}\right). \tag{A.24}$$

Then, we have

$$c(n)\mathbb{P}\left(\exp\left(-\hat{\sigma}(n)\mathbf{E}_x\right)\geq g(n)u\right)=c(n)\mathbb{P}\left(\mathbf{E}_x\leq\frac{-\log u-\log g(n)}{\hat{\sigma}(n)}\right)$$
$$\overset{\text{symmetry}}{=}c(n)\mathbb{P}\Big(\mathbf{E}_x\geq\frac{\log u+\log g(n)}{\hat{\sigma}(n)}\Big).$$

Using the asymptotics derived in eq. (A.24), the above probability will behave as

$$(1+o(1))\frac{1}{\sqrt{2\pi}}c(n)\exp\left(-\frac{\log^2 u}{2\hat{\sigma}(n)^2}\right)\exp\left(-\frac{\log u\log g(n)}{\hat{\sigma}(n)^2}\right)\times$$
$$\times\exp\left(-\frac{\log^2 g(n)}{2\hat{\sigma}(n)^2}\right)\frac{\hat{\sigma}(n)}{\log g(n)+\log u}.$$

We set $c(n)=\sqrt{2\pi}\exp\left(\frac{\log^2 g(n)}{2\hat{\sigma}(n)^2}\right)\frac{\log g(n)}{\hat{\sigma}(n)}$, arriving at

$$(1+o(1))\exp\left(-\frac{\log^2 u}{2\hat{\sigma}(n)^2}\right)\exp\left(-\frac{\log u\log g(n)}{\hat{\sigma}(n)^2}\right)\frac{\log g(n)}{\log g(n)+\log u}.$$

The middle term can be rewritten as

$$\exp\left(-\frac{\log u \log g(n)}{\hat{\sigma}(n)^2}\right) = \exp\left(\log u\left(-\frac{\log g(n)}{\hat{\sigma}(n)^2}\right)\right) = u^{\left(-\frac{\log g(n)}{\hat{\sigma}(n)^2}\right)}.$$

We choose $g(n) := \exp(\alpha\hat{\sigma}(n)^2)$, whence the expression simply becomes $u^{-\alpha}$. The remaining two expressions are

$$\frac{\alpha\hat{\sigma}(n)^2}{\log u + \alpha\hat{\sigma}(n)^2}\exp\left(-\frac{\log^2 u}{2\hat{\sigma}(n)^2}\right) \overset{\hat{\sigma}(n)\to\infty}{\longrightarrow} 1.$$

□

Remark. The above proof works for any sequence $\hat{\sigma}(n)$ resp. $a(n)$, which goes to infinity as $n \to \infty$.

Writing down the dependence of $c(n)$ on $\hat{\sigma}(n)$ and $g(n)$ explicitly

$$\begin{aligned}
c(n) &= \sqrt{2\pi}\exp\left(\frac{\log^2 g(n)}{2\hat{\sigma}(n)^2}\right)\frac{\log g(n)}{\hat{\sigma}(n)} \\
&= \sqrt{2\pi}\exp\left(\frac{\log^2 \exp(\alpha\hat{\sigma}(n)^2)}{2\hat{\sigma}(n)^2}\right)\frac{\log \exp(\alpha\hat{\sigma}(n)^2)}{\hat{\sigma}(n)} \\
&= \sqrt{2\pi}\exp\left(\frac{\alpha^2\hat{\sigma}(n)^4}{2\hat{\sigma}(n)^2}\right)\frac{\alpha\hat{\sigma}(n)^2}{\hat{\sigma}(n)} \\
&= \sqrt{2\pi}\alpha\hat{\sigma}(n)\exp\left(\frac{1}{2}\alpha^2\hat{\sigma}(n)^2\right),
\end{aligned}$$

and since $g(n) = \exp(\alpha\hat{\sigma}(n)^2)$,

$$c(n) = \sqrt{2\pi}\alpha\hat{\sigma}(n)g(n)^{\alpha/2}.$$

Thus, we get the result in the form we will use it throughout the rest of this work.

Corollary A.6.3. *For $g(n) := \exp(\alpha\hat{\sigma}(n)^2)$, $\mathbf{E}_x \sim \mathcal{N}(0,1)$ and any sequence $\hat{\sigma}(n)$, which goes to ∞ as $n \to \infty$, it holds that,*

$$\mathbb{P}\left(\exp\left(-\hat{\sigma}(n)\mathbf{E}_x\right) \geq g(n)u\right) = (1 + L(n,u))\frac{1}{\sqrt{2\pi}\alpha}g(n)^{-\alpha/2}\hat{\sigma}(n)^{-1}u^{-\alpha},$$

where $L(n,u) \overset{u\to\infty}{\longrightarrow} 0$ uniformly in n.

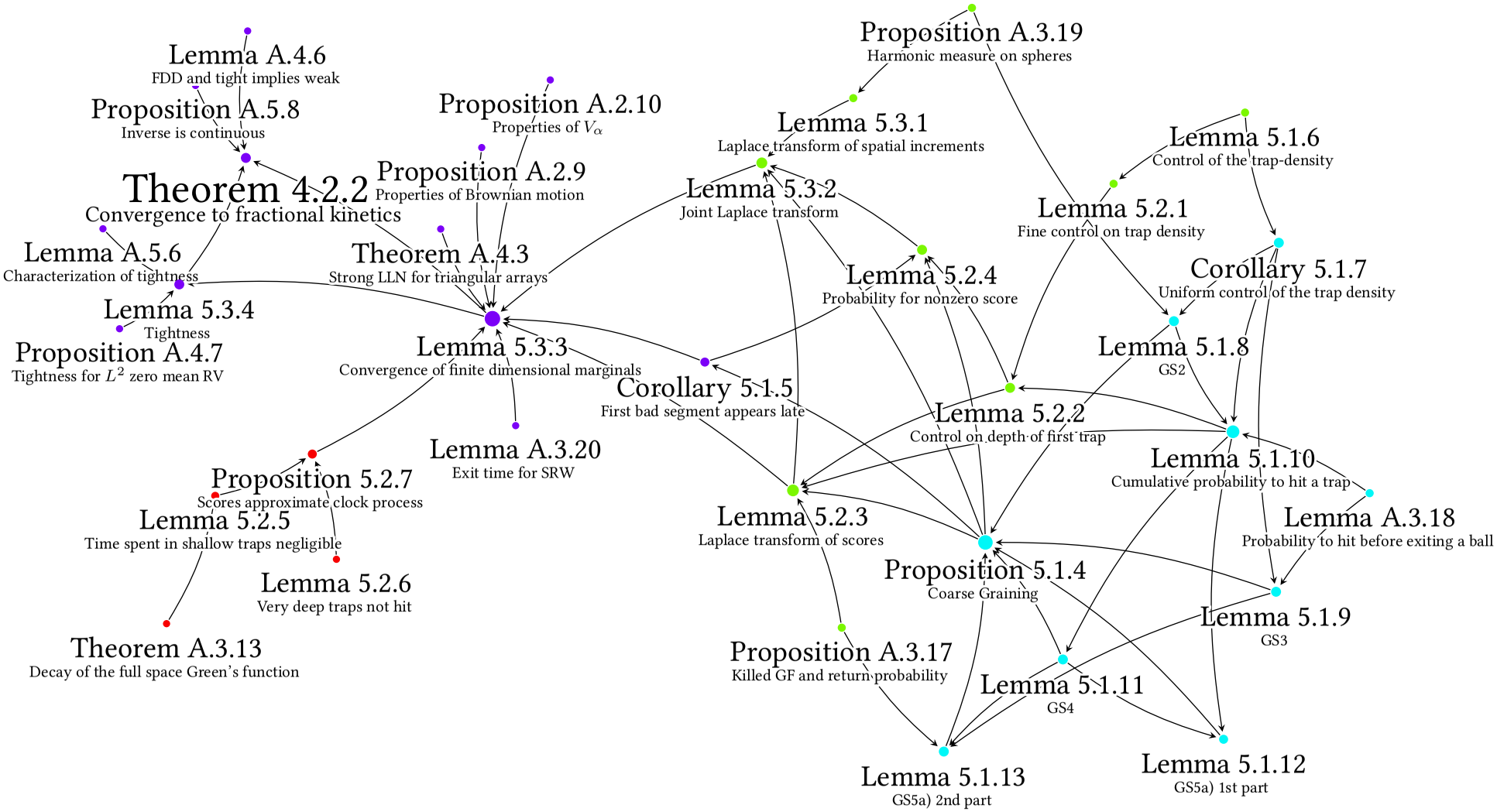

Figure B.1: Flowchart of the proof. Coloring of nodes done by modularity score according to Blondel *et al.* [21]

APPENDIX C

Device simulation models

The diffusive model used in this work (e.g. sec. 7.2) can be formulated as a system of coupled partial differential equations. We will give here the full model, including boundary conditions and a model for the mobility and diffusion coefficient, starting with the conventional drift diffusion device model.

C.1 The drift diffusion device model

For our simulations we use the steady state version of the equations, in the general case, the 0 on the left side of the continuity equation (eq. (C.1)) is replaced by a time derivative. The equations for *charge carrier density* ρ and *electrostatic potential* ψ read

$$0 = \nabla \cdot (-\mu\rho\nabla\psi + D\nabla\rho)\,, \tag{C.1}$$

$$0 = \epsilon_0\epsilon_r\Delta\psi + eN_{\text{sites}}\rho, \tag{C.2}$$

$$\rho(0) = \int_{-\infty}^{\infty} \text{DOS}(E) f_{E_{\text{Electrode},0}}(E)\mathrm{d}E, \tag{C.3}$$

$$\rho(L) = \int_{-\infty}^{\infty} \text{DOS}(E) f_{E_{\text{Electrode},L}}(E)\mathrm{d}E, \tag{C.4}$$

$$\psi(0) = 0, \tag{C.5}$$

$$\psi(L) = V_{\text{appl}} - V_{\text{bi}}. \tag{C.6}$$

Here V_{appl} is the *applied voltage*, V_{bi} the *built in voltage*, $E_{\text{Electrode,0/L}}$, the *workfunction* of the electrode at 0 resp. L and N_{sites} is the volume density of hopping sites.

We assume a Gaussian density of states:

$$\text{DOS}(E) = \frac{N}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{E^2}{2\sigma^2}\right), \tag{C.7}$$

and *Fermi-Dirac-statistics*

$$f_{E_F}(E) = \left(\exp\left(\frac{E - E_F}{k_b T}\right) - 1\right)^{-1}. \tag{C.8}$$

Using this kind of occupation statistics instead of Maxwell-Boltzmann, we have to consider the *generalized Einstein relation*

$$D = \mu \frac{\rho}{\frac{\partial \rho}{\partial E_F}}, \tag{C.9}$$

instead of the conventional *Einstein relation*

$$D = \mu k_B T. \tag{C.10}$$

The derivative $\frac{\partial \rho}{\partial E_F}$ can be computed from the relationship

$$\rho = \int_{-\infty}^{\infty} \mathrm{DOS}(E) f_{E_F}(E) \mathrm{d}E. \tag{C.11}$$

C.1.1 The EGDM mobility function

In the EGDM [88], the mobility $\mu = \mu_0 g_1 g_2$, where μ_0 is a constant and

$$g_1(T,\rho) = \exp\left(-0.44\hat{\sigma}^2\right) \exp\left(\frac{1}{2}(\hat{\sigma}^2 - \hat{\sigma})(2\rho)^\delta\right) \mathbf{1}_{\{\rho \leq 0.1\}} + g_1(T, 0.1)\, \mathbf{1}_{\{\rho > 0.1\}} \tag{C.12}$$

$$g_2(T, -\nabla\psi) = \exp\left\{0.44(\hat{\sigma}^{\frac{3}{2}} - 2.2)\left[\sqrt{1 + 0.8\frac{ea - \nabla\phi}{\sigma^2}} - 1\right]\right\} \mathbf{1}_{\left\{-\nabla\psi \leq \frac{2\sigma}{ea}\right\}} \tag{C.13}$$

$$+ g_2(T, \frac{2\sigma}{ea})\, \mathbf{1}_{\left\{-\nabla\psi > \frac{2\sigma}{ea}\right\}}, \tag{C.14}$$

with

$$\delta = 2\frac{\ln(\hat{\sigma}^2 - \hat{\sigma}) - \ln(\ln(4))}{\hat{\sigma}^2}. \tag{C.15}$$

Note that the dependence on T is via $\hat{\sigma} = \sigma / k_B T$.

C.1.2 The ECDM mobility function

The ECDM mobility [24] is similar, but has several different regimes. Here,

$$g_1(T, c) = \exp\left(-0.29\hat{\sigma}^2\right) \exp\left(0.25\hat{\sigma}^2\right) \times \tag{C.16}$$

$$\times \exp\left(0.7\hat{\sigma}(2\rho)^\delta\right) \mathbf{1}_{\{\rho \leq 0.025\}} + g_1(T, 0.025)\, \mathbf{1}_{\{\rho > 0.025\}}, \tag{C.17}$$

where

$$\delta = 2.3\frac{\ln\left(0.5\hat{\sigma}^2 + 1.4\hat{\sigma}\right) - \ln(\ln(4))}{\hat{\sigma}}. \tag{C.18}$$

The field dependence g_2 uses the field parameter $F = \frac{\nabla\psi}{\hat{\sigma}a}$, then, we let

$$h = \begin{cases} e^{\left((4/0.48)F\left(\left(1.05-1.2\rho^{0.7/\hat{\sigma}^{0.7}}\right)\left(\hat{\sigma}^{3/2}-2\right)\left(\sqrt{1+2F}\right)-1\right)\right)} & \text{if } F \in [0, 0.08), \\ e^{\left((1-4/3)(F/0.16-1)^2\left(\left(1.05-1.2\rho^{0.7/\hat{\sigma}^{0.7}}\right)\left(\hat{\sigma}^{3/2}-2\right)\left(\sqrt{1+2F}\right)-1\right)\right)} & \text{if } F \in [0.08, 0.16), \\ e^{\left(\left(\left(1.05-1.2\rho^{0.7/\hat{\sigma}^{0.7}}\right)\left(\hat{\sigma}^{3/2}-2\right)\left(\sqrt{1+2F}\right)-1\right)\right)} & \text{if } F \in [0.16, \infty). \end{cases} \tag{C.19}$$

Finally,

$$g_2(T, -\nabla\psi) = \left(h^{-2.4/(\hat{\sigma}-1)} + \left(Fg_1(T, -\nabla\psi)/(1-\rho)\right)^{2.4/(\hat{\sigma}-1)}\right)^{-1}. \tag{C.20}$$

C.2 Augmented model

In order to take into account that in short devices, even in the steady state, a certain fraction of carriers is trapped in atypically deep traps, we split the density ρ into ρ_{free} and ρ_{trapped} according to

$$\rho_{\text{free}} = \int_{-\infty}^{\infty} \mathrm{DOS}(E) f_{E_F}(E) \mathrm{d}E, \quad \rho_{\text{trapped}} = \int_{-\infty}^{\infty} \mathrm{DOS}_{\text{trap}}(E) f_{E_F}(E) \mathrm{d}E. \tag{C.21}$$

Then, only the free carriers ρ_{free} are used in the continuity equation (eq. (C.1)), while the sum $\rho_{\text{free}} + \rho_{\text{trapped}}$ is used in the Poisson equation (eq. (C.2)).

This is completely equivalent to the situation when charge carriers can reach equilibrium and actual traps are present. However, one must be aware, that the states introduced in $\mathrm{DOS}_{\text{trap}}$ are not physical states, but merely a way to introduce dispersive behavior into a non-dispersive model. A typical choice for $\mathrm{DOS}_{\text{trap}}$ is a simple delta function (more precisely, point measure) at some trap energy E_{trap} which is sufficiently deep and an intensity N_{trap}. Using these two parameters (*i.e.* trap depth and concentration), dispersive behavior can be emulated by the drift diffusion model as shown in sec. 7.2.

APPENDIX D

Some details on the kinetic Monte Carlo implementation

As we have mentioned, the kMC code used throughout chap. 7 has been specifically written to study the dispersive nature of charge transport without the finite-size effect artifacts. While it is not the core of this thesis, some original work has also been done to achieve this. More precisely, the author could not find any of these tools or methods.

D.1 General remarks on the implementation

The implementation is done in C++, using the MATLAB/C++ interface mex to seamlessly integrate the method with a powerful visualization tool. We have used an object-oriented approach, the main object being the `Site`, where information on the neighbors, the energy and the rates are stored. When the rates do not only depend on the initial site (as in our simplified model) but also on the target and the edge (as it is the case for MA/Marcus rates), we additionally need an `Edge` class. In this case, we collect the site and edge objects in a `Morphology` class.

The simulation is initialized with a set of visited sites, the only member of which is a site at the origin (and the neighbors in case the rates depend on them). Then a random variable is drawn to decide to which neighbor the charge carrier is transferred. In the simple model all probabilities are $1/6$, while in the MA/Marcus model these probabilities are computed from the Energies and transfer integrals of the participating sites and edges. For all random number generators, the Mersenne Twister implementation of the boost library is used.

After each step, we check if the site we now visit is contained in the set of visited sites. If it is not, a new site is generated and added. Again, the procedure is more complicated in the case of MA/Marcus rates, because additionally the edges and all neighbors need to be checked for (they may already exist) and generated if they do not exist. Since this requires many searches through the set of sites and edges, it is crucial to have a very fast way to search through the sites. This is achieved via an identification of $\mathbb{Z}^3$ with

$\mathbb{N}$ described in sec. D.2.

Since this method generates the morphology along the trajectory of the charge carrier, we do not need to worry about the numerical finite size effect interferring with the physical finite size effect, *i.e.* dispersion, which we want to study.

D.2 An efficient way to search in a set of $\mathbb{Z}^3$ indexed objects

In order to efficiently check, if a certain site is already visited we use a way to identify each point on the integer lattice $\mathbb{Z}^3$ with a natural number. While it is common konwledge that there exists a bijection between $\mathbb{Z}^3$ and $\mathbb{N}$, the author did not find any literature which actually contained an example for this – though we can not imagine this has not been done before.

What we do is enumerating the points in $\mathbb{Z}^d$ on increasingly large cube surfaces. It is implemented as follows: Given $x = (x_1, x_2, x_3) \in \mathbb{Z}^d$

- Find $r = \max\{x_1, x_2, x_3\}$.
- x is on the surface of a cube with sidelength $2r + 1$ centered at the origin. We now use the convention that
  - we start enumerating at the left ($x = -r$) surface of the cube. Enumerating on a finite surface is easy, we just go row-by-row.
  - next we consider the right ($x = r$ surface) in the same fashion.
  - the remaining "open cube" is partitioned into rings. We enumerate the rings in ascending x, first treating the case $y = r$, then $y = -r$, and finally $z = r$ resp. $z = -r$. While in each of these cases we only enumerate lines, we must be careful not do double-count the points on the edge, thus we only enumerate points where $|y| < r$ in the last two cases.

This function can now be used as (perfect) hash function in order to find a newly visited site in the set of already visited sites.

D.3 Identification and treatment of computational traps

In the general case, when dealing with the full physical rate expressions, there is a case of trapping which is physically not of importance, but severely slows down computations. Namely the case when two neighboring molecules have a very good connection, while the connection to the other molecules is not extraordinary good.

In this case, the charge carrier will hop back and forth between these two molecules very often before going to another one. While, due to the good connection, this happens in very short physical time, the cost for computing a step in the kinetic Monte Carlo method is the same as for any other. Therefore, a lot of computation time will be required.

This can easily be prevented. When our algorithm has fully discovered a molecule at z_1 and its neighbor at z_2, we compare the rates bewteen z_1 and z_2 with those of z_1 and z_2 to their other neighbors. If we find that the probability to go from z_1 to z_2 and the probability to go from z_2 to z_1 are both greater than 95%, we mark both z_1 and z_2 a trap and handle them differently when a charge carrier arrives at either of them.

Instead of performing a normal kMC step, we compute the number of back-and-forth hops as a geometric random variable with success probability (minimum of escape attempts at z_1 and z_2)

$$\mathbb{P}\left(\text{escape from trap } z_1 z_2\right) = 1 - \left(1 - \mathbb{P}\left(\text{escape from } z_1\right)\right)\left(1 - \mathbb{P}\left(\text{escape from } z_2\right)\right). \tag{D.1}$$

Then we compute the time those steps needed as the sum of two Erlang distributions (sum of exponential distributions) with $r_{z_1 \to z_2}$ resp. $r_{z_2 \to z_1}$ as rate parameter and the result of the geometric random variable as the number of trials.

This way of computing this special case is not an approximation, but is exact because we only used the properties of the exponential, Erlang and geometric distribution to get an easier way of computing steps when we already know that most of the steps will occur between two sites. A similar approach has been proposed very recently by Brereton *et al.* [28].

Bibliography

1. Abramowitz, M. & Stegun, I. A. *Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables* 9th ed. (Dover Publications, New York, 1964).
2. Ambegaokar, V., Halperin, B. I. & Langer, J. S. *Hopping Conductivity in Disordered Systems.* Phys. Rev. B **4,** 2612–2620 (1971).
3. Andres, S., Deuschel, J.-D. & Slowik, M. *Invariance principle for the random conductance model in a degenerate ergodic environment.* ArXiv e-prints (2013).
4. Arkhipov, V. I. & Rudenko, A. I. *Drift and diffusion in materials with traps.* Philos. Mag. B **45,** 189–207 (1982).
5. Arkhipov, V. I., Heremans, P., Emelianova, E. V., Adriaenssens, G. J. & Bässler, H. *Weak-field carrier hopping in disordered organic semiconductors: the effects of deep traps and partly filled density-of-states distribution.* J. Phys.: Condens. Matter **14,** 9899 (2002).
6. Ash, R. & Doléans-Dade, C. *Probability and Measure Theory* 2nd ed. (Academic Press, New York, 2000).
7. Ashcroft, N. W. & Mermin, D. N. *Solid state physics* 33rd ed. (Brooks/Coole Cengage Learning, Belmont, Calif., 2006).
8. Baeumer, B. & Meerschaert, M. M. *Stochastic Solutions for Fractional Cauchy Problems.* Calc. Appl. Anal **4,** 481–500 (2001).
9. Baranovskii, S. D., Cordes, H., Hensel, F. & Leising, G. *Charge-carrier transport in disordered organic solids.* Phys. Rev. B **62,** 7934–7938 (2000).
10. Baranovskii, S. D., Rubel, O. & Thomas, P. *On the concentration and field dependences of hopping mobility in disordered organic solids.* J. Non-Cryst. Solids **352,** 1644–1647 (2006).
11. Barlow, M. T. & Černý, J. *Convergence to fractional kinetics for random walks associated with unbounded conductances.* Probab. Theory Relat. Fields **149,** 639–673 (2011).
12. Barlow, M. T. & Deuschel, J.-D. *Invariance principle for the random conductance model with unbounded conductances.* Ann. Probab. **38,** 234–276 (2010).

13. Bässler, H. *Charge transport in disordered organic photoconductors - A Monte Carlo simulation study.* Phys. Stat. Sol. (b) **15,** 16–56 (1993).
14. Baumeier, B., Stenzel, O., Poelking, C., Andrienko, D. & Schmidt, V. *Stochastic modeling of molecular charge transport networks.* Phys. Rev. B **86,** 184202 (2012).
15. Ben Arous, G. & Černý, J. *Scaling limit for trap models on $\mathbb{Z}^d$.* Ann. Probab. **35,** 2356–2384 (2007).
16. Bertoin, J. *Lévy processes* 1st ed. **121** (Cambridge Univ. Pr., Cambridge, 1998).
17. Bertoin, J. in. 27, 1–91 (1999).
18. Billingsley, P. *Convergence of probability measures* 2nd ed. (Wiley, New York, 1999).
19. Billingsley, P. *Probability and Measure* (Wiley, 2012).
20. Bingham, N. H. *Limit theorems for occupation times of Markov processes.* Z. Wahrscheinlichkeit. **17,** 1–22 (1971).
21. Blondel, V. D., Guillaume, J.-L., Lambiotte, R. & Lefebvre, E. *Fast unfolding of communities in large networks.* J. Stat. Mech.-Theory E. P10008 (2008).
22. Born, M. & Oppenheimer, R. *Zur Quantentheorie der Molekeln.* Ann. phys. **84,** 457–484 (1927).
23. Bouchaud, J. P. *Weak ergodicity breaking and aging in disordered systems.* J. Phys. I France **2,** 1705–1713 (1992).
24. Bouhassoune, M., van Mensfoort, S., Bobbert, P. & Coehoorn, R. *Carrier-density and field-dependent charge-carrier mobility in organic semiconductors with correlated Gaussian disorder.* Org. Elec. **10,** 437–445 (2009).
25. Bovier, A. *Extremes, Sums, Lévy Processes and Ageing* tech. rep. (Lectures given in 2010 at the Technion, Haifa, 2010).
26. Braaksma, B. L. J. *Asymptotic expansions and analytic continuations for a class of Barnes-integrals.* Compos. math. **15,** 239–341 (1962-1964).
27. Breiman, L. *Probability* (Society for Industrial and Applied Mathematics, Philadelphia, 1992).
28. Brereton, T. *et al. Efficient Simulation of Markov Chains Using Segmentation.* Methodology and Computing in Applied Probability **16,** 465–484 (2014).
29. Chiang, C. K. *et al. Electrical Conductivity in Doped Polyacetylene.* Phys. Rev. Lett. **39,** 1098–1101 (1977).
30. Child, C. D. *Discharge From Hot CaO.* Phys. Rev. (Series I) **32,** 492–511 (1911).

31. Ciesielski, Z. & Taylor, S. J. *First Passage times and Sojourn Times for Brownian Motion in Space and the Exact Hausdorff Measure of the Sample Path.* T. Am. Math. Soc. **103,** 434–450 (1962).

32. Coehoorn, R. *Hopping mobility of charge carriers in disordered organic host-guest systems: Dependence on the charge-carrier concentration.* Phys. Rev. B **75,** 155203 (2007).

33. Coehoorn, R., Pasveer, W. F., Bobbert, P. A. & Michels, M. A. J. *Charge-carrier concentration dependence of the hopping mobility in organic materials with Gaussian disorder.* Phys. Rev. B **72,** 115206 (2005).

34. Cottaar, J., Koster, L. J. A., Coehoorn, R. & Bobbert, P. A. *Scaling Theory for Percolative Charge Transport in Disordered Molecular Semiconductors.* Phys. Rev. Lett. **107,** 136601 (2011).

35. Cottaar, J., Coehoorn, R. & Bobbert, P. A. *Scaling theory for percolative charge transport in molecular semiconductors: Correlated versus uncorrelated energetic disorder.* Phys. Rev. B **85,** 245205 (2012).

36. Dirac, P. A. M. *The principles of quantum mechanics* 4th ed. **27** (Clarendon Press, Oxford University Press, Oxford, 2010).

37. Dunlap, D. H., Parris, P. E. & Kenkre, V. M. *Charge-Dipole Model for the Universal Field Dependence of Mobilities in Molecularly Doped Polymers.* Phys. Rev. Lett. **77,** 542–545 (1996).

38. Fontes, L. R. G., Isopi, M. & Newman, C. M. *Random walks with strongly inhomogeneous rates and singular diffusions: convergence, localization and aging in one dimension.* Ann. Probab. **30,** 579–604 (2002).

39. Frenkel, J. *On Pre-Breakdown Phenomena in Insulators and Electronic Semi-Conductors.* Phys. Rev. **54,** 647–648 (1938).

40. Gadella, M. & Gòmez, F. *A Unified Mathematical Formalism for the Dirac Formulation of Quantum Mechanics.* Found. phys. **32,** 815–869 (2002).

41. Gartstein, Y. N. & Conwell, E. M. *High-field hopping mobility in molecular systems with spatially correlated energetic disorder.* Chem. Phys. Lett. **245,** 351 –358 (1995).

42. Gayrard, V. *Convergence of clock process in random environments and aging in Bouchaud's asymmetric trap model on the complete graph.* Electron. J. Probab. **17,** no. 58, 1–33 (2012).

43. Gillespie, D. T. *A general method for numerically simulating the stochastic time evolution of coupled chemical reactions.* J. Comput. Phys. **22,** 403 –434 (1976).

44. Gruenewald, M., Pohlmann, B., Movaghar, B. & Würtz, D. *Theory of non-equilibrium diffusive transport in disordered materials.* Philos. Mag. B **49,** 341–356 (1984).

45. Hohenberg, P. & Kohn, W. *Inhomogeneous Electron Gas.* Phys. Rev. **136,** B864–B871 (1964).

46. Hu, T.-C., Móricz, F. & Taylor, R. *Strong laws of large numbers for arrays of rowwise independent random variables.* Acta Math. Hung. **54,** 153–162 (1989).

47. Kantor, Y. & Kardar, M. *Anomalous dynamics of forced translocation.* Phys. Rev. E **69,** 021806 (2004).

48. Karatzas, I. & Shreve, S. E. *Brownian motion and stochastic calculus* **113** (Springer, New York, 1988).

49. Kepler, R. G., Bierstedt, P. E. & Merrifield, R. E. *Electronic Conduction and Exchange Interaction in a New Class of Conductive Organic Solids.* Phys. Rev. Lett. **5,** 503–504 (1960).

50. Knapp, E. & Ruhstaller, B. *Numerical analysis of steady-state and transient charge transport in organic semiconductor devices.* Opt. Quant. Electron **42,** 667–677 (2011).

51. Knapp, E., Häusermann, R., Schwarzenbach, H. & Ruhstaller, B. *Numerical Simulation of Charge Transport in Disordered Organic Semiconductor Devices.* J. Appl. Phys. **108,** 054504 (2010).

52. Koezuka, H., Tsumura, A. & Ando, T. *Field-effect transistor with polythiophene thin film.* Synth. Metals **18,** 699 –704 (1987).

53. Kohn, W. & Sham, L. J. *Self-Consistent Equations Including Exchange and Correlation Effects.* Phys. Rev. **140,** A1133–A1138 (1965).

54. Kordt, P., Stenzel, O., Baumeier, B., Schmidt, V. & Andrienko, D. *Parametrization of Extended Gaussian Disorder Models from Microscopic Charge Transport Simulations.* J. Chem. Theory Comput. **10,** 2508–2513 (2014).

55. Kotulski, M. *Asymptotic distributions of continuous-time random walks: A probabilistic approach.* J. Stat. Phys. **81,** 777–792 (1995).

56. Lawler, G. *Intersections of Random Walks* (Birkhäuser Boston, 1996).

57. Letheby, H. *XXIX.-On the production of a blue substance by the electrolysis of sulphate of aniline.* J. Chem. Soc. **15,** 161–163 (1862).

58. Levine, I. N. *Quantum chemistry* 7th ed. (Pearson, Boston, 2014).

59. Liu, F., Anh, V., Turner, I. & Zhuang, P. *Time fractional advection-dispersion equation.* Journal of Applied Mathematics and Computing **13,** 233–245 (2003).

60. Loève, M. *Probability theory 1* 4th ed. **45** (Springer, New York, 1977).

61. Lukyanov, A. & Andrienko, D. *Extracting nondispersive charge carrier mobilities of organic semiconductors from simulations of small systems.* Phys. Rev. B **82,** 193202 (2010).

62. Mainardi, F. & Pagnini, G. *The Wright Functions As Solutions of the Time-fractional Diffusion Equation.* Appl. Math. Comput. **141,** 51–62 (2003).

63. Mantegna, R. N. & Stanley, H. E. *Stochastic Process with Ultraslow Convergence to a Gaussian: The Truncated Lévy Flight.* Phys. Rev. Lett. **73,** 2946–2949 (1994).

64. Marcus, R. A. *On the Theory of Oxidation & Reduction Reactions Involving Electron Transfer. I.* J. Chem. Phys. **24,** 966–978 (1956).

65. Martinez, F. S., Pachepsky, Y. A. & Rawls, W. J. in *Advances in Fractional Calculus* 199–212 (2007).

66. Massé, A., Coehoorn, R. & Bobbert, P. A. *Universal Size-Dependent Conductance Fluctuations in Disordered Organic Semiconductors.* Phys. Rev. Lett. **113,** 116604 (2014).

67. May, F. *Charge transport simulations in organic semiconductors* PhD thesis (Max Planck institute for polymer research, 2012).

68. McNeill, R., Suidak, R., H., W. J. & Weiss, D. E. *Electronic Conduction in Polymers. I. The Chemical Structure of Polypyrrole.* Aust. J. chem. **16,** 1056 –1075 (1963).

69. Meerschaert, M. M., Benson, D. A., Scheffler, H.-P. & Baeumer, B. *Stochastic solution of space-time fractional diffusion equations.* Phys. Rev. E **65,** 041103 (2002).

70. Meerschaert, M. & Scheffler, H. *Limit theorems for continuous time random walks with infinite mean waiting times* tech. rep. (2004).

71. Mette, H. & Pick, H. *Elektronenleitfähigkeit von Anthracen-Einkristallen.* Z. phys. **134,** 566–575 (1953).

72. Metzler, R. & Klafter, J. *The random walk's guide to anomalous diffusion: a fractional dynamics approach.* Physics Reports 339 **339,** 1–77 (2000).

73. Miller, A. & Abrahams, E. *Impurity Conduction at Low Concentrations.* Phys. Rev. **120,** 745–755 (1960).

74. Miller, K. S. & Ross, B. *An Introduction to the Fractional Calculus and Fractional Differential Equations* 1st ed. (Wiley, New York, 1993).

75. Mommer, M. & Lebiedz, D. *Modeling Subdiffusion Using Reaction Diffusion Systems.* SIAM J. appl. math. **70,** 112–132 (2009).

76. Monroe, D. *Hopping in Exponential Band Tails.* Phys. Rev. Lett. **54,** 146–149 (1985).

77. Montroll, E. W. & Weiss, G. H. *Random Walks on Lattices. II.* J. Math. Phys. **6,** 167–181 (1965).

78. Mott, N. F. & Gurney, R. W. *Electronic processes in ionic crystals* 2nd ed. (Clarendon Press, Oxford, 1948).

79. Movaghar, B & Schirmacher, W. *On the theory of hopping conductivity in disordered systems.* J. Phys. C-Soild State **14,** 859 (1981).

80. Movaghar, B., Pohlmann, B. & Schirmacher, W. *Theory of electronic hopping transport in disordered materials.* Philos. Mag. B **41,** 29–62 (1980).

81. Movaghar, B., Grünewald, M., Pohlmann, B., Würtz, D. & Schirmacher, W. *Theory of hopping and multiple-trapping transport in disordered systems.* J. Stat. Phys. **30,** 315–334 (1983).

82. Movaghar, B., Grünewald, M., Ries, B., Baessler, H. & Würtz, D. *Diffusion and relaxation energy in disordered organic and inorganic materials.* Phys. Rev. B **33,** 5545–5554 (1986).

83. Nitsche, R., Kurpiers-Guenther, M., Thomschke, M., Schober, M. & Leo, K. *29.3: Combined Electrical and Optical Simulation of OLED Devices.* SID Int. Symp. Dig. Tec. **39,** 411–414 (2008).

84. Nobelprize.org. *The Nobel Prize in Chemistry 2000* Nobel Media AB 2014. `http://www.nobelprize.org/nobel_prizes/chemistry/laureates/2000/`.

85. Noolandi, J. *Equivalence of multiple-trapping model and time-dependent random walk.* Phys. Rev. B **10,** 4474–4479 (1977).

86. Oelerich, J. O., Jansson, F., Nenashev, A. V., Gebhard, F. & Baranovskii, S. D. *Energy position of the transport path in disordered organic semiconductors.* J. Phys.: Condens. Matter **26,** 255801 (2014).

87. Oldham, K. B. & Spanier, J. *The fractional calculus. theory and applications of differentiation and integration to arbitrary order* **111** (Academic Press, New York, 1974).

88. Pasveer, W. *et al. Unified Description of Charge-Carrier Mobilities in Disordered Semiconducting Polymers.* Phys. Rev. Lett. **94,** 206601 (2005).

89. Perline, R. *Strong, weak and false inverse power laws.* Stat. Sci. **20,** 68–88 (2005).

90. Pope, M., Kallmann, H. P. & Magnante, P. *Electroluminescence in Organic Crystals.* J. Chem. Phys. **38,** 2042–2043 (1963).

91. Reed, M. & Simon, B. *Methods of modern mathematical physics: Analysis of operators* (Academic Press, New York, 1978).

92. Ridley, J. & Zerner, M. *An intermediate neglect of differential overlap technique for spectroscopy: Pyrrole and the azines.* Theor. Chim. acta **32,** 111–134 (1973).

93. Rudenko, A. I. & Arkhipov, V. I. *Drift and diffusion in materials with traps.* Philos. Mag. B **45,** 177–187 (1982).

94. Rudenko, A. I. & Arkhipov, V. I. *Drift and diffusion in materials with traps.* Philos. Mag. B **45,** 209–226 (1982).

95. Rühle, V., Junghans, C., Lukyanov, A., Kremer, K. & Andrienko, D. *Versatile Object-Oriented Toolkit for Coarse-Graining Applications.* J. Chem. Theory Comput. **5,** 3211–3223 (2009).

96. Samoradnitsky, G. & Taqqu, S. *Stable Non-Gaussian Random Processes: Stochastic Models with Infinite Variance* (Chapman & Hall, New York, 1994).

97. Sato, K. I. *Lévy Processes and Infinitely Divisible Distributions* (Cambridge University Press, Cambridge, 1999).

98. Scher, H. & Lax, M. *Stochastic transport in a disordered solid. I. Theory.* Phys. Rev. B **7,** 4491–4502 (1973).

99. Scher, H. & Montroll, E. *Anomalous transit-time dispersion in amorphous solids.* Phys. Rev. B **12,** 2455–2477 (1975).

100. Schober, M. *et al. Quantitative description of charge-carrier transport in a white organic light-emitting diode.* Phys. Rev. B **84,** 165326 (2011).

101. Schönherr, G., Bässler, H. & Silver, M. *Dispersive hopping transport via sites having a Gaussian distribution of energies.* Philos. Mag. B **44,** 47–61 (1981).

102. Schrøder, T. B. & Dyre, J. C. *ac Hopping Conduction at Extreme Disorder Takes Place on the Percolating Cluster.* Phys. Rev. Lett. **101,** 025901 (2008).

103. Shankar, R. *Principles of quantum mechanics* 2nd ed. (Springer, New York, 2010).

104. Shirakawa, H., Louis, E. J., MacDiarmid, A. G., Chiang, C. K. & Heeger, A. J. *Synthesis of electrically conducting organic polymers: halogen derivatives of polyacetylene, (CH).* J. Chem. Soc., Chem. Commun. 578–580 (1977).

105. Sibatov, R. T. & Uchaikin, V. V. *Fractional differential kinetics of charge transport in unordered semiconductors.* Semiconductors **41,** 335–340 (2007).

106. Sibatov, R. T. & Uchaikin, V. V. *Fractional differential approach to dispersive transport in semiconductors.* Physics - Usp+ **52,** 1019–1043 (2009).

107. Sibatov, R. T. & Uchaikin, V. V. *Truncated Levy statistics for transport in disordered semiconductors.* Commun. Nonlinear Sci. Numer. Simul. **16,** 4564ÂŰ4572 (2011).

108. Skorokhod, A. *Limit Theorems for Stochastic Processes.* Theory Probab. Appl. **1,** 261–290 (1956).

109. Stodtmann, S., Lee, R. M., Weiler, C. K. F. & Badinski, A. *An efficient numerical method for doped organic semiconductor devices.* J. Appl. Phys **112,** 114909 (2012).

110. Sze, S. M. *Physics of semiconductor devices* 3rd ed. (Wiley, New York, 1981).

111. Szymanski, M. Z., Kulszewicz-Bajer, I., Faure-Vincent, J. & Djurado, D. *Comparison of simulations to experiment for a detailed analysis of space-charge-limited transient current measurements in organic semiconductors.* Phys. Rev. B **85,** 195205 (2012).

112. Tang, C. W. *Two-layer organic photovoltaic cell.* Appl. Phys. Lett. **48,** 183–185 (1986).

113. Tang, C. W. & VanSlyke, S. A. *Organic electroluminescent diodes.* Appl. Phys. Lett. **51,** 913–915 (1987).

114. Tro, N. J. *Chemistry. a molecular approach* 3rd ed. (Pearson, Harlow, 2014).

115. Van Mensfoort, S., Shabro, V., de Vries, R., Janssen, R. & Coehoorn, R. *Hole transport in the organic small molecule material α-NPD: evidence for the presence of correlated disorder.* J. Appl. Phys. **107,** 113710–113710 (2010).

116. Van Roosbroeck, W. *Theory of the flow of electrons and holes in germanium and other semiconductors.* Bell System Tech. J. **29,** 560–607 (1950).

117. Vissenberg, M. & Matters, M. *Theory of the field-effect mobility in amorphous organic transistors.* Phys. Rev. B **57** (1998).

118. Wedemeier, A., Zhang, T., Merlitz, H., Wu, C.-X. & Langowski, J. *The role of chromatin conformations in diffusional transport of chromatin-binding proteins: Cartesian lattice simulations.* J. Chem. Phys. **128,** – (2008).

119. Whitt, W. *Stochastic-Process Limits: An Introduction to Stochastic-Process Limits and Their Application to Queues* (Springer, New York, 2002).

120. Yoshida, K. *Functional analysis* 6th ed. **123** (Springer, Berlin, 1980).