Maximilian Hofer

Branching Brownian Motion at the Tip

Master's Thesis
Joint Master High Energy Physics
Institut Polytechnique de Paris
Eidgenössische Technische Hochschule Zürich

Supervised by
Prof. Dr. Stéphane Munier

Submitted on 10th July 2023

INSTITUT
POLYTECHNIQUE
DE PARIS

ETH zürich

Abstract

This thesis is concerned with the tip (i.e. the furthest populated region) of the one-dimensional branching Brownian motion conditioned to have its rightmost particle arrive at a predefined position X at a certain time T. In recent work conjectures were made about the probability distribution of particle densities in the tip when T tends to infinity and X is significantly larger than the average position m_T of the lead particle. We report here on numerical results that are in tension with the predictions. Although it is possible that the disagreement is merely a consequence of the limited parametric range of the simulation, we also propose a new and alternative computation of the probability distribution that is in better accordance with the data. The necessary background to understand the results is reviewed from the ground up, including basic properties of stochastic processes.

Contents

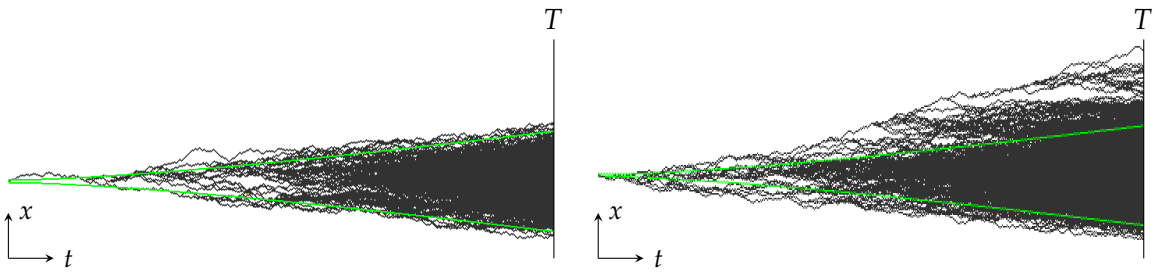
Chapter 1

Introduction

The *branching Brownian motion (BBM)* is a particularly accessible example of a random process in which the number of constituents increases over time. As the name suggests, it is obtained by considering a particle performing a Brownian motion and adding the possibility for it to decay into two (or more) particles that subsequently realize independent BBMs. In this work the decay will always result in two descendants and happen after an exponentially distributed random time. Samples are shown in Figure 1.1a.

Since its introduction [1] by Ikeda, Nagasawa and Watanabe in 1968, the BBM has been the subject of intense study in mathematics [2] and found applications in a wide range of scientific disciplines such as physics [3], biology [4], chemistry [5] and computer science [6]. The motivation underlying the present work comes from particle physics and more precisely from high-energy quantum chromodynamics (QCD). A brief explanation of the context in which the kind of process considered here is relevant is given below. However, the remainder of the report will only be concerned with the mathematical model and not with applications to QCD. The reasoning is in the spirit of a commonly adopted attitude in statistical physics, in which complex systems are replaced by simple toy models that share certain universal properties.

Outline. In the next chapter we will introduce elementary concepts related to diffusion processes to provide the language in which more advanced results can be formulated. In particular, the Wiener process is defined in Section 2.1 and used to describe a broad class of diffusion processes in terms of stochastic differential equations (SDEs). Section 2.2



(a) Unconditioned BBM *(b) Stretched tip*

Figure 1.1: Two examples of branching Brownian motions or more precisely of their discrete counterparts called branching random walks. The green lines correspond to the average positions of the outermost particles and behave for large enough times as described by (3.4)*. Events as on the right where the maximal particle is far beyond the expected position are rare yet of significant impact to the asymptotic behaviour.*

demonstrates how these SDEs relate to certain partial differential equations (PDEs) called Fokker-Planck equations that govern the dynamics of a process's probability density function (PDF). Infinitesimal generators are introduced as well. These tools are then used to derive basic properties of the recurring three-dimensional Bessel process in Section 2.3. By definition, the latter is realized by the modulus of three-dimensional Brownian motion.

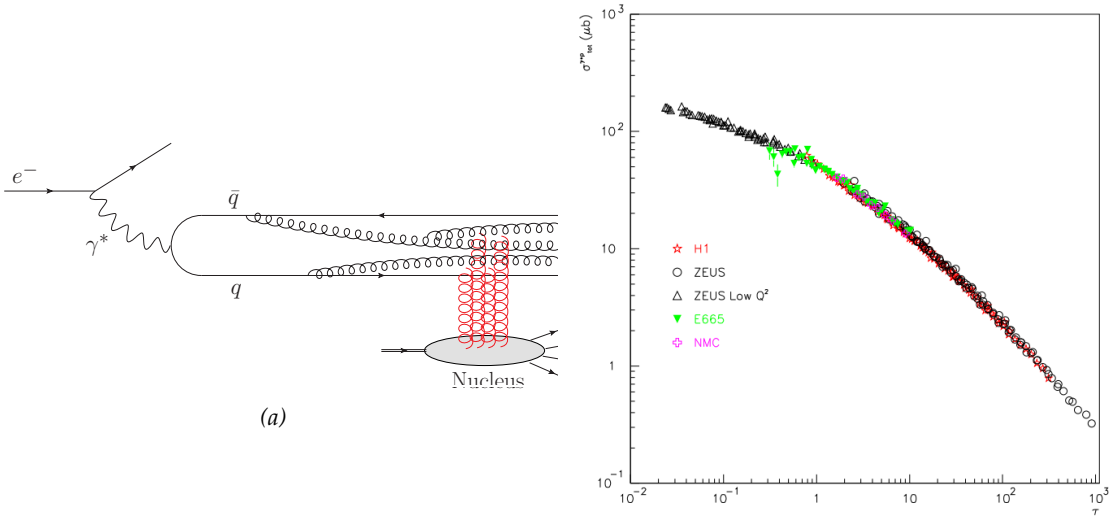
In the third chapter, the BBM is finally defined and PDEs for the average particle number densities and the PDF of the "rightmost" (i.e. with the largest position $x \in \mathbb{R}$) particle are derived, the latter one being called the Fisher-Kolmogorov-Petrovsky-Piscounov or FKPP equation (Sec. 3.1). The region close to the rightmost particle — we call it the *tip* — is sparsely populated and hence the most sensitive one to stochastic effects. Understanding of the asymptotic behaviour of the BBM for large times relies therefore on a comprehension of the statistics of the tip. To study those, the framework of [7, 8] is presented in Section 3.2. It is based on conditioning the BBM to have its rightmost particle at a large time T to be found at a specific position X. The same technique is applied and slightly extended in the discrete setting of a branching random walk (BRW) in Section 3.3.

Chapter 4 is concerned with *stretched* tips that extend far beyond the position at which they are found on average (cf. Figure 1.1b). What makes this case interesting is that particle densities in typical configurations are significantly suppressed in comparison to the average behaviour, which implies the existence of rare but impactful events. A recent study [8] of particle densities in such stretched tips is reviewed in Section 4.1. It conjectures an asymptotic formula (4.8) for the PDF of the particle number density at a fixed distance Δ from X. The bulk of the work leading up to this report consisted in writing a C++ library for the generation of conditioned BBMs with the algorithm described in Section 3.3 to numerically investigate this formula. This is explained in Section 4.2. There we also present the results — which clearly contradict (4.8). Although the reason might just be that the evolution time in the simulation was not sufficiently high to reflect the asymptotic behaviour, a new and different approach leading to the analogous formula (4.13) is seen to model the data much better. This alternative approach is explained in Section 4.3 and constitutes our main result. In the final Section 4.4 we make suggestions on how to proceed from here.

Motivation from QCD. To put the present work into context, we briefly sketch the role of branching processes in QCD. A detailed review of the correspondence can be found in [9]. For a general treatment of the relevant concepts of high energy QCD (like the BK equation mentioned below) we refer the reader to [10].

We consider deep inelastic scattering in which electrons or positrons are scattered off a nucleus by exchanging a virtual photon in the t-channel with some four-momentum q^μ and virtuality $Q^2 \equiv -q^2 > 0$ that couples to a quark line. This setup was for example realized at the HERA accelerator[1] (DESY) where it could be studied at high precision over a wide parameter range. In particular, it enabled cross-section measurements at small values (down to 10^{-4}, with $Q^2 \in [0, \sim 100\,\mathrm{GeV}^2]$) of the Bjorken variable $x = \frac{Q^2}{2p\cdot q} \in [0, 1]$ which has an interpretation as the proton's momentum fraction carried by the quark coupling to the virtual photon. This regime is interesting because the hadron's content becomes dominated by soft gluons as x gets smaller while Q can be kept fix so that the coupling $\alpha_s(Q)$ does not run and can be chosen small enough to enable perturbative computations.

[1] A more recent example is the Electron-Ion Collider in Brookhaven that is underway.



(a) (b)

Figure 1.2: Motivation from QCD. (a) Deep inelastic scattering in the rest frame of the target nucleus in the case of small values of the Bjorken scale x. The gluon production can be modelled by a branching process that shares universal properties with the BBM. (b) Geometric scaling [11]. Although the cross-section $\sigma(Q^2, x)$ (y-axis) of deep inelastic scattering a priori depends on two independent variables, it was surprisingly found that the data can be given in terms of a single one, $\tau = Q^2/Q_s^2(x)$ (x-axis; the function $Q_s^2(x)$ is called the saturation scale). This could be explained by properties of the FKPP equation. Figure from [12].

When looking at this setting in the rest frame of the proton, the picture in Figure 1.2a emerges in which the virtual photon γ^* is seen as a quark-antiquark pair. As one lowers x, the gluon density increases as fluctuations of shorter lifetime become visible to the nucleus. The splitting of gluons can be interpreted as a branching process that turns out to lie in the same *universality class* (see below) as the BBM. Since the virtual photon is color neutral, the nucleus can essentially interact only with gluons of a transverse momentum $k_\perp$ that is smaller than the nuclear saturation scale $Q_A \approx \Lambda_{\text{QCD}} A^{1/6}$ (so that the "resolution" $\sim \frac{1}{k_\perp}$ becomes large enough to "see" the color structure; atomic number A). This turns out to be very analogous to the question of whether the rightmost particle in a BBM is found to the right of a given position. While this can be discussed by studying the non-linear FKPP equation, the corresponding problem in QCD is described by the *Balitsky-Kovchegov* or *BK equation*. In fact, the FKPP and the BK equation share the same universality class. This means that results about the simple processes studied in this work can in principle be translated to statements about high energy QCD. Probably the most impressive example of this correspondence is related to the observation (dubbed *geometric scaling*; see [11]) that the cross-sections $\sigma(Q^2, x)$ in the HERA data can be seen as a function of a single variable $\tau = Q^2/Q_s^2(x)$, see Figure 1.2b. It was shown that this property can be viewed as a consequence of the fact that the solutions of the FKPP equation are *travelling waves* (cf. Sec. 3.1).

Acknowledgements. I am very thankful to Prof. Munier for many insightful discussions, helpful advice and even a desk in his office. I also want to express gratitude to the ETH Foundation and its donors — in particular Reto Jenatsch (†) — who funded my master's studies via the Excellence Scholarship & Opportunity Programme. Credit goes as well to the technical support team of CPHT for maintaining the cluster "Cholesky" of the PHYMATH mesocentre that was used to produce the data presented in Section 4.2.

Chapter 2

Basic Notions of Stochastic Processes

A *stochastic* or *random process* X is broadly speaking a map that assigns a random variable X_t to each time $t \in T$ in a discrete or continuous index set $T \subseteq \mathbb{R}$. In the following sections we review properties of various random processes that are essential for the study of the branching Brownian motion.

We use the notations $\mathbb{P}(A)$ for the probability that the statement A is true, $\mathbb{P}(A \mid B)$ for the probability that A is true given that B holds, $\mathbb{P}(A; B)$ for the probability that A and B hold simultaneously and $\mathbb{E}(X)$ for the expectation value of a random variable X. Moreover, we say that A holds *almost surely* if $\mathbb{P}(A) = 1$.

2.1 The Wiener process

In mathematics, Brownian motion is identified with a specific stochastic process in continuous time ($T = \mathbb{R}_{\geq 0}$) called the *Wiener process* W, characterized by the properties in the following box [13].

Wiener process.

- $W_0 = 0$ almost surely.
- For $t \geq s > 0$, the increment $W_t - W_s$ is a normally distributed random variable of mean 0 and variance $t - s$ that is independent of previous values W_r, $0 < r \leq s$.
- $t \mapsto W_t$ is almost surely continuous.

It must be ensured that this definition is consistent with the fact that $W_t - W_s = (W_t - W_r) + (W_r - W_s)$ for $t > r > s$. To this end let $X \coloneqq W_t - W_r$, $Y \coloneqq W_r - W_s$, $Z \coloneqq W_t - W_s$ and denote the normal distribution of mean μ and variance σ^2 by

$$f(x; \mu, \sigma^2) \coloneqq \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right).$$

According to the above definition X, Y and Z have distributions $f(x; 0, r-s)$, $f(y; 0, t-r)$ and $f(z; 0, t-s)$, respectively. The crucial addition property $X + Y = Z$ can now be shown by verifying that

$$\int_{-\infty}^{\infty} \mathrm{d}x \int_{-\infty}^{\infty} \mathrm{d}y\, f(x; 0, t-r) f(y; 0, r-s) \delta(x + y - z) = f(z; 0, t-s).$$

The evaluation of the left-hand side is straightforward and shall not be presented.

From the above it follows that the increment $\mathrm{d}W_t := W_{t+\mathrm{d}t} - W_t$ in an infinitesimally short time period $\mathrm{d}t$ has distribution $f(0, \mathrm{d}t)$. This notation is commonly used for writing stochastic differential equations (SDEs) to describe random processes related to the Brownian motion. All processes in this chapter are so-called *Itō processes*, meaning that they are of the form

$$\mathrm{d}X_t = \alpha(t, X_t)\,\mathrm{d}t + \beta(t, X_t)\,\mathrm{d}W_t, \tag{2.1}$$

where $\alpha(t, x)$ is a space and time dependent drift velocity and $\beta(t, x)$ is related to the diffusion coefficient (as seen in the next section) that may also be spacetime dependent.

The d-dimensional Wiener process is characterized by the same items as before, or even simpler as a collection of d independent one-dimensional Wiener processes that constitute the components in an orthonormal basis. One can then convince oneself that this does not depend on the choice of basis as the d-dimensional normal distribution is rotationally invariant.

2.2 The Fokker-Planck equation

For a process X as in (2.1) one is usually interested in the probability density function (PDF) $p(t, x) := \frac{\mathrm{d}}{\mathrm{d}x}\mathbb{P}(X_t \le x)$ at a given time t. Fokker-Planck equations are partial differential equations giving the dynamical laws of these densities. We illustrate this first in the case of the Wiener process, $X = W$.

Example (Wiener process). Given the PDF $p(t, y)$ of the Wiener process at a time t, the PDF at a slightly later time $t + \mathrm{d}t$ is by definition of $\mathrm{d}W_t$ determined by the convolution

$$p(t + \mathrm{d}t, x) = \int_{-\infty}^{\infty} \frac{\mathrm{d}y}{\sqrt{2\pi\,\mathrm{d}t}} e^{-\frac{(x-y)^2}{2\,\mathrm{d}t}} p(t, y) = \int_{-\infty}^{\infty} \frac{\mathrm{d}u}{\sqrt{2\pi}} e^{-u^2/2} p(t, x + \sqrt{\mathrm{d}t}u), \tag{2.2}$$

where in the second step we substituted $y = x + \sqrt{\mathrm{d}t}u$. In a Taylor expansion of p in space only the terms with even powers of u survive the integration, such that

$$\begin{aligned} p(t + \mathrm{d}t, x) &= \int_{-\infty}^{\infty} \frac{\mathrm{d}u}{\sqrt{2\pi}} e^{-u^2/2} \left(p(t, x) + u^2 \frac{\mathrm{d}t}{2} \frac{\partial^2 p}{\partial x^2}(t, x) \right) + O(\mathrm{d}t^2) \\ &= p(t, x) + \frac{\mathrm{d}t}{2} \frac{\partial^2 p}{\partial x^2}(t, x) + O(\mathrm{d}t^2). \end{aligned}$$

Rewriting this as a differential equation yields the *Fokker-Planck equation of the Wiener process*

$$\frac{\partial p}{\partial t}(t, x) = \frac{1}{2} \frac{\partial^2 p}{\partial x^2}(t, x) \tag{2.3}$$

which happens to be the infamous *diffusion equation* (with diffusion coefficient $\frac{1}{2}$).

We want to emphasize that (2.2) does not require $\mathrm{d}t$ to be small. Thus, for this simple case we can propagate the initial condition $p(0, x) = \delta(x)$ directly to any future time to find (set $t = 0$ and replace $\mathrm{d}t$ by t in (2.2))

$$p(t, x) = \int_{-\infty}^{\infty} \frac{\mathrm{d}y}{\sqrt{2\pi t}} e^{-\frac{(x-y)^2}{2t}} \delta(y) = \frac{1}{\sqrt{2\pi t}}\, e^{-x^2/2t}. \tag{2.4}$$

Fokker-Planck equation. For a more general stochastic process satisfying $\mathrm{d}X_t = \alpha(t, X_t)\,\mathrm{d}t + \beta(t, X_t)\,\mathrm{d}W_t$ we can proceed in the same way, although the calculation is a bit more involved. To first order in the time step $\mathrm{d}t$ the PDF changes according to

$$p(t + \mathrm{d}t, x) = \int_{-\infty}^{\infty} \frac{\mathrm{d}y}{\sqrt{2\pi\,\mathrm{d}t}\beta(t, y)} \exp\left(-\frac{(x - y - \alpha(t, y)\,\mathrm{d}t)^2}{2\beta(t, y)^2\,\mathrm{d}t}\right) p(t, y). \tag{2.5}$$

A substitution as before is now more difficult because of the y-dependence in α and β. Instead, the application of a Fourier transformation in space turns out to be very useful. A short calculation[1] shows that

$$\hat{p}(t + \mathrm{d}t, k) \equiv \int_{-\infty}^{\infty} \mathrm{d}x\, e^{-ikx} p(t + \mathrm{d}t, x) = \int_{-\infty}^{\infty} \mathrm{d}y e^{-ik(y+\alpha(t,y)\,\mathrm{d}t) - \frac{1}{2}k^2\beta(t,y)^2\,\mathrm{d}t} p(t, y),$$

which can be expanded in $\mathrm{d}t$ to obtain

$$\partial_t \hat{p}(t, k) = \int_{-\infty}^{\infty} \mathrm{d}y\, e^{-iky} \left(-ik\alpha(t, y)p(t, y) - k^2 \frac{1}{2}\beta(t, y)^2 p(t, y)\right)$$

so that after reversing the Fourier transformation we finally arrive at the *Fokker-Planck equation* (also known as *Kolmogorov forward equation*)

$$\boxed{\frac{\partial}{\partial t} p(t, x) = -\frac{\partial}{\partial x}\big(\alpha(t, x)p(t, x)\big) + \frac{\partial^2}{\partial x^2}\left(\frac{\beta(t, x)^2}{2} p(t, x)\right).} \tag{2.6}$$

In analogy to (2.3) we call $\frac{\beta(t,x)^2}{2}$ the diffusion coefficient. By multiplying (2.6) with x, integrating over $x \in \mathbb{R}$ and assuming that $\alpha(t, x)p(t, x)$ and $\beta(t, x)p(t, x)$ vanish sufficiently fast for $x \to \pm\infty$, we find that[2] $\partial_t \mathbb{E}(X_t) = \mathbb{E}(\alpha(t, X_t))$, which illuminates the meaning of the drift $\alpha(t, x)$.

Infinitesimal generator. It is worth noting that in the derivation (2.5)–(2.6) we never made use of any specific properties of $p(t, y)$ — except for the integration kernel which actually is nothing but $p(t + \mathrm{d}t, x \mid t, y)$. The point is that we could have derived in the same way a formula for the convolution with any (well-behaved) function $f(t, x)$,

$$\int_{-\infty}^{\infty} \mathrm{d}y\, p(t + \mathrm{d}t, x \mid t, y) f(t, y) = f(t, x) + \mathrm{d}t\mathcal{G}^\dagger f(t, x) + O(\mathrm{d}t^2), \tag{2.7}$$

where $\mathcal{G}^\dagger f(t, x) = -\partial_x\big(\alpha(t, x)f(t, x)\big) + \partial_x^2\big(\frac{\beta(t,x)^2}{2} f(t, x)\big)$. As suggested by the notation, $\mathcal{G}^\dagger$ is the adjoint (with respect to the $L^2(\mathbb{R})$ scalar product $\langle f, g\rangle = \int_{\mathbb{R}} \mathrm{d}x\, f(x)g(x)$) of another operator $\mathcal{G} = \alpha(t, x)\partial_t + \frac{\beta(t,x)^2}{2}\partial_x^2$ called the *infinitesimal generator* of the process X. We now relate (2.7) to a similar equation for $\mathcal{G}$. Multiplying (2.7) with a test function $g(t, x)$, integrating over x, using that $\langle g, f\rangle = \langle f, g\rangle$ and $\langle g, \mathcal{G}^\dagger f\rangle = \langle f, \mathcal{G}g\rangle$, we obtain

$$\left\langle f, \int_{-\infty}^{\infty} \mathrm{d}x\, p(t + \mathrm{d}t, x \mid t, \cdot) g(t, x)\right\rangle = \langle f, g\rangle + \mathrm{d}t\langle f, \mathcal{G}g\rangle + O(\mathrm{d}t^2).$$

[1] Use $\int_{-\infty}^{\infty} \mathrm{d}x\, e^{-ikx} e^{-(x-A)^2/B} = \sqrt{\pi B} e^{-ikA - \frac{1}{4}Bk^2}$ with $A = y + \alpha(t, y)\,\mathrm{d}t$, $B = 2\beta(t, y)^2\,\mathrm{d}t$.

[2] Use $\mathbb{E}(X_t) = \int_{\mathbb{R}} \mathrm{d}x\, xp(t, x)$ and partial integration.

Since this holds for every test function f this leads to

$$\int_{-\infty}^{\infty} \mathrm{d}x\, p(t+\mathrm{d}t, x \mid t, y) g(t,x) = g(t,y) + \mathrm{d}t \mathcal{G} g(t,y) + O(\mathrm{d}t^2). \tag{2.8}$$

Observing that

$$\mathbb{E}_y(g(t, X_{t+\mathrm{d}t})) = \int_{-\infty}^{\infty} \mathrm{d}x\, p(t+\mathrm{d}t, x \mid t, y) g(t,x),$$

with $\mathbb{E}_y$ being the expectation value with respect to a random process X with $X_t = y$, it follows that

$$\mathcal{G} g(t,y) = \lim_{\mathrm{d}t \searrow 0} \frac{\mathbb{E}_y(g(t, X_{t+\mathrm{d}t})) - g(t,y)}{\mathrm{d}t}. \tag{2.9}$$

In the mathematical literature it is usually via this formula that the infinitesimal generator is defined (e.g. [14, ch. VII]).

2.3 The 3D Bessel process

The 3D Bessel process naturally appears in the analysis of branching Brownian motion (see Section 3.2) and will play a crucial role in Chapter 4. In this section we define it and derive a few properties that will be useful later. This is also a good opportunity to apply the tools introduced in the previous two sections.

Definition and SDE. The *n-dimensional Bessel process* B is the process obeyed by the radial component of the n-dimensional Brownian motion $\boldsymbol{W}$. The infinitesimal generator can be obtained by convolving the kernel of $\boldsymbol{W}$ with a test function $f(t,r)$ that depends only on the norm $r = \|\vec{x}\|$.

$$\mathbb{E}(f(t, B_{t+\mathrm{d}t})) \equiv \int_{\mathbb{R}^n} \frac{\mathrm{d}^n y\, e^{\frac{-(\vec{x}-\vec{y})^2}{2\,\mathrm{d}t}}}{(2\pi\,\mathrm{d}t)^{n/2}} f(t, \|\vec{y}\|) \overset{\vec{y}=\vec{x}+\sqrt{\mathrm{d}t}\vec{u}}{=} \int_{\mathbb{R}^n} \frac{\mathrm{d}^n u\, e^{-\vec{u}^2/2}}{(2\pi)^{n/2}} f(t, \|\vec{x}+\sqrt{\mathrm{d}t}\vec{u}\|).$$

After expanding[3] f in $\sqrt{\mathrm{d}t}$ and solving the Gaussian integrals[4] we find

$$\mathbb{E}(f(t, B_{t+\mathrm{d}t})) = f(t,r) + \mathrm{d}t \left(\frac{n-1}{2r} \partial_r + \frac{1}{2} \partial_r^2 \right) f(t,r) + O(\mathrm{d}t^2),$$

from which we read of the infinitesimal generator $\mathcal{G}$ using (2.9) and from that the diffusion coefficient $\beta^2/2 \equiv 1/2$ and the drift $\alpha(t,r) = \frac{1}{2r}$. The SDE of the n-dimensional Bessel process is thus

$$\mathrm{d}B_t = \frac{n-1}{2B_t}\,\mathrm{d}t + \mathrm{d}W_t. \tag{2.10}$$

One could compute the SDEs of other scalar functions $F(\boldsymbol{W})$ (instead of $F = \|\cdot\|$) of the Wiener process using the same method. Like this, a general rule for the computation of differentials $\mathrm{d}\big(F(\boldsymbol{W})\big)$ can be obtained, called *Itō's formula* [15, Sec. 4.3.3], which will however not be needed here.

[3] $f(t, \|\vec{x}+\sqrt{\mathrm{d}t}\vec{u}\|) = f(t,r) + \sqrt{\mathrm{d}t}\frac{\vec{x}\cdot\vec{u}}{r}\partial_r f(t,r) + \frac{\mathrm{d}t}{2}\left(\frac{\vec{u}^2}{r} - \frac{(\vec{x}\cdot\vec{u})^2}{r^3}\right)\partial_r f(t,r) + \frac{\mathrm{d}t}{2}\frac{(\vec{x}\cdot\vec{u})^2}{r^2}\partial_r^2 f(t,r) + O(\mathrm{d}t^{\frac{3}{2}})$

[4] Use $(2\pi)^{-n/2}\int_{\mathbb{R}^n} \mathrm{d}^n u\, e^{-\vec{u}^2/2} u_i = 0$ and $(2\pi)^{-n/2}\int_{\mathbb{R}^n} \mathrm{d}^n u\, e^{-\vec{u}^2/2} u_i u_j = \delta_{i,j}$

PDF. The SDE (2.10) allows to naturally extend the class of Bessel processes to non-integer $n \in \mathbb{R}$. This level of generality will not be needed for our purposes. In fact, from now on we will only consider the case $n = 3$. The transition probability $p(t, x \mid s, y)\,\mathrm{d}x := \mathbb{P}(B_t \in \mathrm{d}x \mid B_s = y)$ is the same as that of a three-dimensional Brownian motion started at some point $\vec{y}$ of norm y and moving into a spherical shell $xS^2 \times [x, x + \mathrm{d}x]$ of radius x and thickness $\mathrm{d}x$ during a time $t - s$. Without loss of generality we set $s = 0$ and $\vec{y} = y\vec{e}_z$ and compute

$$\begin{aligned} p(t, x \mid 0, y) &= \int_{S^2} \frac{x^2\,\mathrm{d}\Omega(\vec{e})}{(2\pi t)^{3/2}} e^{-\|x\vec{e} - y\vec{e}_z\|^2/2t} = 2\pi x^2 \int_{-1}^{1} \frac{\mathrm{d}\cos\theta}{(2\pi t)^{3/2}} e^{-(x^2+y^2-2xy\cos\theta)/2t} \\ &= \sqrt{\frac{2}{\pi}} \frac{x}{y\sqrt{t}} \exp\left(-\frac{x^2+y^2}{2t}\right) \sinh\frac{xy}{t}. \end{aligned} \tag{2.11}$$

In the special case of $y \to 0$ we find the PDF of the Bessel process started at 0,

$$p(t, x) = \sqrt{\frac{2}{\pi}} \frac{x^2}{t^{3/2}} e^{-x^2/2t}. \tag{2.12}$$

Bessel bridge. For our purposes at least equally important as the Bessel process is the *Bessel bridge* which is obtained by conditioning the Bessel process to arrive at a fixed position X at a certain time T. Using Bayes' theorem and the Markov property $p(T, X \mid t, x; s, y) = p(T, X \mid t, x)$ for $t > s$ the transition probability can be related to (2.11) via

$$p(t, x \mid T, X; s, y) = \frac{p(T, X \mid t, x)\, p(t, x \mid s, y)}{p(T, X \mid s, y)}.$$

Of particular interest is the case $X = y = s = 0$ for which we find

$$p(t, x \mid T, 0; 0, 0) = \sqrt{\frac{2}{\pi}} \left(\frac{T}{t}\right)^{3/2} \frac{x^2}{(T-t)^{3/2}} \exp\left(-\frac{T}{t}\frac{x^2}{2(T-t)}\right). \tag{2.13}$$

Denoting this process by Q, a tedious[5] but straightforward evaluation of (2.9) with the kernel

$$\begin{aligned} &p(t + \mathrm{d}t, x \mid T, 0; t, y) \\ &\qquad = \sqrt{\frac{2}{\pi\,\mathrm{d}t}} \left(1 - \tfrac{\mathrm{d}t}{T-t}\right)^{-\frac{3}{2}} \frac{x}{y} \exp\left(\frac{y^2}{2(T-t)} - \frac{x^2}{2(T-t-\mathrm{d}t)} - \frac{x^2+y^2}{2\,\mathrm{d}t}\right) \sinh\frac{xy}{\mathrm{d}t} \end{aligned}$$

shows that its SDE is given by

$$\mathrm{d}Q_t = \left(\frac{1}{Q_t} - \frac{Q_t}{T-t}\right)\mathrm{d}t + \mathrm{d}W_t. \tag{2.14}$$

Notice how in the limit $T \to \infty$ this just recovers the ordinary Bessel process (2.10). Other important observations for later are the invariance of (2.13) under time reversal $t \to T - t$ and the quadratic prefactor in (2.13) that makes it unlikely to reach zero (of course apart from the final time).

[5] There is actually a more systematic way to study the conditioning of a diffusion process by applying something known as *Doob's h-transform* [16].

Chapter 3

Branching processes

In this chapter we introduce the main process of interest in this work, the one-dimensional *branching Brownian motion (BBM)*. In the first section the process is introduced and partial differential equations are established for the study of two relevant quantities, namely the average particle number density and the distribution of the "rightmost" particle. The importance of the latter is motivated by the fact that sparsely populated regions at the "tip" are most sensitive to stochastic effects. The movement of these rightmost particles is studied in the second section by relating global constraints on their trajectories to local drifts. The strategy that is used was in the given context first proposed in [7] as an algorithm for numerical simulations. We will turn to this use for a slightly different discrete model in the third section and provide a simple generalization.

3.1 Branching and the FKPP equation

In everything that follows, the word *particle* is to be understood in a broad sense as any entity characterized by a single real number. With that in mind, we provide a definition of the one-dimensional

Branching Brownian motion (BBM).

- Start with one particle *produced* at time $t = 0 \in \mathbb{R}_{\geq 0}$ and position $x = 0 \in \mathbb{R}$.
- At an exponentially distributed random time τ (i.e. $\mathbb{P}(\tau > t) = e^{-t}$) after a particle's production, it *decays* or *splits* into two, meaning that it vanishes while at the same point in spacetime two new independent particles are produced.
- Between production and decay, the movement of each particle realizes an independent Brownian motion.

Among the main quantities of interest in the BBM is the (random) particle number density $N(t, x)$, i.e. $N(t, x)\,\mathrm{d}x$ is the number of particles in $[x, x + \mathrm{d}x]$ at time t. While the final chapter is concerned with the PDF of $N(t, x)$ near the maximum of the BBM, we will now compute the average particle number density $\bar{n}(t, x) := \mathbb{E}(N(t, x))$ by relating it to the solution of a partial differential equation.

Average particle number density. To this end let $p(t, x) = (2\pi t)^{-1/2} e^{-x^2/2t}$ be again the PDF (2.4) of the Wiener process. By definition, there is only one particle in the beginning.

With probability $\mathrm{d}t$, it will split into two during the time interval $[0, \mathrm{d}t]$. In this case, each decay product will contribute on average $\bar{n}\big(t, x + O\big(\sqrt{t}\big)\big)$ to the density $\bar{n}(t + \mathrm{d}t, x)$ at time $t + \mathrm{d}t$. The correction $O\big(\sqrt{t}\big)$ coming from the particles' movements can be neglected in this case because in combination with the branching it only leads to terms of order $O(\mathrm{d}t^{3/2})$. On the other hand, if the particle does not branch during $[0, \mathrm{d}t]$, the particle can move a significant distance y (with probability density $p(\mathrm{d}t, y)$) from where it will contribute on average $\bar{n}(t, x - y)$ to $\bar{n}(t + \mathrm{d}t, x)$. In total, the expected particle number density at $t + \mathrm{d}t$ is given by

$$\begin{aligned}\bar{n}(t + \mathrm{d}t, x) &= \mathrm{d}t \cdot 2\bar{n}(t, x) + (1 - \mathrm{d}t) \cdot \int_{-\infty}^{\infty} \mathrm{d}y\, p(\mathrm{d}t, x - y)\bar{n}(t, y) + O(\mathrm{d}t^2), \\ &= \mathrm{d}t \cdot 2\bar{n}(t, x) + (1 - \mathrm{d}t) \cdot [\bar{n}(t, x) + \mathrm{d}t\, \mathcal{G}^\dagger \bar{n}(t, x)] + O(\mathrm{d}t^2) \\ &= \bar{n}(t, x) + [\bar{n}(t, x) + \mathcal{G}^\dagger \bar{n}(t, x)]\, \mathrm{d}t + O(\mathrm{d}t^2),\end{aligned}$$

where we used (2.7) for the second equality. Hence, $\partial_t \bar{n} = \mathcal{G}^\dagger \bar{n} + \bar{n}$ and since we are looking at Brownian motion, $\mathcal{G} = \mathcal{G}^\dagger = \frac{1}{2}\partial_x^2$, we have

$$\frac{\partial \bar{n}}{\partial t}(t, x) = \frac{1}{2}\frac{\partial^2 \bar{n}}{\partial x^2}(t, x) + \bar{n}(t, x). \tag{3.1}$$

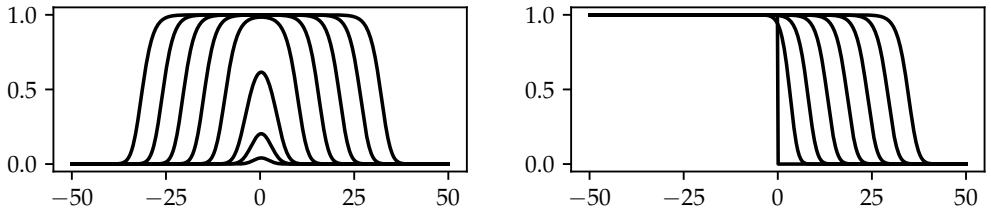
Given the initial condition $\bar{n}(t, x) = \delta(x)$, one can check by direct calculation that the solution is

$$\boxed{\bar{n}(t, x) = \frac{1}{\sqrt{2\pi t}} e^{t - x^2/2t}.} \tag{3.2}$$

In particular, the expected total number of particles is $\int_{\mathbb{R}} \mathrm{d}x\, \bar{n}(t, x) = e^t$. Furthermore, we see from (3.2) that also locally the particle density at any point x is growing exponentially with time (up to an asymptotically negligible correction factor $1/\sqrt{2\pi t}$). A practical consequence of this is that naive numerical implementations are unfeasible for large evolution times. A theoretically important implication is that once a region gets populated, its evolution will soon follow the deterministic equation (3.1) by the law of large numbers. Stochastic effects are therefore predominantly due to sparsely populated regions, which are almost surely located near the outermost particles. The region around the furthest advanced particle is called the *tip* and a full characterization of its statistical properties is the overarching goal towards which this work tries to contribute.

FKPP equation. To make these ideas more precise, let $X_1(t), X_2(t), \ldots, X_{N(t)}(t)$ be the positions of the $N(t)$ particles present at time t, where $N(t) \coloneqq \int_{\mathbb{R}} \mathrm{d}x\, N(t, x)$. Furthermore, let $M_t \coloneqq \max\{X_i(t) \mid 1 \le i \le N(t)\}$ denote the position of the (almost surely unique) *rightmost* or *maximal* particle and let $u(t, x) \coloneqq \mathbb{P}(M_t > x)$ be the probability that a particle ends up to the right of x at time t. We can find a partial differential equation for u in the same way as for $\bar{n}$. With probability $\mathrm{d}t$ the first splitting happens before time $\mathrm{d}t$. If this is the case, the (conditional) probability that an offspring at time $t + \mathrm{d}t$ is at a position $> x$ is given by the probability $2u(t, x)[1 - u(t, x)]\, \mathrm{d}t$ that exactly one of the decay products has offspring $> x$ at t plus the probability $u(t, x)^2\, \mathrm{d}t$ that both decay products have offspring $> x$ at t. On the other hand, if the first splitting occurs after $\mathrm{d}t$, the initial particle may move during the time $\mathrm{d}t$ to a position y, from which it produces with probability $u(t, y)$ an offspring to the right of x at time $t + \mathrm{d}t$. This probability is

$$\int_{-\infty}^{\infty} \mathrm{d}y\, p(\mathrm{d}t, y)u(t, x - y) \overset{(2.7)}{=} u(t, x) + \mathrm{d}t\mathcal{G}^\dagger u(t, x) + O(\mathrm{d}t^2).$$



(a) $u(t,x)$ if $u(0,x) = \varepsilon\delta(x)$ *(b) $u(t,x)$ if $u(0,x) = \mathbb{1}_{x<0}$*

Figure 3.1: Overlaid plots of the solution $u(t,x)$ of the FKPP equation with two different initial conditions. They were created by evolving the discretized equation (3.10) *rescaled with $\Delta t = 10^{-2}$ to times $t = 4, 8, \dots, 28$ (and $t = 6$ in (a) and $t = 0$ in (b)). The initial condition on the left is $u[0,x] = \varepsilon\delta_{x,0}$ with $\varepsilon = 10^{-2}$ (cf. notation in Sec. 3.3).*

In symbols instead of words, the last paragraph can be summarized as

$$\begin{aligned} u(t+\mathrm{d}t,x) &= \mathrm{d}t\cdot\big(2u(t,x)[1-u(t,x)]+u(t,x)^2\big) \\ &\quad + (1-\mathrm{d}t)\int_{-\infty}^{\infty}\mathrm{d}y\, p(\mathrm{d}t,y)u(t,x-y) + O(\mathrm{d}t^2) \\ &= [2u(t,x)-u(t,x)^2]\,\mathrm{d}t + (1-\mathrm{d}t)[u(t,x)+\mathrm{d}t\mathcal{G}^\dagger u(t,x)] + O(\mathrm{d}t^2) \\ &= u(t,x) + [\mathcal{G}^\dagger u(t,x) + u(t,x) - u(t,x)^2]\,\mathrm{d}t + O(\mathrm{d}t^2), \end{aligned}$$

hence $\partial_t u = \mathcal{G}^\dagger u + u - u^2$. In our case of Brownian motion, where $\mathcal{G}^\dagger = \frac{1}{2}\partial_x^2$, we obtain the *Fisher-Kolmogorov-Petrovsky-Piscunov* or *FKPP equation* [17, 18]

$$\boxed{\frac{\partial u}{\partial t}(t,x) = \frac{1}{2}\frac{\partial^2 u}{\partial x^2}(t,x) + u(t,x) - u(t,x)^2.} \tag{3.3}$$

In our setup it is supplemented by the initial condition $u(0,x) = \mathbb{1}_{x<0}$. Indeed, since the initial particle is located at 0, it is to the right of x if and only if $x < 0$.

To qualitatively understand (3.3), we first notice that there are two constant solutions $u \equiv 0, 1$. Expansion around zero, $u(t,x) = \varepsilon(t,x)$, turns (3.3) to first order in ε into the linear equation (3.1) with the solution being the convolution of $\varepsilon(0,\cdot)$ with (3.2). Since the solution is exponentially growing, the fixed point is unstable. The growth will continue until u is of order $O(1)$, which is when the quadratic term $-u^2$ will set in and ensure that u stays within $[0,1]$.

Expanding around one, $u(t,x) = 1 - \varepsilon(t,x)$, yields instead the equation $\partial_t\varepsilon = \frac{1}{2}\partial_x^2\varepsilon - \varepsilon + O(\varepsilon^2)$. The difference in sign as opposed to (3.1) leads to the solutions being exponentially suppressed with time, meaning that $u \equiv 1$ is a stable fixed point. An important consequence is that regions in which $u(t,x)$ is close to one act as barriers in the sense that fluctuations on one side of the barrier do not perceptibly affect the other side.

Travelling waves. Based on these brief considerations we can understand the evolution of u in Figure 3.1 for the initial condition $u(0,x) = \varepsilon\delta(x)$ as well as for the actual initial condition $u(0,x) = \mathbb{1}_{x<0}$. In particular, notice that in both cases the qualitative behaviour on the positive x-axis $\mathbb{R}_{>0}$ is identical as soon as the solutions have grown (close enough) to 1. After a short time a wave front forms that seems to uniformly move without changing shape. This observation is almost accurate and was made precise by Bramson [19, 20].

The details depend on the asymptotic behaviour of the initial condition $u(0,x)$ as $x \to \infty$. In our case the result is that

$$u(t, \xi + m_t) \overset{t\to\infty}{\longrightarrow} \omega(\xi),$$

where $\omega(\xi) \simeq \text{const.} \cdot \xi e^{-\sqrt{2}\xi}$ for large ξ and

$$m_t \equiv \mathbb{E}(M_t) = \sqrt{2}t - \frac{3}{2\sqrt{2}} \log t + \text{const.} + o(1). \tag{3.4}$$

Determination of the constant terms requires knowledge of the initial condition at finite x and are irrelevant for our purposes. The presence of a logarithmic correction to m_t can already be understood from the linear equation (3.1) by following the position x_t defined e.g. by $\bar{n}(t, x_t) = 1$. From (3.2) it follows immediately that

$$x_t = \sqrt{2}t\sqrt{1 - \frac{\log 2\pi t}{2t}} = \sqrt{2}t - \frac{1}{2\sqrt{2}} \log t + O(1) \quad \text{as } t \to \infty.$$

Interestingly, the non-linear term $-u^2$ changes the coefficient of the logarithmic correction according to (3.4), but does not affect the leading order. It is insightful to look at the shape of $\bar{n}$ around x_t, i.e.

$$\bar{n}(t, \xi + x_t) = \exp\left(-\frac{\xi^2}{2t} - \sqrt{2}\xi - \xi \cdot O(t^{-1} \log t)\right) \quad \text{as } t \to \infty. \tag{3.5}$$

Notice that for $\xi = O(\sqrt{t})$ the leading order behaviour is simply $\bar{n}(t, \xi + x_t) \simeq \text{const.} \cdot \exp(-\sqrt{2}\xi)$ with the constant being 1 unless $\xi \propto \sqrt{t}$. Since the transition from $\bar{n} = 1$ to $\bar{n} \approx 0$ happens exponentially fast in ξ it is reasonable to think that the non-linearity of (3.3) should not significantly change the picture for $\xi \gg 1$. Recalling that the main effect of the non-linear term is to confine u to 1 once it gets close enough, it is a fruitful attempt to model (3.3) by the linearized equation $\partial_t u = \frac{1}{2}\partial_x^2 u + u$ together with a moving absorptive boundary condition $u(t, m_t) = 1$, $\partial_x u(t, m_t) = 0$. In fact, this heuristic approach can be used (see e.g. [9, Sec. 4]) to derive (3.4) and the analogue of (3.5) for u. In the next section we will need actually a formula for $\tilde{u} = -\partial_x u$ instead of u, so we rather provide that (cf. [21, Appendix] and [8, eq. (5)]),

$$\tilde{u}(t, x) \simeq \text{const.} \cdot (x - m_t) e^{\sqrt{2}(x - m_t)} \exp\left(-\frac{(x - m_t)^2}{2t}\right), \tag{3.6}$$

which holds asymptotically for $t \gg 1$ within what we call the *scaling region* $1 \ll x - m_t = O(\sqrt{t})$.

3.2 Conditioning to rare events

As mentioned in the discussion of (3.2), the exponential growth of particle numbers makes a numerical generation of the BBM difficult for large times. However, since the main interest lies in the tip of the BBM, this problem can be mitigated by only keeping track of particles that have offspring near the tip. This can be achieved with the strategy presented in [7] which we will adopt now for its utility as a theoretical tool. In Section 3.3 we will review the usage of the same idea for simulations and give a slight generalization that was used in the numerical implementation explained in Section 4.2. The goal of the current section is to establish a relation between the BBM and the 3D Bessel bridge that was introduced in Section 2.3. This connection was suggested in the article [8] that forms the background for Chapter 4.

Conditioning (red particles). Let $\varepsilon > 0$ and fix $X \in \mathbb{R}$ and $T \in \mathbb{R}_{>0}$, where T is thought to be "large" and X comparable to m_T (although the results of this section hold independently of that). We call a particle *red* if at time T its (almost surely unique) rightmost offspring is in $I_\varepsilon := [X, X+\varepsilon]$. The probability that a particle at spacetime-position (t,x) has offspring to the right of X at time T is $U(x,t) := u(T-t, X-x)$ and the probability that it is red is $\varepsilon\tilde{U}(t,x) \equiv U(t,x) - U(t,x-\varepsilon) = \varepsilon\partial_x U(t,x) + O(\varepsilon^2)$. We will see that in all expressions below the leading epsilons cancel and one can simply replace $\tilde{U}$ by $\partial_x U$.

From now on we restrict ourselves to BBMs whose initial particle is red. This implies that there are red particles at time T and that they are all located in I_ε. Since each particle has a unique parent, one can trace back almost surely continuous red paths from the final particles to the initial one[1]. Non-red particles can be traced back to a branching point with a red particle. Overall, the system can be viewed as a branching process of red particles that can split into two red particles and additionally branch off non-red particles which realize a different BBM conditioned to have their rightmost offspring at time T to lie in $(-\infty, X)$.

Decay of red particles. Let $R : [0,T] \to \mathbb{R} : t \mapsto R_t$ be one of the red paths. We want to find the transition probability $p(t+dt, x \,|\, t, y)\,\mathrm{d}x := \mathbb{P}(R_{t+\mathrm{d}t} \in \mathrm{d}x \,|\, R_t = y)$. Since the redness of an unconstrained particle can only change through splitting during $\mathrm{d}t$, we first note that the chance of a red particle (■) decaying into another red and a non-red particle (□) is

$$\begin{aligned}\mathbb{P}(■ \to ■□) = \mathbb{P}(\bullet \to ■□ \,|\, \bullet \text{ is red}) &= \frac{\mathbb{P}(\bullet \to ■□;\ \bullet \text{ is red})}{\mathbb{P}(\bullet \text{ is red})} \\ &= \mathrm{d}t\,\frac{2\varepsilon\tilde{U}\big(t+\mathrm{d}t, y+O(\sqrt{\mathrm{d}t})\big)\big[1-U\big(t+\mathrm{d}t, y+O(\sqrt{\mathrm{d}t})\big)\big]}{\varepsilon\tilde{U}(t,y)}.\end{aligned}$$

Here, $\bullet$ is an unconstrained particle, the first $\mathrm{d}t$ in the second line gives the general splitting probability during $\mathrm{d}t$, the factor 2 accounts for the equivalence ■□ = □■ and the corrections $O(\sqrt{\mathrm{d}t})$ remind us that the space coordinate can slightly change during $\mathrm{d}t$. Expanding this in $\mathrm{d}t$ yields

$$\mathbb{P}(■ \to ■□) = \mathrm{d}t\, 2[1-U(t,y)] + O(\mathrm{d}t^2). \tag{3.7}$$

Red particle trajectory. From (3.7) we learn that the chances of splitting are effectively doubled in regions where $U(t,y) \ll 1$. Even then the probability is only of order $O(\mathrm{d}t)$, so when looking at the movement of the red particle over an infinitesimal time we can ignore the possibility of a branching during $\mathrm{d}t$. However, there is a first-order contribution from the possibility of branching off a non-red particle without moving. The chance of a particle moving from (t,y) to $(t+\mathrm{d}t, x)$ knowing that it is red, is schematically $\mathbb{P}(y \to x \,|\, \text{red}) = \mathbb{P}(y \to x;\ \text{red})/\mathbb{P}(\text{red})$ and precisely

$$p(t+\mathrm{d}t, x \,|\, t, y) = \frac{e^{-\frac{(x-y)^2}{2\,\mathrm{d}t}}}{\sqrt{2\pi\,\mathrm{d}t}}\left((1-\mathrm{d}t)\frac{\varepsilon\tilde{U}(t+\mathrm{d}t, x)}{\varepsilon\tilde{U}(t,y)} + \mathrm{d}t\, 2[1-U(t,y)]\right) + O(\mathrm{d}t^2).$$

For a more compact equation, the terms related to branching (those with prefactors $\mathrm{d}t$) shall be collected in the symbol

$$p_{\mathrm{br}} := \frac{e^{-\frac{(x-y)^2}{2\,\mathrm{d}t}}}{\sqrt{2\pi\,\mathrm{d}t}}\,\mathrm{d}t\left(-\frac{\tilde{U}(t+\mathrm{d}t, x)}{\tilde{U}(t,y)} + 2[1-U(t,y)]\right).$$

[1] Notice that a path that is red at $t=0$ and $t=T$ must also be entirely red in-between.

Due to the factor $\mathrm{d}t$, the kernel p_{br} has at order $O(\mathrm{d}t)$ only support at $x = y$. More precisely, convolving it with a test function $f(t,x)$ and substituting $x = y + \sqrt{\mathrm{d}t}u$ yields

$$\begin{aligned}\int_{-\infty}^{\infty} \mathrm{d}x\, p_{\mathrm{br}} f(t,x) &= \int_{-\infty}^{\infty} \frac{\mathrm{d}u\, e^{-u^2/2}}{\sqrt{2\pi}}\, \mathrm{d}t \left(-\tfrac{\tilde{U}(t,y)}{\tilde{U}(t,y)} + 2[1-U(t,y)]\right) f(t,y) + O(\mathrm{d}t^2) \\ &= \mathrm{d}t\, \big(1 - 2U(t,y)\big) f(t,y).\end{aligned}$$

Proceeding in the same fashion for the remaining term in p gives

$$\begin{aligned}&\int_{-\infty}^{\infty} \mathrm{d}x\, \big(p(t+\mathrm{d}t, x \,|\, t, y) - p_{\mathrm{br}}\big)\, f(t,x) \\ &= \int_{-\infty}^{\infty} \frac{\mathrm{d}u\, e^{-u^2/2}}{\sqrt{2\pi}} \frac{\tilde{U}(t+\mathrm{d}t, y+\sqrt{\mathrm{d}t}u)}{\tilde{U}(t,y)} f(t, y+\sqrt{\mathrm{d}t}u) + O(\mathrm{d}t^2) \\ &= \int_{-\infty}^{\infty} \frac{\mathrm{d}u\, e^{-u^2/2}}{\sqrt{2\pi}} \left(f + \frac{\mathrm{d}t}{\tilde{U}}\left(\partial_t \tilde{U} + \tfrac{1}{2}\partial_y^2 \tilde{U}\right) f + \frac{u^2\, \mathrm{d}t}{\tilde{U}} (\partial_y \tilde{U} \partial_y f + \tfrac{1}{2}\tilde{U}\partial_y^2 f)\right) + O(\mathrm{d}t^2) \\ &= f + \mathrm{d}t \left(\partial_t \log \tilde{U} + \frac{1}{2\tilde{U}}\partial_y^2 \tilde{U} + (\partial_y \log \tilde{U})\partial_y + \frac{1}{2}\partial_y^2\right) f + O(\mathrm{d}t^2),\end{aligned}$$

where the second equality comes from expanding in $\mathrm{d}t$ and discarding odd integrals and the third one from evaluating the Gaussian integrals. In the last two lines every function is evaluated at (t,y). By adding the contribution for p_{br} and comparing the result with (2.8) one can read off the infinitesimal generator of the process R. To simplify it, we finally commit to the limit $\varepsilon \to 0^+$, so that $\tilde{U}(t,x) = \partial_x U(t,x)$, allowing us to obtain the equation

$$\partial_t \log \tilde{U}(t,x) + \frac{1}{2\tilde{U}}\partial_x^2 \tilde{U}(t,x) + 1 - 2U(t,x) = 0$$

from differentiating the FKPP equation (3.3). Putting everything together, we find the infinitesimal generator for the red particles to be $\mathcal{G} = (\partial_x \log \tilde{U})\partial_x + \frac{1}{2}\partial_x^2$. From this we read of the diffusion coefficient $\frac{1}{2}$ and the drift $\alpha(t,x) = \partial_x \log \tilde{U}(t,x)$, which enables us to write an SDE and a Fokker-Planck equation for the (now unique) red particle trajectory,

$$\boxed{\mathrm{d}R_t = \partial_x \log \tilde{U}\, \mathrm{d}t + \mathrm{d}W_t, \qquad \partial_t p = \frac{1}{2}\partial_x^2 p - \partial_x(\partial_x \log \tilde{U}\; p).} \tag{3.8}$$

This result was already stated in words in [7] together with a brief computation of the drift. For that calculation (see [7, before eq. (11)]) the contribution p_{br} was neglected. Even though this is sufficient for finding the correct drift term, the derivation given above shows that p_{br} must be incorporated to arrive at the correct Fokker-Planck equation.

What can we say about $\partial_x \log \tilde{U}$? In the scaling region[2] $X - x - m_{T-t} = O(\sqrt{T-t})$ and for large T we can use (3.6) to find that

$$\partial_x \log \tilde{U} \simeq \frac{1}{x - X + m_{T-t}} + \sqrt{2} + \frac{X - x - m_{T-t}}{T-t}. \tag{3.9}$$

Approximating $m_{T-t} \simeq \sqrt{2}(T-t)$ (cf. (3.4)) and letting $Q_t = X - \sqrt{2}(T-t) - R_t$, we find that the SDE of Q_t is exactly that of the Bessel bridge (2.14) introduced earlier! This relation will be extensively used in Chapter 4 in the study of particle densities in the tip.

[2] cf. end of the previous section

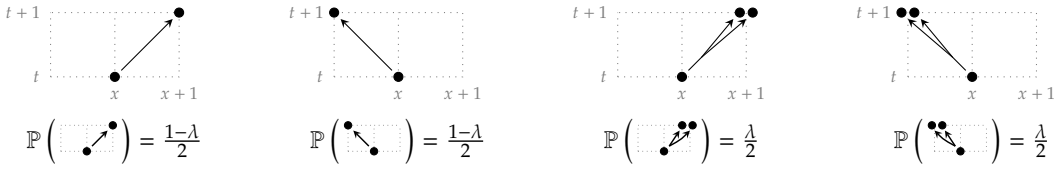


Figure 3.2: The possible actions of one particle in one time step in the considered discrete model.

3.3 A discrete model for the BBM

In this section we describe a discrete model proposed by Éric Brunet in private communication that was used for numerical simulations. Consider a two-dimensional spacetime lattice with spacings Δt and Δx. For a simplified notation we work on $\mathbb{Z}_{\geq 0} \times \mathbb{Z}$, i.e. set $\Delta t = 1$ and $\Delta x = 1$. To avoid confusion, we will use square brackets for functions on the lattice with unit spacing. Via dimensional analysis the results can be generalized to arbitrary spacings which will be useful when computing scaling limits. For instance, a function $f : \mathbb{R}_{\geq 0} \times \mathbb{R} \to \mathbb{R} : (t, x) \mapsto f(t, x)$ is described on our lattice by a function $f : \mathbb{Z}_{\geq 0} \times \mathbb{Z} \to \mathbb{R} : (t, x) \mapsto f[t, x]$. In the limit $\Delta t = \Delta x^2 \to 0$, the original function should asymptotically be recoverable as $f(t, x) \simeq f\left[\frac{t}{\Delta t}, \frac{x}{\Delta x}\right]$.

Discrete model.

- At time $t = 0$, there is exactly one particle which is located at $x = 0$.
- During each time step $t \to t + 1$, each particle independently moves with probability $\frac{1}{2}$ one step to the left, otherwise it moves one step to the right.
- After arriving at the new site, each particle independently and with probability $\lambda \in [0, 1]$ splits into two without moving.

The options are visualized in Figure 3.2 to introduce the notation we will use.

Discrete FKPP equation. As before, let M_t be the position of the rightmost particle and let $u[t, x] = \mathbb{P}(M_t \geq x)$. Then

$$\begin{aligned} u[t+1, x] = \frac{1-\lambda}{2}(u[t, x+1] + u[t, x-1]) + \frac{\lambda}{2}\left(u[t, x+1]^2 + u[t, x-1]^2\right) \\ + 2 \cdot \frac{\lambda}{2} u[t, x+1](1 - u[t, x+1]) + 2 \cdot \frac{\lambda}{2} u[t, x-1](1 - u[t, x-1]), \end{aligned}$$

where in the first line the linear terms correspond to movement without splitting and the non-linear ones to the case of splitting into two particles that both have offspring $\geq x$ at t. The second line describes decay to one particle with offspring $\geq x$ at t and one particle with no such offspring. In summary, we obtain the discrete FKPP equation

$$u[t+1, x] = \frac{1+\lambda}{2}\big(u[t, x+1] + u[t, x-1]\big) - \frac{\lambda}{2}\big(u[t, x+1]^2 + u[t, x-1]^2\big). \tag{3.10}$$

The initial condition is $u[0, x] = \mathbb{1}_{x \leq 0}$.

Continuum limit. To check that (3.3) can be recovered from (3.10) we go back to arbitrary spacings and replace e.g. $u[t + 1, x]$ by $u(t + \Delta t, x) = u(t, x) + \Delta t \partial_t u(t, x) + O(\Delta t^2)$. Next,

notice that we have to set $\lambda = \Delta t$ to obtain the same branching rate as before. Similarly, we will need $\Delta t \simeq \Delta x^2$ in order to obtain the correct diffusion constant (see below). At least, it should be assumed that Δt and Δx^2 are of the same order, so that the sub-leading terms in the left- and right-hand sides of (3.10) are equally relevant. Taking this into account, the expansion of (3.10) reads

$$u(t,x) + \Delta t \partial_t u(t,x) = \frac{1+\Delta t}{2}\big(2u(t,x) + \Delta x^2 \partial_x^2 u(t,x)\big) - \Delta t u(t,x)^2 + O(\Delta t^2),$$

or equivalently

$$\partial_t u(t,x) = \frac{\Delta x^2}{2\Delta t}\partial_x^2 u(t,x) + u(t,x) - u(t,x)^2 + O(\Delta t).$$

Here we see that indeed (3.3) is recovered if $\Delta t \simeq \Delta x^2$ for $\Delta t \to 0$.

Conditioning (red particles). Next, we reintroduce red particles and conditioning. Let $X \in \mathbb{Z}, T \in \mathbb{Z}_{>0}$ such that $X \equiv T \mod 2$. A particle is called *red* if its rightmost offspring at time T are located exactly at site X. If we write $U[t,x] := u[T-t, X-x]$, the probability that a given particle at (t,x) is red is $\tilde{U}[t,x] := U[t,x] - U[t,x-1]$. As in the previous section, we condition the initial particle to be red and use conditional probabilities to understand the movement of red particles. For instance, the chances of a red particle (■) moving to the right without splitting is

$$\mathbb{P}\Big(\quad\Big) \equiv \mathbb{P}\Big(\quad \Big|\ \bullet \text{ is red}\Big) = \frac{\mathbb{P}\Big(\quad ;\ \bullet \text{ is red}\Big)}{\mathbb{P}(\bullet \text{ is red})} = \frac{1-\lambda}{2}\frac{\tilde{U}[t+1,x+1]}{\tilde{U}[t,x]}.$$

Similarly, the probability for a red particle to jump to the right and split into a red (■) and a non-red particle (□) is

$$\mathbb{P}\Big(\quad\Big) = \frac{\lambda}{2}\frac{2\tilde{U}[t+1,x+1]\big(1-U[t+1,x+1]\big)}{\tilde{U}[t,x]}.$$

If we are just interested in following the red particles, we do not care whether a non-red particle was spawned. Thus, it makes sense to combine the two probabilities just given into $\mathbb{P}($ $) \equiv \mathbb{P}($ $) + \mathbb{P}($ $)$. Analogously, one can obtain the probabilities for the other possible actions of a red particle, the result being

$$\begin{aligned}
\mathbb{P}\Big(\quad\Big) &= \tfrac{\tilde{U}[t+1,x+1]}{\tilde{U}[t,x]}\left(\tfrac{1+\lambda}{2} - \lambda U[t+1,x+1]\right), \\
\mathbb{P}\Big(\quad\Big) &= \tfrac{\tilde{U}[t+1,x-1]}{\tilde{U}[t,x]}\left(\tfrac{1+\lambda}{2} - \lambda U[t+1,x-1]\right), \qquad (3.11)\\
\mathbb{P}\Big(\quad\Big) &= \tfrac{\lambda\tilde{U}[t+1,x+1]^2}{2\tilde{U}[t,x]} \quad \text{and} \quad \mathbb{P}\Big(\quad\Big) = \tfrac{\lambda\tilde{U}[t+1,x-1]^2}{2\tilde{U}[t,x]}.
\end{aligned}$$

Together with the discrete FKPP equation (3.10) for determining U, this is all the information needed to generate trees of red particles only, i.e. without the need to save any other particle. The price to pay is that the probabilities become spacetime dependent. One might also object that the algorithm only permits the generation of biased BRWs for which the final destination X of the eventually rightmost particle is known. However, the procedure can easily be extended to obtain unbiased branching walks by randomly drawing X in advance according to the distribution $u(T,x) = \mathbb{P}(M_T \geq x)$ that has to be computed anyway.

Generalization (yellow particles). It is possible to have a finer specification of non-red particles. For example in [7], particles ending in $[X-\Delta, X)$ were coloured in orange and in the same fashion as shown, conditional probabilities for movement and splitting were determined. We will now give a simple but novel generalization of this idea by tracking particles ending at specific distances from X. To this end, fix n pairwise distinct site numbers $\{x_i\}_{i=1}^n$ in $\mathbb{Z}_{>0}$ and let $X_i = X - x_i$. We call a particle *yellow* (△) if at time T it has offspring in $\{X_i \mid 1 \le i \le n\}$ but none in $[X, \infty)$. The probability $Y[t,x]$ that an arbitrarily picked particle found at (t,x) is yellow is then related via $Y(t,x) = y(T-t, X-x)$ to the probability $y[t,x]$ of an unconditioned BRW to have at time t offspring in $\{x_i \mid 1 \le i \le n\}$ but none in $[x, \infty)$. In the same way as for the FKPP equation (3.10), a recurrence relation can be obtained,

$$\begin{aligned} y[t+1,x] = \frac{1-\lambda}{2}\big(y[t,x+1]+y[t,x-1]\big) - \frac{\lambda}{2}\big(y[t,x+1]^2 + y[t,x-1]^2\big) \\ + 2\cdot\frac{\lambda}{2}y[t,x+1]\,(1-u[t,x+1]) + 2\cdot\frac{\lambda}{2}y[t,x-1]\,(1-u[t,x-1])\,. \end{aligned}$$

The linear terms in the first row give the contribution from moving without splitting. The second and third for account for decay into a yellow and a non-red particle (with symmetry factors of 2) and the non-linear term in the first row corrects for double counting because the non-red particles could also be yellow, in which case there should not be a symmetry factor. In short, the recursion reads

$$y[t+1,x] = \sum_{\sigma=\pm1}\left(\frac{1+\lambda}{2} - \lambda u[t,x+\sigma] - \frac{\lambda}{2}y[t,x+\sigma]\right) y[t,x+\sigma] \tag{3.12}$$

and is meant to be solved with the initial condition $y[0,x] = \sum_{i=1}^n \delta_{x_i,x}$. Computing the scaling limit ($\Delta x^2 = \Delta t \to 0$) in the same way as above for the FKPP equation gives

$$\partial_t y(t,x) = \frac{1}{2}\partial_x^2 y(t,x) + \big(1-2u(t,x)\big)y(t,x) - y(t,x)^2 \tag{3.13}$$

with initial condition $y(t,0) = \sum_{i=1}^n \delta(x-x_i)$. Notice that for $x - m_t \gg \sqrt{t}$, the coefficient $1-2u(t,x)$ approaches $+1$ and (3.13) just becomes the FKPP equation (3.3). On the other hand, for $m_t - x \gg \sqrt{t}$, the coefficient $1-2u(t,x)$ becomes -1 and the linear term will cause y to die off exponentially with time.

After having discussed the dynamics of y, we explain how this can be used to generate red and yellow particles. This works very analogous to the simpler ■□-colouring (3.11), except that we additionally colour certain non-red particles in yellow. The remaining particles (neither red nor yellow) shall be called *white* (○). The possible right-moving actions of red and yellow particles with their respective probabilities are

$$\begin{aligned} &\mathbb{P}\Big(\quad\Big) = \tfrac{\lambda\tilde U[t+1,x+1]^2}{2\tilde U[t,x]}, \qquad \mathbb{P}\Big(\quad\Big) = \tfrac{\lambda\tilde U[t+1,x+1]}{2\tilde U[t,x]}Y[t+1,x+1] \\ &\mathbb{P}\Big(\quad\Big) = \tfrac{\tilde U[t+1,x+1]}{\tilde U[t,x]}\big(\tfrac{1+\lambda}{2} - \lambda U[t+1,x+1] - \lambda Y[t+1,x+1]\big) \\ &\mathbb{P}\Big(\quad\Big) = \tfrac{Y[t+1,x+1]}{Y[t,x]}\big(\tfrac{1+\lambda}{2} - \lambda U[t+1,x+1] - \lambda Y[t+1,x+1]\big) \\ &\mathbb{P}\Big(\quad\Big) = \tfrac{\lambda Y[t+1,x+1]^2}{2Y[t,x]}. \end{aligned} \tag{3.14}$$

The probabilities of the left-moving actions are obtained by replacing $x+1$ by $x-1$ above.

Chapter 4

Particle Density in Stretched Tips

In this chapter we describe particle number densities in *stretched tips* (i.e. with $M_T - m_T \gg 1$) for large evolution times T based on the picture introduced in [8] that will be reviewed in the first section. We show that the model of the paper can be summarized in the compact form of equation (4.5). This integral and the final result (4.8) of [8] are then the subject of numerical investigations in Section 4.2. A few technical details on the software that was specifically developed for this purpose are given as well. Since the data and (4.8) are in strong tension, a new approach is proposed in the third section that not only models the data better, but also provides hints at the origins of the discrepancies. In the final section we give an outlook on possible avenues that might be taken from here.

4.1 Additional background

As mentioned already in the introduction, the stochastic character of the BBM manifests itself predominantly in the tip, which is why a proper understanding of the latter is desired. In [21] it was shown that the particle densities in the tip are at leading order well modelled by a deterministic evolution together with a cut-off that accounts for the fact that there is a rightmost particle (cf. (4.3) below). This is for example in contrast to the average number density (3.2) that for $t > 0$ is non-zero everywhere. Later in [22] this model was extended to account for the next-to-leading order correction. The relevant stochastic effect is captured in a single effective particle that moves significantly ahead (by a distance $\delta \sim O(\log t)$) of the average position m_t (where the cut-off is placed), before it causes a shower that is again modelled in the same way. This picture correctly predicts corrections in m_t beyond the logarithmic order (which is already a consequence of the cut-off) shown in (3.4). A crucial assumption is that the particle density in the range $(m_t, m_t + \delta)$ is indeed significantly lower than those at m_t. The desire to properly justify this is the origin of our motivation for studying stretched tips. Meanwhile, these developments have also caused interest among mathematicians who hope that such studies can be relevant to identify the point measure that asymptotically describes the BBM at infinite time (cf. [23, 24]).

Notable progress in this direction has been made in [25]. Therein, a comparison of typical (n_{typ}) and average ($\bar{n} = \mathbb{E}(N)$) particle number densities at a distance Δ of the maximal particle at a given time T is carried out, for the unconstrained BBM as well as for the case in which the maximal particle is conditioned to arrive at a position $M_T = X$, where X is to be chosen such that $1 \ll X - m_T = O(\sqrt{T})$ as $T \to \infty$. In the unconstrained case, the result is

$$\bar{n}(T, M_T - \Delta) \propto n_{\mathrm{typ}}(T, M_T - \Delta) \propto \Delta e^{\sqrt{2}\Delta}, \tag{4.1}$$

where the equivalence relation "$\propto$" indicates that the ratio of its left- and right-hand sides tends to a finite constant as $\Delta \to \infty$. For the conditioned BBM with stretched tip one has instead

$$\bar{n}(T, M_T - \Delta \,|\, M_T = X) \propto e^{\sqrt{2}\Delta}, \qquad n_{\text{typ}}(T, M_T - \Delta \,|\, M_T = X) \propto e^{\sqrt{2}\Delta - \zeta\Delta^{2/3}}. \tag{4.2}$$

The constant ζ could not be determined in [25] but within the picture of [8] and Section 4.3, predictions can be made. More important than the precise value is however the observation that in stretched tips typical densities are significantly below the average ones. A more precise study of typical realizations in these stretched tips is the goal of this chapter and of the article [8] which we review in the remainder of this section.

Setup. The description of the particle number density N in the tip is carried out in [8] from the point of view introduced in Section 3.2, in which the system consists of a single red particle moving according to (3.8) and branching off non-red particles on its way with the rate given in (3.7). Each of these non-red particles develops into an exponentially growing shower after a sufficient amount of time that contributes an amount $N^{(i)}$ to the number density $N(T, X - \Delta)$ at a distance Δ from $M_T = X$. Two simplifications are then made in [8] that are reasonable in the case where $T, \Delta \gg 1$ and X is in the scaling region, i.e. of distance $O(\sqrt{T})$ from m_T (as given in (3.4)). Firstly, the conditioning of $N^{(i)}$ to only contain particles in $(-\infty, X)$ at time T is lifted because for large Δ and $X - m_T \geq \Delta$ there are anyway only few particles with offspring to the right of X — at least in comparison with the exponential growth of particle numbers at $X - \Delta$. For the same reason one might neglect in a second step the fluctuations in $N^{(i)}$ between different i and just use a common typical number density n_{typ} (see below) for all showers. Those are produced at a rate (3.7). In summary, these approximations amount to the statement [8, eq. (13)]

$$N(T, X - \Delta \,|\, \{R_T = X, R_{t<T} = r_t\}) \simeq \int_0^T \mathbb{P}(\blacksquare \to \blacksquare\square \text{ during } \mathrm{d}t)\, n_{\text{typ}}(T - t, X - \Delta - r_t),$$

where R_t is the trajectory of the red particle as in the previous chapter. Because of the quadratic prefactor in the Bessel bridge PDF (2.13), r_t is usually smaller than $X - m_{T-t}$ so that (3.7) lets us further simplify $\mathbb{P}(\blacksquare \to \blacksquare\square \text{ during } \mathrm{d}t) \approx 2\,\mathrm{d}t$. Hence, the branching along the red path becomes a Poisson point process of constant intensity 2 in time[1] and

$$N(T, X - \Delta \,|\, \{R_T = X, R_{t<T} = r_t\}) \simeq 2 \int_0^T \mathrm{d}t\, n_{\text{typ}}(T - t, X - \Delta - r_t),$$

cf. [8, eq. (20)]. A formula for n_{typ} was derived in [22, eq. (6), Sec. 4.1] based on an idea introduced in [21] to use a cut-off to model the stochastic behaviour of the tip in a BBM. Adapted to the considered problem it reads [8, eq. (23)]

$$n_{\text{typ}}(T - t, y - x) \simeq \text{const} \cdot (x + m_{T-t} - y) e^{\sqrt{2}(x + m_{T-t} - y)} \exp\left(-\frac{(x + m_{T-t} - y)^2}{2(T - t)}\right) \mathbb{1}_{\{x + m_{T-t} \geq y\}}. \tag{4.3}$$

Setting $y = m_{T-t} + X - \Delta$ and $x = r_t$, simplifying $m_{T-t} \simeq \sqrt{2}(T - t)$ and introducing $Q_t = X - 2\sqrt{2}(T - t) - r_t$ which to good approximation (cf. (3.9)) is a 3D Bessel bridge

[1] A Poisson point process of intensity $I(x)$ in $x \in \mathbb{R}$ randomly assigns points to the real line such that in each infinitesimal interval $\mathrm{d}x$ the probability that it contains such a point is $I(x)\,\mathrm{d}x$.

(with SDE (2.14)), we arrive at an expression for N in terms of a functional of a sample process Q_t ([8, eq. (24)]),

$$N(T, X - \Delta \,|\, \{Q_t\}) \simeq \text{const.} \cdot \int_0^T (\Delta - Q_t) e^{\sqrt{2}(\Delta - Q_t)} \exp\left(-\frac{(\Delta - Q_t)^2}{2\sqrt{T-t}}\right) \mathbb{1}_{\{\Delta \geq Q_t\}} \frac{\mathrm{d}t}{(T-t)^{3/2}}. \tag{4.4}$$

Simplification of the integral. In the final part of the analysis in [8], a saddle point approximation is applied to evaluate the above integral for large Δ. For this, the exponent gets first simplified by neglecting Q_t in the quadratic term only,

$$\sqrt{2}(\Delta - Q_t) - \frac{(\Delta - Q_t)^2}{2(T-t)} \approx -\sqrt{2}Q_t - \frac{\Delta^2}{2(T-t)}.$$

This is reasonable because the cut-off ensures $\Delta \geq Q_t$ and in the case $\Delta \approx Q_t$ the additional factor $(T-t)^{-1}$ in the quadratic term (which is comparable to Q_t^{-1} according to the PDF (2.13)) suppresses it with respect to the linear one. The factors in the integrand (4.4) that are not in the exponent turn out to be sub-leading for large Δ and are eventually neglected in [8]. Thus, after substituting $t \to T - t$, which leaves the Bessel bridge Q invariant, the computation of N boils down to the evaluation of

$$N e^{-\sqrt{2}\Delta} \simeq \int_0^T \mathrm{d}t \, e^{-\sqrt{2}Q_t - \Delta^2/2t} \mathbb{1}_{\Delta \geq Q_t}.$$

In the limit $T \to \infty$, the Bessel bridge behaves like a Bessel process B_t started at zero which typically has a magnitude $\sim \sqrt{t}$. Therefore, when the cut-off sets in, the term $-\sqrt{2}Q_t$ makes the exponential small, so we might ignore the cut-off as well. We arrive thus at the even simpler integral

$$N e^{-\sqrt{2}\Delta} \simeq \int_0^\infty \mathrm{d}t \, e^{-\sqrt{2}B_t - \Delta^2/2t}. \tag{4.5}$$

For small t, the $-\Delta^2/2t$ will explode and pull the integrand to zero while for large t the growth of $B_t \sim \sqrt{t}$ will dominate and let the integrand vanish again. In between, the almost surely continuous function $S(t) \coloneqq \sqrt{2}B_t + \Delta^2/2t$ will almost surely have a minimum at some time t_s.

Saddle point approximation. For a saddle point approximation, derivatives of S need to be taken. We therefore rescale $B_t = \Xi_t \sqrt{t}$, making the distribution of Ξ_t independent of time,

$$\mathbb{P}(\Xi_t \in \mathrm{d}\xi) = \sqrt{\frac{2}{\pi}} \xi^2 e^{-\xi^2/2} \, \mathrm{d}\xi. \tag{4.6}$$

To differentiate S, it is necessary to assume but difficult to justify that the sample process ξ_t does not vary significantly around t_s. Instead of repeating the discussion in [8] here, we will just take it for granted now and see how it turns out after looking at the numerical results in the next section. With that in mind, the saddle point can be found by setting $\dot{S}(t_s) \equiv \xi_{t_s}/\sqrt{2t_s} - \Delta^2/2t_s^2 = 0$, which is solved by

$$t_s = \frac{\Delta^{4/3}}{2^{1/3} \xi_{t_s}^{2/3}}. \tag{4.7}$$

One can further compute $S(t_s) = 3(\xi_{t_s}\Delta/2)^{2/3}$ and $\ddot{S}(t_s) \equiv -\frac{3}{4}(\xi_{t_s}/\Delta)^2$, which we use to rewrite the integral (4.5) as

$$Ne^{-\sqrt{2}\Delta} \simeq \int_0^\infty \mathrm{d}t\, \exp\left(-\frac{3}{2^{2/3}}(\xi_{t_s}\Delta)^{2/3} - \frac{3}{4}\left(\frac{\xi_{t_s}}{\Delta}\right)^2 (t-t_s)^2 + O\big((t-t_s)^3\big)\right).$$

The result of this integral is a prefactor times $\exp(-S(t_s))$. Ignoring the prefactor (in line with previous omissions) and introducing $L = -\log(Ne^{-\sqrt{2}\Delta})\Delta^{-2/3}$, we obtain the relation $L \simeq 3(\xi_{t_s}/2)^{2/3}$ which enables us to evaluate the PDF of L from (4.6), yielding (cf. [8, eq. (28)])

$$\mathbb{P}(L \in \mathrm{d}l) \simeq \frac{4}{27}\sqrt{\frac{2}{3\pi}}\, l^{7/2} e^{-2l^3/27}\, \mathrm{d}l. \tag{4.8}$$

Despite the numerous approximations and simplifications that were necessary to arrive at this rather simple formula, the next section will show that it nonetheless captures some features of the true distribution, most notably the non-trivial $\Delta^{2/3}$ scaling. Moreover, a prediction for the constant ζ in (4.2) can be made now as the median of the distribution (4.8), i.e. (cf. [8, eq. (29)])

$$\frac{\Gamma(\frac{3}{2}, \frac{2}{27}\zeta^3)}{\Gamma(\frac{3}{2})} = \frac{1}{2} \quad \Rightarrow \quad \zeta \approx 2.5182, \tag{4.9}$$

with the incomplete gamma function $\Gamma(s,x) \coloneqq \int_x^\infty \mathrm{d}t\, t^{s-1}e^{-t}$.

4.2 Numerical studies

Here we present numerical results comparing equations (4.5) and (4.8) with a simulation of the full (discretized) model presented in Section 3.3. For this purpose a C++ library[2] was written from scratch. We shall outline a few of the technical details in the following paragraph, but the results can be understood without them.

Comments on the simulations. Great care has to be taken in the implementation of $u[t,x]$ (respectively y, $\tilde{U}$ etc.) and its evolution via the discrete FKPP equation (3.10). The challenges are caused by the fact that the evolution equation for $U[t,x] = u[T-t, X-x]$ (which is needed in (3.14) rather than y) is backward in time compared to the evolution of particles. Therefore, U has to be computed in advance up to the final time T. Trying to save all values of u that are not exactly 1 or 0 would imply memory requirements growing like $O(T^2)$ for large T. As seen in Figure 3.1b and the surrounding discussion, the transition from 1 to 0 effectively happens over a small distance of order $O(1)$, yet it turns out that a number of sites of order $O(\sqrt{T})$ must be kept, making the total memory needed go like $O(T^{3/2})$. Indeed, it is important that sufficiently many sites (we used $\sim 4\sqrt{T}$) are taken into account to cover the whole scaling region until the Gaussian factor in (3.5) sets in. The reason is that the exponential decrease $\sim e^{-\sqrt{2}\xi}$ in the same equation competes with the exponential growth of small fluctuations visualized in Figure 3.1a. More concretely, it was shown in [21] that cutting off the scaling region leads to a change in the velocity of the travelling wave. In contrast, in the region where u is close to 1 only a few sites must be saved because fluctuations are damped by the non-linear term in the

[2]The code can be found at `https://github.com/malphax/brw`, at the time of writing there is however no documentation.

FKPP equation (3.3) (cf. discussion thereafter). The memory requirements can be further reduced to order $O(T)$ by saving $u[t, \cdot]$ only at certain checkpoints t (e.g. after every $\sqrt{t}$-th step). Then, for the particle generation between two checkpoints one reevaluates u just between the checkpoints and discards the values once the later checkpoint has been reached. Performance-wise, the computation of u was optimized sufficiently so that the time usage is irrelevant compared to that needed for the particle propagation.

Even though the algorithm described in 3.3 allows to only keep track of the relevant particles and thereby reduces the complexity from exponential to polynomial, the generation of random numbers is relatively slow and limits the maximum final time. The only data set that is presented gives the evolution over 500000 time steps of size $\Delta t = 10^{-2}$ and with splitting rate $\lambda = 10^{-2}$. The final position was chosen so that $X - m_T = 100$. The only approximation in place was that particle numbers were evolved deterministically instead of stochastically when the expected particle number at the target site is larger than 1000. For instance, if there are n yellow particles on a given site and p is the probability for a given yellow particle on this site to move to the left without splitting, then usually a binomially distributed random number with these parameters n and p is drawn to decide how many particles move to the left. If however $n \cdot p \geq 1000$, then just np particles are moved. This is justified by the realization that for sufficiently large times the statistics effectively depend only on the stochastic nature of very few particles in the tip[3]. Despite this simplification and the usage of more than a hundred CPU cores, only about 150 samples could be generated per day. The bottleneck seems to lie in the generation of random numbers (done with C++'s Mersenne-Twister engine `std::mt19337`), of which multiple billions are needed for each sample.

Results. Figure 4.1 compares the simulation data to the result of the saddle point approximation in (4.8). There seem to be two main discrepancies. Firstly, a shift of about 1.5 separates both distributions. Secondly, the true distribution appears less symmetric than formula (4.8). While the tails on the right-hand side are comparable, the data looks cut-off near zero. On a positive note, (4.8) apparently does predict the correct scaling $\Delta^{2/3}$ as clearly reflected in the data. Without these scaling factors the lines for different values of Δ would be well separated as demonstrated in Figure 4.2. Even apart from the mysterious shift, one should probably not read too much into the similarities between (4.8) and the data, because we also numerically evaluated (4.5) and its measured PDF seems to match neither the real data nor the saddle point approximation, see Figure 4.3. For the given value of Δ one might speculate that the assumption of constant ξ in the derivation of (4.8) was too rough. For this reason, we will attempt a more direct approach to minimizing $S(t)$ in the next section. This will lead to a one-parameter family (4.13) of approximative PDFs for L. Interestingly, Figure 4.3 suggests that a proper choice of r may also serve as a model for the simulated data. The discussion of this shall be postponed to the end of the next section. To finish off the current section, we want to emphasize that the discrepancy to the simulation data does not rule out the usefulness of the approximations (4.5) and (4.8). It might just be that T and Δ have not been chosen sufficiently large. For this reason it would be desirable to increase the performance of the code, probably through a higher degree of parallelization.

[3]This claim has been investigated for instance in [26].

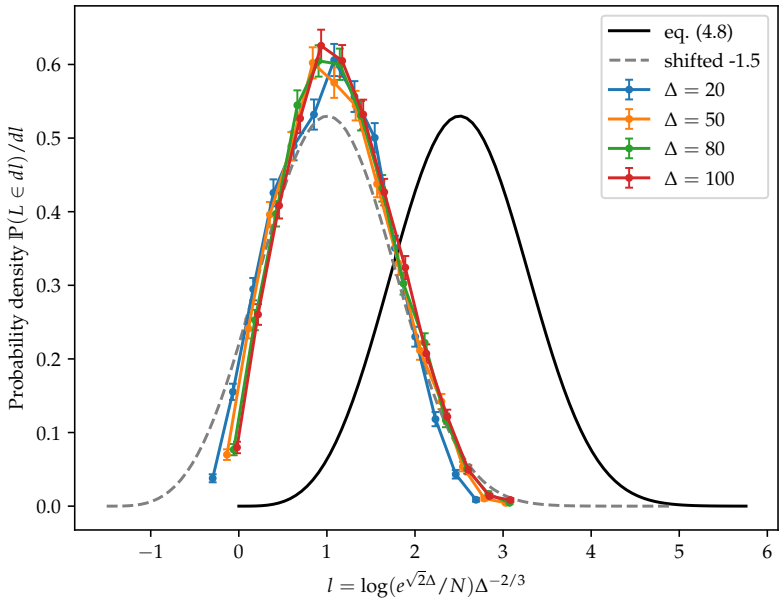


Figure 4.1: Comparison of (4.8) *(solid black line) with a Monte Carlo (MC) simulation of the full model (Sec. 3.3) for* $T = 5000$ *and* $\Delta t = 10^{-2}$*. Each of the coloured lines corresponds to a different choice of* Δ *(in the units of the continuous model). Although the shape seems to broadly match, there is evidently a shift between the data and the formula* (4.8)*. Horizontally translating the formula by 1.5 to the left (dashed gray line), we see furthermore that the tails on the right-hand side look more or less aligned, while on the left-hand side the real data falls of faster at zero.*

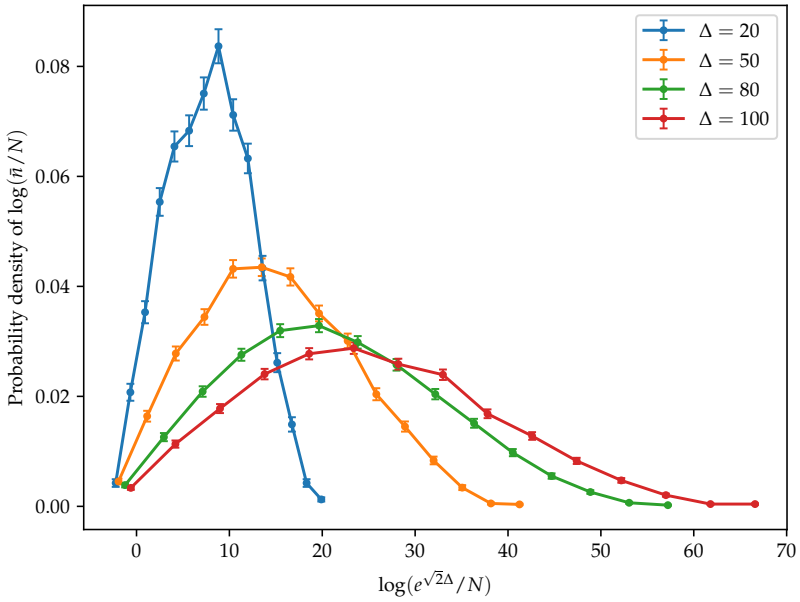


Figure 4.2: The same data as in Figure 4.1 is presented but without scaling away the Δ*-dependence. This should give appreciation of the fact that the data and* (4.8) *indeed share the correct* $\Delta^{2/3}$ *scaling.*

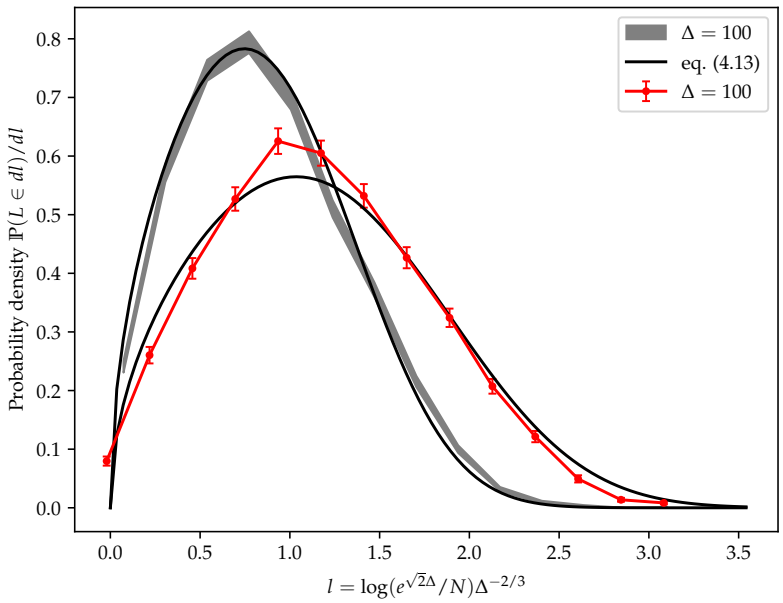


Figure 4.3: Numerical evaluation of the simplified integral (4.5) *(gray band) compared to the data of the full simulation (red line with statistical error bars). The approximate PDF* (4.13) *derived in the next section is displayed in black for* $r = 0.2$ *(higher curve) and* $r = 0.7$ *(lower curve).*

4.3 Analysis

Our starting point is (4.5) and as before we try to minimize $S(t) = \sqrt{2}B_t + \Delta^2/2t$, but this time we will not approximate the Bessel process B_t by a constant times $\sqrt{t}$. Instead, we will directly search for an (approximate) distribution of the minimum by relating it to a hitting probability. The idea is that for any process X_t started at $x_0 > m$ for some number m, the minimum $\min_{t\in[0,T]} X_t$ is less than or equal to m if and only if X_t hits the boundary m during the interval $[0, T]$. Thus, for each $m > 0$ we need to find the probability $\mathbb{P}(\text{hit} \,|\, m, \Delta)$ that the Bessel process B_t *hits* (i.e. intersects the graph of) the function $g(t) := -\frac{1}{2\sqrt{2}t}\Delta^2 + m$. Since $B_t > 0$ for $t > 0$ almost surely (cf. (2.12)), the boundary function can also be replaced by its positive part $g_+(t) = g(t)\mathbb{1}_{\{g(t)\geq 0\}}(t)$. Figure 4.4a illustrates the problem.

Approximation with a single step. Because the time-dependence is difficult to handle, we start by replacing g_+ with a step function $g_r^{(1)}(t) = m\Theta(t - t_r)$ consisting of a single step of height m at position t_r (cf. Figure 4.4b), parametrized by $r \in [0, 1)$ such that $g(t_r) = rm$, i.e.

$$t_r = \frac{\Delta^2}{2\sqrt{2}(1-r)m}. \tag{4.10}$$

For $a > b \geq 0$, let $\tau_{ab} = \inf\{t > 0 \,|\, B_t = b \text{ given } B_0 = a\}$ be the random first hitting time of the (immobile) boundary b by a Bessel process launched at a, in particular $\tau_{ab} = \infty$ if b is never reached. Denoting as earlier by $p(t, x)$ the PDF (2.12) of the Bessel process (started

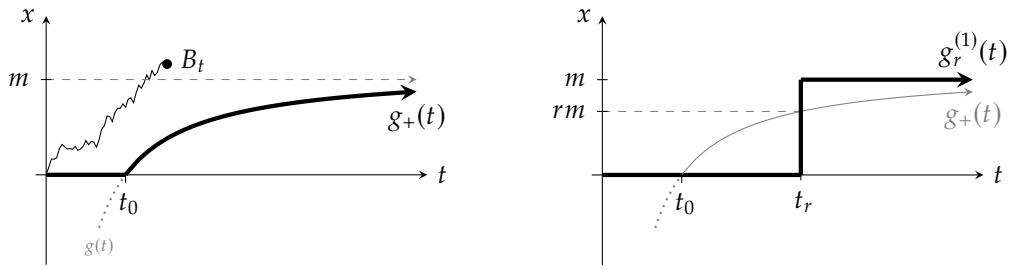


Figure 4.4: (a) What is the probability that a Bessel process B_t launched at zero hits the moving boundary $g(t) = m - \Delta^2/2t$? The answer would provide the distribution of the minimum of $S(t)$. (b) The function $g(t)$ is approximated by the step function $g_r^{(1)}(t) = m\Theta(t - t_r)$ where $r \in [0, 1)$ is a parameter and t_r is defined such that $g(t_r) = rm$ or explicitly via (4.10).

at zero), the replacement of g by $g_r^{(1)}$ amounts to the approximation

$$\mathbb{P}(\text{no hit} \,|\, m, \Delta) \approx \mathbb{P}^{(1)}(\text{no hit} \,|\, m, \Delta) \equiv \int_m^\infty \mathrm{d}x \, p(t_r, x)\mathbb{P}(\tau_{xm} = \infty). \tag{4.11}$$

The reason is that B_t cannot hit zero anyway, so during $[0, t_r]$ the only requirement is to end up at a position $x > m$. The probability to arrive there is then to be multiplied with the probability to not hit the boundary m starting from x. The first hitting time distribution of the considered three-dimensional Bessel process can be given in terms of elementary functions (see [27]),

$$\mathbb{P}(\tau_{ab} \in \mathrm{d}t) = \frac{b}{a}\frac{a-b}{\sqrt{2\pi t^3}} \exp\left(-\frac{(b-a)^2}{2t}\right) \mathrm{d}t. \tag{4.12}$$

To prove this, one can start by showing that for fixed b the Laplace transform of (4.12)

$$u(x) := \mathbb{E}(e^{-\lambda\tau_{ab}}) \equiv \int_0^\infty \mathbb{P}(\tau_{xb} \in \mathrm{d}t)e^{-\lambda t}$$

is an eigenfunction of the infinitesimal generator $\mathcal{G} = \frac{1}{2}\frac{\mathrm{d}^2}{\mathrm{d}x^2} + \frac{1}{x}\frac{\mathrm{d}}{\mathrm{d}x}$ of the Bessel process (cf. Sec. 2.3), i.e. $\mathcal{G}u = \lambda u$. For a Bessel process of arbitrary dimension this can be brought to the form of a modified Bessel equation. For the case at hand it is easier to just check that the unique solution satisfying $u(b) = 0$ and $u(+\infty) = 0$ is $u(x) = \frac{b}{x}e^{\sqrt{2\lambda}(b-x)}$. The inverse Laplace transform of this function provides the PDF (4.12). Integration of the latter from $t = 0$ to ∞ yields then simple formula $\mathbb{P}(\tau_{xm} = \infty) = 1 - x/m$. Substituting $x = u\sqrt{t_r}$ in (4.11), we rewrite the latter as

$$\mathbb{P}^{(1)}(\text{no hit} \,|\, m, \Delta) = \sqrt{\frac{2}{\pi}} \int_{m/\sqrt{t_t}}^\infty \mathrm{d}u \, u^2 e^{-u^2/2} \left(1 - \frac{m}{u\sqrt{t_r}}\right).$$

Writing $M := \min_{t>0}\{B_t + \Delta^2/(2\sqrt{2}t)\}$, we use (4.10) and the substitution $z = u^2/2$ to compute

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}m}\mathbb{P}(M < m) = \frac{\mathrm{d}}{\mathrm{d}m}[1 - \mathbb{P}(\text{no hit} \mid m, \Delta)] &\approx \sqrt{\frac{2}{\pi}} \int_{m/\sqrt{t_r}}^{\infty} \mathrm{d}u\, \frac{3}{2}\frac{u}{\sqrt{t_r}} e^{-u^2/2} \\ &= \frac{3}{\sqrt{2\pi t_r}} \int_{m^2/2t_r}^{\infty} \mathrm{d}z\, e^{-z} = \frac{3}{\sqrt{2\pi t_r}} e^{-m^2/2t_r} \\ &= \frac{3}{\Delta}\sqrt{\frac{\sqrt{2}m(1-r)}{\pi}} \exp\left(-\frac{\sqrt{2}m^3}{\Delta^2}(1-r)\right),\end{aligned}$$

from which we retrieve an approximate formula for the distribution of M. Realizing that $L \equiv -\log(Ne^{-\sqrt{2}\Delta})\Delta^{-2/3} \approx \sqrt{2}M\Delta^{-2/3}$, we find the pendant of (4.8),

$$\mathbb{P}^{(1)}(L \in \mathrm{d}l) = 3\sqrt{\frac{l(1-r)}{2\pi}} e^{-(1-r)l^3/2}\, \mathrm{d}l. \tag{4.13}$$

In analogy to (4.9) we can make a prediction for ζ_r for each choice of r by solving the equation

$$\frac{\Gamma(\frac{1}{2}, \frac{1}{2}(1-r)\zeta_r^3)}{\Gamma(\frac{1}{2})} = \frac{1}{2}.$$

Numerically we find $\zeta_0 = 0.7691\cdots$ from which the other ζ_r can be invoked in virtue of

$$\zeta_r = (1-r)^{-\frac{1}{3}}\zeta_0.$$

4.4 Discussion and outlook

We may now have a second look at Figure 4.3. Since the single step is a pretty rough approximation of the function of interest, it is somewhat surprising that (4.13) and (4.5) are in such good agreement. A heuristic explanation comes from the fact that B_t grows roughly like $\sqrt{t}$. Therefore, (4.10) leads to expect that a step is a good approximation except when $B_{t_r} \sim \sqrt{t_r} \approx m$, i.e. $m \approx \Delta^{2/3}(1-r)^{-1/3}\sqrt{2}$ or $l \approx 1$. This would explain why the tails for large l and the cut-off at 0 are generally well approached. Furthermore, notice how this basic argument already predicts the correct $\Delta^{2/3}$-scaling, which appears to be fundamental to the problem. In (4.8) and (4.13) it is even manifest in that Δ does not appear in the PDFs. In the case of (4.13) a rescaling of Δ is equivalent to a variation of r. Since Figure 4.3 surprisingly indicates that an appropriate choice of this parameter permits to model the full simulation data using (4.13), it would be important to see if this still holds at substantially higher values of T and Δ and if the best-fit parameters r for the full model and the integral 4.5 will approach each other. This leads us to suggest the following possible continuations to the present work.

- Given the discrepancies between the data and the predictions it would be important to identify exactly by which approximation(s) in the derivation of (4.5) they are caused.
- Highly related is whether those differences vanish as T and Δ increase. Testing this in the full model will require significant effort to increase the performance of the code. One promising approach to do so is by making better use of parallelization. The

code already supports multithreading, so that different samples can be processed on different cores with shared usage of the solution to the discrete FKPP equation (3.10). However, since the respective evolutions of u and the particle numbers at a given site and time are independent of the computations happening at other sites at the same time, the problem would be ideally suited for general purpose GPU programming. In fact, we already could solve (3.10) on the GPU using the `CUDA` based Thrust library. For the performance-wise much more relevant particle propagation, the implementation is more challenging and even if successful, it is unclear if this will suffice to make significantly larger values of T accessible.

- A different attempt to increase the range in T is offered by a backward evolution, i.e. by generating first only the computationally cheap red trajectory (via (3.11)) and then tracing back its way from (T, X) to $(0, 0)$ while branching off yellow particles. The advantage is that one starts already at the final time, so that the most significant contributions to the particle density at $X - \Delta$ are made sooner, so that once it seems to converge, one might stop the generation early. The risk is however that untypically impactful particles that were branched off early will be neglected.

- The numerical treatment of (4.5) should be much easier to extend to larger values of Δ. The only obstacle is that $\exp(-\Delta^{2/3})$ soon becomes too tiny for the usual data types. In the worst case this can however be remedied by using a library for arbitrary precision floating point numbers.

- On the analytical side, it will be interesting to see if the approximation (4.13) can be improved, for example by adding a second step (and then a third etc.). First tries lead us however to believe that the resulting expressions will quickly become very unhandy. A different attempt is offered by the *Feynman-Kac formula* [28] which allows to relate the distribution of the integral (4.5) to a linear partial differential equation with space and time dependent coefficients.

Bibliography

[1] N. Ikeda, M. Nagasawa and S. Watanabe. 'Branching Markov processes II'. In: *Journal of Mathematics of Kyoto University* 8.3 (1968), pp. 365–410. DOI: 10.1215/kjm/1250524059.

[2] A. Bovier. *Gaussian Processes on Trees: From Spin Glasses to Branching Brownian Motion*. Cambridge Studies in Advanced Mathematics. Cambridge University Press, 2016. DOI: 10.1017/9781316675779.

[3] B. Derrida and H. Spohn. 'Polymers on disordered trees, spin glasses, and traveling waves'. In: *Journal of Statistical Physics* 51 (1988), pp. 817–840. DOI: 10.1007/BF01014886.

[4] J. Murray. *Mathematical Biology: I. An Introduction*. Interdisciplinary Applied Mathematics. Springer New York, 2007. ISBN: 9780387224374. DOI: 10.1007/b98868.

[5] D. G. Aronson and H. F. Weinberger. 'Nonlinear diffusion in population genetics, combustion, and nerve pulse propagation'. In: *Partial Differential Equations and Related Topics*. Ed. by J. A. Goldstein. Berlin, Heidelberg: Springer Berlin Heidelberg, 1975, pp. 5–49. ISBN: 978-3-540-37440-4. DOI: 10.1007/BFb0070595.

[6] S. N. Majumdar, D. S. Dean and P. L. Krapivsky. 'Understanding search trees via statistical physics'. In: *Pramana* 64.1175 (2005).

[7] É. Brunet, A. D. Le, A. H. Mueller and S. Munier. 'How to generate the tip of branching random walks evolved to large times'. In: *Europhysics Letters* 131.40002 (Sept. 2020). DOI: 10.1209/0295-5075/131/40002.

[8] A. D. Le, A. H. Mueller and S. Munier. 'Probabilistic picture for particle number densities in stretched tips of the branching Brownian motion'. In: *Europhysics Letters* 140.51003 (Dec. 2022). DOI: 10.1209/0295-5075/aca699.

[9] S. Munier. 'Quantum chromodynamics at high energy and statistical physics'. In: *Physics Reports* 473.1 (2009), pp. 1–49. ISSN: 0370-1573. DOI: https://doi.org/10.1016/j.physrep.2009.02.001.

[10] Y. V. Kovchegov and E. Levin. *Quantum Chromodynamics at High Energy*. Vol. 33. Oxford University Press, 2013. ISBN: 978-1-00-929144-6, 978-1-00-929141-5, 978-1-00-929142-2, 978-0-521-11257-4, 978-1-139-55768-9. DOI: 10.1017/9781009291446.

[11] A. M. Stasto, K. J. Golec-Biernat and J. Kwiecinski. 'Geometric scaling for the total gamma* p cross-section in the low x region'. In: *Phys. Rev. Lett.* 86 (2001), pp. 596–599. DOI: 10.1103/PhysRevLett.86.596. arXiv: hep-ph/0007192.

[12] C. Marquet and L. Schoeffel. 'Geometric scaling in diffractive deep inelastic scattering'. In: *Phys. Lett. B* 639 (2006), pp. 471–477. DOI: 10.1016/j.physletb.2006.07.004. arXiv: hep-ph/0606079.

[13] R. Durrett. 'Brownian Motion'. In: *Probability: Theory and Examples*. 4th ed. Cambridge Series in Statistical and Probabilistic Mathematics. Cambridge University Press, 2010. Chap. 8, pp. 353–400. DOI: 10.1017/CBO9780511779398.009.

[14] D. Revuz and M. Yor. *Continuous Martingales and Brownian Motion*. Grundlehren der mathematischen Wissenschaften. Springer Berlin Heidelberg, 2004. ISBN: 978-3-662-06400-9. DOI: 10.1007/978-3-662-06400-9.

[15] C. Gardiner. *Stochastic Methods. A Handbook for the Natural and Social Sciences*. Springer Berlin Heidelberg, 2010. ISBN: 978-3-642-08962-6. DOI: 10.1007/978-3-662-06400-9.

[16] J. L. Doob. 'Conditional Brownian Motion'. In: *Classical Potential Theory and Its Probabilistic Counterpart*. Berlin, Heidelberg: Springer Berlin Heidelberg, 2001, pp. 668–702. ISBN: 978-3-642-56573-1. DOI: 10.1007/978-3-642-56573-1_29.

[17] R. A. Fisher. 'The wave of advance of advantageous genes'. In: *Annals of Eugenics* 7.4 (1937), pp. 355–369. DOI: https://doi.org/10.1111/j.1469-1809.1937.tb02153.x.

[18] A. Kolmogorov, I. Petrovskii and N. Piskunov. 'A Study of the Diffusion Equation with Increase in the Amount of Substance, and its Application to a Biological Problem'. In: (1937). Ed. by M. U. M. M. 1.

[19] M. D. Bramson. 'Maximal displacement of branching brownian motion'. In: *Communications on Pure and Applied Mathematics* 31.5 (1978), pp. 531–581. DOI: https://doi.org/10.1002/cpa.3160310502.

[20] M. Bramson. 'Convergence of solutions of the Kolmogorov equation to travelling waves'. In: *Memoirs of the American Mathematical Society* 44 (1983).

[21] É. Brunet and B. Derrida. 'Shift in the velocity of a front due to a cutoff'. In: *Phys. Rev. E* 56 (3 Sept. 1997), pp. 2597–2604. DOI: 10.1103/PhysRevE.56.2597. URL: https://link.aps.org/doi/10.1103/PhysRevE.56.2597.

[22] A. H. Mueller and S. Munier. 'Phenomenological picture of fluctuations in branching random walks'. In: *Phys. Rev. E* 90.042143 (4 Aug. 2014). DOI: 10.1103/PhysRevE.90.042143.

[23] É. Brunet and B. Derrida. 'A Branching Random Walk Seen from the Tip'. In: *Journal of Statistical Physics* 143 (May 2011), pp. 420–446. DOI: 10.1007/s10955-011-0185-z.

[24] E. Aïdékon, J. Berestycki and É. Brunet. 'Branching Brownian motion seen from its tip'. In: *Probab. Theory Relat. Fields* 157 (2013), pp. 405–451. DOI: 10.1007/s00440-012-0461-0.

[25] A. H. Mueller and S. Munier. 'Particle-number distribution in large fluctuations at the tip of branching random walks'. In: *Phys. Rev. E* 102 (2 Aug. 2020), p. 022104. DOI: 10.1103/PhysRevE.102.022104.

[26] É. Brunet and B. Derrida. 'Effect of Microscopic Noise on Front Propagation'. In: *Journal of Statistical Physics* 103 (2001), pp. 269–282. DOI: https://doi.org/10.1023/A:1004875804376.

[27] Y. Hamana and H. Matsumoto. *The probability distributions of the first hitting times of Bessel processes*. 2012. arXiv: 1106.6132 [math.PR].

[28] M. Kac. 'On distributions of certain Wiener functionals'. In: *Trans. Amer. Math. Soc.* 65 (1949), pp. 1–13. DOI: 10.2307/1990512.