Non-Parametric Estimation of the Diffusion Coefficient of a Branching Diffusion with Immigration

Dissertation zur Erlangung des Grades
„Doktor der Naturwissenschaften“

am Fachbereich Physik, Mathematik und Informatik
an der Johannes Gutenberg-Universität in Mainz

Tobias Berg
geboren in Frankfurt am Main

Mainz, den 31. März 2015

1. Berichterstatter:
2. Berichterstatter:

Datum der mündlichen Prüfung: 30. Juni 2015

D77 - Mainzer Dissertation

Abstract

We consider finite systems of branching particles where the particles move independently of each other according to one-dimensional diffusions

$$dX_t = b(X_t)\,dt + \sigma(X_t)\,dW_t.$$

Particles die at a position-dependent rate and leave a random number of offspring located in space according to some transition kernel. In addition, new particles immigrate at a constant rate. A process with these properties is called *branching diffusion with immigration* (BDI). Observing a BDI at discrete points in time, it is not evident which discretely observed points belong to which path. Therefore, we develop an algorithm for reconstructing the underlying trajectory. With the aid of this algorithm, we construct a non-parametric estimator for the squared diffusion coefficient $\sigma^2(\cdot)$, essentially by filling a classical regression scheme. We prove consistency and a central limit theorem.

Zusammenfassung

Wir betrachten Systeme von endlich vielen Partikeln, wobei die Partikel sich unabhängig voneinander gemäß eindimensionaler Diffusionen

$$dX_t = b(X_t)\,dt + \sigma(X_t)\,dW_t$$

bewegen. Die Partikel sterben mit positionsabhängigen Raten und hinterlassen eine zufällige Anzahl an Nachkommen, die sich gemäß eines Übergangskerns im Raum verteilen. Zudem immigrieren neue Partikel mit einer konstanten Rate. Ein Prozess mit diesen Eigenschaften wird *Verzweigungsprozess mit Immigration* genannt. Beobachten wir einen solchen Prozess zu diskreten Zeitpunkten, so ist zunächst nicht offensichtlich, welche diskret beobachteten Punkte zu welchem Pfad gehören. Daher entwickeln wir einen Algorithmus, um den zugrundeliegenden Pfad zu rekonstruieren. Mit Hilfe dieses Algorithmus konstruieren wir einen nichtparametrischen Schätzer für den quadrierten Diffusionskoeffizienten $\sigma^2(\cdot)$, wobei die Konstruktion im Wesentlichen auf dem Auffüllen eines klassischen Regressionsschemas beruht. Wir beweisen Konsistenz und einen zentralen Grenzwertsatz.

Introduction

In this thesis, we consider finite systems of branching diffusions with immigration and the random branching of particles. Our underlying model can be described as follows:
Every particle of a finite system of particles moves independently of other particles in $\mathbb{R}$ according to a one-dimensional diffusion

$$dX_t = b(X_t)\,dt + \sigma(X_t)\,dW_t,$$

where $W = (W_t)_{t\geq 0}$ is a one-dimensional Brownian motion and both the drift coefficient $b(\cdot)$ and the diffusion coefficient $\sigma(\cdot)$ are Lipschitz continuous functions. Each particle is "killed" at a position-dependent rate $\kappa(\cdot)\colon \mathbb{R} \to \mathbb{R}_+$, which means that a particle situated at $x \in \mathbb{R}$ at time $t \geq 0$ dies during a short time interval $(t, t+\Delta]$ with probability

$$\kappa(x) \cdot \Delta + o(\Delta), \quad \text{as} \quad \Delta \to 0.$$

At its time of death, the particle is replaced by a random number of offspring $k \in \mathbb{N}_0$ with probability $p_k(x)$, where the k newborn particles are distributed in $\mathbb{R}$ according to the law

$$Q_k(x;\cdot)\colon \mathcal{B}(\mathbb{R}^k) \to [0,1].$$

These newborn particles move and branch according to the same mechanism as the parent particle. In addition, new particles immigrate at a constant rate $c > 0$, i.e., at each immigration event exactly one new particle is added to the system of pre-existing particles and is distributed in $\mathbb{R}$ according to the law $\nu(\cdot)$.

The resulting process $(\eta_t)_{t\geq 0}$ of particle configurations is a strong Markov process and is called *branching diffusion with immigration* (BDI). BDI processes and their properties have been studied in several papers, see [2], [10], [12], [13], [14], [15], [28] and [29]. However, in these papers there may be some differences in the model as described above. For example, a modification of the model was investigated in [10], [28] and [29]: In these models, particles live in $\mathbb{R}^d$, $d \in \mathbb{N}$, and interactions between the particles in both their spatial motion and the branching/reproduction/immigration mechanisms are allowed, i.e., the quantities defining the model may also depend on the configuration of co-existing particles. Furthermore, the models in [2], [12], [13], [15] and [29] are based on the assumption that branching particles reproduce at their position of death, whereas in our model offspring particles are scattered in space according to some law $Q_k(\cdot,\cdot)$.

During the last years, the existence, boundedness and continuity of the invariant measure $m(\cdot)$ on the configuration space $\mathcal{S} := \bigcup_{\ell\in\mathbb{N}_0} \mathbb{R}^{d\ell}$ and of the occupation measure

$$\overline{m}(B) := \int_S \mathbf{x}(B)\, m(d\mathbf{x}), \quad B \in \mathcal{B}(\mathbb{R}^d),$$

have extensively been investigated. In [13], an interesting aspect was discussed: It becomes apparent that – as soon as branching particles reproduce at their position of death $x \in \mathbb{R}^d$ – the density of the invariant measure $m(\cdot)$ may exist but it is neither bounded nor continuous. In [29], Löcherbach proved the existence of a bounded and continuous density of $\overline{m}(\cdot)$ on $\mathbb{R}^d$ by assuming uniform ellipticity and strong smoothness resp. boundedness conditions both on the drift coefficient $b(\cdot)$ and the diffusion coefficient $\sigma(\cdot)$. For this, Löcherbach makes use of Malliavin calculus. However, in her paper she assumes that branching particles reproduce – either zero or two offspring – at their position of death. As we have noticed before, this may preclude the existence of a bounded and continuous Lebesgue density of $m(\cdot)$. Also, in [15] Höpfner and Löcherbach discuss results about the existence of a density of $\overline{m}(\cdot)$ and its regularity properties. In their framework, offspring particles start their spatial motion at their parents' position of death, too.
In Hammer's thesis [10], the issue of the existence of bounded and continuous Lebesgue densities both of $m(\cdot)$ and $\overline{m}(\cdot)$ is approached. Granting that the BDI η is recurrent in the sense of Harris, Hammer stated four assumptions which are sufficient for this existence, namely:

1. Continuous Transition Density and Heat Kernel Estimate of the "Killed" Particle Motion

2. Absolutely Continuity of Offspring and Immigration Laws

3. Fixed Bound of Possible Offspring

4. Exponential Decay of $m(\mathbb{R}^\ell)$, $\ell \in \mathbb{N}$.

In particular, Hammer does not allow branching particles to reproduce at their position of death $x \in \mathbb{R}$, but rather that offspring particles are scattered in space according to some law $Q_k(x;\cdot)$ which fulfils certain assumptions. We remark that Hammer exposed his assumptions in the "interactive" framework where the quantities defining the model may also depend on the configuration of co-existing particles.

Concerning statistical applications, BDI processes have been examined in [2], [12] and [14]. In [2], Brandt constructed a non-parametric estimator for the squared diffusion coefficient $\sigma^2(\cdot)$ of a BDI. For this, Brandt developed a reconstruction algorithm of the trajectory of a discretely observed BDI and combined this algorithm with the Nadaraya-Watson-Estimator for the squared diffusion coefficient $\sigma^2(\cdot)$ of one-dimensional diffusions, which was developed in [7]. Apart from that, in [12] resp. [14] an estimator for the branching rate $\kappa(\cdot)$ of a BDI was constructed.

In the first chapter of this thesis, we give a mathematical definition of the BDI process and we explain which basic assumptions and notations are used. In particular, by assuming the void configuration (the state of no existing particle) as a recurrent atom, the process $(\eta_t)_{t\geq 0}$ becomes positive recurrent in the sense of Harris, i.e., it allows for a finite invariant measure $m(\cdot)$ on $\mathcal{S}$. Furthermore, we assume that our BDI model is based on Hammer's framework from [10] in order to make use of the bounded and continuous Lebesgue densities of $m(\cdot)$ and of $\overline{m}(\cdot)$. As in our framework the quantities which define the model only depend on the position of the particles, Hammer's four assumptions slightly simplify. The reason for restricting ourselves to the "position-dependent" case and to the case that particles live in $\mathbb{R}$ is that in the last chapter statistical applications are applied. They are troublesome for diffusions which live in $\mathbb{R}^d$, $d \in \mathbb{N}\backslash\{1\}$, since their occupation time may be small or even zero in certain regions.

In the second chapter, we examine some properties of BDI's. As a BDI contains many one-dimensional diffusion paths, we consider their properties first. Particularly, for statistical applications in the last chapter, their behaviour during a short time interval $(t, t+\Delta]$ is analysed. Afterwards, some results are extended for branching diffusions. Denoting for some $\mathbf{x} \in \mathcal{S}$ the length of $\mathbf{x} = (x^1, ..., x^\ell) \in \mathbb{R}^\ell$ by $\ell = \ell(\mathbf{x})$ and defining

$$S_\varepsilon := \Big\{\mathbf{x} \in \mathcal{S} \,\big|\, \ell(\mathbf{x}) \geq 2, \exists i \neq j \in \big\{1, ..., \ell(\mathbf{x})\big\} : |x^i - x^j| < \varepsilon\Big\},$$

as our main result of this chapter we state a rate of convergence for $m(S_\varepsilon)$, namely

$$m(S_\varepsilon) = \mathcal{O}(\varepsilon), \quad \text{as} \quad \varepsilon \to 0,$$

see Theorem 2.11. For this result, it is important that we act on Hammer's framework. The reason for this is the phenomenon we have mentioned before: In the framework where branching particles reproduce at their position of death, under certain assumptions a density of $m(\cdot)$ exists, but it takes the value $+\infty$ on a non-empty subset of

$$\mathcal{N} := \Big\{\mathbf{x} \in \mathcal{S} \,\big|\, \ell(\mathbf{x}) \geq 2, \exists i \neq j \in \big\{1, ..., \ell(\mathbf{x})\big\} : x^i = x^j\Big\},$$

c.f. [10, p. 25f] resp. [13].

The third chapter consists of a reconstruction algorithm for the trajectory of a BDI $(\eta_t)_{t\geq 0}$, provided that we consider the BDI process at discrete points in time $i\Delta$, $i \in \mathbb{N}_0$, where $\Delta > 0$. For this, we define a rule for a pair $(\eta_{i\Delta}, \eta_{(i+1)\Delta})$ being "interpretable", i.e., we demand that there exists an arrangement $(\beta_{i\Delta}, \beta_{(i+1)\Delta})$ of $(\eta_{i\Delta}, \eta_{(i+1)\Delta})$ with the following properties: Both $\beta_{i\Delta}$ and $\beta_{(i+1)\Delta}$ have particles which do not have close neighbours and the distance of each particle of $\beta_{i\Delta}$ to exactly one other particle of $\beta_{(i+1)\Delta}$ is not more than Δ^λ, for given $0 < \lambda < \frac{1}{2}$. We will show that the expected quota of "properly interpretable pairs" (pairs whose assignment rule is correct) up to deterministic time horizons $T := T_\Delta > 0$ converges to 1, as $\Delta \to 0$, with a rate of convergence being set by the rate of $m(S_\varepsilon)$, see Theorem 3.4. Our algorithm extends the partial reconstruction algorithm developed by Brandt in [2].

In the last chapter, which consists of four sections, statistical applications for the BDI process are applied. First, we present known estimators and their properties for the squared diffusion coefficient $\sigma^2(\cdot)$ of one-dimensional diffusions. In the second section, we construct a non-parametric estimator $\hat{\sigma}^2_\Delta(\cdot)$ for the squared diffusion coefficient $\sigma^2(\cdot)$ of a BDI, provided that the trajectory of the BDI is considered at discrete points in time. The idea of this estimator relies on Hoffmann's publications [17] and [18] for estimating the squared diffusion coefficient $\sigma^2(\cdot)$ of one-dimensional diffusions

$$dX_t = b(X_t)\,dt + \sigma(X_t)\,dW_t.$$

Hoffmann constructs wavelet-estimators for $\sigma^2(\cdot)$ which are optimal in the minimax sense (for integrated errors) over Besov balls, essentially by filling a classical regression scheme. Hoffmann's procedure is as follows: Initially, sub-boxes of a compact set $D \subseteq \mathbb{R}$ are filled with a finite number of observed points $X_{i\Delta}$, $i \in \mathbb{N}_0$. Then, Hoffmann makes use of these observed points in order to apply the regression identity

$$\left(\frac{X_{(i+1)\Delta} - X_{i\Delta}}{\sqrt{\Delta}}\right)^2 = \sigma^2(X_{i\Delta}) + \varepsilon_{i\Delta} + \mathcal{O}_P\big(\sqrt{\Delta}\big), \quad \text{as} \quad \Delta \to 0,$$

where $\varepsilon_{i\Delta}$ are centered martingale increments. Due to this identity, $\sigma^2(x)$ is estimated by a wavelet-estimator and this estimator attains the classical minimax rate of convergence $\Delta^{\frac{r}{1+2r}}$, $r \in \mathbb{N}$, where r is the regularity of the diffusion coefficient $\sigma(\cdot)$. For the construction of our estimator, we adapt some of Hoffmann's ideas:

- We fill sub-boxes of the compact set $D := [0,1]$ with a finite number of observed "interpretable pairs" $(\eta_{i\Delta}, \eta_{(i+1)\Delta})$, $i \in \mathbb{N}_0$, up to a certain time horizon $T_\Delta > 0$.
- We use some particles of the "interpretable pairs" in order to apply the following regression identity for one-dimensional diffusions $(X_t)_{t\geq 0}$

  $$\left(\frac{X_{t+\Delta} - X_t}{\sqrt{\Delta}}\right)^2 = \sigma^2(X_t)\cdot\big(1 + U_{(t,\Delta)}\big) + \mathcal{O}_P\big(\sqrt{\Delta}\big), \quad \text{as} \quad \Delta \to 0,$$

  where $U_{(t,\Delta)}$ is a $\mathcal{F}'_{t+\Delta}$-measurable random variable being independent of $\mathcal{F}'_t$ and satisfying

  $$U_{(t,\Delta)} \stackrel{d}{=} 2\cdot\int_0^1 W_s\,dW_s$$

  for every $t \geq 0$, $\Delta > 0$ (note that $(\mathcal{F}'_t)_{t\geq 0}$ is the filtration generated by X). This regression identity relies on ideas and estimates developed in [8] and [25, p. 356f].

It is not obvious that there are enough "interpretable pairs" in every sub-box of $[0,1]$ up to a certain time horizon $T_\Delta > 0$. However, we will solve this issue by applying the Harris recurrence of the BDI, the rate of the reconstruction algorithm and the continuity of the density of $\overline{m}(\cdot)$.

In the third section, we show consistency of the estimator $\hat{\sigma}^2_\Delta(\cdot)$ for the class of non-negative Lipschitz continuous diffusion coefficients $\sigma(\cdot)$ being bounded and bounded away from zero. More exactly, we prove for every $x \in [0,1]$

$$\left|\hat{\sigma}^2_\Delta(x) - \sigma^2(x)\right| = \mathcal{O}_{\mathbf{P}_m}\big(\Delta^{\frac{1}{3}}\big), \quad \text{as} \quad \Delta \to 0,$$

see Theorem 4.11. In the fourth section, we state a central limit theorem, so we show for every $x \in [0,1]$ and for every $0 < \varepsilon < \frac{1}{3}$

$$\sqrt{\Delta^{-\frac{2}{3}}} \cdot \Delta^{\varepsilon} \cdot \left(\frac{\hat{\sigma}^2_{\Delta,\varepsilon}(x)}{\sigma^2(x)} - 1 \right) \xrightarrow{\Delta \to 0} Z \quad \text{in} \quad \mathbf{P}_m\text{-distribution},$$

where Z is a standard normal distributed random variable, see Theorem 4.13. Finally, we discuss our results and verify how they fit to well-known classical regression results: For consistency, we attain the classical minimax rate of convergence $\Delta^{\frac{1}{3}}$ by choosing the classical optimal bandwidth $h_\Delta := \Delta^{\frac{1}{3}}$. However, for the central limit theorem our estimator depends on both $0 < \varepsilon < \frac{1}{3}$ and $\Delta > 0$ and we receive a rate of convergence which is slightly weaker than $\sqrt{\Delta^{-\frac{2}{3}}}$.

For reasons of privacy protection, it is not allowed to state the names of others here. The acknowledgments are therefore left blank in this electronic version of the thesis.

Contents

Introduction **5**

1 The Model: Branching Diffusions with Immigration **13**

1.1 Basic Assumptions and Notations 14

1.2 Ergodicity and Invariant Measure 19

1.3 Hammer's Framework 22

2 Properties of Branching Diffusions with Immigration **29**

2.1 Properties of One-Dimensional Diffusions 29

2.2 Properties of Branching Diffusions 35

3 A Reconstruction Algorithm for Branching Diffusions with Immigration **45**

3.1 Notation 45

3.2 A Partial Reconstruction Algorithm 46

4 Non-Parametric Estimation of the Diffusion Coefficient of Branching Diffusions with Immigration **57**

4.1 Estimators for One-Dimensional Diffusions 57

4.2 Construction of the Estimator 60

4.3 Consistency of the Estimator 66

4.4 A Central Limit Theorem for the Estimator 71

Bibliography **75**

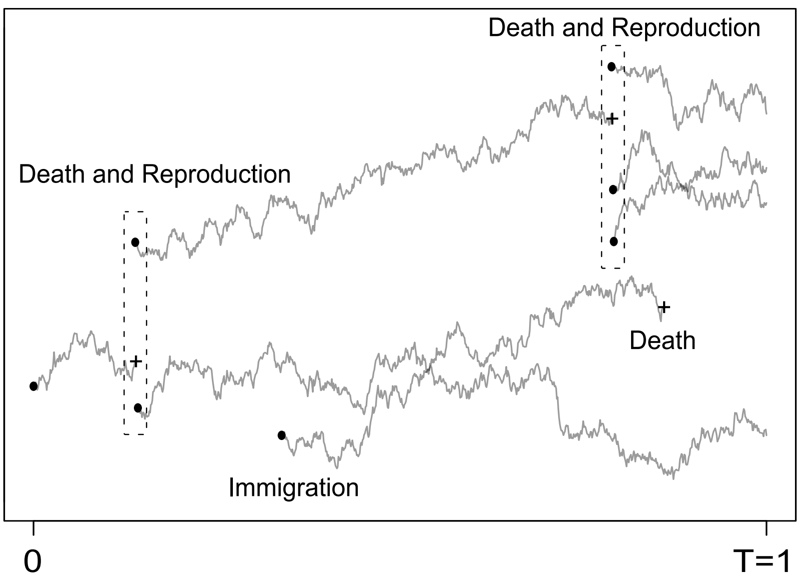
Chapter 1

The Model: Branching Diffusions with Immigration

In this first chapter, we are going to introduce the underlying model. We consider systems of finitely many particles (each living in $\mathbb{R}$) travelling independently of each other according to a solution of a diffusion

$$dX_t = b(X_t)\,dt + \sigma(X_t)\,dW_t.$$

Every particle branches according to a position-dependent rate. When a particle branches, it dies and produces, depending on its position in $\mathbb{R}$, a random number of offspring. The newborn particles are distributed randomly in space, depending on the position of the branching parent particle. In addition, immigration occurs at a constant rate. At each immigration event, exactly one new particle is added to the system at a position according to some law. The resulting stochastic process of finite particle configurations is called *branching diffusion with immigration*, henceforth BDI. The following graphic demonstrates a typical BDI path with its branching, reproduction and immigration mechanisms during the time $[0, 1]$.



1.1 Basic Assumptions and Notations

We will write $E := \mathbb{R}$ for the single particle space. A BDI as described before is a strong Markov process $\eta = (\eta_t)_{t\geq 0}$ taking values in the space

$$\mathcal{S} := \dot{\bigcup_{\ell\in\mathbb{N}_0}} E^\ell$$

of finite ordered particle configurations, where $E^0 := \{\delta\}$ denotes the void configuration (the state of no existing particle). We want to distinguish between elements of the single particle space E and the configuration space $\mathcal{S}$ by using standard letters x, y, z for elements of E and bold letters $\mathbf{x}, \mathbf{y}, \mathbf{z}$ for elements of $\mathcal{S}$. The length of $\mathbf{x} \in \mathcal{S}$ is denoted by $\ell(\mathbf{x})$, i.e., $\mathbf{x} \in E^\ell$ if and only if $\ell = \ell(\mathbf{x})$. We now state the first assumption which governs the motion of particles between branching or immigration events.

Assumption 1.1 (Particle Motion)

1. *For every $\ell \in \mathbb{N}$ the ℓ-particle motion $X^\ell = (X^{1,\ell}, ..., X^{\ell,\ell})$ on E^ℓ is given by the stochastic differential equations*

$$dX_t^{j,\ell} = b\big(X_t^{j,\ell}\big)\,dt + \sigma\big(X_t^{j,\ell}\big)\,dW_t^j, \quad 1 \leq j \leq \ell, \tag{1.1}$$

with independent one-dimensional standard Brownian motions $W^1, ..., W^\ell$ driving the motion of every particle and drift and diffusion coefficients

$$b(\cdot)\colon E \to E \quad \text{and} \quad \sigma(\cdot)\colon E \to E.$$

2. *Both the drift and the diffusion coefficients are Lipschitz continuous functions.*

$\Diamond$

Remark 1.2

1. If $\ell = 0$, we use the convention $X^0 \equiv \delta$.

2. According to [21, p. 178f], a sufficient condition for a unique strong solution of (1.1) is that both $b(\cdot)$ and $\sigma(\cdot)$ are globally Lipschitz, i.e., there is a constant $L > 0$ such that

$$|b(x) - b(x')| + |\sigma(x) - \sigma(x')| \leq L \cdot |x - x'| \tag{1.2}$$

for every $x, x' \in E$. As this is granted in Assumption 1.1, the stochastic differential equation (1.1) has a unique strong solution.

3. The stochastic basis (the sample space) on which the ℓ-particle motion happens to be defined is not specified. As Hammer suggests in [10, p. 2], it could be the canonical path space $\mathcal{C}(E_+; E^\ell)$ or any other suitable space. Furthermore, the probability measure corresponding to the diffusion X^ℓ starting at $\mathbf{x} = (x^1, ..., x^\ell) \in E^\ell$ is denoted by $P_\mathbf{x}$, corresponding expectations by $E_\mathbf{x}$ and its semigroup by

$$P_t f(\mathbf{x}) := E_\mathbf{x}\big(f(X_t^\ell)\big), \tag{1.3}$$

where $f(\cdot) \in \mathcal{B}(E^\ell)$.

4. For every $1 \leq j \leq \ell$ the diffusion $X^{j,\ell}$ is a one-dimensional diffusion which takes values in E. Its generator is given by
$$\mathscr{A} f(x) = b(x) \cdot \frac{\partial f}{\partial x}(x) + \frac{1}{2}\sigma^2(x) \cdot \frac{\partial^2 f}{\partial x^2}(x)$$
for twice continuously differentiable functions $f(\cdot)$ with compact support in E, $x \in E$, c.f. [21, p. 212f]. Because of the independence of the diffusions $X^{j,\ell}$, $1 \leq j \leq \ell$, according to [21, p. 214f], the generator $\mathscr{A}^{(\ell)}$ of X^ℓ is the sum of the generators of every single diffusion $X^{j,\ell}$, i.e.,
$$\mathscr{A}^{(\ell)} f(\mathbf{x}) = \sum_{i=1}^{\ell} b(x^i) \cdot \frac{\partial f}{\partial x^i}(\mathbf{x}) + \frac{1}{2}\sum_{i=1}^{\ell} \sigma^2(x^i) \cdot \frac{\partial^2 f}{\partial (x^i)^2}(\mathbf{x}) \tag{1.4}$$
for twice continuously differentiable functions $f(\cdot)$ with compact support in E^ℓ, $\mathbf{x} \in E^\ell$.
5. Also, $E := \mathbb{R}^d$, $d \in \mathbb{N}$, could be set for the single particle space, i.e., each solution $X_t^{j,\ell}$ takes values in $\mathbb{R}^d$. However, we focus on the case $d = 1$ since in the last chapter we will use statistical applications. They are troublesome for diffusions which live in $\mathbb{R}^d$, $d \in \mathbb{N}\backslash\{1\}$, since their occupation time may be small or even zero in certain regions.

$\Diamond$

Now, we show how the branching, reproduction and immigration mechanisms work.

Assumption 1.3 (Branching and Reproduction Mechanism)
We are given a non-negative measurable function
$$\kappa(\cdot)\colon E \to E_+$$
which is bounded and bounded away from zero, i.e., there are constants $\underline{\kappa}, \overline{\kappa} > 0$ *such that*
$$\underline{\kappa} \leq \kappa(x) \leq \overline{\kappa} \tag{1.5}$$
for every $x \in E$. *Moreover, we are given measurable functions*
$$p_k(\cdot)\colon E \to [0,1], \quad k \in \mathbb{N}_0,$$
such that $\sum_{k\in\mathbb{N}_0} p_k(\cdot) \equiv 1$ *and transition probabilities*
$$Q_k(\cdot;\cdot)\colon E \times \mathcal{B}(E^k) \to [0,1], \quad k \in \mathbb{N}.$$
We put
$$Q_0(x;\cdot) := \epsilon_\delta(\cdot), \quad x \in E.$$
A particle belonging to the configuration $\mathbf{x} = (x^1, ..., x^\ell) \in E^\ell$ *and situated at position* $x^i \in E$ *at time* $t > 0$ *branches at a position-dependent rate* $\kappa(x^i)$, *i.e., it dies during a small time interval* $(t, t + \Delta]$ *with probability*
$$\kappa(x^i) \cdot \Delta + o(\Delta), \tag{1.6}$$

as $\Delta \to 0$. *At its death time, it is replaced by a random number* $k \in \mathbb{N}_0$ *of offspring particles with probability* $p_k(x^i)$. *The* k *offspring particles are distributed in* E^k *according to the law*

$$Q_k\big(x^i; dv^1 \dots dv^k\big) \quad on \quad (E^k, \mathcal{B}(E^k)).$$

We will refer to $\kappa(\cdot)$ *as the branching rate, to* $(p_k(\cdot))_{k\in\mathbb{N}_0}$ *as the reproduction law and to* $(Q_k(\cdot;\cdot))_{k\in\mathbb{N}}$ *as the spatial branching distribution.*

$\Diamond$

Assumption 1.4 (Immigration Mechanism)
New particles immigrate at a constant rate $c > 0$: *If there are* $\ell \in \mathbb{N}_0$ *particles at positions* $\mathbf{x} = (x^1, ..., x^\ell) \in E^\ell$ *at time* $t > 0$, *one new particle immigrates during a small time interval* $(t, t+\Delta]$ *with probability*

$$c \cdot \Delta + o(\Delta), \tag{1.7}$$

as $\Delta \to 0$. *The immigrating particle is distributed in* E *according to*

$$\nu(\cdot) \colon \mathcal{B}(E) \to [0,1],$$

to which we refer as the immigration law.

$\Diamond$

Remark 1.5

1. Combining both Assumptions 1.3 and 1.4,

$$\alpha^{(\ell)}(\mathbf{x}) := \sum_{i=1}^{\ell} \kappa(x^i) + c \tag{1.8}$$

gives the rate at which a branching or immigration event happens, starting from a ℓ-particle configuration $\mathbf{x} \in E^\ell$.

2. According to (1.5), (1.6) and (1.7), there are two constants $\underline{c}$, $\overline{c} > 0$ such that

$$\underline{c} \leq \min\{\underline{\kappa}, c\} \leq \max\{\overline{\kappa}, c\} \leq \overline{c}.$$

Thus, using (1.8) for every $\ell \in \mathbb{N}$ and $\mathbf{x} \in E^\ell$ it holds

$$\underline{c} \cdot (\ell+1) \leq \alpha^{(\ell)}(\mathbf{x}) \leq \overline{c} \cdot (\ell+1), \tag{1.9}$$

i.e.,

$$\alpha^{(\ell)}(\cdot) \asymp \ell, \quad \ell \to \infty.$$

3. Denoting τ for the first branching/immigration time and using equation (1.8), the semi-group of the ℓ-particle motion X^ℓ killed at rate $\alpha^{(\ell)}(\cdot)$ is given by

$$\begin{aligned} P_t^\alpha f(\mathbf{x}) &:= E_{\mathbf{x}}\left(f(X_t^\ell) \cdot \mathbb{1}_{\{t<\tau\}} \right) \\ &= E_{\mathbf{x}}\left(f(X_t^\ell) \, \exp\left(- \int_0^t \alpha^{(\ell)}(X_s^\ell) \, ds \right) \right), \end{aligned} \tag{1.10}$$

where $f(\cdot) \in \mathcal{B}(E^\ell)$ and $\mathbf{x} \in E^\ell$. Furthermore, the occupation times of the killed ℓ-particle motion are given by the generalized resolvent

$$R_\alpha^{(\ell)}(\mathbf{x}; B) := E_\mathbf{x}\left(\int_0^\tau \mathbb{1}_B(X_t^\ell)dt\right) \tag{1.11}$$

for every $\mathbf{x} \in E^\ell$, $B \in \mathcal{B}(E^\ell)$. By partial integration, it holds

$$\begin{aligned} R_\alpha^{(\ell)}(\mathbf{x}; B) &= E_\mathbf{x}\left(\int_0^\infty \mathbb{1}_B(X_t^\ell)\ \exp\left(-\int_0^t \alpha^{(\ell)}(X_s^\ell)\, ds\right) dt\right) \\ &= \int_0^\infty P_t^\alpha(\mathbf{x}, B)\, dt \end{aligned} \tag{1.12}$$

for every $\mathbf{x} \in E^\ell$, $B \in \mathcal{B}(E^\ell)$. Setting $B := E^\ell$, we receive in (1.11)

$$R_\alpha^{(\ell)}(\mathbf{x}, E^\ell) = E_\mathbf{x}(\tau) \tag{1.13}$$

and combining (1.9) with (1.12), for every $\mathbf{x} \in E^\ell$ it holds

$$\frac{1}{\overline{c} \cdot (\ell+1)} \leq R_\alpha^{(\ell)}(\mathbf{x}; E^\ell) \leq \frac{1}{\underline{c} \cdot (\ell+1)},$$

i.e.,

$$R_\alpha^{(\ell)}(\cdot; E^\ell) \asymp \frac{1}{\ell}, \quad \ell \to \infty. \tag{1.14}$$

Denoting P_t^κ the semigroup of the killed particle motion X^ℓ *without* immigration

$$P_t^\kappa f(\mathbf{x}) := E_\mathbf{x}\left(f(X_t^\ell)\ \exp\left(-\sum_{j=1}^{\ell}\int_0^t \kappa(X_s^{j,\ell})\, ds\right)\right), \tag{1.15}$$

because of (1.8) it holds

$$P_t^\alpha f(\mathbf{x}) = \exp(-ct) \cdot P_t^\kappa f(\mathbf{x}) \tag{1.16}$$

for every $\mathbf{x} \in E^\ell, f(\cdot) \in \mathcal{B}(E^\ell)$.

4. Originally, in Hammer's model the branching/reproduction rate resp. the spatial offspring distribution may both depend on the branching parent and the configuration of co-existing particles, c.f. [10, p. 2f]. Apart from that, in [2], [12], [13], [15] and [29], it is assumed that offspring particles start their spatial motion at their parents' position of death, which means

$$Q_k(x; \cdot) = \epsilon_x(\cdot)^{\otimes k} \tag{1.17}$$

on E^k for every $k \in \mathbb{N}$ and $x \in E$. However, in the next section we will give reasons why we do not allow for spatial offspring distributions such as (1.17) and we will specify the spatial offspring distribution.

$\Diamond$

Now, we describe how we can create a BDI by using the previous assumptions. As Hammer mentions in [10, p. 6], a BDI η with the desired properties relies on the "killing and restarting"-procedure for Markov processes developed by Ikeda, Nagasawa and Watanabe in [20] (see also [30]). Let us outline this construction: Let $\mathbf{X} = (\mathbf{X})_{t\geq 0}$ denote the $\mathcal{S}$-valued process describing a finite system of particles such that for $\ell \in \mathbb{N}$, starting from ℓ particles, $\mathbf{X}$ evolves as the given process X^ℓ on E^ℓ, without any branching or immigration. In other words, $\mathbf{X}$ is the direct sum process of the given ℓ-particle motions X^ℓ, $\ell \in \mathbb{N}$. For $\ell = 0$ (starting from the void configuration δ), by convention we have $\mathbf{X}_t = \delta$ for all $t > 0$. The process $\mathbf{X}$ is now stopped with a configuration-dependent rate $\alpha(\cdot)\colon \mathcal{S} \to E_+$ defined layer-wise as in (1.8). At its death time, it is restarted with a new initial configuration chosen by a jump kernel $K(\cdot;\cdot)\colon \mathcal{S}\times\mathcal{B}(\mathcal{S}) \to [0,1]$ which is defined as follows: For each $\ell \in \mathbb{N}$, $1 \leq i \leq \ell$ and $k \in \mathbb{N}_0$ define a mapping $\Pi_{\ell,k,j}(\cdot;\cdot)\colon E^\ell \times E^k \to E^{\ell-1+k}$ by

$$\begin{aligned}\Pi_{\ell,k,i}(\mathbf{x};\mathbf{v}) &:= (x^1, ..., x^{i-1}, v^1, ..., v^k, x^{i+1}, ..., x^\ell), \quad k \in \mathbb{N},\\ \Pi_{\ell,0,i}(\mathbf{x}) &:= \Pi_{\ell,0,i}(\mathbf{x};\delta) := (x^1, ..., x^{i-1}, x^{i+1}, ..., x^\ell), \quad k = 0.\end{aligned}$$

It can be interpreted as a mapping which replaces the i-th particle x^i of a given ℓ-particle configuration $\mathbf{x} = (x^1, ..., x^\ell) \in E^\ell$ by $k \in \mathbb{N}_0$ particles at positions $v^1, ..., v^k$. Also, for $\ell \in \mathbb{N}_0$, $\mathbf{x} = (x^1, ..., x^\ell) \in E^\ell$ and $v \in E$ we write

$$(\mathbf{x}, v) := (x^1, ..., x^\ell, v) \in E^{\ell+1}$$

for the configuration obtained by concatenation (with the understanding that $(\delta, v) = v$ if $\ell = 0$). The jump kernel is then defined as

$$\begin{aligned}K(\mathbf{x};\cdot) := \frac{1}{\alpha^{(\ell)}(\mathbf{x})}\cdot\Bigg(\sum_{k\in\mathbb{N}_0}\sum_{i=1}^{\ell} \kappa(x^i)p_k(x^i)\int_{E^k} Q_k(x^i; dv^1 ... dv^k)\, \epsilon_{\Pi_{\ell,k,i}(\mathbf{x};\mathbf{v})}(\cdot)\\ + c\int_E \nu(dv)\,\epsilon_{(\mathbf{x},v)}(\cdot)\Bigg)\end{aligned} \tag{1.18}$$

for $\mathbf{x} = (x^1, ..., x^\ell) \in E^\ell$, $\ell \in \mathbb{N}_0$.

The procedure described above can be made by using the so-called "Revival Theorem" for Markov processes, see [20]. Applying it under our assumptions, the resulting process of particle configurations can be constructed as a strong Markov process

$$\eta = \big(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, (\mathbf{P}_\mathbf{x})_{\mathbf{x}\in\mathcal{S}}, (\eta_t)_{t\geq 0}, (\theta_t)_{t\geq 0}\big) \tag{1.19}$$

on some suitable stochastic basis with a right-continuous filtration. The process η takes values in $\mathcal{S}_\partial = \mathcal{S} \cup \{\partial\}$, where ∂ is an extra point adjoined to the space $\mathcal{S}$ as a "cemetery" in order to account for the possibility of explosion of the process (accumulation of branching or immigration events in finite time). Writing

$$\tau_\infty := \inf\big\{t > 0\colon \eta_t \notin \mathcal{S}\big\} \leq \infty$$

for the (possibly finite) life-time (in the sense of explosion time), η has càdlàg paths before time τ_∞, and we have an increasing sequence

$$0 := \tau_0 < \tau_1 < \tau_2 ... \uparrow \sup_{n\in\mathbb{N}} \tau_n = \tau_\infty \tag{1.20}$$

of $(\mathcal{F}_t)_{t\geq 0}$- stopping times given by

$$\tau_n = \tau_{n-1} + \tau_1 \circ \theta_{\tau_{n-1}}, \quad n \in \mathbb{N}, \tag{1.21}$$

corresponding to branching or immigration events in the process η, c.f. [10, p. 6].

Remark 1.6

1. Concerning the BDI process η, we want to write expectations w.r.t. the probability measure $\mathbf{P_x}$ by $\mathbf{E_x}$, where $\mathbf{x} \in \mathcal{S}$. The transition semigroup of η is denoted by

$$\mathbf{P}_t(\mathbf{x}; F) := \mathbf{E_x}\left(\mathbb{1}_F(\eta_t) \cdot \mathbb{1}_{\{t<\tau_\infty\}}\right). \tag{1.22}$$

   Furthermore, we interpret (1.12) as a transition kernel on $\mathcal{S} \times \mathcal{B}(\mathcal{S})$ such that, if $\mathbf{x} \in E^\ell$, the measure $R_\alpha(\mathbf{x}, \cdot)$ charges only the layer E^ℓ, i.e.,

$$R_\alpha(\mathbf{x}, \cdot) := R_\alpha^{(\ell)}(\mathbf{x}; \cdot \cap E^\ell), \tag{1.23}$$

   where $\mathbf{x} \in E^\ell$, $F \in \mathcal{B}(E^\ell)$.

2. Remember the generator $\mathscr{A}^{(\ell)}$ from the ℓ-particle motion in (1.4) and the jump kernel $K(\mathbf{x}, \cdot)$ from (1.18). We denote by $\mathscr{A}$ the generator of the direct sum process $\mathbf{X}$. It acts on functions $f(\cdot) = (f^{(\ell)}(\cdot))_{\ell\in\mathbb{N}_0}$ on $\mathcal{S}$ layerwise via $(\mathscr{A}f)^{(\ell)}(\mathbf{x}) = \mathscr{A}^{(\ell)} f^{(\ell)}(\mathbf{x})$ for $\mathbf{x} \in E^\ell$. Then, the generator $\mathcal{A}$ of the BDI η is given by

$$\mathcal{A}f(\mathbf{x}) = \mathscr{A}f(\mathbf{x}) + \alpha(\mathbf{x}) \cdot \int_{\mathcal{S}} K(\mathbf{x}; d\mathbf{y})(f(\mathbf{y}) - f(\mathbf{x})), \quad \mathbf{x} \in \mathcal{S}.$$

3. Regarding the jump kernel in (1.18), in contrast to [2], [13] and [15], branching particles are not inserted at the end of the configuration but substitute the dying particle.

$\Diamond$

So far, it is not evident that η has an infinite life-time. In the next section, we will see that assuming positive Harris recurrence of the process η with δ as a recurrent atom, we obtain infinite life-time of η.

1.2 Ergodicity and Invariant Measure

As mentioned before, we will work, in addition to Assumptions 1.1, 1.3 and 1.4, under the following assumption, c.f. [10, p. 11].

Assumption 1.7 (Recurrence)
We assume that the process η admits the void configuration δ as a recurrent atom with finite expected return time. Defining

$$R := \inf_{n\in\mathbb{N}} \{\tau_n \mid \eta_{\tau_n} = \delta\},$$

we suppose that

$$\mathbf{E}_{\mathbf{x}}(R) < \infty, \quad \mathbf{x} \in \mathcal{S}. \tag{1.24}$$

$\Diamond$

Remark 1.8
It is clear that Assumption 1.7 entails non-explosion of the BDI η. Thus, in (1.20) we have

$$0 =: \tau_0 < \tau_1 < \tau_2 ... \uparrow \infty = \tau_\infty \quad \mathbf{P}_{\mathbf{x}}\text{-a.s., } \mathbf{x} \in \mathcal{S}.$$

$\Diamond$

Definition 1.9
Under Assumption 1.7, a finite measure $m(\cdot)$ on $(\mathcal{S}, \mathcal{B}(\mathcal{S}))$ is defined via

$$m(F) := \mathbf{E}_\delta \left(\int_0^R \mathbb{1}_F(\eta_s)\, ds \right), \quad F \in \mathcal{B}(\mathcal{S}). \tag{1.25}$$

$\Diamond$

Note that $m(F)$ gives the expected occupation time of a Borel set $F \in \mathcal{B}(\mathcal{S})$ during one life cycle of the BDI η. Condition (1.24) is sufficient to ensure recurrence of the BDI η in a strong sense, and the measure $m(\cdot)$ defined in (1.25) turns out to be the (essentially unique) invariant measure for η on $(\mathcal{S}, \mathcal{B}(\mathcal{S}))$ as the following proposition shows.

Proposition 1.10
Grant Assumption 1.7. Then, the BDI η is positive recurrent in the sense of Harris, and its invariant measure (unique up to normalisation) coincides with $m(\cdot)$ defined in (1.25).

Proof
We can decompose the trajectory of the BDI η into independent identically distributed life cycles on account of the successive re-entry times into the void configuration δ. Then, by applying the strong law of large numbers (see for instance [33, p. 150]) and (1.24), one can verify one (of several different ones) definition of the Harris recurrence, i.e., for every $\mathbf{x} \in \mathcal{S}$ and $m(F) > 0$

$$\int_0^\infty \mathbb{1}_F(\eta_s)\, ds = \infty \quad \mathbf{P}_{\mathbf{x}}\text{-a.s.}$$

(see for instance [1, p. 24]). Finally, by using the ratio limit theorem (see [1, p. 30]) for Harris recurrent processes, one obtains that the invariant measure coincides with $m(\cdot)$ defined in (1.25). For the detailed proof, see [10, p. 12f].

$\square$

Remark 1.11

1. There is a series representation of the invariant measure $m(\cdot)$ on the configuration space $\mathcal{S}$. For every $F \in \mathcal{B}(\mathcal{S})$ we have

$$m(F) = \sum_{n \in \mathbb{N}_0} \mathbf{E}_\delta \left(\mathbb{1}_{\{\tau_n < R\}} \cdot R_\alpha(\eta_{\tau_n}; F) \right), \tag{1.26}$$

c.f. [10, p. 17].

2. Assumption 1.7 also implies that the void configuration δ is a recurrent atom for the skeleton chain $(\eta_{i\Delta})_{i \in \mathbb{N}_0}$ for every $\Delta > 0$. To prove this, we assume the contrary. Then, there is $\Delta > 0$ such that the probability of the event

$$\Big\{ \exists i_0 \in \mathbb{N}_0 \text{ such that for every } i \geq i_0 \colon \eta_{i\Delta} \notin \delta \Big\}$$

is greater than zero. As δ is a recurrent atom for the time-continuous chain $(\eta_t)_{t \geq 0}$, there is a sequence $(R_n)_{n \in \mathbb{N}}$ of finite successive re-entry times into the void configuration such that $\eta_{R_n} \in \delta$ and $R_n \geq i_0$ for every $n \in \mathbb{N}$. As $\eta_{i\Delta} \notin \delta$ for every $i \geq i_0$, with probability greater than zero immigration occurs during the time $(R_n, R_n + \Delta]$ for every $n \in \mathbb{N}$. Now, for every $n \in \mathbb{N}$ we define

$$A_{n,\Delta} := \Big\{ \text{Immigration occurs during the time } (R_n, R_n + \Delta] \Big\}.$$

Denoting $A^c_{n,\Delta}$ the complement of $A_{n,\Delta}$, it holds

$$\sum_{n \in \mathbb{N}} \mathbf{P}(A^c_{n,\Delta}) = \sum_{n \in \mathbb{N}} \exp(-c\Delta) = \infty$$

and because of the pairwise independence of $(A^c_{n,\Delta})_{n \in \mathbb{N}}$, by the Borell-Cantelli lemma (see for instance [24, p. 53]) we receive

$$\mathbf{P}\big(A_{n,\Delta} \text{ finally} \big) = 0,$$

which is a contradiction. Hence, the skeleton chain $(\eta_{i\Delta})_{i \in \mathbb{N}_0}$ is recurrent in the sense of Harris and its invariant measure $m'(\cdot)$ can be identified with the invariant measure $m(\cdot)$ from the time-continuous Harris chain $(\eta_t)_{t \geq 0}$, according to [1]. Out of the finiteness of $m(\cdot)$, the finiteness of $m'(\cdot)$ immediately follows, i.e., $(\eta_{i\Delta})_{i \in \mathbb{N}_0}$ is positive recurrent in the sense of Harris, too.

$\Diamond$

We introduce some additional notation: For a function $f(\cdot) \colon E \to E$, let $\overline{f}(\cdot) : \mathcal{S} \to E$ denote the function on the configuration space defined by

$$\overline{f}(\mathbf{x}) := \sum_{i=1}^{\ell} f(x^i) \quad \text{if} \quad \mathbf{x} = (x^1, ..., x^\ell) \in E^\ell,\ \ell \in \mathbb{N}, \quad \overline{f}(\delta) := 0.$$

If $f(\cdot)$ is of the form $f(x) = \mathbb{1}_B(x)$ for some Borel set $B \in \mathcal{B}(E)$, we also write

$$\mathbf{x}(B) := \sum_{i=1}^{\ell(\mathbf{x})} \mathbb{1}_B(x^i), \quad \mathbf{x} \in \mathcal{S}, \tag{1.27}$$

for the number of particles in the configuration $\mathbf{x}$ with position in B. This notation is motivated by the measure-valued point of view, where $\mathbf{x}(B)$ is just the total mass of the Borel set B under the finite point measure $\mathbf{x} = \sum_{i=1}^{\ell} \epsilon_{x^i}$.

Definition 1.12

Under Assumption 1.7 and making use of (1.27), *we define a measure* $\overline{m}(\cdot)$ *on* $(E, \mathcal{B}(E))$ *by*

$$\overline{m}(B) := \int_{\mathcal{S}} \mathbf{x}(B)\, m(d\mathbf{x}) = \mathbf{E}_\delta \left(\int_0^R \eta_s(B)\, ds \right), \quad B \in \mathcal{B}(E).$$

The measure $\overline{m}(\cdot)$ *is called the invariant occupation measure or intensity measure of* $m(\cdot)$.

$\diamondsuit$

The measure $\overline{m}(\cdot)$ describes (up to normalization) the expected occupation time of a subset $B \in \mathcal{B}(E)$ by all particles whose life span is contained in one life cycle of η. We emphasise that under Assumption 1.7 alone, it is generally not assured that $\overline{m}(\cdot)$ is a finite measure on $(E, \mathcal{B}(E))$, i.e., finiteness of $\overline{m}(\cdot)$ is a strictly stronger condition than (1.24). Further, finiteness of $\overline{m}(\cdot)$ means

$$\overline{m}(E) = \int_{\mathcal{S}} \ell(\mathbf{x})\, m(d\mathbf{x}) = \sum_{\ell \in \mathbb{N}} \ell \cdot m(E^\ell) < \infty \tag{1.28}$$

and thus concerns the decay of $m(E^\ell)$, as $\ell \to \infty$. In the next section, we will give sufficient conditions in order to achieve finiteness and good properties of $\overline{m}(\cdot)$.

1.3 Hammer's Framework

In his thesis [10], Hammer found sufficient conditions for the existence of a locally bounded and continuous Lebesgue density of the invariant measure $m(\cdot)$. Even though under condition (1.17) (branching particles reproduce at their position of death) a Lebesgue density may exist, it can have, even under best conditions, "strangely shaped" densities. Let us outline this phenomenon, c.f. [10, p. 25f] resp. [13]: Consider a binary branching Brownian motion with immigration in E^d, $d \geq 2$, where particles move independently of each other on Brownian paths. Furthermore, particles branch at a constant rate $\kappa > 0$ and leave either zero or two offspring at their position of death with probability p_0 or $p_2 = 1 - p_0 < \frac{1}{2}$. Immigration occurs at a constant rate $c > 0$ with an immigration distribution having a strictly positive density

$$p(\cdot) \in \mathcal{C}_b^\infty(E^d) := \Big\{ f(\cdot) \colon E^d \to E \,\Big|\, f(\cdot) \text{ is bounded and infinitely often differentiable} \\ \text{with bounded derivates of all orders} \Big\}.$$

First, Hammer shows that both the invariant measure $m(\cdot)$ and the occupation measure $\overline{m}(\cdot)$ are finite. Secondly, he proves that there is a non-negative function $\underline{\gamma}(\cdot)$ on $\mathcal{S}$ and a set

$$\mathcal{N} := \Big\{ \mathbf{x} \in \mathcal{S} \,\big|\, \ell(\mathbf{x}) \geq 2, \exists i \neq j \in \big\{1, ..., \ell(\mathbf{x})\big\} : x^i = x^j \Big\} \tag{1.29}$$

such that $\underline{\gamma}(\cdot)$ is continuous on $\mathcal{N}^c$, has singularities at all points of a non-empty subset $\widetilde{\mathcal{N}} \subseteq \mathcal{N}$ and the invariant density $\gamma(\cdot)$ is minorized by $\underline{\gamma}(\cdot)$. In particular, there can be no version of the density $\gamma(\cdot)$ which is continuous and locally bounded.
In order to avoid this effect, it is reasonable to take Hammer's framework as a basis. In his thesis, Hammer stated two assumptions ([10, p. 30f]) which are sufficient for the existence of a locally bounded and continuous Lebesgue density of $m(\cdot)$. Before we outline these assumptions, remark that $\mathcal{C}_0(E)$ denotes the set of real valued continuous functions $f(\cdot)$ vanishing for $x \to \pm\infty$, i.e.,

$$\mathcal{C}_0(E) := \Big\{ f(\cdot)\colon E \to E \,\big|\, f(\cdot) \text{ is continuous and } \lim_{x\to\infty} f(x) = \lim_{x\to-\infty} f(x) = 0 \Big\}.$$

Assumption 1.13 (Heat Kernel Estimate)
We assume that there exist $t_0 > 0$ and $C > 0$ such that the one-dimensional diffusion with killing rate $\kappa(\cdot)$ and $0 < \underline{\kappa} \leq \kappa(\cdot) \leq \overline{\kappa}$ admits a density function $p_t^{(\kappa)}(x, y)$ being continuous in the variable $y \in E$ and satisfying a heat kernel estimate, i.e.,

$$p_t^{(\kappa)}(x; y) \leq C \cdot \exp(-\underline{\kappa} t) \cdot t^{-\frac{1}{2}} \exp\left(-\frac{(x-y)^2}{2Ct} \right), \quad t \in (0, t_0],\, x, y \in E. \tag{1.30}$$

$\Diamond$

Assumption 1.14 (Absolutely Continuity of Offspring and Immigration Laws)

1. *For all $k \in \mathbb{N}$ we assume the following: We have*

$$Q_k(x; dv^1 ... dv^k) = q_k(x; v^1, ..., v^k)\, dv^1 ... dv^k \quad \text{on} \quad (E^k, \mathcal{B}(E^k))$$

for all $x \in E$, where $q_k(x; \cdot)\colon E^k \to E_+$ is continuous for each fixed $x \in E$. Moreover, there is a function $\hat{q}_k(\cdot) \in \mathcal{C}_0(E) \cap \mathcal{L}^1(E)$ such that for all $x \in E$ and $(v^1, ..., v^k) \in E^k$ it holds

$$q_k(x; v^1, ..., v^k) \leq \prod_{j=1}^{k} \hat{q}_k(x - v^j). \tag{1.31}$$

For $k = 0$ we write

$$q_0(x; \delta) := \hat{q}_0(x) := 1, \quad x \in E.$$

In addition, for the Fourier transform of the upper bound $\hat{q}_k(\cdot)$ we require

$$\mathcal{F}(\hat{q}_k)(\cdot) \in \mathcal{L}^1(E), \quad k \in \mathbb{N}. \tag{1.32}$$

2. *We assume the following: The immigration law $\nu(\cdot)$ can be written as*

$$\nu(dv) = \pi(v)\, dv \quad \textit{on} \quad (E, \mathcal{B}(E)),$$

where $\pi(\cdot)\colon E \to E_+$ is a function in $\mathcal{C}_0(E) \cap \mathcal{L}^1(E)$. Finally, we require

$$\mathcal{F}(\pi)(\cdot) \in \mathcal{L}^1(E). \tag{1.33}$$

$\Diamond$

Using these assumptions, Hammer was able to prove the following theorem.

Theorem 1.15 (Locally Bounded and Continuous Lebesgue Density of $m(\cdot)$)
Under Assumptions 1.13 and 1.14, the invariant measure $m(\cdot)$ on the configuration space $(\mathcal{S}, \mathcal{B}(\mathcal{S}))$ admits a Lebesgue density $\gamma(\cdot) = \left(\gamma^{(\ell)}(\cdot)\right)_{\ell \in \mathbb{N}_0}$ which is locally $\mathcal{C}_0$, i.e.,

$$\gamma^{(\ell)}(\cdot) \in \mathcal{C}_0(E^\ell), \quad \ell \in \mathbb{N}_0.$$

Proof
For the proof, see [10, p. 35f].

$\square$

Remark 1.16

1. At first glance, Assumptions 1.13 and 1.14 may seem artificial. However, considering that as soon as branching particles reproduce at their position of death the Lebesgue density of $m(\cdot)$ can have bad properties, some stronger conditions have to balance this. An example for $Q_k(x, \cdot)$ being a normal distribution with small variance and fulfilling the first assumption in Assumption 1.14 is given in [10, p. 34f].

2. The heat kernel estimation in Assumption 1.13 is based on the idea that the transition density of a one-dimensional diffusion can be estimated by a transition density of a Brownian motion which is
$$\frac{1}{\sqrt{2\pi t}} \exp\left(-\frac{(x-y)^2}{2t}\right).$$
It is a classical result from partial differential equation theory (see for instance [6, p. 228f] resp. [26]) that Assumption 1.13 is fulfilled if the following two points are granted:

   - The diffusion coefficient $\sigma(\cdot)$ is bounded away from zero.
   - The drift coefficient $b(\cdot)$, the diffusion coefficient $\sigma(\cdot)$ and the branching rate $\kappa(\cdot)$ are bounded and continuous in the sense of Hölder.

   Apart from that, the density $p_t^{(\kappa)}(\cdot;\cdot)$ does not vanish if the diffusion coefficient $\sigma(\cdot)$ is bounded away from zero.

3. In [2], [12] and [29], it assumed that $p_1(x) = 0$ for every $x \in E$ since in these frameworks offspring particles start their spatial motion at their parents' position of death, which means that distinguishing between a dying and a non-dying particle would not be possible if $p_1(x) > 0$. By Assumption 1.14, offspring particles are scattered around their parents' position of death $x \in E$ according to a specified law $Q_k(x, \cdot)$, which is why $p_1(x) > 0$ for $x \in E$ is allowed.

4. As the particles move independently of each other, there is a product structure of

$$p_t^\kappa(\mathbf{x};\mathbf{y}) = \prod_{i=1}^{\ell(\mathbf{x})} p_t^{(\kappa)}(x^i; y^i), \quad \mathbf{x}, \mathbf{y} \in \mathcal{S}. \tag{1.34}$$

We want to emphasise this fact because it will be an important tool for our proofs later. Originally, in Hammer's thesis both the drift coefficient and the diffusion coefficient may also depend on the configuration of co-existing particles, i.e., a product structure of $p_t^\kappa(\mathbf{x};\mathbf{y})$ is generally not given. Apart from that, in the original version of Assumption 1.13 Hammer demands that $P_t^\alpha(\mathbf{x};\mathbf{y})$ has a transition density $p_t^\alpha(\mathbf{x};\mathbf{y})$ and he assumes an estimate for $p_t^\alpha(\mathbf{x};\mathbf{y})$ similarly as in (1.30) (whereas C is replaced by constants C_ℓ, $\ell \in \mathbb{N}$). Because of our position-dependent case and the independence of the particles, our assumption reduces as stated in Assumption 1.13.

5. As Hammer mentions in [10, p. 32], under Assumption 1.13 the generalized resolvent kernel $R_\alpha^{(\ell)}(\mathbf{x};\cdot)$ (equation (1.12)) has the density

$$r_\alpha^{(\ell)}(\mathbf{x};\mathbf{y}) := \int_0^\infty p_t^\alpha(\mathbf{x};\mathbf{y})\, dt \tag{1.35}$$

which is integrable in $\mathbf{y}$ for every fixed $\mathbf{x} \in \mathcal{S}$ since (1.13) and (1.20) give

$$R_\alpha(\mathbf{x};\mathcal{S}) = R_\alpha^{(\ell)}(\mathbf{x}; E^\ell) = \mathbf{E}_\mathbf{x}(\tau_1) < \infty.$$

Furthermore, defining for $\mathbf{x}, \mathbf{y} \in \mathcal{S}$

$$r_{\alpha,\varepsilon}^{(\ell)}(\mathbf{x};\mathbf{y}) := \int_0^\varepsilon p_t^\alpha(\mathbf{x};\mathbf{y})\, dt \tag{1.36}$$

and using the Chapman-Kolmogorov identity (see for instance [33, p. 228]), we can rewrite $r_\alpha^{(\ell)}(\mathbf{x};\mathbf{y})$ to

$$r_\alpha^{(\ell)}(\mathbf{x};\mathbf{y}) = r_{\alpha,\varepsilon}^{(\ell)}(\mathbf{x};\mathbf{y}) + \int_{E^\ell} r_\alpha^{(\ell)}(\mathbf{x};\mathbf{z})\, p_\varepsilon^\alpha(\mathbf{z};\mathbf{y})\, d\mathbf{z}. \tag{1.37}$$

By applying this identity, one can show that (1.30) extends to every interval $(0, t_0]$, where $t_0 > 0$, c.f. [10, p. 30]. In this case, the constant $C > 0$ in (1.13) changes and depends on t_0.

6. Using the series representation of the invariant measure $m(\cdot)$ in (1.26), the density of $\gamma(\cdot)$ in restriction to each layer E^ℓ is given by

$$\gamma^{(\ell)}(\cdot) = \sum_{n \in \mathbb{N}_0} \mathbf{E}_\delta \left(\mathbb{1}_{\{\tau_n < R\}} \cdot r_\alpha^{(\ell)}(\eta_{\tau_n}; \cdot) \right). \tag{1.38}$$

   If the diffusion coefficient $\sigma(\cdot)$ is bounded away from zero, $p_t^{(\kappa)}(\cdot, \cdot)$ does not vanish according to the second remark before. Then, $r_\alpha^{(\ell)}(\cdot; \cdot)$ from (1.35) does not vanish either and the density in (1.38) is strictly positive.

7. As we can see from (1.26) and (1.38) combined with (1.37), to ensure that the invariant measure $m(\cdot)$ admits a Lebesgue density $\gamma(\cdot)$ (without any boundedness or continuity properties), Assumption 1.14 is not needed, but rather Assumption 1.13. We mention this fact since it occurs in the next example and in a proof later.

$\diamondsuit$

Example 1.17
For every $\varepsilon > 0$ we define the set of configurations

$$S_\varepsilon := \Big\{ \mathbf{x} \in \mathcal{S} \,\big|\, \ell(\mathbf{x}) \geq 2, \exists i \neq j \in \{1, ..., \ell(\mathbf{x})\} : |x^i - x^j| < \varepsilon \Big\} \tag{1.39}$$

which consists of configurations $\mathbf{x} \in \mathcal{S}$ having at least two components with a distance of less than ε and let

$$D_\varepsilon := \mathcal{S} \backslash S_\varepsilon = \Big\{ \mathbf{x} \in \mathcal{S} \,\big|\, \ell(\mathbf{x}) \geq 2, \forall i \neq j \in \{1, ..., \ell(\mathbf{x})\} : |x^i - x^j| \geq \varepsilon \Big\} \tag{1.40}$$

its complement. Defining

$$\begin{aligned} S_0 := \bigcap_{\varepsilon > 0} S_\varepsilon &= \Big\{ \mathbf{x} \in \mathcal{S} \,\big|\, \ell(\mathbf{x}) \geq 2, \exists i \neq j \in \{1, ..., \ell(\mathbf{x})\} : x^i = x^j \Big\} \\ &= \bigcup_{\ell \geq 2} \Big\{ \mathbf{x} \in E^\ell \,\big|\, \exists i \neq j \colon x^i = x^j \Big\}, \end{aligned} \tag{1.41}$$

S_0 coincides with $\mathcal{N}$ from (1.29) and consists of a union of hyperplanes which each has Lebesgue measure zero. As $m(\cdot)$ admits a Lebesgue density because of Assumption 1.13 and because of descending continuity, it holds

$$m(S_\varepsilon) \xrightarrow{\varepsilon \to 0} m(S_0) = 0.$$

By this procedure, it is not evident how fast the convergence of $m(S_\varepsilon)$ to zero is, i.e., we do not have a rate of convergence. Particularly, in the case where offspring particles start their spatial motion at their parents' position of death, we have mentioned in the beginning of this section that the density of $m(\cdot)$ may take the value $+\infty$ on $\mathcal{N}$. Therefore, this fact can preclude a rate of convergence for $m(S_\varepsilon)$. However, by using assumptions from Hammer's framework, we are able to give a rate of convergence for $m(S_\varepsilon)$ and we will show that it is of order $\mathcal{O}(\varepsilon)$, as $\varepsilon \to 0$. For this, we still need a further assumption which is mentioned herafter.

$\diamondsuit$

In the section before, we have mentioned that we also want to guarantee good properties for the occupation measure $\overline{m}(\cdot)$, i.e., $\overline{m}(\cdot)$ shall be finite and shall have a bounded Lebesgue density on E. For this case, Hammer also stated sufficient conditions (two assumptions) we want to enumerate now, c.f. [10, p. 58f].

Assumption 1.18 (Fixed Bound of Possible Offspring)
We require that Assumption 1.14 holds (except that we do not assume (1.32) *and* (1.33)*). In addition, grant the following: There exists a fixed upper bound $k_0 \in \mathbb{N}$ for the possible number of offspring, i.e., there is $k_0 \in \mathbb{N}$ such that*

$$p_k(x) = 0 \tag{1.42}$$

for every $x \in E$ and every $k > k_0$.

$\Diamond$

Assumption 1.19 (Exponential Decay of $m(E^\ell)$)
We assume exponential decay of $(m(E^\ell))_{\ell \in \mathbb{N}}$, i.e., there exist constants $C > 0$ and $0 < q < 1$ such that

$$m(E^\ell) \le C q^\ell$$

for every $\ell \in \mathbb{N}$.

$\Diamond$

Using these assumptions, Hammer was able to show the following theorem.

Theorem 1.20 (Bounded and Continuous Lebesgue Density of $\overline{m}(\cdot)$)
Grant Assumptions 1.18 and 1.19. Then, the occupation measure $\overline{m}(\cdot)$ on $(E, \mathcal{B}(E))$ admits a Lebesgue density of class $\mathcal{C}_0(E)$.

Proof
For the proof see [10, p. 63f].

$\square$

Remark 1.21

1. Using the density $\gamma(\cdot) = \left(\gamma^{(\ell)}(\cdot)\right)_{\ell \in \mathbb{N}_0}$ of the invariant measure $m(\cdot)$, the density of the occupation measure $\overline{m}(\cdot)$ is given by

$$\frac{d\overline{m}}{d\lambda\!\!\lambda}(\cdot) = \sum_{\ell \in \mathbb{N}} \sum_{i=1}^{\ell} \int_{E^{\ell-1}} \gamma^{(\ell)}\left(x^1, ..., x^{i-1}, \cdot, x^{i+1}, ..., x^\ell\right) dx^1 ... \, dx^{i-1} \, dx^{i+1} \, ... \, dx^\ell.$$

   If the diffusion coefficient $\sigma(\cdot)$ is bounded away from zero, for every $\ell \in \mathbb{N}_0$ the density $\gamma^{(\ell)}(\cdot)$ is strictly positive according to the sixth remark in Remark 1.16. Then, the density $\frac{d\overline{m}}{d\lambda\!\!\lambda}(\cdot)$ does not vanish either.

2. Because of Assumption 1.19, for every $k \in \mathbb{N}$ it holds

$$\int_{\mathcal{S}} \ell^k(\mathbf{x})\, m(d\mathbf{x}) = \sum_{\ell \in \mathbb{N}} \ell^k \cdot m(E^\ell) \le C \cdot \sum_{\ell \in \mathbb{N}} \ell^k \cdot q^\ell < \infty. \tag{1.43}$$

   Setting $k = 1$ implies finiteness of the occupation measure $\overline{m}(\cdot)$, i.e.,

$$\overline{m}(E) = \int_{\mathcal{S}} \ell(\mathbf{x})\, m(d\mathbf{x}) < \infty. \tag{1.44}$$

3. Hammer stated reasonable conditions for fulfilling Assumption 1.19. Assuming additionally to (1.42) that the branching rate and the reproduction rate are constants, i.e.,

$$\kappa(x) \equiv \kappa \quad \text{and} \quad p_k(x) \equiv p_k$$

   for every $x \in E$ with $p_0 < 1$ and "subcritical reproduction", i.e.,

$$\rho := \sum_{k \in \mathbb{N}} k p_k < 1, \tag{1.45}$$

   there exist constants $C > 0$ and $0 < q < 1$ such that

$$m(E^\ell) \le C \cdot \frac{q^\ell}{\ell} \le C q^\ell,$$

   c.f. [10, p. 20].

4. Originally, Hammer demands that there is an increasing sequence of constants

$$1 \le K_1 \le K_2 \le ... \le K_\ell \le K_{\ell+1} \le ... < \infty$$

   growing at most polynomially in $\ell \in \mathbb{N}$ such that the marginals of $p_t^\alpha(\mathbf{x};\mathbf{y})$

$$\int_{E^{\ell-1}} p_t^\alpha(\mathbf{x};\mathbf{y})\, dy^1 ... \, dy^{i-1}\, dy^{i+1} ... \, dy^\ell$$

   can be estimated similarly as in (1.30) with C replaced by K_ℓ, $\ell \in \mathbb{N}$. As in our case particles only depend on the position and move independently of each other, this condition is fulfilled because of the product structure (1.34) and $K_\ell := C$ for every $\ell \in \mathbb{N}$.

5. In [29], Löcherbach proved the existence of a bounded and continuous density of $\overline{m}(\cdot)$ on E^d, $d \in \mathbb{N}$, by assuming uniform ellipticity and strong smoothness resp. boundedness conditions both on the drift and the diffusion coefficients. For this, Löcherbach makes use of Malliavin calculus. However, in her paper she assumes that branching particles reproduce – either zero or two offspring – at their position of death. As we have remarked in the beginning of this section, this may preclude the existence of a locally bounded and continuous density of $m(\cdot)$.
   Furthermore, in [15] Höpfner and Löcherbach discuss results about the existence of a density of $\overline{m}(\cdot)$ and its regularity properties. In their framework, offspring particles start their spatial motion at their parents' position of death, too.

$\Diamond$

Chapter 2

Properties of Branching Diffusions with Immigration

In this chapter, we want to present some results about BDI processes $\eta = (\eta_t)_{t\geq 0}$. Before we show these properties, it is helpful to examine the properties of one-dimensional diffusions, i.e., solutions of equation (1.1)

$$dX_t = b(X_t)\,dt + \sigma(X_t)\,dW_t.$$

Particularly, we are interested in their behaviour during a small time interval $(t, t+\Delta]$, where $t \geq 0$.

2.1 Properties of One-Dimensional Diffusions

In this subsection, let $(\Omega', \mathcal{F}', P)$ be the probability space of the (strong) solution $X = X^{j,\ell}$ of equation (1.1)

$$dX_t = b(X_t)\,dt + \sigma(X_t)\,dW_t$$

and let $(\mathcal{F}'_t)_{t\geq 0}$ be the filtration generated by X.

Assumption 2.1

(a) We assume that both the drift coefficient $b(\cdot)$ and the diffusion coefficient $\sigma(\cdot)$ are bounded, i.e., there are constants $\overline{b}$, $\overline{\sigma} > 0$ such that for every $x \in E$

$$|b(x)| \leq \overline{b} \quad \text{and} \quad |\sigma(x)| \leq \overline{\sigma}.$$

(b) The diffusion coefficient $\sigma(\cdot)$ is bounded away from zero, i.e., there is a constant $\underline{\sigma} > 0$ such that for every $x \in E$

$$0 < \underline{\sigma} \leq \sigma(x).$$

$\Diamond$

Proposition 2.2

Let $X = X^{j,\ell}$ be the solution of (1.1) *and let Assumption 2.1(a) hold. Then, for every $t \geq 0$*

1. $E\left(\left|\int_t^{t+\Delta} b(X_s) - b(X_t)\, ds\right|\right) = \mathcal{O}\big(\Delta^{\frac{3}{2}}\big), \quad as \quad \Delta \to 0.$
2. $E\left(\left|\int_t^{t+\Delta} (X_s - X_t)\, b(X_s)\, ds\right|\right) = \mathcal{O}\big(\Delta^{\frac{3}{2}}\big), \quad as \quad \Delta \to 0.$
3. $E\left(\left|\int_t^{t+\Delta} \sigma^2(X_s) - \sigma^2(X_t)\, ds\right|\right) = \mathcal{O}\big(\Delta^{\frac{3}{2}}\big), \quad as \quad \Delta \to 0.$
4. $E\left(\left|\int_t^{t+\Delta} \sigma(X_s) - \sigma(X_t)\, dW_s\right|\right) = \mathcal{O}(\Delta), \quad as \quad \Delta \to 0.$
5. $\int_t^{t+\Delta} (X_s - X_t) \cdot \sigma(X_s)\, dW_s = \sigma^2(X_t) \cdot \int_t^{t+\Delta} W_s\, dW_s + \mathcal{O}_P\big(\Delta^{\frac{3}{2}}\big), \quad as \quad \Delta \to 0.$

Proof

The following estimations rely on ideas from [8, p. 129f] and [25, p. 356f]. Because of the Markov property, it is enough to consider the case $t = 0$. Let $0 < s \leq \Delta$.

1. Because of the boundedness of $b(\cdot)$, it holds

$$\begin{aligned} |X_s - X_0| &= \Big|\int_0^s \sigma(X_v)\, dW_v + \int_0^s b(X_v)\, dv\Big| \\ &\leq \Big|\int_0^s \sigma(X_v)\, dW_v\Big| + \int_0^s |b(X_v)|\, dv \\ &\leq \sup_{r\in[0,\Delta]} \Big|\int_0^r \sigma(X_v)\, dW_v\Big| + \overline{b} \cdot \Delta. \end{aligned} \tag{2.1}$$

Burkholder-Davis-Gundy's inequality (see [32, p. 93]), the quadratic variation

$$\Big\langle \int_0^{\cdot} \sigma(X_v)\, dW_v \Big\rangle_r = \int_0^r \sigma^2(X_v)\, dv$$

and the boundedness of $\sigma(\cdot)$ imply that for every $p \geq 1$ there is a constant $K_p > 0$ such that

$$\begin{aligned} E\bigg(\Big(\sup_{r\in[0,\Delta]} \Big|\int_0^r \sigma(X_v)\, dW_v\Big|\Big)^p\bigg) &= K_p \cdot E\bigg(\Big(\int_0^\Delta \sigma^2(X_v)\, dv\Big)^{\frac{p}{2}}\bigg) \\ &\leq K_p \cdot \overline{\sigma}^p \cdot \Delta^{\frac{p}{2}}. \end{aligned} \tag{2.2}$$

Therefore, with $p = 1$ in (2.2) we get in (2.1) for sufficiently small $\Delta > 0$

$$\begin{aligned} E\Big(\sup_{s\in[0,\Delta]} |X_s - X_0|\Big) &\leq E\Big(\sup_{r\in[0,\Delta]} \Big|\int_0^r \sigma(X_v)\, dW_v\Big|\Big) + \overline{b} \cdot \Delta \\ &\leq K_1 \cdot E\bigg(\Big(\int_0^\Delta \sigma^2(X_v)\, dv\Big)^{\frac{1}{2}}\bigg) + \overline{b} \cdot \Delta \\ &\leq K_1 \overline{\sigma} \cdot \sqrt{\Delta} + \overline{b} \cdot \Delta \\ &\leq C' \cdot \sqrt{\Delta}, \end{aligned} \tag{2.3}$$

where $C' := \max\big\{K_1\overline{\sigma},\, \overline{b}\big\}$. As $b(\cdot)$ is Lipschitz continuous, we gain with (2.3)

$$\begin{aligned} E\left(\Big|\int_0^\Delta b(X_s)-b(X_0)\,ds\Big|\right) &\le L\cdot E\left(\int_0^\Delta |X_s-X_0|\,ds\right)\\ &\le L\cdot\Delta\cdot E\Big(\sup_{s\in[0,\Delta]}|X_s-X_0|\Big)\\ &\le L\cdot C'\cdot\Delta^{\frac{3}{2}}=\mathcal{O}\big(\Delta^{\frac{3}{2}}\big),\quad\text{as}\quad\Delta\to 0. \end{aligned}\tag{2.4}$$

2. As $b(\cdot)$ is bounded, we receive

$$E\left(\Big|\int_0^\Delta (X_s-X_0)\cdot b(X_s)\,ds\Big|\right)\le \overline{b}\cdot E\left(\int_0^\Delta |X_s-X_0|\,ds\right)$$

and with the same estimation from (2.4) we obtain

$$E\left(\Big|\int_0^\Delta (X_s-X_0)\cdot b(X_s)\,ds\Big|\right)=\mathcal{O}\big(\Delta^{\frac{3}{2}}\big),\quad\text{as}\quad\Delta\to 0.$$

3. As $\sigma(\cdot)$ is bounded and Lipschitz continuous, it holds

$$\begin{aligned} \big|\sigma^2(X_s)-\sigma^2(X_0)\big| &= \big|\sigma(X_s)+\sigma(X_0)\big|\cdot\big|\sigma(X_s)-\sigma(X_0)\big|\\ &\le 2\overline{\sigma}\cdot L\cdot|X_s-X_0|\\ &\le 2\overline{\sigma}\cdot L\cdot\sup_{s\in[0,\Delta]}|X_s-X_0|. \end{aligned}$$

Proceeding as in (2.3) and (2.4), we get

$$E\left(\Big|\int_0^\Delta \sigma^2(X_s)-\sigma^2(X_0)\,ds\Big|\right)=\mathcal{O}\big(\Delta^{\frac{3}{2}}\big),\quad\text{as}\quad\Delta\to 0.$$

4. By (2.1), it holds

$$\begin{aligned} \big(X_s-X_0\big)^2 &\le \left(\sup_{r\in[0,\Delta]}\Big|\int_0^r\sigma(X_v)\,dW_v\Big|+\overline{b}\cdot\Delta\right)^2\\ &=\left(\sup_{r\in[0,\Delta]}\Big|\int_0^r\sigma(X_v)\,dW_v\Big|\right)^2+2\cdot\sup_{r\in[0,\Delta]}\Big|\int_0^r\sigma(X_v)\,dW_v\Big|\cdot\overline{b}\cdot\Delta+\overline{b}^2\cdot\Delta^2. \end{aligned}\tag{2.5}$$

Setting $p=1$ and $p=2$ in (2.2), we receive in (2.5)

$$E\Big(\sup_{s\in[0,\Delta]}\big(X_s-X_0\big)^2\Big)=\mathcal{O}(\Delta),\quad\text{as}\quad\Delta\to 0.\tag{2.6}$$

Now, because of Jensen's inequality (see for instance [33, p. 112]), Itô's isometry (see for instance [32, p. 48]), the Lipschitz continuity of $\sigma(\cdot)$ and (2.6), we finally receive

$$
\begin{aligned}
E\left(\Big|\int_0^\Delta \sigma(X_s)-\sigma(X_0)\,dW_s\Big|\right) &\leq \left(E\left(\Big(\int_0^\Delta \sigma(X_s)-\sigma(X_0)\,dW_s\Big)^2\right)\right)^{\frac{1}{2}} \\
&= \left(E\left(\int_0^\Delta \big(\sigma(X_s)-\sigma(X_0)\big)^2\,ds\right)\right)^{\frac{1}{2}} \\
&\leq L\cdot\left(E\Big(\int_0^\Delta \sup_{s\in[0,\Delta]}\Big(X_s-X_0\Big)^2\,ds\Big)\right)^{\frac{1}{2}} \\
&= L\cdot\left(E\Big(\sup_{s\in[0,\Delta]}\Big(X_s-X_0\Big)^2\Big)\right)^{\frac{1}{2}}\cdot\sqrt{\Delta} \\
&= \mathcal{O}(\Delta), \quad \text{as} \quad \Delta\to 0.
\end{aligned}
$$

5. Defining

$$\bar{X}_s := \int_0^s \sigma(X_v)-\sigma(X_0)\,dW_v + \int_0^s b(X_v)\,dv,$$

we decompose

$$X_s - X_0 = \sigma(X_0)\cdot W_s + \bar{X}_s. \tag{2.7}$$

Because of the previous results, it holds

$$E\Big(\sup_{s\in[0,\Delta]}\bar{X}_s^2\Big) = \mathcal{O}(\Delta^2), \quad \text{as} \quad \Delta\to 0.$$

Using this, by Jensen's inequality, Itô's isometry, Fubini's theorem (see for instance [33, p. 103]) and the boundedness of $\sigma(\cdot)$, we receive

$$
\begin{aligned}
E\left(\Big|\sigma(X_0)\cdot\int_0^\Delta \bar{X}_s\,dW_s\Big|\right) &\leq \overline{\sigma}\cdot\left(\int_0^\Delta E\Big(\sup_{s\in[0,\Delta]}\bar{X}_s^2\Big)\,ds\right)^{\frac{1}{2}} \\
&= \mathcal{O}\big(\Delta^{\frac{3}{2}}\big), \quad \text{as} \quad \Delta\to 0.
\end{aligned} \tag{2.8}
$$

Similarly, we gain

$$
\begin{aligned}
&E\Big(\Big|\int_0^\Delta \big(X_s-X_0\big)\cdot\big(\sigma(X_s)-\sigma(X_0)\big)\,dW_s\Big|\Big) \\
&\qquad\leq L\cdot\left(\int_0^\Delta E\Big(\sup_{s\in[0,\Delta]}\big(X_s-X_0\big)^4\Big)\,ds\right)^{\frac{1}{2}}.
\end{aligned}
$$

Squaring (2.5) and using (2.2) with $p=1$, $p=2$, $p=3$ and $p=4$, it holds

$$E\Big(\sup_{s\in[0,\Delta]}\big(X_s-X_0\big)^4\Big) = \mathcal{O}(\Delta^2), \quad \text{as} \quad \Delta\to 0.$$

Therefore, we receive

$$E\left(\Big|\int_0^\Delta \big(X_s - X_0\big)\cdot\big(\sigma(X_s)-\sigma(X_0)\big)\,dW_s\Big|\right) = \mathcal{O}\big(\Delta^{\frac{3}{2}}\big), \quad \text{as} \quad \Delta\to 0. \tag{2.9}$$

Finally, we obtain by (2.7), (2.8) and (2.9)

$$\begin{aligned}
\int_0^\Delta (X_s - X_0)\cdot\sigma(X_s)\,dW_s &= \sigma(X_0)\cdot\int_0^\Delta (X_s - X_0)\,dW_s \\
&\quad + \int_0^\Delta \big(X_s - X_0\big)\cdot\big(\sigma(X_s)-\sigma(X_0)\big)\,dW_s \\
&= \sigma(X_0)\cdot\int_0^\Delta \big(\sigma(X_0)\cdot W_s + \bar{X}_s\big)\,dW_s \\
&\quad + \int_0^\Delta \big(X_s - X_0\big)\cdot\big(\sigma(X_s)-\sigma(X_0)\big)\,dW_s \\
&= \sigma^2(X_0)\cdot\int_0^\Delta W_s\,dW_s + \sigma(X_0)\cdot\int_0^\Delta \bar{X}_s\,dW_s \\
&\quad + \int_0^\Delta \big(X_s - X_0\big)\cdot\big(\sigma(X_s)-\sigma(X_0)\big)\,dW_s \\
&= \sigma^2(X_0)\cdot\int_0^\Delta W_s\,dW_s + \mathcal{O}_P\big(\Delta^{\frac{3}{2}}\big), \quad \text{as} \quad \Delta\to 0.
\end{aligned}$$

□

The following regression identity for one-dimensional diffusions will play a crucial role in the last chapter in which we apply statistical applications for the BDI process.

Theorem 2.3
Let $X = X^{j,\ell}$ be the solution of (1.1) *and let Assumption 2.1(a) hold. Then, for every $t \geq 0$ it holds*

$$\left(\frac{X_{t+\Delta}-X_t}{\sqrt{\Delta}}\right)^2 = \sigma^2(X_t)\cdot\big(1+U_{(t,\Delta)}\big) + \mathcal{O}_P\big(\sqrt{\Delta}\big), \quad \text{as} \quad \Delta\to 0,$$

where $U_{(t,\Delta)}$ is a $\mathcal{F}'_{t+\Delta}$-measurable random variable being independent of $\mathcal{F}'_t$ and satisfying

$$U_{(t,\Delta)} \overset{d}{=} 2\cdot\int_0^1 W_s\,dW_s$$

for every $t \geq 0$ and $\Delta > 0$.

Proof
As the previous proof, this proof is based on ideas from [8, p. 129f] and [25, p. 356f]. Let $t \geq 0$. We define the functions $\check{b}(x) := 2x\cdot b(x)$ and $\check{\sigma}(x) := 2x\cdot\sigma(x)$. Then, by using Itô's formula (see for instance [32, p. 60]), we receive

$$X_{t+\Delta}^2 - X_t^2 = \int_t^{t+\Delta}\check{b}(X_s)\,ds + \int_t^{t+\Delta}\check{\sigma}(X_s)\,dW_s + \int_t^{t+\Delta}\sigma^2(X_s)\,ds.$$

Using this and

$$X_{t+\Delta} - X_t = \int_t^{t+\Delta} b(X_s)\,ds + \int_t^{t+\Delta} \sigma(X_s)\,dW_s,$$

it holds

$$\begin{aligned}
\big(X_{t+\Delta} - X_t\big)^2 &= \big(X_{t+\Delta}^2 - X_t^2\big) - 2X_t \cdot \big(X_{t+\Delta} - X_t\big) \\
&= \int_t^{t+\Delta} \check{b}(X_s) - \check{b}(X_t)\,ds \\
&\quad + \int_t^{t+\Delta} \check{\sigma}(X_s) - \check{\sigma}(X_t)\,dW_s \\
&\quad + \int_t^{t+\Delta} \sigma^2(X_s) - \sigma^2(X_t)\,ds + \sigma^2(X_t)\cdot\Delta \\
&\quad - \int_t^{t+\Delta} 2X_t \cdot \big(b(X_s) - b(X_t)\big)\,ds \\
&\quad - \int_t^{t+\Delta} 2X_t \cdot \big(\sigma(X_s) - \sigma(X_t)\big)\,dW_s \\
&= \sigma^2(X_t)\cdot\Delta \\
&\quad + \int_t^{t+\Delta} \check{\sigma}(X_s) - \check{\sigma}(X_t)\,dW_s - \int_t^{t+\Delta} 2X_t \cdot \big(\sigma(X_s) - \sigma(X_t)\big)\,dW_s \\
&\quad + \int_t^{t+\Delta} \check{b}(X_s) - \check{b}(X_t)\,ds - \int_t^{t+\Delta} 2X_t \cdot \big(b(X_s) - b(X_t)\big)\,ds \\
&\quad + \int_t^{t+\Delta} \sigma^2(X_s) - \sigma^2(X_t)\,ds \\
&= \sigma^2(X_t)\cdot\Delta \\
&\quad + 2\cdot\int_t^{t+\Delta} (X_s - X_t)\cdot\sigma(X_s)\,dW_s \\
&\quad + 2\cdot\int_t^{t+\Delta} (X_s - X_t)\cdot b(X_s)\,ds \\
&\quad + \int_t^{t+\Delta} \sigma^2(X_s) - \sigma^2(X_t)\,ds.
\end{aligned}$$

Hence, by using the fifth, second and third part of the previous proposition, we receive

$$\left(\frac{X_{t+\Delta} - X_t}{\sqrt{\Delta}}\right)^2 = \sigma^2(X_t)\cdot\big(1 + U_{(t,\Delta)}\big) + \mathcal{O}_P\big(\sqrt{\Delta}\big), \quad \text{as} \quad \Delta \to 0,$$

where $U_{(t,\Delta)}$ is a $\mathcal{F}'_{t+\Delta}$-measurable random variable being independent of $\mathcal{F}'_t$ and satisfying

$$U_{(t,\Delta)} \overset{d}{=} 2\cdot\int_0^1 W_s\,dW_s \quad \text{for every} \quad t \geq 0,\ \Delta > 0.$$

□

A further important tool will be the exponential inequality (by Brandt, c.f. [2, p. 23f]) which states that for given $0 < \lambda < \frac{1}{2}$ a diffusion stays in a neighbourhood of a size Δ^λ during the time period $(t, t+\Delta]$ with high probability.

Lemma 2.4
Let $X = X^{j,\ell}$ be the solution of (1.1) *and let Assumption 2.1(a) hold. Then, for $t \geq 0$, $\Delta > 0$ and $0 < \lambda < \frac{1}{2}$ it holds*

$$P_y\left(\left\{\sup_{s\in[0,\Delta]} |X_{s+t} - X_t| > \Delta^\lambda\right\}\right) \leq c_1 \exp\left(-c_2\Delta^{2\lambda-1}\right) =: g_\lambda(\Delta),$$

where c_1, c_2 are positive constants being independent of t and $y \in E$.

Proof
For the proof see [2, p. 26f].

□

Remark 2.5
There is a version of Lemma 2.4 which does not use the boundedness of the drift coefficient $b(\cdot)$. By using (1.2), there is a constant $K > 0$ such that

$$|b(x)| \leq K \cdot (1 + |x|)$$

for every $x \in E$ (notice that K depends on the Lipschitz-constant L from (1.2)), i.e., the drift coefficient $b(\cdot)$ has a linear growth rate. Then, according to [16, p. 206f], the following holds: For $0 < \lambda < \frac{1}{2}$ and $\frac{1}{2} < \eta' < 1 - \lambda$ there is $\Delta_0 > 0$ such that

$$P_y\left(\left\{\sup_{s\in[0,\Delta]} |X_{s+t} - X_t| > \Delta^\lambda\right\} \cap \left\{|X_t| \leq \Delta^{-\eta'}\right\}\right) \leq g_\lambda(\Delta)$$

for every $t \geq 0$ and every $0 < \Delta < \Delta_0$. However, in the last chapter some applications demand that the drift coefficient is bounded, therefore we restrict ourselves to Assumption 2.1(a).

◊

2.2 Properties of Branching Diffusions

Using the previous lemma, our aim is to show the following: For given $0 < \lambda < \frac{1}{2}$, the probability that a subprocess of a branching diffusion *without* immigration, starting in x^k, leaves a neighbourhood of a size Δ^λ around x^k during a time period of length Δ is of order $\mathcal{O}(\Delta)$, as $\Delta \to 0$. For this, let $s > 0$, $y \in E$ and

$$\left(\eta_t^{(s,y)}\right)_{t\geq s} \tag{2.10}$$

a subprocess of a BDI, which describes a branching diffusion *without* immigration starting at time s with a single particle situated at $y \in E$. Let

$$B_{\Delta^\lambda}(y) := \left\{z \in E \,\middle|\, |z - y| \leq \Delta^\lambda\right\} \tag{2.11}$$

denote the points of E which have a distance to y of less than Δ^λ. We define

$$a^\Delta(s,y) := \left\{ \int_s^{s+\Delta} \eta_u^{(s,y)}(B^c_{\Delta^\lambda}(y))\, du > 0 \right\} \tag{2.12}$$

which represents the event that a subprocess *without* immigration $(\eta_t^{(s,y)})_{t\geq s}$ starts at time s with a single particle located in $y \in E$ and leaves a neighbourhood of a size Δ^λ around y during a short time period $(s, s+\Delta]$. Using (2.12), we define

$$A^\Delta(s,\mathbf{x}) := \bigcup_{k=1}^{\ell(\mathbf{x})} a^\Delta(s,x^k). \tag{2.13}$$

Taking the previous notation, we can now formulate the proposition.

Proposition 2.6

Let Assumption 2.1(a) hold. Then, for $s \geq 0$ and $y \in E$

$$\mathbf{P}\left(a^\Delta(s,y)\right) = \mathcal{O}(\Delta),$$

as $\Delta \to 0$, independently of the initial position.

Proof

The proof is based on [2, p. 32], though it has some differences since branching particles do not reproduce at their position of death. As the subprocess $\eta^{(s,y)}$ from (2.10) is a Markov process, it is enough to consider the case $s = 0$. Up to the first branching at time τ_1, the process $(\eta_t^{(0,y)})_{0\leq t\leq \tau_1}$ is a one-dimensional diffusion, as in (1.1). Remember that according to Assumption 1.3, it holds

$$\mathbf{P}\big(\{\tau_1 < \Delta\}\big) = \mathcal{O}(\Delta), \tag{2.14}$$

as $\Delta \to 0$. We define the event

$$A_y := \left\{ \sup_{0\leq u\leq \Delta\wedge\tau_1} \left|\eta_u^{(0,y)} - y\right| > \Delta^\lambda \right\}$$

which describes that the diffusion $(\eta_t^{(0,y)})_{0\leq t\leq\tau_1}$ leaves the Δ^λ-neighbourhood of its inital position $y \in E$ before time $\Delta \wedge \tau_1$. Then, because of Lemma 2.4 and (2.14) it holds

$$\begin{aligned}
\mathbf{P}\left(a^\Delta(0,y)\right) &= \mathbf{P}\left(a^\Delta(0,y)\cap A_y\right) + \mathbf{P}\left(a^\Delta(0,y)\cap A_y^c\right) \\
&\leq \mathbf{P}(A_y) + \mathbf{P}\left(A_y^c, \tau_1 < \Delta, \bigcup_{k=1}^{\ell(\eta_{\tau_1})} \left\{ \int_{\tau_1}^{\Delta} \eta_u^{(0,\eta^k_{\tau_1})}(B^c_{\Delta^\lambda}(y))\, du > 0 \right\} \right) \\
&\leq \mathbf{P}(A_y) + \mathbf{P}\big(\{\tau_1 < \Delta\}\big) \\
&\leq g_\lambda(\Delta) + \mathcal{O}(\Delta) \\
&= \mathcal{O}(\Delta),
\end{aligned} \tag{2.15}$$

as $\Delta \to 0$.

□

Remark 2.7
In (2.15) we have used the inclusion

$$\left\{A_y^c, \tau_1 < \Delta, \bigcup_{k=1}^{\ell(\eta_{\tau_1})} \left\{ \int_{\tau_1}^{\Delta} \eta_u^{(0,\eta_{\tau_1}^k)}(B_{\Delta^\lambda}^c(y))\, du > 0 \right\} \right\} \subseteq \{\tau_1 < \Delta\} \tag{2.16}$$

in order to show that the probability of (2.16) is of order $\mathcal{O}(\Delta)$, as $\Delta \to 0$. At first glance, this estimation seems rough because we have not used anything about the spatial offspring distribution and we have only applied the fact that branching events occur with probability of order $\mathcal{O}(\Delta)$ in a time period of length Δ, as $\Delta \to 0$. In his thesis, Brandt was able to state an exponential inequality (similar to Lemma 2.4) for the probability of the left side of (2.16), assuming that "subcritical reproduction" (see (1.45)) holds and that offspring particles start their spatial motion at their parents' position of death, c.f. [2, p. 32f]. However, by granting Assumptions 1.14 and 1.18 from Hammer's framework, we are not able to reach such good estimations because the probability that at the branching time τ_1 every offspring starts its motion in a neighbourhood of length Δ^λ around y (conditioned on the event $A_y^c \cap \{\tau_1 < \Delta\}$) is of order $\mathcal{O}(\Delta^\lambda)$, as $\Delta \to 0$, as we will prove below. Hence, we receive for the probability that at the branching time τ_1 at least one offspring starts its motion outside of $B_{\Delta^\lambda}(y)$ (conditioned on the event $A_y^c \cap \{\tau_1 < \Delta\}$)

$$\begin{aligned}
&\mathbf{P}\left(\bigcup_{k=1}^{\ell(\eta_{\tau_1})} \left\{ \int_{\tau_1}^{\Delta} \eta_u^{(0,\eta_{\tau_1}^k)}(B_{\Delta^\lambda}^c(y))\, du > 0 \right\} \,\middle|\, A_y^c \cap \{\tau_1 < \Delta\} \right) \\
&\geq \mathbf{P}\left(\bigcup_{k=1}^{\ell(\eta_{\tau_1})} \left\{ \eta_{\tau_1}^k \in B_{\Delta^\lambda}^c(y) \right\} \,\middle|\, A_y^c \cap \{\tau_1 < \Delta\} \right) \\
&= 1 - \mathbf{P}\left(\bigcap_{k=1}^{\ell(\eta_{\tau_1})} \left\{ \eta_{\tau_1}^k \in B_{\Delta^\lambda}(y) \right\} \,\middle|\, A_y^c \cap \{\tau_1 < \Delta\} \right) \\
&\geq 1 - \mathcal{O}(\Delta^\lambda),
\end{aligned} \tag{2.17}$$

as $\Delta \to 0$. So, inclusion (2.16) is not a severe restriction. We want to finish this remark by showing

$$\mathbf{P}\left(\bigcap_{k=1}^{\ell(\eta_{\tau_1})} \left\{ \eta_{\tau_1}^k \in B_{\Delta^\lambda}(y) \right\} \,\middle|\, A_y^c \cap \{\tau_1 < \Delta\} \right) = \mathcal{O}(\Delta^\lambda),$$

as $\Delta \to 0$. For this, w.l.o.g. we assume $\ell(\eta_{\tau_1}) > 0$. Then, according to the absolutely continuity of the offspring law (Assumption 1.14) and using that there cannot exist more than k_0 offspring (Assumption 1.18), it holds

$$\begin{aligned}
&\mathbf{P}\left(\bigcap_{k=1}^{\ell(\eta_{\tau_1})} \left\{ \eta_{\tau_1}^k \in B_{\Delta^\lambda}(y) \right\} \,\middle|\, A_y^c \cap \{\tau_1 < \Delta\} \right) \\
&\qquad = \mathbb{1}_{B_{\Delta^\lambda}(y)}(\eta_{\tau_1^-}) \cdot \sum_{k=1}^{k_0} p_k(\eta_{\tau_1^-}) \cdot Q_k\left(\eta_{\tau_1^-}; \left[-\Delta^\lambda + y; y + \Delta^\lambda\right]^k\right)
\end{aligned}$$

$$
\begin{aligned}
&\leq \mathbb{1}_{B_{\Delta^\lambda}(y)}(\eta_{\tau_1^-}) \cdot \sum_{k=1}^{k_0} p_k(\eta_{\tau_1^-}) \cdot \prod_{i=1}^{k} \int_{-\Delta^\lambda+y}^{y+\Delta^\lambda} \hat{q}_k\left(\eta_{\tau_1^-} - v^i\right) dv^i \\
&\leq \mathbb{1}_{B_{\Delta^\lambda}(y)}(\eta_{\tau_1^-}) \cdot \sum_{k=1}^{k_0} p_k(\eta_{\tau_1^-}) \cdot \prod_{i=1}^{k} \left(2\Delta^\lambda \cdot ||\hat{q}_k||_\infty\right) \\
&\leq \sum_{k=1}^{k_0} \Delta^{\lambda k} \cdot 2^k \cdot ||\hat{q}_k||_\infty^k \\
&\leq \Delta^\lambda \cdot 2^{k_0} \cdot k_0 \cdot \max_{1\leq k\leq k_0} \left\{||\hat{q}_k||_\infty^k\right\} \\
&= \mathcal{O}(\Delta^\lambda), \qquad (2.18)
\end{aligned}
$$

as $\Delta \to 0$.

$\Diamond$

Using the previous proposition, we have the following corollary.

Corollary 2.8

Let Assumption 2.1(a) hold. Then, there is a constant $C > 0$ such that for sufficiently small $\Delta > 0$ and for every $s \geq 0$, $\mathbf{x} \in \mathcal{S}$

$$\mathbf{P}\left(A^\Delta(s, \mathbf{x})\right) \leq C\Delta \cdot \ell(\mathbf{x}).$$

Proof

This is a direct consequence of the definition of $A^\Delta(s, \mathbf{x})$ from (2.13) and Proposition 2.6.

$\Box$

Now, we introduce some notation. Given $\eta_{i\Delta}$, we write $\eta_{(i+1)\Delta}^{(i\Delta,\eta_{i\Delta}^k)}$ for the positions of particles belonging to $\eta_{(i+1)\Delta}$ and being offspring of a particle situated at $\eta_{i\Delta}^k$, i.e.,

$$\eta_{(i+1)\Delta}^{(i\Delta,\eta_{i\Delta}^k)} := \Big\{\eta_{(i+1)\Delta}^j \in \eta_{(i+1)\Delta} \,\Big|\, 1 \leq j \leq \ell(\eta_{(i+1)\Delta}),\ \text{the particle situated at } \eta_{(i+1)\Delta}^j \text{ is the offspring of a particle situated at } \eta_{i\Delta}^k\Big\} \qquad (2.19)$$

(if there are no offspring, set $\eta_{(i+1)\Delta}^{(i\Delta,\eta_{i\Delta}^k)} := \emptyset$). Moreover, $\eta_{(i+1)\Delta}^I$ denotes the positions of all particles belonging to $\eta_{(i+1)\Delta}$ and being offspring of particles which have immigrated during the time interval $(i\Delta, (i+1)\Delta]$ (if there are no immigrants, set $\eta_{(i+1)\Delta}^I := \emptyset$), i.e.,

$$\eta_{(i+1)\Delta}^I := \Big\{\eta_{(i+1)\Delta}^j \in \eta_{(i+1)\Delta} \,\Big|\, 1 \leq j \leq \ell(\eta_{(i+1)\Delta}),\ \text{the particle situated at } \eta_{(i+1)\Delta}^j \text{ is the offspring of a particle which immigrated during } (i\Delta, (i+1)\Delta]\Big\}. \qquad (2.20)$$

Using this notation, we can split each configuration $\eta_{(i+1)\Delta}$ into two disjoint parts, namely

$$\eta_{(i+1)\Delta} = \bigcup_{k=1}^{\ell(\eta_{i\Delta})} \eta_{(i+1)\Delta}^{(i\Delta,\eta_{i\Delta}^k)} \mathbin{\dot\cup} \eta_{(i+1)\Delta}^I. \qquad (2.21)$$

Furthermore, we remember the set of configurations with close and far neighbours from (1.39) and (1.40). In our case, we set $\varepsilon := 2\Delta^\lambda$ resp. $\varepsilon := 4\Delta^\lambda$ and consider

$$S_{2\Delta^\lambda} = \Big\{ \mathbf{x} \in \mathcal{S} \,\big|\, \ell(\mathbf{x}) \geq 2, \exists i \neq j \in \big\{1, ..., \ell(\mathbf{x})\big\} : |x^i - x^j| < 2\Delta^\lambda \Big\}$$

resp.

$$D_{4\Delta^\lambda} = \Big\{ \mathbf{x} \in \mathcal{S} \,\big|\, \ell(\mathbf{x}) \geq 2, \forall i \neq j \in \big\{1, ..., \ell(\mathbf{x})\big\} : |x^i - x^j| \geq 4\Delta^\lambda \Big\}.$$

Proposition 2.9

Let Assumption 2.1(a) hold. Then, there is a constant $C > 0$ such that for sufficiently small $\Delta > 0$ and for every $\mathbf{x} \in \mathcal{S}$

$$\mathbb{1}_{D_{4\Delta^\lambda}}(\mathbf{x}) \cdot \mathbf{P}_{\mathbf{x}}\Big(\big\{\eta_\Delta \in S_{2\Delta^\lambda}\big\}\Big) \leq C\Delta \cdot \ell(\mathbf{x}).$$

Proof

Depending on the event (2.13), for $\mathbf{x} \in \mathcal{S}$ it holds

$$\begin{aligned} \mathbb{1}_{D_{4\Delta^\lambda}}(\mathbf{x}) \cdot \mathbf{P}_{\mathbf{x}}\Big(\big\{\eta_\Delta \in S_{2\Delta^\lambda}\big\}\Big) &= \mathbb{1}_{D_{4\Delta^\lambda}}(\mathbf{x}) \cdot \mathbf{P}_{\mathbf{x}}\Big(\big\{\eta_\Delta \in S_{2\Delta^\lambda}\big\} \cap A^\Delta(0, \mathbf{x})\Big) \\ &\quad + \mathbb{1}_{D_{4\Delta^\lambda}}(\mathbf{x}) \cdot \mathbf{P}_{\mathbf{x}}\Big(\big\{\eta_\Delta \in S_{2\Delta^\lambda}\big\} \cap \big(A^\Delta(0, \mathbf{x})\big)^c\Big). \end{aligned} \tag{2.22}$$

The first term can be estimated by Corollary 2.8, i.e., there is a constant $C' > 0$ such that

$$\mathbf{P}_{\mathbf{x}}\Big(\big\{\eta_\Delta \in S_{2\Delta^\lambda}\big\} \cap A^\Delta(0, \mathbf{x})\Big) \leq C'\Delta \cdot \ell(\mathbf{x}) \tag{2.23}$$

for sufficiently small $\Delta > 0$. Consider the second term in (2.22). Because of the event $\big(A^\Delta(0, \mathbf{x})\big)^c$, every subprocess

$$\big(\eta_t^{(0,x^k)}\big)_{t \geq 0}$$

that starts in x^k (with $\mathbf{x} \in D_{4\Delta^\lambda}$) at time zero stays in a neighbourhood of a size Δ^λ around x^k during the time $[0, \Delta)$. As every particle belonging to $\mathbf{x}$ has a distance of more than $4\Delta^\lambda$ to its adjacent particles and η_Δ has at least two particles with a distance of less than $2\Delta^\lambda$, immigration must occur during the time $[0, \Delta)$. Denoting τ_1 the first immigration time after time zero, this means

$$\begin{aligned} &\mathbb{1}_{D_{4\Delta^\lambda}}(\mathbf{x}) \cdot \mathbf{P}_{\mathbf{x}}\Big(\big\{\eta_\Delta \in S_{2\Delta^\lambda}\big\} \cap \big(A^\Delta(0, \mathbf{x})\big)^c\Big) \\ &\qquad \leq \mathbb{1}_{D_{4\Delta^\lambda}}(\mathbf{x}) \cdot \mathbf{P}_{\mathbf{x}}\Big(\big\{\eta_\Delta^I \neq \emptyset\big\} \cap \big\{\tau_1 < \Delta\big\}\Big). \end{aligned} \tag{2.24}$$

According to (1.7), the probability that immigration occurs in a time period of length Δ is of order $\mathcal{O}(\Delta)$, as $\Delta \to 0$. This is why there is a constant $C'' > 0$ such that for sufficiently small $\Delta > 0$

$$\mathbb{1}_{D_{4\Delta^\lambda}}(\mathbf{x}) \cdot \mathbf{P}_{\mathbf{x}}\Big(\big\{\eta_\Delta \in S_{2\Delta^\lambda}\big\} \cap \big(A^\Delta(0, \mathbf{x})\big)^c\Big) \leq C''\Delta. \tag{2.25}$$

Altogether, using (2.23) and (2.25), we receive for sufficiently small $\Delta > 0$

$$\begin{aligned}
\mathbb{1}_{D_{4\Delta^\lambda}}(\mathbf{x}) \cdot \mathbf{P}_{\mathbf{x}}\Big(\{\eta_\Delta \in S_{2\Delta^\lambda}\}\Big) &= \mathbb{1}_{D_{4\Delta^\lambda}}(\mathbf{x}) \cdot \mathbf{P}_{\mathbf{x}}\Big(\{\eta_\Delta \in S_{2\Delta^\lambda}\} \cap A^\Delta(0,\mathbf{x})\Big) \\
&\quad + \mathbb{1}_{D_{4\Delta^\lambda}}(\mathbf{x}) \cdot \mathbf{P}_{\mathbf{x}}\Big(\{\eta_\Delta \in S_{2\Delta^\lambda}\} \cap \big(A^\Delta(0,\mathbf{x})\big)^c\Big) \\
&\le C'\Delta \cdot \ell(\mathbf{x}) + C''\Delta \\
&\le C'''\Delta \cdot \big(\ell(\mathbf{x}) + 1\big) \\
&\le C\Delta \cdot \ell(\mathbf{x}),
\end{aligned}$$

where $C''' := \max\{C', C''\} > 0$ and $C := 2C''' > 0$.

□

Remark 2.10
In the previous proposition, we have not mentioned the position where the immigrated particle starts its motion. If we grant Assumption 1.14, the immigration law has a Lebesgue density in $\mathcal{C}_0(E) \cap \mathcal{L}^1(E)$. Considering this, we could get a better estimation in (2.25), namely that (2.25) is of order $\mathcal{O}(\Delta^{1+\lambda})$, as $\Delta \to 0$. However, since the first term of (2.22) is smaller than or equal to $C'\Delta \cdot \ell(\mathbf{x})$ (and this estimation is not that rough according to Remark 2.7), the order of (2.22) does not really improve even if (2.25) is of order $\mathcal{O}(\Delta^{1+\lambda})$, as $\Delta \to 0$. This is why we discontinue inquiry into this line.

◊

As already mentioned in Example 1.17, we want to conclude this chapter by giving a rate of convergence for $m(S_\varepsilon)$, as $\varepsilon \to 0$, where

$$S_\varepsilon = \Big\{\mathbf{x} \in \mathcal{S} \,\big|\, \ell(\mathbf{x}) \ge 2, \exists i \ne j \in \big\{1, ..., \ell(\mathbf{x})\big\} : |x^i - x^j| < \varepsilon\Big\}$$

is the set of configurations in $\mathcal{S}$ in which at least two components have a distance of less than $\varepsilon > 0$. In this chapter, this rate of convergence is our main result and it will be an important tool for the next proofs.

Theorem 2.11
Grant Assumptions 1.13, 1.19 and 2.1(a). Then, it holds

$$m(S_\varepsilon) = \mathcal{O}(\varepsilon), \quad \text{as} \quad \varepsilon \to 0.$$

Proof
According to the seventh remark in Remark 1.16, we know that a sufficient criterion for a Lebesgue density of the invariant measure $m(\cdot)$ is Assumption 1.13. Therefore, it holds

$$m(S_\varepsilon) = \sum_{\ell \in \mathbb{N}_0} \int_{S_\varepsilon \cap E^\ell} \gamma^{(\ell)}(\mathbf{x})\, d\mathbf{x}, \tag{2.26}$$

where $\gamma^{(\ell)}(\cdot)$ is given by equation (1.38)

$$\gamma^{(\ell)}(\mathbf{x}) = \sum_{n\in\mathbb{N}_0} \mathbf{E}_\delta \left(\mathbb{1}_{\{\tau_n<R\}} \cdot r_\alpha^{(\ell)}(\eta_{\tau_n};\mathbf{x}) \right) \tag{2.27}$$

with

$$r_\alpha^{(\ell)}(\eta_{\tau_n};\mathbf{x}) = r_{\alpha,1}^{(\ell)}(\eta_{\tau_n};\mathbf{x}) + \int_{E^\ell} r_\alpha^{(\ell)}(\eta_{\tau_n};\mathbf{z})\, p_1^\alpha(\mathbf{z};\mathbf{x})\, d\mathbf{z} \tag{2.28}$$

according to (1.37). Because of Fubini's theorem, (2.27) and (2.28), it holds

$$\begin{aligned} \int\limits_{S_\varepsilon\cap E^\ell} \gamma^{(\ell)}(\mathbf{x})\, d\mathbf{x} &= \sum_{n\in\mathbb{N}_0} \mathbf{E}_\delta \Bigg(\mathbb{1}_{\{\tau_n<R\}} \cdot \int\limits_{S_\varepsilon\cap E^\ell} r_{\alpha,1}^{(\ell)}(\eta_{\tau_n};\mathbf{x})\, d\mathbf{x} \Bigg) \\ &\quad + \sum_{n\in\mathbb{N}_0} \mathbf{E}_\delta \Bigg(\mathbb{1}_{\{\tau_n<R\}} \cdot \int\limits_{S_\varepsilon\cap E^\ell} d\mathbf{x} \int_{E^\ell} r_\alpha^{(\ell)}(\eta_{\tau_n};\mathbf{z})\, p_1^\alpha(\mathbf{z};\mathbf{x})\, d\mathbf{z} \Bigg). \end{aligned} \tag{2.29}$$

Now, we estimate the single terms which appear in (2.29).

1. By applying (1.36) and Fubini's theorem, it holds

$$\int\limits_{S_\varepsilon\cap E^\ell} r_{\alpha,1}^{(\ell)}(\eta_{\tau_n};\mathbf{x})\, d\mathbf{x} = \int_0^1 dt \int\limits_{S_\varepsilon\cap E^\ell} p_t^\alpha(\eta_{\tau_n};\mathbf{x})\, d\mathbf{x}.$$

Using the identity in (1.16) and (1.34), we receive

$$\int\limits_{S_\varepsilon\cap E^\ell} p_t^\alpha(\eta_{\tau_n};\mathbf{x})\, d\mathbf{x} \le \int\limits_{S_\varepsilon\cap E^\ell} \prod_{i=1}^{\ell} p_t^{(\kappa)}(\eta_{\tau_n}^i;x^i)\, dx^1 ... dx^\ell \tag{2.30}$$

(remember that we write x^k for a component of the configuration $\mathbf{x}\in\mathcal{S}$, where $1\le k\le \ell(\mathbf{x})$). For fixed $\ell\in\mathbb{N}$, in the set $S_\varepsilon\cap E^\ell$ there are $\binom{\ell}{2}$ possibilities to arrange components $i\neq j$ with $|x^i-x^j|<\varepsilon$. Using this, Fubini's theorem and the heat kernel estimate for a single motion with branching (according to the fifth remark in Remark 1.16, w.l.o.g. we may set $t_0 := 1$ in (1.30), i.e., there is is a constant $C_1>0$ such that

$$p_t^{(\kappa)}(\eta_{\tau_n}^1;x) \le C_1\cdot t^{-\frac{1}{2}} \exp\left(-\frac{(\eta_{\tau_n}^1 - x)^2}{2C_1 t} \right)$$

for every $0<t\le 1$), we can estimate (2.30) by

$$\begin{aligned} &\int\limits_{S_\varepsilon\cap E^\ell} \prod_{i=1}^{\ell} p_t^{(\kappa)}(\eta_{\tau_n}^i;x^i)\, dx^1 ... dx^\ell \\ &\le \binom{\ell}{2} \int_E p_t^{(\kappa)}(\eta_{\tau_n}^\ell;x^\ell)\, dx^\ell ... \int_E p_t^{(\kappa)}(\eta_{\tau_n}^2;x^2)\, dx^2 \int\limits_{-\varepsilon+x^2}^{x^2+\varepsilon} p_t^{(\kappa)}(\eta_{\tau_n}^1;x^1)\, dx^1 \\ &\le C_1\cdot\sqrt{2\pi C_1}\cdot \binom{\ell}{2} \int_E p_t^{(\kappa)}(\eta_{\tau_n}^\ell;x^\ell)\, dx^\ell ... \int_E p_t^{(\kappa)}(\eta_{\tau_n}^2;x^2)\, dx^2 \\ &\qquad\qquad \times \int\limits_{-\varepsilon+x^2}^{x^2+\varepsilon} \frac{1}{\sqrt{2\pi C_1 t}} \exp\left(-\frac{(\eta_{\tau_n}^1 - x^1)^2}{2C_1 t} \right) dx^1. \end{aligned} \tag{2.31}$$

For $t > 0$ the function

$$\varphi_{\eta^1_{\tau_n},C_1 t}(x) := \frac{1}{\sqrt{2\pi C_1 t}} \exp\left(-\frac{(\eta^1_{\tau_n} - x)^2}{2C_1 t}\right)$$

is the density of a normal distribution with expectation $\eta^1_{\tau_n}$ and variance $C_1 t$. Hence, for every $x \in E$ it holds

$$\int\limits_{-\varepsilon+x}^{x+\varepsilon} \frac{1}{\sqrt{2\pi C_1 t}} \exp\left(-\frac{(\eta^1_{\tau_n} - y)^2}{2C_1 t}\right) dy \leq 2\varepsilon \cdot \frac{1}{\sqrt{2\pi C_1 t}}. \tag{2.32}$$

Furthermore, for every $1 \leq i \leq \ell$ it holds

$$\int_E p_t^{(\kappa)}(\eta^i_{\tau_n}, x)\, dx \leq 1. \tag{2.33}$$

Sticking together (2.32) and (2.33), we receive in (2.31)

$$\begin{aligned}
&C_1 \cdot \sqrt{2\pi C_1} \cdot \binom{\ell}{2} \int_E p_t^{(\kappa)}(\eta^\ell_{\tau_n}; x^\ell)\, dx^\ell ... \int_E p_t^{(\kappa)}(\eta^2_{\tau_n}; x^2)\, dx^2 \\
&\qquad\qquad \times \int\limits_{-\varepsilon+x^2}^{x^2+\varepsilon} \frac{1}{\sqrt{2\pi C_1 t}} \exp\left(-\frac{(\eta^1_{\tau_n} - x^1)^2}{2C_1 t}\right) dx^1 \\
&\leq C_1 \cdot \binom{\ell}{2} \cdot 2\varepsilon \cdot t^{-\frac{1}{2}} \cdot \\
&\leq C_1 \cdot \ell^2 \cdot \varepsilon \cdot t^{-\frac{1}{2}}.
\end{aligned} \tag{2.34}$$

Hence, combining (2.30) with (2.34), we have

$$\int\limits_{S_\varepsilon \cap E^\ell} p_t^\alpha(\eta_{\tau_n}; \mathbf{x})\, d\mathbf{x} \leq C_1 \cdot \ell^2 \cdot \varepsilon \cdot t^{-\frac{1}{2}}. \tag{2.35}$$

Thus, we obtain in (2.35)

$$\begin{aligned}
\int_0^1 dt \int\limits_{S_\varepsilon \cap E^\ell} p_t^\alpha(\eta_{\tau_n}; \mathbf{x})\, d\mathbf{x} &\leq C_1 \cdot \ell^2 \cdot \varepsilon \cdot \int_0^1 t^{-\frac{1}{2}}\, dt \\
&= 2C_1 \cdot \ell^2 \cdot \varepsilon
\end{aligned} \tag{2.36}$$

2. Now, we want to estimate the expression

$$\int\limits_{S_\varepsilon \cap E^\ell} d\mathbf{x} \int_{E^\ell} r_\alpha^{(\ell)}(\eta_{\tau_n}; \mathbf{z})\, p_1^\alpha(\mathbf{z}; \mathbf{x})\, d\mathbf{z} \tag{2.37}$$

from equation (2.29) similarly as above. We rewrite (2.37) by Fubini's theorem to

$$\int\limits_{S_\varepsilon \cap E^\ell} d\mathbf{x} \int_{E^\ell} r_\alpha^{(\ell)}(\eta_{\tau_n}; \mathbf{z})\, p_1^\alpha(\mathbf{z}; \mathbf{x})\, d\mathbf{z} = \int_{E^\ell} r_\alpha^{(\ell)}(\eta_{\tau_n}; \mathbf{z})\, d\mathbf{z} \int\limits_{S_\varepsilon \cap E^\ell} p_1^\alpha(\mathbf{z}; \mathbf{x})\, d\mathbf{x}. \tag{2.38}$$

Proceeding as in the lines before (2.35), we receive with $t = 1$

$$\int\limits_{S_\varepsilon \cap E^\ell} p_1^\alpha(\mathbf{z};\mathbf{x})\, d\mathbf{x} \leq C_1 \cdot \ell^2 \cdot \varepsilon. \tag{2.39}$$

Hence, (2.38) can be estimated by

$$\int\limits_{S_\varepsilon \cap E^\ell} d\mathbf{x} \int_{E^\ell} r_\alpha^{(\ell)}(\eta_{\tau_n};\mathbf{z})\, p_1^\alpha(\mathbf{z};\mathbf{x})\, d\mathbf{z} \leq C_1 \cdot \ell^2 \cdot \varepsilon \cdot \int_{E^\ell} r_\alpha^{(\ell)}(\eta_{\tau_n};\mathbf{z})\, d\mathbf{z}. \tag{2.40}$$

We remember (1.14) which states an asymptotic behaviour for the occupation times of the killed ℓ-particle motion, i.e.,

$$\int_{E^\ell} r_\alpha^{(\ell)}(\eta_{\tau_n};\mathbf{z})\, d\mathbf{z} = R_\alpha^{(\ell)}(\eta_{\tau_n};E^\ell) \asymp \frac{1}{\ell}, \quad \ell \to \infty. \tag{2.41}$$

Thus, there is a constant $C_1' > 0$ such that (2.40) becomes

$$\int\limits_{S_\varepsilon \cap E^\ell} d\mathbf{x} \int_{E^\ell} r_\alpha^{(\ell)}(\eta_{\tau_n};\mathbf{z})\, p_1^\alpha(\mathbf{z};\mathbf{x})\, d\mathbf{z} \leq C_1' \cdot \ell \cdot \varepsilon \tag{2.42}$$

for sufficiently large ℓ.

3. According to (1.26), it holds

$$m(E^\ell) = \sum_{n\in\mathbb{N}_0} \mathbf{E}_\delta \left(\mathbb{1}_{\{\tau_n < R\}} \cdot \mathbb{1}_{\left\{\eta_{\tau_n} \in E^\ell\right\}} \cdot R_\alpha^{(\ell)}(\eta_{\tau_n};E^\ell) \right). \tag{2.43}$$

Combining (2.43) with (2.41), we achieve that

$$\sum_{n\in\mathbb{N}_0} \mathbf{E}_\delta \left(\mathbb{1}_{\{\tau_n < R\}} \cdot \mathbb{1}_{\left\{\eta_{\tau_n} \in E^\ell\right\}} \right) \asymp \ell \cdot m(E^\ell), \quad \ell \to \infty. \tag{2.44}$$

As we assume exponential decay of $(m(E^\ell))_{\ell\in\mathbb{N}}$ (Assumption 1.19), there exist $C' > 0$ and $0 < q < 1$ such that

$$m(E^\ell) \leq C' q^\ell$$

and in (2.44)

$$\sum_{n\in\mathbb{N}_0} \mathbf{E}_\delta \left(\mathbb{1}_{\{\tau_n < R\}} \cdot \mathbb{1}_{\left\{\eta_{\tau_n} \in E^\ell\right\}} \right) \leq C' \cdot \ell \cdot q^\ell \tag{2.45}$$

for sufficiently large ℓ.

4. Now, we combine the results. In (2.29) we have, because of (2.36), (2.42) and (2.45)

$$\begin{aligned} \int\limits_{S_\varepsilon \cap E^\ell} \gamma^{(\ell)}(\mathbf{x})\, d\mathbf{x} \leq{}& C' \cdot \ell \cdot q^\ell \cdot 2C_1 \cdot \ell^2 \cdot \varepsilon \\ &+ C' \cdot \ell \cdot q^\ell \cdot C_1' \cdot \ell \cdot \varepsilon \end{aligned} \tag{2.46}$$

resp. if we set $C'' := \max\left\{2C_1 \cdot C',\, C' \cdot C_1'\right\}$ we receive in (2.46)

$$\int\limits_{S_\varepsilon \cap E^\ell} \gamma^{(\ell)}(\mathbf{x})\, d\mathbf{x} \leq C'' \cdot \ell^3 \cdot q^\ell \cdot \varepsilon \tag{2.47}$$

for sufficiently large ℓ. The right-hand side of (2.47) is summable in $\ell \in \mathbb{N}_0$, i.e.,

$$C''' := \sum_{\ell \in \mathbb{N}_0} \ell^3 \cdot q^\ell < \infty. \tag{2.48}$$

Writing $C := C'' \cdot C''' > 0$, we get by equations (2.47) and (2.48), regarding to our expression (2.26) in the very beginning,

$$\begin{aligned} m(S_\varepsilon) &= \sum_{\ell \in \mathbb{N}_0} \int\limits_{S_\varepsilon \cap E^\ell} \gamma^{(\ell)}(\mathbf{x})\, d\mathbf{x} \\ &\leq C'' \cdot C''' \cdot \varepsilon \\ &= C \cdot \varepsilon, \end{aligned} \tag{2.49}$$

i.e.,

$$m(S_\varepsilon) = \mathcal{O}(\varepsilon),$$

as $\varepsilon \to 0$. Our proof is complete.

□

Remark 2.12

1. Regarding (2.32), we notice that because of this inequality we finally obtain a rate of order $\mathcal{O}(\varepsilon)$, as $\varepsilon \to 0$. If it were possible to estimate (2.32) more accurately, we could attain a better order. However, a Taylor expansion on the left side of (2.32) shows that the order of this expression cannot be improved. This is why we are not able to get a better rate of convergence than $\mathcal{O}(\varepsilon)$, as $\varepsilon \to 0$.

2. As we can see in the previous proof, Assumptions 1.13 and 1.19 from Hammer's framework play a crucial role in order to obtain a rate of convergence. However, we do not make use of Assumptions 1.14 or 1.18 from Hammer's framework.

◊

Chapter 3

A Reconstruction Algorithm for Branching Diffusions with Immigration

In this chapter, our aim is to reconstruct the underlying trajectory of a BDI assuming that we observe the BDI process at discrete points in time. This reconstruction algorithm connects ideas from Brandt's algorithm ([2, p. 39]) to some assumptions from Hammer's framework (see subsection 1.3).
Before we can specify this algorithm, we will introduce some notation and want to remark that we write $\Delta := \Delta_n := 1/n$, where $n \in \mathbb{N}$ and $T := T_\Delta > 0$ for a time horizon which may depend on $\Delta > 0$ (if it depends on Δ, then $T_\Delta \to \infty$ and $T_\Delta/\Delta \to \infty$ for $\Delta \to 0$ must be fulfilled).

3.1 Notation

Consider the path of a BDI η at a regular grid of time $i\Delta$, where $i \in \mathbb{N}_0$. We assume that we are able to see only the positions of the particles but not their pedigree, which means that at each time $i\Delta$ the configuration

$$\eta_{i\Delta} = \left(\eta_{i\Delta}^1, ..., \eta_{i\Delta}^{\ell(\eta_{i\Delta})}\right) \tag{3.1}$$

is observed in form of the point measure $\sum_{k=1}^{\ell(\eta_{i\Delta})} \epsilon_{\eta_{i\Delta}^k}$. Let

$$\beta_{i\Delta} = \left(\beta_{i\Delta}^1, ..., \beta_{i\Delta}^{\ell(\eta_{i\Delta})}\right) \tag{3.2}$$

denote an arbitrary arrangement of $\eta_{i\Delta}$ viewed as point measure $\sum_{k=1}^{\ell(\eta_{i\Delta})} \epsilon_{\eta_{i\Delta}^k}$. By introducing the equivalence relation for $\mathbf{x}, \mathbf{y} \in \mathcal{S}$ via

$$\mathbf{x} =_p \mathbf{y} :\Longleftrightarrow \begin{cases} \ell(\mathbf{x}) = \ell(\mathbf{y}) \text{ and there exists a permutation } \pi' \\ \text{of } \{1, ..., \ell(\mathbf{x})\} \text{ such that } \pi'(\mathbf{x}) = \mathbf{y}, \end{cases}$$

it holds

$$\beta_{i\Delta} =_p \eta_{i\Delta}. \tag{3.3}$$

We are now faced with the following problem: Based on the observations $\beta_{i\Delta}$ and $\beta_{(i+1)\Delta}$, we cannot determine

$$\eta_{(i+1)\Delta}^{(i\Delta,\beta_{i\Delta}^k)} := \Big\{\eta_{(i+1)\Delta}^j \in \eta_{(i+1)\Delta} \,\Big|\, 1 \le j \le \ell(\eta_{(i+1)\Delta}),\ \text{the particle situated at } \eta_{(i+1)\Delta}^j \text{ is the offspring of a particle situated at } \beta_{i\Delta}^k\Big\} \tag{3.4}$$

since we do not know how the BDI process behaves during the time $(i\Delta, (i+1)\Delta]$ as branching or immigration could occur. However, we are able to approximate this expression as we will see in the next section. Before we outline this approximation, remember the set D_ε from (1.40) which describes the set of configurations with far neighbours, i.e.,

$$D_\varepsilon = \Big\{\mathbf{x} \in \mathcal{S} \,\Big|\, \ell(\mathbf{x}) \ge 2, \forall i \ne j \in \big\{1, ..., \ell(\mathbf{x})\big\} : |x^i - x^j| \ge \varepsilon\Big\}.$$

3.2 A Partial Reconstruction Algorithm

Let $0 < \lambda < \frac{1}{2}$ be fixed and let $\beta_{i\Delta}$ be an arrangement of $\eta_{i\Delta}$ with $i \in \mathbb{N}_0$ and $1 \le k \le \ell(\eta_{i\Delta})$. Define

$$\beta_{(i+1)\Delta}^{[i\Delta,\beta_{i\Delta}^k]} := \Big\{\beta_{(i+1)\Delta}^m \in \beta_{(i+1)\Delta} \,\Big|\, 1 \le m \le \ell(\beta_{(i+1)\Delta}),\ |\beta_{(i+1)\Delta}^m - \beta_{i\Delta}^k| \le \Delta^\lambda\Big\} \tag{3.5}$$

which is the set of the positions of particles (belonging to $\beta_{(i+1)\Delta}$) whose distance to $\beta_{i\Delta}^k$ is less than or equal to Δ^λ. Moreover, let

$$\beta_{(i+1)\Delta}^I := \beta_{(i+1)\Delta} \Big\backslash \left(\bigcup_{j=1}^{\ell(\beta_{i\Delta})} \beta_{(i+1)\Delta}^{[i\Delta,\beta_{i\Delta}^j]}\right)$$

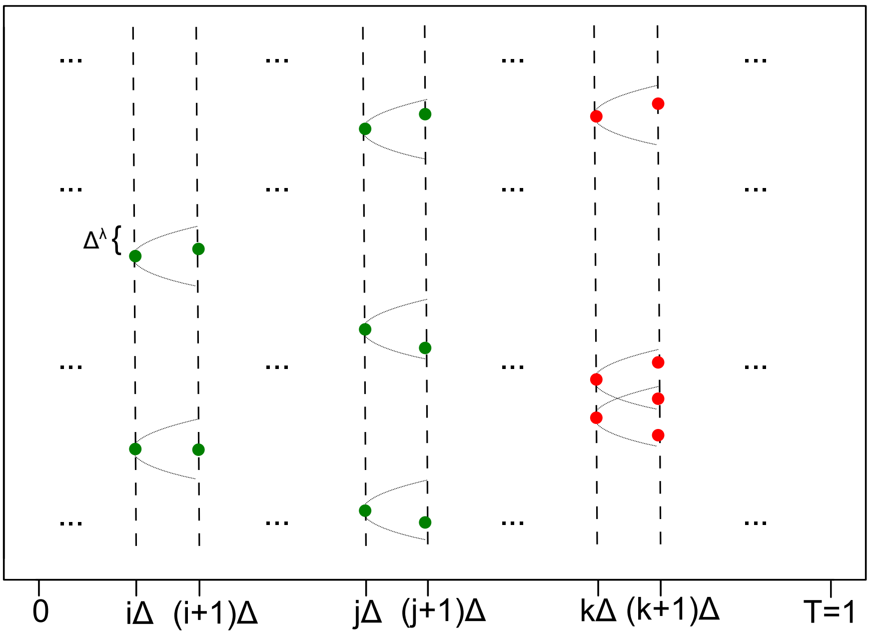
denote the set of the positions of particles (belonging to $\beta_{(i+1)\Delta}$) whose distance to every particle of $\beta_{i\Delta}$ is greater than Δ^λ.

Definition 3.1 (Interpretable Pair)
A pair $(\eta_{i\Delta}, \eta_{(i+1)\Delta})$ is called interpretable if there exists an arrangement $(\beta_{i\Delta}, \beta_{(i+1)\Delta})$ with the following properties:

$$\begin{cases} \beta_{i\Delta} \in D_{4\Delta^\lambda} \quad \textit{and} \quad \beta_{(i+1)\Delta} \in D_{2\Delta^\lambda}, \\ \big|\beta_{(i+1)\Delta}^{[i\Delta,\beta_{i\Delta}^k]}\big| = 1 \quad \textit{for every} \quad 1 \le k \le \ell(\beta_{i\Delta}). \end{cases}$$

$\Diamond$

The previous definition focuses on pairs $(\eta_{i\Delta}, \eta_{(i+1)\Delta})$ which have an arrangement $(\beta_{i\Delta}, \beta_{(i+1)\Delta})$ with good properties: The arrangement $(\beta_{i\Delta}, \beta_{(i+1)\Delta})$ has to fulfil that both $\beta_{i\Delta}$ and $\beta_{(i+1)\Delta}$ have particles which do not have close neighbours, i.e., every particle of $\beta_{i\Delta}$ has to have a distance of greater than at least $4\Delta^\lambda$ to its adjacent particles resp. every particle of $\beta_{(i+1)\Delta}$ has to have a distance of greater than $2\Delta^\lambda$ to its adjacent particles. Furthermore, for every particle situated at $\beta^k_{i\Delta}$ (and belonging to $\beta_{i\Delta}$) we consider a Δ^λ-neighbourhood of $\beta^k_{i\Delta}$. Then, arrangements which have exactly one particle of $\beta_{(i+1)\Delta}$ in a Δ^λ-neighbourhood of $\beta^k_{i\Delta}$ for every $1 \leq k \leq \ell(\beta_{i\Delta})$ are filtered.
The following graphic demonstrates our idea. The green dots are particles of interpretable pairs, whereas the red dots are particles of a pair which is not interpretable.



Remark 3.2

1. It is obvious that not every pair $(\eta_{i\Delta}, \eta_{(i+1)\Delta})$ is interpretable and that we are not interested in every arrangement $(\beta_{i\Delta}, \beta_{(i+1)\Delta})$. This is why we have to show that there are enough interpretable pairs resp. the case that there is a non-interpretable pair does not occur often.

2. If we compare our reconstruction algorithm with Brandt's algorithm in [2, p. 39], there are the following differences: Our focus lies on pairs $(\eta_{i\Delta}, \eta_{(i+1)\Delta})$ which have to fulfil certain conditions. Whereas Brandt demands that an arrangement $\beta_{i\Delta}$ of $\eta_{i\Delta}$ has to have far neighbours for considering a Δ^λ-neighbourhood around every particle of $\beta_{i\Delta}$, we require that both the particles of $\beta_{i\Delta}$ and of $\beta_{(i+1)\Delta}$ have far neighbours. In addition, we demand that there is exactly one particle of $\beta_{(i+1)\Delta}$ in a Δ^λ-neighbourhood of $\beta^k_{i\Delta}$ for every $1 \leq k \leq \ell(\beta_{i\Delta})$. So, using our algorithm we filter more configurations.

3. For every $i \in \mathbb{N}_0$ the event

$$\Big\{(\eta_{i\Delta}, \eta_{(i+1)\Delta}) \text{ is interpretable}\Big\} \tag{3.6}$$

is measurable concerning

$$\mathcal{H}_{i\Delta} := \sigma\left(\eta_0, \eta_\Delta, ..., \eta_{(i+1)\Delta}\right)$$

(with $\mathbb{H} := (\mathcal{H}_{i\Delta})_{i\in\mathbb{N}_0}$) and the set of all interpretable pairs is denoted by

$$\widetilde{G} := \Big\{(\eta_{i\Delta}, \eta_{(i+1)\Delta}) \in \mathcal{S} \times \mathcal{S} \,\Big|\, i \in \mathbb{N}_0,\ (\eta_{i\Delta}, \eta_{(i+1)\Delta}) \text{ is interpretable}\Big\}. \tag{3.7}$$

$\Diamond$

For every $i \in \mathbb{N}_0$ and every $1 \leq k \leq \ell(\beta_{i\Delta})$ we define the event

$$b^\Delta(i\Delta, \beta_{i\Delta}^k) := \Big\{\text{The particle situated at } \beta_{i\Delta}^k \text{ at time } i\Delta \text{ dies during the time } (i\Delta, (i+1)\Delta]\Big\}.$$

We remember the sets (3.4) resp. (3.5) and for every $i \in \mathbb{N}_0$ we consider the event

$$\bigcap_{k=1}^{\ell(\beta_{i\Delta})} \Big\{\beta_{(i+1)\Delta}^{[i\Delta, \beta_{i\Delta}^k]} =_p \eta_{(i+1)\Delta}^{(i\Delta, \beta_{i\Delta}^k)}\Big\} \cap \left(b^\Delta(i\Delta, \beta_{i\Delta}^k)\right)^c. \tag{3.8}$$

The event (3.8) is measurable concerning

$$\begin{aligned}\mathcal{G}_{i\Delta} &:= \sigma\left(\eta_s \,\middle|\, 0 \leq s \leq (i+1)\Delta\right) \\ &= \mathcal{F}_{(i+1)\Delta}\end{aligned}$$

(with $\mathbb{G} := (\mathcal{G}_{i\Delta})_{i\in\mathbb{N}_0}$) and is fulfilled if for every $1 \leq k \leq \ell(\beta_{i\Delta})$ the sets

$$\beta_{(i+1)\Delta}^{[i\Delta, \beta_{i\Delta}^k]} \quad \text{and} \quad \eta_{(i+1)\Delta}^{(i\Delta, \beta_{i\Delta}^k)}$$

coincide and no particle belonging to $\beta_{i\Delta}$ dies during the time $(i\Delta, (i+1)\Delta]$. Particularly, this means that the assignment in (3.5) is correct. If we combine the event (3.8) with (3.6), we receive

$$\begin{aligned}&\Big\{(\eta_{i\Delta}, \eta_{(i+1)\Delta}) \text{ is interpretable}\Big\} \\ &\cap \bigcap_{k=1}^{\ell(\beta_{i\Delta})} \Big\{\beta_{(i+1)\Delta}^{[i\Delta, \beta_{i\Delta}^k]} =_p \eta_{(i+1)\Delta}^{(i\Delta, \beta_{i\Delta}^k)}\Big\} \cap \left(b^\Delta(i\Delta, \beta_{i\Delta}^k)\right)^c\end{aligned} \tag{3.9}$$

which is measurable concerning the sigma-field $\mathcal{G}_{i\Delta}$. (3.9) describes the event that the pair $(\eta_{i\Delta}, \eta_{(i+1)\Delta})$ is interpretable, the assignment in (3.5) is correct and every particle belonging to $\beta_{i\Delta}$ neither dies nor it makes "big excursions" during the time $(i\Delta, (i+1)\Delta]$. We want to emphasise that (3.9) is not measurable concerning the sigma-field $\mathcal{H}_{i\Delta}$ since we do not know what happens during the time $(i\Delta, (i+1)\Delta]$ by observing the path only at discrete points in time $i\Delta$, $i \in \mathbb{N}_0$.

Definition 3.3 (Properly Interpretable Pair)
A pair $(\eta_{i\Delta}, \eta_{(i+1)\Delta})$ *which fulfils* (3.9) *is called properly interpretable pair.*

$\Diamond$

The set of all properly interpretable pairs is denoted by

$$G := \Big\{(\eta_{i\Delta}, \eta_{(i+1)\Delta}) \in \mathcal{S} \times \mathcal{S} \,\Big|\, i \in \mathbb{N}_0,\ (\eta_{i\Delta}, \eta_{(i+1)\Delta}) \text{ is properly interpretable}\Big\}. \tag{3.10}$$

In the next theorem, we will see that the expected quota of properly interpretable pairs during the time $[0, T]$ converges to 1, as $\Delta \to 0$, which is why our procedure makes sense.

Theorem 3.4
Grant Assumptions 1.13, 1.19 and 2.1(a). Then, the expected quota of pairs $(\eta_{i\Delta}, \eta_{(i+1)\Delta})$, $0 \le i \le \lfloor T/\Delta \rfloor - 1$, *being properly interpretable converges to* 1. *It even holds*

$$\begin{aligned} 1 \ge \mathbf{E}_m &\left(\frac{1}{\lfloor T/\Delta \rfloor} \sum_{i=0}^{\lfloor T/\Delta \rfloor - 1} \mathbb{1}_G \left((\eta_{i\Delta}, \eta_{(i+1)\Delta}) \right) \right) \\ &= \mathbf{P}_m \Big(\big\{ (\eta_0, \eta_\Delta) \in G \big\} \Big) \\ &\ge 1 - \mathcal{O}(\Delta^\lambda), \end{aligned}$$

as $\Delta \to 0$.

In order to show this theorem, we must prove the following proposition.

Proposition 3.5
Grant Assumptions 1.13, 1.19 and 2.1(a) and let $\beta_{i\Delta}$ *be observations of the BDI* η, *where* $0 \le i \le \lfloor T/\Delta \rfloor - 1$. *Then, it holds*

1. $\mathbf{E}_m \left(\frac{1}{\lfloor T/\Delta \rfloor} \sum_{i=0}^{\lfloor T/\Delta \rfloor - 1} \mathbb{1}_{S_{4\Delta^\lambda}}(\beta_{i\Delta}) \right) = \mathcal{O}(\Delta^\lambda), \quad \text{as} \quad \Delta \to 0.$
2. $\mathbf{E}_m \left(\frac{1}{\lfloor T/\Delta \rfloor} \sum_{i=0}^{\lfloor T/\Delta \rfloor - 1} \mathbb{1}_{\bigcup_{k=1}^{\ell(\beta_{i\Delta})} \left\{ \left| \eta_{(i+1)\Delta}^{(i\Delta, \beta_{i\Delta}^k)} \right| \neq 1 \right\}} \right) = \mathcal{O}(\Delta), \quad \text{as} \quad \Delta \to 0.$
3. $\mathbf{E}_m \left(\frac{1}{\lfloor T/\Delta \rfloor} \sum_{i=0}^{\lfloor T/\Delta \rfloor - 1} \mathbb{1}_{\bigcup_{k=1}^{\ell(\beta_{i\Delta})} b^\Delta(i\Delta, \beta_{i\Delta}^k)} \right) = \mathcal{O}(\Delta), \quad \text{as} \quad \Delta \to 0.$
4. $\mathbf{E}_m \left(\frac{1}{\lfloor T/\Delta \rfloor} \sum_{i=0}^{\lfloor T/\Delta \rfloor - 1} \mathbb{1}_{D_{4\Delta^\lambda}}(\beta_{i\Delta}) \cdot \mathbb{1}_{S_{2\Delta^\lambda}}(\beta_{(i+1)\Delta}) \right) = \mathcal{O}(\Delta), \quad \text{as} \quad \Delta \to 0.$
5. $\mathbf{E}_m \left(\frac{1}{\lfloor T/\Delta \rfloor} \sum_{i=0}^{\lfloor T/\Delta \rfloor - 1} \mathbb{1}_{D_{4\Delta^\lambda}}(\beta_{i\Delta}) \cdot \mathbb{1}_{\bigcup_{k=1}^{\ell(\beta_{i\Delta})} \left\{ \beta_{(i+1)\Delta}^{[i\Delta, \beta_{i\Delta}^k]} \neq_p \eta_{(i+1)\Delta}^{(i\Delta, \beta_{i\Delta}^k)} \right\}} \right) = \mathcal{O}(\Delta), \quad \text{as} \quad \Delta \to 0.$

Proof

1. As $m(\cdot)$ is the invariant measure of the BDI process, it holds

$$\begin{aligned}\mathbf{E}_m\left(\frac{1}{\lfloor T/\Delta\rfloor}\sum_{i=0}^{\lfloor T/\Delta\rfloor-1}\mathbb{1}_{S_{4\Delta^\lambda}}(\beta_{i\Delta})\right) &= \frac{1}{\lfloor T/\Delta\rfloor}\sum_{i=0}^{\lfloor T/\Delta\rfloor-1}\mathbf{P}_m\Big(\{\beta_{i\Delta}\in S_{4\Delta^\lambda}\}\Big)\\ &= \frac{1}{\lfloor T/\Delta\rfloor}\cdot\lfloor T/\Delta\rfloor\cdot m(S_{4\Delta^\lambda})\\ &= m(S_{4\Delta^\lambda}).\end{aligned}$$

Now, we can use Theorem 2.11 which states

$$m(S_{4\Delta^\lambda}) = \mathcal{O}(\Delta^\lambda),$$

as $\Delta\to 0$.

2. We can split the $\mathcal{F}_{(i+1)\Delta}$-measurable event

$$\Big\{|\eta_{(i+1)\Delta}^{(i\Delta,\beta_{i\Delta}^k)}|\neq 1\Big\} = \Big\{|\eta_{(i+1)\Delta}^{(i\Delta,\beta_{i\Delta}^k)}| = 0\Big\}\,\dot\cup\,\Big\{|\eta_{(i+1)\Delta}^{(i\Delta,\beta_{i\Delta}^k)}| > 1\Big\}$$

into two disjoint parts. The first part describes that at time $(i+1)\Delta$ no offspring of a particle situated at $\beta_{i\Delta}^k$ is left. Hence, the particle starting in $\beta_{i\Delta}^k$ must have died during the time $(i\Delta,(i+1)\Delta]$. The second part describes the event that at time $(i+1)\Delta$ two or more offspring of a particle situated at $\beta_{i\Delta}^k$ are left. This means that the particle starting in $\beta_{i\Delta}^k$ must have branched during the time $(i\Delta,(i+1)\Delta]$. Thus,

$$\Big\{|\eta_{(i+1)\Delta}^{(i\Delta,\beta_{i\Delta}^k)}|\neq 1\Big\}\subseteq b^\Delta(i\Delta,\beta_{i\Delta}^k). \tag{3.11}$$

According to Assumption 1.3, the probability that the particle situated at $\beta_{i\Delta}^k$ dies during the time $(i\Delta,(i+1)\Delta]$ is of order $\mathcal{O}(\Delta)$, as $\Delta\to 0$, so we receive in (3.11)

$$\mathbf{P}_m\left(\Big\{\,|\eta_{(i+1)\Delta}^{(i\Delta,\beta_{i\Delta}^k)}|\neq 1\Big\}\;|\,\mathcal{F}_{i\Delta}\right) = \mathcal{O}(\Delta), \tag{3.12}$$

as $\Delta\to 0$. Using (3.12), the invariance of $m(\cdot)$ and the finiteness of the occupation measure $\overline{m}(\cdot)$ according to (1.44), we obtain

$$\begin{aligned}&\mathbf{E}_m\left(\frac{1}{\lfloor T/\Delta\rfloor}\sum_{i=0}^{\lfloor T/\Delta\rfloor-1}\mathbb{1}_{\bigcup_{k=1}^{\ell(\beta_{i\Delta})}\Big\{|\eta_{(i+1)\Delta}^{(i\Delta,\beta_{i\Delta}^k)}|\neq 1\Big\}}\right)\\ &\leq \mathbf{E}_m\left(\frac{1}{\lfloor T/\Delta\rfloor}\sum_{i=0}^{\lfloor T/\Delta\rfloor-1}\sum_{k=1}^{\ell(\beta_{i\Delta})}\mathbf{P}_m\left(\Big\{|\eta_{(i+1)\Delta}^{(i\Delta,\beta_{i\Delta}^k)}|\neq 1\Big\}\;|\,\mathcal{F}_{i\Delta}\right)\right)\end{aligned}$$

$$
\begin{aligned}
&= \mathcal{O}(\Delta) \cdot \mathbf{E}_m\left(\frac{1}{\lfloor T/\Delta \rfloor} \sum_{i=0}^{\lfloor T/\Delta \rfloor - 1} \ell(\beta_{i\Delta}) \right) \\
&= \mathcal{O}(\Delta) \cdot \frac{1}{\lfloor T/\Delta \rfloor} \cdot \lfloor T/\Delta \rfloor \int_{\mathcal{S}} \ell(\mathbf{x}) \, m(d\mathbf{x}) \\
&= \mathcal{O}(\Delta) \int_{\mathcal{S}} \ell(\mathbf{x}) \, m(d\mathbf{x}) \\
&= \mathcal{O}(\Delta),
\end{aligned}
$$

as $\Delta \to 0$.

3. Because of Assumption 1.3, for every $1 \le k \le \ell(\beta_{i\Delta})$ it holds

$$\mathbf{P}_m\left(b^{\Delta}(i\Delta, \beta_{i\Delta}^k) \,\middle|\, \mathcal{F}_{i\Delta} \right) = \mathcal{O}(\Delta),$$

as $\Delta \to 0$. Proceeding exactly as in the second part of this proof, it follows

$$\mathbf{E}_m\left(\frac{1}{\lfloor T/\Delta \rfloor} \sum_{i=0}^{\lfloor T/\Delta \rfloor - 1} \mathbb{1}_{\bigcup_{k=1}^{\ell(\beta_{i\Delta})} b^{\Delta}(i\Delta, \beta_{i\Delta}^k)} \right) = \mathcal{O}(\Delta),$$

as $\Delta \to 0$.

4. As $\beta_\Delta =_p \eta_\Delta$, we can use Proposition 2.9 which states that there is a constant $C > 0$ such that for sufficiently small $\Delta > 0$ and for every $\mathbf{x} \in \mathcal{S}$

$$\mathbb{1}_{D_{4\Delta^\lambda}}(\mathbf{x}) \cdot \mathbf{P}_{\mathbf{x}}\Big(\big\{ \beta_\Delta \in S_{2\Delta^\lambda} \big\} \Big) \le C\Delta \cdot \ell(\mathbf{x}). \tag{3.13}$$

Using that $m(\cdot)$ is invariant, (3.13) and the finiteness of the occupation measure $\overline{m}(\cdot)$ from (1.44), we receive for sufficiently small $\Delta > 0$

$$
\begin{aligned}
&\mathbf{E}_m\left(\frac{1}{\lfloor T/\Delta \rfloor} \sum_{i=0}^{\lfloor T/\Delta \rfloor - 1} \mathbb{1}_{D_{4\Delta^\lambda}}(\beta_{i\Delta}) \cdot \mathbb{1}_{S_{2\Delta^\lambda}}(\beta_{(i+1)\Delta}) \right) \\
&= \frac{1}{\lfloor T/\Delta \rfloor} \cdot \lfloor T/\Delta \rfloor \cdot \mathbf{E}_m\Big(\mathbb{1}_{D_{4\Delta^\lambda}}(\beta_0) \cdot \mathbb{1}_{S_{2\Delta^\lambda}}(\beta_\Delta) \Big) \\
&= \int_{\mathcal{S}} \mathbb{1}_{D_{4\Delta^\lambda}}(\mathbf{x}) \cdot \mathbf{P}_{\mathbf{x}}\Big(\big\{ \beta_\Delta \in S_{2\Delta^\lambda} \big\} \Big) \, m(d\mathbf{x}) \\
&\le \int_{\mathcal{S}} C\Delta \cdot \ell(\mathbf{x}) \, m(d\mathbf{x}) \\
&= C\Delta \int_{\mathcal{S}} \ell(\mathbf{x}) \, m(d\mathbf{x}) \\
&= \mathcal{O}(\Delta),
\end{aligned}
$$

as $\Delta \to 0$.

5. First of all, we show that there is a constant $C > 0$ such that for sufficiently small $\Delta > 0$ and for every $0 \le i \le \lfloor T/\Delta \rfloor - 1$

$$\mathbb{1}_{D_{4\Delta^\lambda}}(\beta_{i\Delta}) \cdot \mathbf{P}_m\left(\left\{ \beta_{(i+1)\Delta}^{[i\Delta, \beta_{i\Delta}^k]} \neq_p \eta_{(i+1)\Delta}^{(i\Delta, \beta_{i\Delta}^k)} \right\} \,\middle|\, \mathcal{F}_{i\Delta} \right) \le C\Delta \cdot \ell(\beta_{i\Delta}). \tag{3.14}$$

In order to prove (3.14), we proceed similarly as in the proof of Proposition 2.9. We remember the events from (2.12) and (2.13) and write

$$A^{\Delta}(i\Delta, \beta_{i\Delta}) = \bigcup_{k=1}^{\ell(\beta_{i\Delta})} a^{\Delta}(i\Delta, \beta_{i\Delta}^k) \tag{3.15}$$

for the event that there is one particle belonging to $\beta_{i\Delta}$ (and situated at $\beta_{i\Delta}^k$) whose subprocess

$$\left(\eta_t^{(i\Delta, \beta_{i\Delta}^k)}\right)_{t \geq 0}$$

(without immigration) leaves a neighbourhood of a size Δ^λ around $\beta_{i\Delta}^k$ during the time period $(i\Delta, (i+1)\Delta]$. Depending on (3.15), we get in (3.14)

$$\begin{aligned}
&\mathbb{1}_{D_{4\Delta^\lambda}}(\beta_{i\Delta}) \cdot \mathbf{P}_m \left(\left\{ \beta_{(i+1)\Delta}^{[i\Delta, \beta_{i\Delta}^k]} \neq_p \eta_{(i+1)\Delta}^{(i\Delta, \beta_{i\Delta}^k)} \right\} \,\middle|\, \mathcal{F}_{i\Delta} \right) \\
&\leq \mathbb{1}_{D_{4\Delta^\lambda}}(\beta_{i\Delta}) \cdot \mathbf{P}_m \left(\left\{ \beta_{(i+1)\Delta}^{[i\Delta, \beta_{i\Delta}^k]} \neq_p \eta_{(i+1)\Delta}^{(i\Delta, \beta_{i\Delta}^k)} \right\} \cap A^{\Delta}(i\Delta, \beta_{i\Delta}) \,\middle|\, \mathcal{F}_{i\Delta} \right) \\
&\quad + \mathbb{1}_{D_{4\Delta^\lambda}}(\beta_{i\Delta}) \cdot \mathbf{P}_m \left(\left\{ \beta_{(i+1)\Delta}^{[i\Delta, \beta_{i\Delta}^k]} \neq_p \eta_{(i+1)\Delta}^{(i\Delta, \beta_{i\Delta}^k)} \right\} \cap \left(A^{\Delta}(i\Delta, \beta_{i\Delta}) \right)^c \,\middle|\, \mathcal{F}_{i\Delta} \right).
\end{aligned} \tag{3.16}$$

As $\beta_{i\Delta} =_p \eta_{i\Delta}$, the first term in (3.16) can be estimated with Corollary 2.8, i.e., there is a constant $C' > 0$ such that for sufficiently small $\Delta > 0$

$$\mathbb{1}_{D_{4\Delta^\lambda}}(\beta_{i\Delta}) \cdot \mathbf{P}_m \left(\left\{ \beta_{(i+1)\Delta}^{[i\Delta, \beta_{i\Delta}^k]} \neq_p \eta_{(i+1)\Delta}^{(i\Delta, \beta_{i\Delta}^k)} \right\} \cap A^{\Delta}(i\Delta, \beta_{i\Delta}) \,\middle|\, \mathcal{F}_{i\Delta} \right) \leq C'\Delta \cdot \ell(\beta_{i\Delta}). \tag{3.17}$$

Consider the second term in (3.16). Because of $\left(A^{\Delta}(i\Delta, \beta_{i\Delta})\right)^c$, for every $1 \leq k' \leq \ell(\beta_{i\Delta})$ the subprocess

$$\left(\eta_t^{(i\Delta, \beta_{i\Delta}^{k'})}\right)_{t \geq 0}$$

(without immigration) which starts in $\beta_{i\Delta}^{k'}$ (with $\beta_{i\Delta} \in D_{4\Delta^\lambda}$) at time $i\Delta$ stays in a neighbourhood of a size Δ^λ around $\beta_{i\Delta}^{k'}$ during the time $(i\Delta, (i+1)\Delta]$. Particularly, this means

$$\eta_{(i+1)\Delta}^{(i\Delta, \beta_{i\Delta}^k)} \subseteq \beta_{(i+1)\Delta}^{[i\Delta, \beta_{i\Delta}^k]}. \tag{3.18}$$

Because of (3.18) and

$$\left\{ \beta_{(i+1)\Delta}^{[i\Delta, \beta_{i\Delta}^k]} \neq_p \eta_{(i+1)\Delta}^{(i\Delta, \beta_{i\Delta}^k)} \right\}, \tag{3.19}$$

immigration must occur during the time $(i\Delta, (i+1)\Delta]$, i.e., denoting $\tau^{i\Delta}$ the first immigration time after time $i\Delta$, we get for $\beta_{i\Delta} \in D_{4\Delta^\lambda}$

$$\begin{aligned}
&\mathbf{P}_m \left(\left\{ \beta_{(i+1)\Delta}^{[i\Delta, \beta_{i\Delta}^k]} \neq_p \eta_{(i+1)\Delta}^{(i\Delta, \beta_{i\Delta}^k)} \right\} \cap \left(A^{\Delta}(i\Delta, \beta_{i\Delta}) \right)^c \,\middle|\, \mathcal{F}_{i\Delta} \right) \\
&\leq \mathbf{P}_m \left(\left\{ \beta_{(i+1)\Delta}^{[i\Delta, \beta_{i\Delta}^k]} \cap \eta_{(i+1)\Delta}^I \neq \emptyset \right\} \cap \left\{ \tau^{i\Delta} < (i+1)\Delta \right\} \,\middle|\, \mathcal{F}_{i\Delta} \right).
\end{aligned} \tag{3.20}$$

According to (1.7), there is a constant $C'' > 0$ such that the probability of immigration during the time period $(i\Delta, (i+1)\Delta]$ is smaller than or equal to $C''\Delta$, for sufficiently small $\Delta > 0$.

This is why (3.20) can be estimated by

$$\mathbb{1}_{D_{4\Delta^\lambda}}(\beta_{i\Delta}) \cdot \mathbf{P}_m \left(\left\{ \beta_{(i+1)\Delta}^{[i\Delta,\beta_{i\Delta}^k]} \neq_p \eta_{(i+1)\Delta}^{(i\Delta,\beta_{i\Delta}^k)} \right\} \cap \left(A^\Delta(i\Delta,\beta_{i\Delta}) \right)^c \;\middle|\; \mathcal{F}_{i\Delta} \right) \leq C''\Delta \tag{3.21}$$

for sufficiently small $\Delta > 0$. Combining (3.17) and (3.21), we have proved (3.14).

Now, we can finish our proof. Using inequality (3.14), $m(\cdot)$ being an invariant measure and

$$\int_{\mathcal{S}} \ell^2(\mathbf{x})\, m(d\mathbf{x}) < \infty$$

from (1.43), we obtain for sufficiently small $\Delta > 0$

$$\begin{aligned}
&\mathbf{E}_m \left(\frac{1}{\lfloor T/\Delta \rfloor} \sum_{i=0}^{\lfloor T/\Delta \rfloor - 1} \mathbb{1}_{D_{4\Delta^\lambda}}(\beta_{i\Delta}) \cdot \mathbb{1}_{\bigcup_{k=1}^{\ell(\beta_{i\Delta})} \left\{ \beta_{(i+1)\Delta}^{[i\Delta,\beta_{i\Delta}^k]} \neq_p \eta_{(i+1)\Delta}^{(i\Delta,\beta_{i\Delta}^k)} \right\}} \right) \\
&\leq \mathbf{E}_m \left(\frac{1}{\lfloor T/\Delta \rfloor} \sum_{i=0}^{\lfloor T/\Delta \rfloor - 1} \mathbb{1}_{D_{4\Delta^\lambda}}(\beta_{i\Delta}) \cdot \sum_{k=1}^{\ell(\beta_{i\Delta})} \mathbf{P}_m \left(\left\{ \beta_{(i+1)\Delta}^{[i\Delta,\beta_{i\Delta}^k]} \neq_p \eta_{(i+1)\Delta}^{(i\Delta,\beta_{i\Delta}^k)} \right\} \;\middle|\; \mathcal{F}_{i\Delta} \right) \right) \\
&\leq \mathbf{E}_m \left(\frac{1}{\lfloor T/\Delta \rfloor} \sum_{i=0}^{\lfloor T/\Delta \rfloor - 1} \mathbb{1}_{D_{4\Delta^\lambda}}(\beta_{i\Delta}) \cdot \ell(\beta_{i\Delta}) \cdot C\Delta \cdot \ell(\beta_{i\Delta}) \right) \\
&\leq C\Delta \int_{\mathcal{S}} \ell^2(\mathbf{x})\, m(d\mathbf{x}) \\
&= \mathcal{O}(\Delta),
\end{aligned}$$

as $\Delta \to 0$.

$\square$

Remark 3.6

1. In the first part of the proof of Proposition 3.5, because of Theorem 2.11 we are able to obtain a rate of convergence for

$$\mathbf{E}_m \left(\frac{1}{\lfloor T/\Delta \rfloor} \sum_{i=0}^{\lfloor T/\Delta \rfloor - 1} \mathbb{1}_{S_{4\Delta^\lambda}}(\beta_{i\Delta}) \right) = m(S_{4\Delta^\lambda}).$$

   In particular, this is due to Assumptions 1.13 and 1.19 from Hammer's framework.

2. In the fifth part of the proof of Proposition 3.5, in inequality (3.20) we have not made use of the position where the immigrating particle starts its motion. By applying the same arguments as in Remark 2.10, we would not achieve better results if we considered this (for this, Assumption 1.14 must be assumed). Therefore, we discontinue inquiry into this line.

$\Diamond$

Now, using Proposition 3.5 we are able to prove Theorem 3.4.

Proof (of Theorem 3.4)

It holds

$$
\begin{aligned}
&\sum_{i=0}^{\lfloor T/\Delta\rfloor-1} \mathbb{1}_G\left((\eta_{i\Delta},\eta_{(i+1)\Delta})\right)\\
&= \sum_{i=0}^{\lfloor T/\Delta\rfloor-1} \mathbb{1}_{D_{4\Delta}\lambda\times D_{2\Delta}\lambda}\left((\beta_{i\Delta},\beta_{(i+1)\Delta})\right)\cdot \mathbb{1}_{\bigcap_{k=1}^{\ell(\beta_{i\Delta})}\left\{\beta_{(i+1)\Delta}^{[i\Delta,\beta_{i\Delta}^k]}=_p\eta_{(i+1)\Delta}^{(i\Delta,\beta_{i\Delta}^k)}\right\}}\\
&\qquad\times \mathbb{1}_{\bigcap_{k=1}^{\ell(\beta_{i\Delta})}\left\{\left|\eta_{(i+1)\Delta}^{(i\Delta,\beta_{i\Delta}^k)}\right|=1\right\}}\cdot \mathbb{1}_{\bigcap_{k=1}^{\ell(\beta_{i\Delta})}\left(b^\Delta(i\Delta,\beta_{i\Delta}^k)\right)^c}\\
&\geq \sum_{i=0}^{\lfloor T/\Delta\rfloor-1} \mathbb{1}_{D_{4\Delta}\lambda}(\beta_{i\Delta})\cdot \mathbb{1}_{D_{2\Delta}\lambda}(\beta_{(i+1)\Delta})\cdot\Bigg(1-\mathbb{1}_{\bigcup_{k=1}^{\ell(\beta_{i\Delta})}\left\{\left|\eta_{(i+1)\Delta}^{(i\Delta,\beta_{i\Delta}^k)}\right|\neq 1\right\}}\\
&\qquad-\mathbb{1}_{\bigcup_{k=1}^{\ell(\beta_{i\Delta})}\left\{\beta_{(i+1)\Delta}^{[i\Delta,\beta_{i\Delta}^k]}\neq_p\eta_{(i+1)\Delta}^{(i\Delta,\beta_{i\Delta}^k)}\right\}}-\mathbb{1}_{\bigcup_{k=1}^{\ell(\beta_{i\Delta})}b^\Delta(i\Delta,\beta_{i\Delta}^k)}\Bigg)\\
&\geq \sum_{i=0}^{\lfloor T/\Delta\rfloor-1} \mathbb{1}_{D_{4\Delta}\lambda}(\beta_{i\Delta}) - \sum_{i=0}^{\lfloor T/\Delta\rfloor-1} \mathbb{1}_{D_{4\Delta}\lambda}(\beta_{i\Delta})\cdot \mathbb{1}_{\bigcup_{k=1}^{\ell(\beta_{i\Delta})}\left\{\beta_{(i+1)\Delta}^{[i\Delta,\beta_{i\Delta}^k]}\neq_p\eta_{(i+1)\Delta}^{(i\Delta,\beta_{i\Delta}^k)}\right\}}\\
&\quad-\sum_{i=0}^{\lfloor T/\Delta\rfloor-1} \mathbb{1}_{\bigcup_{k=1}^{\ell(\beta_{i\Delta})}\left\{\left|\eta_{(i+1)\Delta}^{(i\Delta,\beta_{i\Delta}^k)}\right|\neq 1\right\}} - \sum_{i=0}^{\lfloor T/\Delta\rfloor-1} \mathbb{1}_{D_{4\Delta}\lambda}(\beta_{i\Delta})\cdot \mathbb{1}_{S_{2\Delta}\lambda}(\beta_{(i+1)\Delta})\\
&\quad-\sum_{k=1}^{\lfloor T/\Delta\rfloor-1} \mathbb{1}_{\bigcup_{k=1}^{\ell(\beta_{i\Delta})}b^\Delta(i\Delta,\beta_{i\Delta}^k)}\\
&= \lfloor T/\Delta\rfloor - \sum_{i=0}^{\lfloor T/\Delta\rfloor-1} \mathbb{1}_{S_{4\Delta}\lambda}(\beta_{i\Delta})\\
&\qquad-\sum_{i=0}^{\lfloor T/\Delta\rfloor-1} \mathbb{1}_{\bigcup_{k=1}^{\ell(\beta_{i\Delta})}\left\{\left|\eta_{(i+1)\Delta}^{(i\Delta,\beta_{i\Delta}^k)}\right|\neq 1\right\}}\\
&\qquad-\sum_{k=1}^{\lfloor T/\Delta\rfloor-1} \mathbb{1}_{\bigcup_{k=1}^{\ell(\beta_{i\Delta})}b^\Delta(i\Delta,\beta_{i\Delta}^k)}\\
&\qquad-\sum_{i=0}^{\lfloor T/\Delta\rfloor-1} \mathbb{1}_{D_{4\Delta}\lambda}(\beta_{i\Delta})\cdot \mathbb{1}_{S_{2\Delta}\lambda}(\beta_{(i+1)\Delta})\\
&\qquad-\sum_{i=0}^{\lfloor T/\Delta\rfloor-1} \mathbb{1}_{D_{4\Delta}\lambda}(\beta_{i\Delta})\cdot \mathbb{1}_{\bigcup_{k=1}^{\ell(\beta_{i\Delta})}\left\{\beta_{(i+1)\Delta}^{[i\Delta,\beta_{i\Delta}^k]}\neq_p\eta_{(i+1)\Delta}^{(i\Delta,\beta_{i\Delta}^k)}\right\}}
\end{aligned}
$$

resp. after dividing by $\lfloor T/\Delta\rfloor$

$$1 \geq \frac{1}{\lfloor T/\Delta\rfloor}\sum_{i=0}^{\lfloor T/\Delta\rfloor-1} \mathbb{1}_G\left((\eta_{i\Delta},\eta_{(i+1)\Delta})\right)$$

$$
\begin{aligned}
&\geq 1 - \frac{1}{\lfloor T/\Delta \rfloor} \sum_{i=0}^{\lfloor T/\Delta \rfloor - 1} \mathbb{1}_{S_{4\Delta^\lambda}}(\beta_{i\Delta}) \\
&\quad - \frac{1}{\lfloor T/\Delta \rfloor} \sum_{i=0}^{\lfloor T/\Delta \rfloor - 1} \mathbb{1}_{\bigcup_{k=1}^{\ell(\beta_{i\Delta})} \left\{ \left| \eta_{(i+1)\Delta}^{(i\Delta, \beta_{i\Delta}^k)} \right| \neq 1 \right\}} \\
&\quad - \frac{1}{\lfloor T/\Delta \rfloor} \sum_{k=1}^{\lfloor T/\Delta \rfloor - 1} \mathbb{1}_{\bigcup_{k=1}^{\ell(\beta_{i\Delta})} b^\Delta(i\Delta, \beta_{i\Delta}^k)} \\
&\quad - \frac{1}{\lfloor T/\Delta \rfloor} \sum_{i=0}^{\lfloor T/\Delta \rfloor - 1} \mathbb{1}_{D_{4\Delta^\lambda}}(\beta_{i\Delta}) \cdot \mathbb{1}_{S_{2\Delta^\lambda}}(\beta_{(i+1)\Delta}) \\
&\quad - \frac{1}{\lfloor T/\Delta \rfloor} \sum_{i=0}^{\lfloor T/\Delta \rfloor - 1} \mathbb{1}_{D_{4\Delta^\lambda}}(\beta_{i\Delta}) \cdot \mathbb{1}_{\bigcup_{k=1}^{\ell(\beta_{i\Delta})} \left\{ \beta_{(i+1)\Delta}^{[i\Delta, \beta_{i\Delta}^k]} \neq_p \eta_{(i+1)\Delta}^{(i\Delta, \beta_{i\Delta}^k)} \right\}} .
\end{aligned}
\tag{3.22}
$$

Using Theorem 3.5, we conclude

$$
1 \geq \mathbf{E}_m \left(\frac{1}{\lfloor T/\Delta \rfloor} \sum_{i=0}^{\lfloor T/\Delta \rfloor - 1} \mathbb{1}_G \left((\eta_{i\Delta}, \eta_{(i+1)\Delta}) \right) \right) \geq 1 - \mathcal{O}(\Delta^\lambda) - \mathcal{O}(\Delta).
$$

Since $0 < \lambda < \frac{1}{2}$, this means

$$
1 \geq \mathbf{E}_m \left(\frac{1}{\lfloor T/\Delta \rfloor} \sum_{i=0}^{\lfloor T/\Delta \rfloor - 1} \mathbb{1}_G \left((\eta_{i\Delta}, \eta_{(i+1)\Delta}) \right) \right) \geq 1 - \mathcal{O}(\Delta^\lambda),
$$

as $\Delta \to 0$. Our proof is complete.

$\square$

Remark 3.7

In inequality (3.22), we could have estimated

$$
-\frac{1}{\lfloor T/\Delta \rfloor} \sum_{i=0}^{\lfloor T/\Delta \rfloor - 1} \mathbb{1}_{D_{4\Delta^\lambda}}(\beta_{i\Delta}) \cdot \mathbb{1}_{S_{2\Delta^\lambda}}(\beta_{(i+1)\Delta}) \geq -\frac{1}{\lfloor T/\Delta \rfloor} \sum_{i=0}^{\lfloor T/\Delta \rfloor - 1} \mathbb{1}_{S_{2\Delta^\lambda}}(\beta_{(i+1)\Delta}). \tag{3.23}
$$

Then, we could renounce using the fourth part of Proposition 3.5 and by applying the first part of Proposition 3.5, we still receive the same result from Theorem 3.4. However, by our previous proof we can see that the only term with weaker order is

$$
\mathbf{E}_m \left(\frac{1}{\lfloor T/\Delta \rfloor} \sum_{i=0}^{\lfloor T/\Delta \rfloor - 1} \mathbb{1}_{S_{4\Delta^\lambda}}(\beta_{i\Delta}) \right) = \mathcal{O}(\Delta^\lambda),
$$

as $\Delta \to 0$, since the other terms are of order $\mathcal{O}(\Delta)$, as $\Delta \to 0$. This is why we do not estimate as in (3.23).

$\Diamond$

Due to the previous results, we receive the following useful corollary.

Corollary 3.8
Grant Assumptions 1.13, 1.19 and 2.1(a).

1. *The expected quota of interpretable pairs* $(\eta_{i\Delta}, \eta_{(i+1)\Delta})$, $0 \leq i \leq \lfloor T/\Delta \rfloor - 1$, *converges to* 1. *It even holds*

$$\begin{aligned} 1 &\geq \mathbf{E}_m\left(\frac{1}{\lfloor T/\Delta \rfloor} \sum_{i=0}^{\lfloor T/\Delta \rfloor - 1} \mathbb{1}_{\widetilde{G}}\left((\eta_{i\Delta}, \eta_{(i+1)\Delta})\right) \right) \\ &= \mathbf{P}_m\Big(\big\{(\eta_0, \eta_\Delta) \in \widetilde{G}\big\}\Big) \\ &\geq 1 - \mathcal{O}(\Delta^\lambda), \end{aligned}$$

   as $\Delta \to 0$.

2. *For every* $s \geq 0$, *it holds*

$$\mathbf{P}_m\Big(\big\{(\eta_s, \eta_{s+\Delta}) \in \widetilde{G} \backslash G\big\}\Big) = \mathcal{O}(\Delta),$$

   as $\Delta \to 0$.

Proof
The first part directly follows by Theorem 3.4 and $G \subseteq \widetilde{G}$. The second part follows by using similar arguments to those in the proof of Proposition 3.5.

$\square$

Chapter 4

Non-Parametric Estimation of the Diffusion Coefficient of Branching Diffusions with Immigration

In this chapter, we have three main goals: First, our aim is to construct an estimator for the squared diffusion coefficient $\sigma^2(\cdot)$ of a BDI from discrete data by using the results from the previous chapters. Secondly, we want to show consistency of this estimator. Thirdly, we show that our estimator fulfils a central limit theorem.
As a BDI process contains many one-dimensional diffusions which branch and immigrate, it is helpful to know how estimation for the diffusion coefficient of a one-dimensional diffusion (as in (1.1)) works. In the case that a one-dimensional diffusion is observed at discrete points in time, non-parametric estimation of the diffusion coefficient is an issue which has been widely elaborated in literature, for example by Dacunha-Castelle and Florens-Zmirou in [3], by Florens-Zmirou in [7], by Comte, Genon-Catalot and Rozenholc in [4] resp. [25, p. 341f], by Kutoyants in [27], by Genon-Catalot and Jacod in [8] and [9], by Jacod in [23], or by Hoffmann in [17], [18] and [19]. We want to take a closer look at some of them.

4.1 Estimators for One-Dimensional Diffusions

A well-known estimator for the squared diffusion coefficient of a one-dimensional diffusion is an estimator by Florens-Zmirou in [7] which is based on the Nadaraya-Watson-Estimator. For this, let $X = X^{j,\ell}$ be the (strong) solution of (1.1), let h_Δ be the bandwidth and $B_{h_\Delta}(x) := \big\{y \in E \,\big|\, |y-x| < h_\Delta\big\}$. Then, for fixed point in time $T := 1$ and for $x \in E$ the estimator is

$$\widetilde{\sigma}^2_\Delta(x) := \frac{\sum\limits_{i=0}^{\lfloor T/\Delta\rfloor -1} \mathbb{1}_{B_{h_\Delta}(x)}(X_{i\Delta}) \cdot \left(\frac{X_{(i+1)\Delta}-X_{i\Delta}}{\sqrt{\Delta}}\right)^2}{\sum\limits_{i=0}^{\lfloor T/\Delta\rfloor -1} \mathbb{1}_{B_{h_\Delta}(x)}(X_{i\Delta})} \cdot \mathbb{1}_{\left\{\sum\limits_{i=0}^{\lfloor T/\Delta\rfloor -1} \mathbb{1}_{B_{h_\Delta}(x)}(X_{i\Delta}) \neq 0\right\}}. \tag{4.1}$$

Theorem 4.1 (Florens-Zmirou)
Let $X = X^{j,\ell}$ *be the solution of* (1.1) *with drift coefficient* $b(\cdot) \in \mathcal{C}_b^2(E)$ *and diffusion coefficient* $\sigma(\cdot) \in \mathcal{C}_b^3(E)$*. Further, let* $\widetilde{\sigma}^2_\Delta(\cdot)$ *be the estimator from* (4.1) *and let* $x \in [0,1]$*.*

1. *If* $\Delta^{-1}h_\Delta^4 \to 0$*, as* $\Delta \to 0$*, then* $\widetilde{\sigma}^2_\Delta(x)$ *is a consistent estimator for* $\sigma^2(x)$*.*

2. *If* $\Delta^{-1}h_\Delta^3 \to 0$*, as* $\Delta \to 0$*, then*

$$\sqrt{\frac{h_\Delta}{\Delta}}\left(\frac{\widetilde{\sigma}^2_\Delta(x)}{\sigma^2(x)} - 1\right) \xrightarrow{\Delta \to 0} \left(L_T(x)\right)^{-\frac{1}{2}} \cdot Z \quad \text{in } P\text{-distribution},$$

   where $L_T(x)$ *denotes the local time of the diffusion* X *at* x *and* Z *is a standard normal distributed random variable independently of* X*.*

Proof
The proof can be found in [7, p. 800f].

□

Remark 4.2
1. The estimator $\widetilde{\sigma}^2_\Delta(\cdot)$ is optimal in the minimax sense under square-error loss for the class of non-negative Lipschitz continuous diffusion coefficients being bounded and bounded away from zero, c.f. [19, p. 342f].

2. For an introduction to local times of stochastic processes, we refer to [31, p. 222f].

◊

Another possibility for estimating the squared diffusion coefficient of a one-dimensional diffusion was examined by Hoffmann in [17] and [18] by using wavelet methods and Besov spaces (for an introduction to wavelets and Besov spaces, we refer to [11, p. 17ff]). Hoffmann constructs wavelet-estimators which are optimal in the minimax sense (for integrated errors) over Besov balls, essentially by filling a classical regression scheme. For this, similarly to Theorem 2.3, under certain conditions on the drift and diffusion coefficients Hoffmann makes use of the regression identity

$$\left(\frac{X_{(i+1)\Delta} - X_{i\Delta}}{\sqrt{\Delta}}\right)^2 = \sigma^2(X_{i\Delta}) + \varepsilon_{i\Delta} + \mathcal{O}_P\big(\sqrt{\Delta}\big), \quad \text{as} \quad \Delta \to 0, \tag{4.2}$$

where $\varepsilon_{i\Delta}$ are martingale increments, c.f. [17, p. 449]. Assuming that the path of the diffusion is observed at discrete points in time $i\Delta$ with $0 \le i \le \lfloor T/\Delta \rfloor - 1$, he recovers $\sigma^2(x)$ from several observations $X_{i\Delta}$ by regarding (4.2), where each observation $X_{i\Delta}$ lies around x in an appropriate small area. More exactly, Hoffmann divides the compact set $D := [0,1]$ into $\lfloor h_\Delta^{-1} \rfloor$ many boxes C_Δ^ℓ and demands that every box has been filled with $\lfloor \Delta^{-1} h_\Delta \rfloor$ observations $X_{i\Delta}$ up to time $T := 1$. For each observed point $X_{i\Delta}$, he sets a mark on a regular grid into the box in which the observed point has fallen. By this, Hoffmann succeeds in constructing wavelet-estimators for the squared diffusion coefficient $\sigma^2(\cdot)$.

Nevertheless, in Hoffmann's procedure the following problem occurs: It is not clear that each box C^{ℓ}_{Δ}, $1 \le \ell \le \lfloor h_{\Delta}^{-1} \rfloor$, has been filled with $\lfloor \Delta^{-1} h_{\Delta} \rfloor$ observations $X_{i\Delta}$ up to time $T := 1$. For example, choose the compact set $[0,1]$ and consider the Ornstein-Uhlenbeck process

$$X_t = X_0 \cdot e^{-t} + 3 \cdot (1 - e^{-t}) + 0.3 \cdot \int_0^t e^{s-t}\, dW_s, \quad X_0 = 0.9 \quad \text{a.s.}$$

It is unlikely that this process takes on values in $[0, 0.5]$ during the time $[0,T]$. Hoffmann handles the problem of having enough observations $X_{i\Delta}$ in each box by conditioning on the event that for every $x \in [0,1]$ the local time

$$L_T(x) := \lim_{\varepsilon \to 0} \frac{1}{2\varepsilon} \int_0^T \mathbb{1}_{\left\{|X_s - x| \le \varepsilon\right\}}\, ds \tag{4.3}$$

is greater than a given threshold $\nu' > 0$, c.f. [17, p. 449] and [18, p. 136]. However, for statistical application this event is not appropriate since by observing the path of a diffusion at discrete points in time we are not able to determine whether this event has been fulfilled.

Remark 4.3

1. We want to outline the results from classical regression scheme, i.e., estimation of a regression function $f(\cdot)$ (by using a local polynomial estimator of order $r \in \mathbb{N}$) from $n \in \mathbb{N}$ observations (X_i, Y_i), where X_i are equally spaced or uniformly distributed on $D := [0,1]$ and

   $$Y_i = f(X_i) + \varepsilon_i$$

   with centered martingale increments ε_i, c.f. [34, p. 34f]. The estimator makes use of nh observations and (under certain assumptions) it has mean-squared risk MSE satisfying for every $f(\cdot)$ with regularity $r \in \mathbb{N}$ and for every $x \in [0,1]$

   $$\text{MSE} = \text{MSE}(x) = \big(\text{Bias}_f(x)\big)^2 + \text{Var}_f(x) \le c_1 \cdot h^{2r} + c_2 \cdot \frac{1}{nh}, \tag{4.4}$$

   where c_1 and c_2 are positive constants, c.f. [34, p. 40]. Thus, by bias-variance-tradeoff, we receive the optimal bandwidth $h := h_n := n^{-\frac{1}{1+2r}}$ and

   $$\text{MSE} = \mathcal{O}\big(n^{-\frac{2r}{1+2r}}\big), \quad \text{as} \quad n \to \infty. \tag{4.5}$$

   Hence, the optimal rate of convergence is $n^{-\frac{r}{1+2r}}$ which coincides with $h = h_n$ for $r = 1$.

2. In [5], the rates from classical regression are achieved for estimating a regression function $f(\cdot)$ by using wavelet methods. Based on this, in [17] and [18] Hoffmann shows that his wavelet-based regression estimators for $\sigma^2(\cdot)$ are minimax (for integrated errors) over Besov balls, i.e., they attain the rate $\Delta^{\frac{r}{1+2r}}$, where $r \in \mathbb{N}$ is the regularity of the diffusion coefficient $\sigma(\cdot)$. Also, Jacod achieves the same rate in [23], where he constructs a non-parametric estimator of kernel type for the diffusion coefficient $\sigma(\cdot)$, both for pointwise estimation and for estimation on a compact subset $D \subseteq E$.

$\Diamond$

4.2 Construction of the Estimator

In [2, p. 57f], Brandt developed an estimator for the squared diffusion coefficient of a BDI resting on Florens-Zmirou's Nadaraya-Watson-Estimator. We now want to construct an estimator for the squared diffusion coefficient of a BDI by approaching Hoffmann's way of filling a classical regression scheme and combining it with our reconstruction algorithm from the third chapter.

We consider the BDI process at discrete points in time $i\Delta$ during the time interval $[0, T_\Delta]$, where $0 \le i \le \lfloor T_\Delta/\Delta \rfloor - 1$ (in our case, the time horizon T_Δ depends on Δ and is specified below). Our aim is to estimate $\sigma^2(x)$ for some $x \in D$, where $D \subseteq E$ is compact. W.l.o.g., we may set $D := [0,1]$. Denoting $h'_\Delta := \left(\lfloor h_\Delta^{-1} \rfloor\right)^{-1}$, we divide $[0,1]$ into h'^{-1}_Δ many boxes

$$C_\Delta^\ell := \left[(\ell - 1) \cdot h'_\Delta,\, \ell \cdot h'_\Delta\right], \quad 1 \le \ell \le h'^{-1}_\Delta.$$

Remember the interpretable pairs from Definition 3.1 and define for every $1 \le \ell \le h'^{-1}_\Delta$

$$\widetilde{G}_\Delta^\ell := \Big\{ (\eta_{i\Delta}, \eta_{(i+1)\Delta}) \in \mathcal{S} \times \mathcal{S} \,\Big|\, i \in \mathbb{N}_0, (\eta_{i\Delta}, \eta_{(i+1)\Delta}) \text{ is interpretable and } \exists k \in \{1, ..., \ell(\eta_{i\Delta})\} \colon \eta_{i\Delta}^k \in C_\Delta^\ell \Big\} \tag{4.6}$$

the set of all interpretable pairs $(\eta_{i\Delta}, \eta_{(i+1)\Delta})$ whose first configuration $\eta_{i\Delta}$ at least contains one particle in C_Δ^ℓ. For every $1 \le \ell \le h'^{-1}_\Delta$ we now define $\mathcal{H}_{i\Delta} = \sigma\left(\eta_0, \eta_\Delta, ..., \eta_{(i+1)\Delta}\right)$ stopping times via

$$\begin{aligned} T_1^\ell &:= \inf \Big\{ i\Delta \ge 0 \,\Big|\, (\eta_{i\Delta}, \eta_{(i+1)\Delta}) \in \widetilde{G}_\Delta^\ell \Big\}, \\ T_j^\ell &:= \inf \Big\{ i\Delta > T_{j-1}^\ell \,\Big|\, (\eta_{i\Delta}, \eta_{(i+1)\Delta}) \in \widetilde{G}_\Delta^\ell \Big\}, \quad j \in \mathbb{N}\backslash\{1\}. \end{aligned} \tag{4.7}$$

Define for $1 \le \ell \le h'^{-1}_\Delta$ and $j \in \mathbb{N}$

$$\beta^*_{T_j^\ell} := \inf \Big\{ \beta^k_{T_j^\ell} \,\Big|\, 1 \le k \le \ell(\beta_{T_j^\ell}),\ \beta^k_{T_j^\ell} \in C_\Delta^\ell \Big\} \tag{4.8}$$

and $\beta^*_{T_j^\ell + \Delta}$ let be the position of the unique particle (belonging to $\beta_{T_j^\ell + \Delta}$) whose distance to $\beta^*_{T_j^\ell}$ is smaller than or equal to Δ^λ (this particle exists because of the property of being interpretable). Further, we define for every $1 \le \ell \le h'^{-1}_\Delta$ and $j \in \mathbb{N}$ the increments

$$Y^*_{T_j^\ell} := \frac{\beta^*_{T_j^\ell + \Delta} - \beta^*_{T_j^\ell}}{\sqrt{\Delta}}. \tag{4.9}$$

Let M_Δ be monotonously increasing with $M_\Delta \to \infty$, as $\Delta \to 0$. Provided that every box C_Δ^ℓ is filled with $\lfloor M_\Delta \rfloor$ points during the time $[0, T_\Delta]$, i.e., by chosing

$$T_\Delta := \sup_{1 \le \ell \le h'^{-1}_\Delta} T^\ell_{\lfloor M_\Delta \rfloor} + \Delta, \tag{4.10}$$

we define the estimator $\hat{\sigma}^2_\Delta(\cdot)$ for $\sigma^2(\cdot)$ as follows.

Definition 4.4 (Estimator for the Squared Diffusion Coefficient of a BDI)
By using the previous notations, the estimator $\hat{\sigma}^2_\Delta(\cdot)$ for the squared diffusion coefficient $\sigma^2(\cdot)$ of a BDI is defined for every $x \in [0,1]$ via

$$\hat{\sigma}^2_\Delta(x) := \sum_{\ell=1}^{h'^{-1}_\Delta} \mathbb{1}_{C^\ell_\Delta}(x) \cdot \frac{1}{\lfloor M_\Delta \rfloor} \sum_{j=1}^{\lfloor M_\Delta \rfloor} \big(Y^*_{T^\ell_j}\big)^2.$$

$\Diamond$

Remark 4.5

1. The estimator works as follows: For given $x \in [0,1]$, the first sum in Definition 4.4 chooses the box $C^\ell_\Delta = \big[(\ell-1) \cdot h'_\Delta,\, \ell \cdot h'_\Delta\big]$, $1 \le \ell \le h'^{-1}_\Delta$, which contains x. Then, the estimator $\hat{\sigma}^2_\Delta(x)$ estimates $\sigma^2(x)$ by making use of $\lfloor M_\Delta \rfloor$ quadratic increments
$$\big(Y^*_{T^\ell_j}\big)^2 = \left(\frac{\beta^*_{T^\ell_j+\Delta} - \beta^*_{T^\ell_j}}{\sqrt{\Delta}}\right)^2$$
and averaging them. By specifying h_Δ and M_Δ appropriately and by applying the reconstruction algorithm from the third chapter together with the following regression identity for one-dimensional diffusions $(X_t)_{t\ge 0}$ (see Theorem 2.3)
$$\left(\frac{X_{t+\Delta} - X_t}{\sqrt{\Delta}}\right)^2 = \sigma^2(X_t) \cdot \big(1 + U_{(t,\Delta)}\big) + \mathcal{O}_P\big(\sqrt{\Delta}\big), \quad \text{as} \quad \Delta \to 0,$$
where $U_{(t,\Delta)}$ is a $\mathcal{F}'_{t+\Delta}$-measurable random variable being independent of $\mathcal{F}'_t$ and satisfying
$$U_{(t,\Delta)} \overset{d}{=} 2 \cdot \int_0^1 W_s \, dW_s$$
for every $t \ge 0$ and $\Delta > 0$, we will prove that our estimator is consistent and that it fulfils a central limit theorem.

2. We have to make sure that T_Δ in (4.10) is finite $\mathbf{P}_m$-a.s. Otherwise, defining the estimator in Definition 4.4 does not make sense since its sum needs $\lfloor M_\Delta \rfloor$ quadratic increments $\big(Y^*_{T^\ell_j}\big)^2$. In the next subsection, we will prove that finiteness of T_Δ can be assured by applying the Harris recurrence of the BDI, the rate of the reconstruction algorithm and the continuity of the density of $\overline{m}(\cdot)$.

3. In (4.8), we could also set
$$\beta^*_{T^\ell_j} := \sup\Big\{\beta^k_{T^\ell_j} \,\big|\, 1 \le k \le \ell(\beta_{T^\ell_j}),\ \beta^k_{T^\ell_j} \in C^\ell_\Delta\Big\}$$
because it does not matter which particle in the box C^ℓ_Δ at time T^ℓ_j is chosen. For filling points in the boxes C^ℓ_Δ, $1 \le \ell \le h'^{-1}_\Delta$, one could also think about using every particle of $\beta_{T^\ell_j}$ which lies in C^ℓ_Δ. However, since the pair $(\eta_{T^\ell_j}, \eta_{T^\ell_j+\Delta})$ is interpretable particles of $\beta_{T^\ell_j}$ have a distance of more than $4\Delta^\lambda$ to each other. This is why we merely take one particle of $\beta_{T^\ell_j}$ which lies in C^ℓ_Δ.

$\Diamond$

4.2.1 Specifications

In this subsection, we want to specify $0 < \lambda < \frac{1}{2}$ from the reconstruction algorithm and h_Δ resp. M_Δ from the previous construction. Furthermore, we want to explain why the time horizon T_Δ as defined in (4.10) is finite $\mathbf{P}_m$-a.s. Because of the considerations from the first remark in Remark 4.3, we know that the optimal bandwidth $h_\Delta = \Delta^{\frac{1}{1+2r}}$ in classical regression framework depends on the regularity $r \in \mathbb{N}$ of the diffusion coefficient $\sigma(\cdot)$. We adapt this bandwidth to our framework and examine the case $r = 1$ which deals with the class of non-negative Lipschitz continuous diffusion coefficients $\sigma(\cdot)$ being bounded and bounded away from zero.

Assumption 4.6
Let $\Delta > 0$ and let the diffusion coefficient $\sigma(\cdot)$ be in the class of non-negative Lipschitz continuous functions being bounded and bounded away from zero. Further, let $0 < \varepsilon < \frac{1}{3}$.

1. *The parameter λ fulfils $0 < \frac{5}{12} < \lambda < \frac{1}{2}$.*
2. *The bandwidth h_Δ is set to $h_\Delta := \Delta^{\frac{1}{3}}$.*
3. *M_Δ is set to either*
   (a) $M_\Delta := \Delta^{-1} h_\Delta = \Delta^{-\frac{2}{3}}$

   or

   (b) $M_\Delta := M_{\Delta,\varepsilon} := \Delta^{-1} h_\Delta \cdot 2\Delta^{2\varepsilon} = 2\Delta^{-\frac{2}{3}+2\varepsilon}$.

$\Diamond$

Remark 4.7

1. Depending on how to set M_Δ, we will write Assumption 4.6(a) resp. Assumption 4.6(b). For showing consistency of our estimator, we will make use of Assumption 4.6(a), i.e., the number of used observations is the same as in classical regression framework (see first remark in Remark 4.3). However, to prove a central limit theorem we will see that we have to use Assumption 4.6(b) for technical reasons (the exact reasons for this will be explained after the proof). Apart from that, later we will outline why we restrict ourselves to the case $r = 1$.

2. By the first two points of Assumption 4.6, it holds

$$\begin{aligned} 0 \leq \frac{\Delta^{2\lambda}}{h_\Delta} \leq \frac{\Delta^{\lambda}}{h_\Delta} \leq \frac{\Delta^{\frac{4}{5}\lambda}}{h_\Delta} \\ = \frac{\Delta^{\frac{4}{5}\lambda}}{\Delta^{\frac{1}{3}}} \\ = \Delta^{\frac{4}{5}\lambda - \frac{1}{3}} \xrightarrow{\Delta \to 0} 0 \end{aligned} \tag{4.11}$$

and

$$\frac{\sqrt{\Delta}}{h_\Delta} = \Delta^{\frac{1}{6}} \xrightarrow{\Delta \to 0} 0. \tag{4.12}$$

Applying Assumption 4.6(a), we receive

$$M_\Delta^{-\frac{1}{2}} = h_\Delta \tag{4.13}$$

and by Assumption 4.6(b), we obtain for every $0 < \varepsilon < \frac{1}{3}$

$$M_{\Delta,\varepsilon}^{-\frac{1}{2}} = \frac{1}{\sqrt{2}} \cdot h_\Delta \cdot \Delta^{-\varepsilon}. \tag{4.14}$$

$\diamondsuit$

We now want to show that the time horizon in (4.10) is finite $\mathbf{P}_m$-a.s. Since the second part of Remark 1.11 provides that the skeleton chain $(\eta_{i\Delta})_{i\in\mathbb{N}_0}$ (resp. $\big((\eta_{i\Delta}, \eta_{(i+1)\Delta})\big)_{i\in\mathbb{N}_0}$) is positive recurrent in the sense of Harris, for verifying the finiteness of the stopping times T_ℓ^j for every $1 \le \ell \le h_\Delta'^{-1}$ and $1 \le j \le \lfloor M_\Delta \rfloor$, we have to show

$$\int \mathbb{1}_{\widetilde{G}}\big((\eta_0, \eta_\Delta)\big) \cdot \eta_0(C_\Delta^\ell)\, dm' > 0,$$

where $m'(\cdot)$ is the invariant measure of the skeleton chain. As we know from the second part of Remark 1.11 that $m'(\cdot)$ can be identified with $m(\cdot)$, we can also check that

$$\int \mathbb{1}_{\widetilde{G}}\big((\eta_0, \eta_\Delta)\big) \cdot \eta_0(C_\Delta^\ell)\, dm > 0. \tag{4.15}$$

We want to verify this in the next proposition.

Proposition 4.8

Let Assumptions 1.13, 1.14, 1.18, 1.19, 2.1 and 4.6 hold. Then, for sufficiently small $\Delta > 0$ it holds

$$\int \mathbb{1}_{\widetilde{G}}\big((\eta_0, \eta_\Delta)\big) \cdot \eta_0(C_\Delta^\ell)\, dm > 0.$$

Proof

According to Theorem 1.20, the occupation measure $\overline{m}(\cdot)$ has a continuous density $\frac{d\overline{m}}{d\lambda\!\!\lambda}(\cdot)$. As $\sigma(\cdot)$ is bounded away from zero, the density $\frac{d\overline{m}}{d\lambda\!\!\lambda}(\cdot)$ is strictly positive according to the first remark in Remark 1.21. This is why

$$C_1 := \min_{x\in[0,1]} \left\{ \frac{d\overline{m}}{d\lambda\!\!\lambda}(x) \right\} > 0 \tag{4.16}$$

and for every $1 \le \ell \le h_\Delta'^{-1}$ we receive

$$\overline{m}\big(C_\Delta^\ell\big) \ge C_1 \cdot h_\Delta. \tag{4.17}$$

Moreover, by the Hölder inequality (see for instance [33, p. 99]) with conjugates $p := \frac{5}{4}$ and $q := 5$, the first part of Corollary 3.8 and (1.43), there is $C_2 > 0$ such that for sufficiently small Δ and every $1 \leq \ell \leq h_\Delta'^{-1}$

$$\begin{aligned}\int \mathbb{1}_{\widetilde{G}^c}\big((\eta_0,\eta_\Delta)\big) \cdot \eta_0(C_\Delta^\ell)\, dm &\leq \Big(\int \mathbb{1}_{\widetilde{G}^c}\big((\eta_0,\eta_\Delta)\big)\, dm\Big)^{\frac{4}{5}} \cdot \Big(\int \ell^5(\eta_0)\, dm\Big)^{\frac{1}{5}} \\ &= \Big(\mathbf{P}_m\Big(\big\{(\eta_0,\eta_\Delta) \notin \widetilde{G}\big\}\Big)\Big)^{\frac{4}{5}} \cdot \Big(\int_{\mathcal{S}} \ell^5(\mathbf{x})\, m(d\mathbf{x})\Big)^{\frac{1}{5}} \\ &\leq C_2 \cdot \Delta^{\frac{4}{5}\lambda}. \end{aligned} \tag{4.18}$$

Because of (4.11), it holds

$$C_2 \cdot \frac{\Delta^{\frac{4}{5}\lambda}}{h_\Delta} \xrightarrow{\Delta \to 0} 0,$$

in particular this means that there is $\Delta_0 > 0$ such that for all $0 < \Delta \leq \Delta_0$

$$C_1 h_\Delta - C_2 \Delta^{\frac{4}{5}\lambda} > 0. \tag{4.19}$$

Now, we combine (4.17), (4.18), (4.19) and we receive that for every $0 < \Delta \leq \Delta_0$

$$\begin{aligned}\int \mathbb{1}_{\widetilde{G}}\big((\eta_0,\eta_\Delta)\big) \cdot \eta_0(C_\Delta^\ell)\, dm &= \int \eta_0(C_\Delta^\ell)\, dm - \int \mathbb{1}_{\widetilde{G}^c}\big((\eta_0,\eta_\Delta)\big) \cdot \eta_0(C_\Delta^\ell)\, dm \\ &= \int_{\mathcal{S}} \mathbf{x}(C_\Delta^\ell)\, m(d\mathbf{x}) - \int \mathbb{1}_{\widetilde{G}^c}\big((\eta_0,\eta_\Delta)\big) \cdot \eta_0(C_\Delta^\ell)\, dm \\ &= \overline{m}(C_\Delta^\ell) - \int \mathbb{1}_{\widetilde{G}^c}\big((\eta_0,\eta_\Delta)\big) \cdot \eta_0(C_\Delta^\ell)\, dm \\ &\geq C_1 h_\Delta - C_2 \Delta^{\frac{4}{5}\lambda} \\ &> 0. \end{aligned}$$

Our proof is complete.

□

Remark 4.9
In the proof of the previous proposition, there are two crucial passages we want to emphasise. First, for verifying (4.17) we make use of the previous equation (4.16), which relies on Assumptions 1.14, 1.18, 1.19 and 2.1. Secondly, we apply our reconstruction algorithm from the third chapter together with (1.43) in (4.18), which relies on Assumptions 1.13, 1.19 and 2.1. In particular, we entirely utilize Hammer's framework from subsection 1.3, i.e., we make use of Assumptions 1.13, 1.14, 1.18 and 1.19.

◊

By using the previous proposition, we receive the following theorem.

Theorem 4.10

Grant Assumptions 1.13, 1.14, 1.18, 1.19, 2.1 and 4.6. Further, for every $1 \leq \ell \leq h_\Delta'^{-1}$ *and* $1 \leq j \leq \lfloor M_\Delta \rfloor$ *let* T_j^ℓ *be defined as in* (4.7). *Then, for sufficiently small* $\Delta > 0$ *it holds*

$$T_\Delta := \left(\sup_{1 \leq \ell \leq h_\Delta'^{-1}} T_{\lfloor M_\Delta \rfloor}^\ell + \Delta \right) < \infty \quad \mathbf{P}_m\text{-a.s.}$$

Proof

As we have mentioned before, because of the Harris recurrence of the chain $(\eta_{i\Delta})_{i\in\mathbb{N}_0}$ (resp. $\big((\eta_{i\Delta}, \eta_{(i+1)\Delta})\big)_{i\in\mathbb{N}_0}$) and the previous proposition, for every $1 \leq \ell \leq h_\Delta'^{-1}$ and $1 \leq j \leq \lfloor M_\Delta \rfloor$ the stopping times T_j^ℓ are finite $\mathbf{P}_m$-a.s. (for sufficiently small $\Delta > 0$). This is why T_Δ is finite $\mathbf{P}_m$-a.s. for sufficiently small $\Delta > 0$, too.

$\square$

Due to the previous theorem, we receive the result that our regression scheme will be filled most certainly, assuming that we observe the BDI process sufficiently long. By this, in comparison to Hoffmann's way of filling a regression scheme (see (4.3)), our regression scheme is filled without conditioning on any local time.

4.3 Consistency of the Estimator

In this section, for every $x \in [0,1]$ we show that the estimator $\hat{\sigma}^2_\Delta(x)$ from Definition 4.4 is consistent for $\sigma^2(x)$. The following theorem holds.

Theorem 4.11

Grant Assumptions 1.13, 1.14, 1.18, 1.19, 2.1 and 4.6(a). Then, for every $x \in [0,1]$ the estimator $\hat{\sigma}^2_\Delta(x)$ is consistent for $\sigma^2(x)$, i.e.,

$$\hat{\sigma}^2_\Delta(x) \xrightarrow{\Delta \to 0} \sigma^2(x) \quad \textit{in } \mathbf{P}_m\textit{-probability.}$$

It even holds

$$\big|\hat{\sigma}^2_\Delta(x) - \sigma^2(x)\big| = \mathcal{O}_{\mathbf{P}_m}\big(\Delta^{\frac{1}{3}}\big), \quad \textit{as} \quad \Delta \to 0.$$

Proof

First of all, we notice that for every $1 \le \ell \le h'^{-1}_\Delta$ and $1 \le j \le \lfloor M_\Delta \rfloor$ the stopping times T^ℓ_j from (4.7) resp. the time horizon $T_\Delta := \sup_{1 \le \ell \le h'^{-1}_\Delta} T^\ell_{\lfloor M_\Delta \rfloor} + \Delta$ fulfil

$$\mathbf{P}_m\Big(\big\{T^\ell_j < \infty\big\}\Big) = \mathbf{P}_m\Big(\big\{T_\Delta < \infty\big\}\Big) = 1$$

for sufficiently small $\Delta > 0$, according to Theorem 4.10. This is why we set $\Delta > 0$ in the way that this condition is fulfilled and we observe the trajectory of the BDI at discrete points in time $i\Delta$, where $0 \le i \le \lfloor T_\Delta/\Delta \rfloor - 1$.

We remember the sets of interpretable pairs $\widetilde{G}$ and properly interpretable pairs G from Definitions 3.1 and 3.3. It holds

$$\begin{aligned}
0 \le \hat{\sigma}^2_\Delta(x) &= \sum_{\ell=1}^{h'^{-1}_\Delta} \mathbb{1}_{C^\ell_\Delta}(x) \cdot \frac{1}{\lfloor M_\Delta \rfloor} \sum_{j=1}^{\lfloor M_\Delta \rfloor} \big(Y^*_{T^\ell_j}\big)^2 \\
&= \sum_{\ell=1}^{h'^{-1}_\Delta} \mathbb{1}_{C^\ell_\Delta}(x) \cdot \frac{1}{\lfloor M_\Delta \rfloor} \sum_{j=1}^{\lfloor M_\Delta \rfloor} \big(Y^*_{T^\ell_j}\big)^2 \cdot \mathbb{1}_{\widetilde{G}}\big((\eta_{T^\ell_j}, \eta_{T^\ell_j+\Delta})\big) \\
&= \sum_{\ell=1}^{h'^{-1}_\Delta} \mathbb{1}_{C^\ell_\Delta}(x) \cdot \frac{1}{\lfloor M_\Delta \rfloor} \sum_{j=1}^{\lfloor M_\Delta \rfloor} \big(Y^*_{T^\ell_j}\big)^2 \cdot \mathbb{1}_{G}\big((\eta_{T^\ell_j}, \eta_{T^\ell_j+\Delta})\big) \\
&\quad + \sum_{\ell=1}^{h'^{-1}_\Delta} \mathbb{1}_{C^\ell_\Delta}(x) \cdot \frac{1}{\lfloor M_\Delta \rfloor} \sum_{j=1}^{\lfloor M_\Delta \rfloor} \big(Y^*_{T^\ell_j}\big)^2 \cdot \mathbb{1}_{\widetilde{G}\setminus G}\big((\eta_{T^\ell_j}, \eta_{T^\ell_j+\Delta})\big),
\end{aligned}$$

hence

$$\begin{aligned}
|\hat{\sigma}^2_\Delta(x) - \sigma^2(x)| &\le \Big| \sum_{\ell=1}^{h'^{-1}_\Delta} \mathbb{1}_{C^\ell_\Delta}(x) \cdot \frac{1}{\lfloor M_\Delta \rfloor} \sum_{j=1}^{\lfloor M_\Delta \rfloor} \Big(\big(Y^*_{T^\ell_j}\big)^2 \cdot \mathbb{1}_{G}\big((\eta_{T^\ell_j}, \eta_{T^\ell_j+\Delta})\big) - \sigma^2(x)\Big)\Big| \\
&\quad + \Big| \sum_{\ell=1}^{h'^{-1}_\Delta} \mathbb{1}_{C^\ell_\Delta}(x) \cdot \frac{1}{\lfloor M_\Delta \rfloor} \sum_{j=1}^{\lfloor M_\Delta \rfloor} \big(Y^*_{T^\ell_j}\big)^2 \cdot \mathbb{1}_{\widetilde{G}\setminus G}\big((\eta_{T^\ell_j}, \eta_{T^\ell_j+\Delta})\big)\Big| \\
&=: (\mathrm{I}) + (\mathrm{II}). \qquad (4.20)
\end{aligned}$$

Consider (I): Here, we show that

$$(\mathrm{I}) = \mathcal{O}_{\mathbf{P}_m}(h_\Delta) + O_{\mathbf{P}_m}\big(M_\Delta^{-\frac{1}{2}}\big), \tag{4.21}$$

as $\Delta \to 0$. For every $1 \leq \ell \leq h_\Delta'^{-1}$ and $1 \leq j \leq \lfloor M_\Delta \rfloor$ each particle situated at $\beta^*_{T_j^\ell}$ does not die during the time interval $(T_j^\ell, T_j^\ell + \Delta]$ because $(\eta_{T_j^\ell}, \eta_{T_j^\ell+\Delta})$ is a properly interpretable pair. This is why we may use Theorem 2.3, i.e., for every $1 \leq \ell \leq h_\Delta'^{-1}$ and $1 \leq j \leq \lfloor M_\Delta \rfloor$ the following identity holds

$$\begin{aligned}\big(Y^*_{T_j^\ell}\big)^2 \cdot \mathbb{1}_G\big((\eta_{T_j^\ell}, \eta_{T_j^\ell+\Delta})\big) &= \big(Y^*_{T_j^\ell}\big)^2 \cdot \mathbb{1}_G\big((\eta_{T_j^\ell}, \eta_{T_j^\ell+\Delta})\big) \cdot \mathbb{1}_{\{T_j^\ell < \infty\}} \\ &= \sigma^2(\beta^*_{T_j^\ell}) \cdot \big(1 + U_j^\ell\big) + \mathcal{O}_{\mathbf{P}_m}\big(\sqrt{\Delta}\big),\end{aligned} \tag{4.22}$$

as $\Delta \to 0$, where U_j^ℓ are independent identically distributed random variables with distribution $U_j^\ell \stackrel{d}{=} 2 \cdot \int_0^1 W_s \, dW_s$. Furthermore, by the boundedness resp. Lipschitz continuity of $\sigma(\cdot)$ and the fact that by construction for $x \in C_\Delta^\ell$ it holds $\big|\beta^*_{T_j^\ell} - x\big| = \mathcal{O}(h_\Delta)$, as $\Delta \to 0$, we receive

$$\big|\sigma^2(\beta^*_{T_j^\ell}) - \sigma^2(x)\big| = \mathcal{O}(h_\Delta), \tag{4.23}$$

as $\Delta \to 0$. Sticking together (4.22) and (4.23), we obtain

$$\begin{aligned}(\mathrm{I}) &= \Big|\sum_{\ell=1}^{h_\Delta'^{-1}} \mathbb{1}_{C_\Delta^\ell}(x) \cdot \frac{1}{\lfloor M_\Delta \rfloor} \sum_{j=1}^{\lfloor M_\Delta \rfloor} \Big(\big(Y^*_{T_j^\ell}\big)^2 \cdot \mathbb{1}_G\big((\eta_{T_j^\ell}, \eta_{T_j^\ell+\Delta})\big) - \sigma^2(x)\Big)\Big| \\ &= \Big|\sum_{\ell=1}^{h_\Delta'^{-1}} \mathbb{1}_{C_\Delta^\ell}(x) \cdot \frac{1}{\lfloor M_\Delta \rfloor} \sum_{j=1}^{\lfloor M_\Delta \rfloor} \Big(\sigma^2(\beta^*_{T_j^\ell}) \cdot \big(1 + U_j^\ell\big) + \mathcal{O}_{\mathbf{P}_m}\big(\sqrt{\Delta}\big) - \sigma^2(x)\Big)\Big| \\ &\leq \mathcal{O}_{\mathbf{P}_m}\big(\sqrt{\Delta}\big) + \sum_{\ell=1}^{h_\Delta'^{-1}} \mathbb{1}_{C_\Delta^\ell}(x) \cdot \frac{1}{\lfloor M_\Delta \rfloor} \sum_{j=1}^{\lfloor M_\Delta \rfloor} \big|\sigma^2(\beta^*_{T_j^\ell}) - \sigma^2(x)\big| \\ &\qquad + \Big|\sum_{\ell=1}^{h_\Delta'^{-1}} \mathbb{1}_{C_\Delta^\ell}(x) \cdot \frac{1}{\lfloor M_\Delta \rfloor} \sum_{j=1}^{\lfloor M_\Delta \rfloor} \sigma^2(\beta^*_{T_j^\ell}) \cdot U_j^\ell\Big| \\ &\leq \mathcal{O}_{\mathbf{P}_m}\big(\sqrt{\Delta}\big) + \mathcal{O}(h_\Delta) + \Big|\sum_{\ell=1}^{h_\Delta'^{-1}} \mathbb{1}_{C_\Delta^\ell}(x) \cdot \frac{1}{\lfloor M_\Delta \rfloor} \sum_{j=1}^{\lfloor M_\Delta \rfloor} \big(\sigma^2(\beta^*_{T_j^\ell}) - \sigma^2(x)\big) \cdot U_j^\ell\Big| \\ &\qquad + \Big|\sum_{\ell=1}^{h_\Delta'^{-1}} \mathbb{1}_{C_\Delta^\ell}(x) \cdot \frac{1}{\lfloor M_\Delta \rfloor} \sum_{j=1}^{\lfloor M_\Delta \rfloor} \sigma^2(x) \cdot U_j^\ell\Big| \\ &\leq \mathcal{O}_{\mathbf{P}_m}\big(\sqrt{\Delta}\big) + \mathcal{O}(h_\Delta) + \sum_{\ell=1}^{h_\Delta'^{-1}} \mathbb{1}_{C_\Delta^\ell}(x) \cdot \frac{1}{\lfloor M_\Delta \rfloor} \sum_{j=1}^{\lfloor M_\Delta \rfloor} \big|\sigma^2(\beta^*_{T_j^\ell}) - \sigma^2(x)\big| \cdot |U_j^\ell| \\ &\qquad + \sigma^2(x) \cdot \Big|\sum_{\ell=1}^{h_\Delta'^{-1}} \mathbb{1}_{C_\Delta^\ell}(x) \cdot \frac{1}{\lfloor M_\Delta \rfloor} \sum_{j=1}^{\lfloor M_\Delta \rfloor} U_j^\ell\Big|\end{aligned}$$

$$
\begin{aligned}
&= \mathcal{O}_{\mathbf{P}_m}\big(\sqrt{\Delta}\big) + \mathcal{O}(h_\Delta) + \mathcal{O}(h_\Delta) \cdot \frac{1}{\lfloor M_\Delta \rfloor} \sum_{j=1}^{\lfloor M_\Delta \rfloor} |U_j^{\ell'}| \\
&\qquad + \sigma^2(x) \cdot \Big| \frac{1}{\lfloor M_\Delta \rfloor} \sum_{j=1}^{\lfloor M_\Delta \rfloor} U_j^{\ell'} \Big|, \qquad (4.24)
\end{aligned}
$$

as $\Delta \to 0$.
Since $|U_j^{\ell'}| \overset{d}{=} 2 \cdot \big| \int_0^1 W_s \, dW_s \big|$ for every $j \in \mathbb{N}$ and $0 < \mathbf{E}_m\big(2 \cdot \big| \int_0^1 W_s \, dW_s \big|\big) \leq 2$, it holds

$$
\frac{1}{\lfloor M_\Delta \rfloor} \sum_{j=1}^{\lfloor M_\Delta \rfloor} |U_j^{\ell'}| = \mathcal{O}_{\mathbf{P}_m}(1), \qquad (4.25)
$$

as $\Delta \to 0$.
As $U_j^{\ell'}$, $j \in \mathbb{N}$, are independent identically distributed random variables with distribution $U_j^{\ell'} \overset{d}{=} 2 \cdot \int_0^1 W_s \, dW_s$, hence

$$
\mathbf{E}_m\bigg(2 \cdot \int_0^1 W_s \, dW_s\bigg) = 0 \quad \text{resp.} \quad \mathrm{Var}_m\bigg(2 \cdot \int_0^1 W_s \, dW_s\bigg) = 2 > 0, \qquad (4.26)
$$

by using the Central Limit Theorem (see for instance [22, p. 416]) it holds

$$
\Big| \frac{1}{\lfloor M_\Delta \rfloor} \sum_{j=1}^{\lfloor M_\Delta \rfloor} U_j^{\ell'} \Big| = \mathcal{O}_{\mathbf{P}_m}\big(M_\Delta^{-\frac{1}{2}}\big), \qquad (4.27)
$$

as $\Delta \to 0$. Applying (4.25), (4.27) and (4.12) to (4.24), we receive

$$
\begin{aligned}
&\mathcal{O}_{\mathbf{P}_m}\big(\sqrt{\Delta}\big) + \mathcal{O}(h_\Delta) + \mathcal{O}(h_\Delta) \cdot \frac{1}{\lfloor M_\Delta \rfloor} \sum_{j=1}^{\lfloor M_\Delta \rfloor} |U_j^{\ell'}| + \sigma^2(x) \cdot \Big| \frac{1}{\lfloor M_\Delta \rfloor} \sum_{j=1}^{\lfloor M_\Delta \rfloor} U_j^{\ell'} \Big| \\
&\qquad = \mathcal{O}_{\mathbf{P}_m}\big(\sqrt{\Delta}\big) + \mathcal{O}(h_\Delta) + \mathcal{O}(h_\Delta) \cdot \mathcal{O}_{\mathbf{P}_m}(1) + \sigma^2(x) \cdot \mathcal{O}_{\mathbf{P}_m}\big(M_\Delta^{-\frac{1}{2}}\big) \\
&\qquad = o_{\mathbf{P}_m}(h_\Delta) + \mathcal{O}_{\mathbf{P}_m}(h_\Delta) + \mathcal{O}_{\mathbf{P}_m}(h_\Delta) + \mathcal{O}_{\mathbf{P}_m}\big(M_\Delta^{-\frac{1}{2}}\big) \\
&\qquad = \mathcal{O}_{\mathbf{P}_m}(h_\Delta) + \mathcal{O}_{\mathbf{P}_m}\big(M_\Delta^{-\frac{1}{2}}\big), \qquad (4.28)
\end{aligned}
$$

as $\Delta \to 0$.

Consider (II): Here, we show that

$$
(\mathrm{II}) = o_{\mathbf{P}_m}(h_\Delta), \qquad (4.29)
$$

as $\Delta \to 0$, by verifying

$$
\mathbf{E}_m\big((\mathrm{II})\big) = \mathcal{O}\big(\Delta^{2\lambda}\big),
$$

as $\Delta \to 0$. Because of the property of being interpretable, for every $1 \leq \ell \leq h_\Delta'^{-1}$ and $1 \leq j \leq \lfloor M_\Delta \rfloor$ it holds $\big|\beta^*_{T_j^\ell+\Delta} - \beta^*_{T_j^\ell}\big| \leq \Delta^\lambda$, this is why

$$\big(Y^*_{T_j^\ell}\big)^2 = \left(\frac{\beta^*_{T_j^\ell+\Delta} - \beta^*_{T_j^\ell}}{\sqrt{\Delta}}\right)^2 \leq \Delta^{2\lambda-1}. \tag{4.30}$$

Using this and the second part of Corollary 3.8, i.e.,

$$\mathbf{E}_m\Big(\mathbb{1}_{\widetilde{G}\backslash G}\big((\eta_{T_j^\ell}, \eta_{T_j^\ell+\Delta})\big)\Big) = \mathbf{P}_m\Big(\big\{(\eta_{T_j^\ell}, \eta_{T_j^\ell+\Delta}) \in \widetilde{G}\backslash G\big\} \,\big|\, \big\{T_j^\ell < \infty\big\}\Big) = \mathcal{O}(\Delta), \tag{4.31}$$

as $\Delta \to 0$, it holds

$$\begin{aligned}
\mathbf{E}_m\big((\mathrm{II})\big) &= \mathbf{E}_m\left(\sum_{\ell=1}^{h_\Delta'^{-1}} \mathbb{1}_{C_\Delta^\ell}(x) \cdot \frac{1}{\lfloor M_\Delta \rfloor} \sum_{j=1}^{\lfloor M_\Delta \rfloor} \big(Y^*_{T_j^\ell}\big)^2 \cdot \mathbb{1}_{\widetilde{G}\backslash G}\big((\eta_{T_j^\ell}, \eta_{T_j^\ell+\Delta})\big)\right) \\
&\leq \Delta^{2\lambda-1} \cdot \sum_{\ell=1}^{h_\Delta'^{-1}} \mathbb{1}_{C_\Delta^\ell}(x) \cdot \frac{1}{\lfloor M_\Delta \rfloor} \sum_{j=1}^{\lfloor M_\Delta \rfloor} \mathbf{E}_m\Big(\mathbb{1}_{\widetilde{G}\backslash G}\big((\eta_{T_j^\ell}, \eta_{T_j^\ell+\Delta})\big)\Big) \\
&= \mathcal{O}\big(\Delta^{2\lambda}\big),
\end{aligned} \tag{4.32}$$

as $\Delta \to 0$. Since $\Delta^{2\lambda} = o(h_\Delta)$, as $\Delta \to 0$, see (4.11), it follows

$$(\mathrm{II}) = o_{\mathbf{P}_m}(h_\Delta),$$

as $\Delta \to 0$.

Combining (4.21) resp. (4.29) and using (4.13), we can complete our proof since

$$\begin{aligned}
\big|\hat{\sigma}^2_\Delta(x) - \sigma^2(x)\big| &\leq (\mathrm{I}) + (\mathrm{II}) \\
&= \mathcal{O}_{\mathbf{P}_m}(h_\Delta) + \mathcal{O}_{\mathbf{P}_m}\big(M_\Delta^{-\frac{1}{2}}\big) + o_{\mathbf{P}_m}(h_\Delta) \\
&= \mathcal{O}_{\mathbf{P}_m}(h_\Delta) + \mathcal{O}_{\mathbf{P}_m}\big(M_\Delta^{-\frac{1}{2}}\big) \\
&= \mathcal{O}_{\mathbf{P}_m}(h_\Delta) = \mathcal{O}_{\mathbf{P}_m}\big(\Delta^{\frac{1}{3}}\big),
\end{aligned} \tag{4.33}$$

as $\Delta \to 0$.

□

Remark 4.12

1. If we consider the result from the previous theorem, there is some analogy to the results from classical regression we have mentioned in the first remark in Remark 4.3. The second to last line of (4.33) gives

$$\big(\hat{\sigma}^2_\Delta(x) - \sigma^2(x)\big)^2 = \mathcal{O}_{\mathbf{P}_m}(h_\Delta^2) + \mathcal{O}_{\mathbf{P}_m}\big(M_\Delta^{-1}\big),$$

as $\Delta \to 0$. As we use Assumption 4.6(a), i.e., $M_\Delta^{-1} = \Delta h_\Delta^{-1}$, this result corresponds to MSE of (4.4) in classical regression and we receive its optimal rate $\Delta^{\frac{2}{3}}$, as $\Delta \to 0$, because we obtain

$$\left(\hat{\sigma}_\Delta^2(x) - \sigma^2(x)\right)^2 = \mathcal{O}_{\mathbf{P}_m}(\Delta^{\frac{2}{3}}),$$

as $\Delta \to 0$. The reason why we do not receive the rate $\mathcal{O}(\Delta^{\frac{2}{3}})$, as $\Delta \to 0$, (after taking expectation) as in (4.4) is that in our decomposition stochastically bounded terms emerge. However, this is an effect which particularly appears due to using the Central Limit Theorem in (4.27) and cannot be avoided.

2. We now want to explain why we restrict ourselves to the case $r = 1$. Actually, our proof for consistency works analogically for the class of $r \in \mathbb{N}\backslash\{1\}$-times continuously differentiable diffusion coefficients $\sigma(\cdot)$, where $\sigma(\cdot)$ is bounded and bounded away from zero and its r derivatives are each bounded. For this, the parameters λ, h_Δ and M_Δ have to be fit, i.e., we choose

$$\frac{3+2r}{4+8r} < \lambda < \frac{1}{2} \tag{4.34}$$

and motivated by classical regression framework

$$h_\Delta := \Delta^{\frac{1}{1+2r}} \quad \text{and} \quad M_\Delta := \Delta^{-1} h_\Delta = \Delta^{-\frac{2r}{1+2r}} \tag{4.35}$$

are set. Then, it holds

$$\left(\hat{\sigma}_\Delta^2(x) - \sigma^2(x)\right)^2 = \mathcal{O}_{\mathbf{P}_m}\left(\Delta^{\frac{2}{1+2r}}\right) + \mathcal{O}_{\mathbf{P}_m}\left(M_\Delta^{-1}\right),$$

as $\Delta \to 0$. As we can see, in comparison to classical regression from the first remark in Remark 4.3, we do not reach the power r in the bias term. The reason for this are equations (4.23) appearing in (4.28) and (4.23) being applied together with (4.25) in (4.28) since the terms

$$\mathcal{O}(h_\Delta) = \mathcal{O}_{\mathbf{P}_m}\left(\Delta^{\frac{1}{1+2r}}\right) \quad \text{resp.} \quad \mathcal{O}(h_\Delta) \cdot \mathcal{O}_{\mathbf{P}_m}(1) = \mathcal{O}_{\mathbf{P}_m}(h_\Delta) = \mathcal{O}_{\mathbf{P}_m}\left(\Delta^{\frac{1}{1+2r}}\right),$$

as $\Delta \to 0$, cannot be further estimated in any power of $r \in \mathbb{N}\backslash\{1\}$. This can be explained by the fact that in our case local polynomial weights of order $r \in \mathbb{N}\backslash\{1\}$ which eliminate each power being smaller than r are not available (in classical regression, such weights exist, c.f. [34, p. 36f]). As we do not have this analogy to classical regression results for $r \in \mathbb{N}\backslash\{1\}$, we restrict ourselves to $r = 1$.

$\Diamond$

4.4 A Central Limit Theorem for the Estimator

In this section, we show that our estimator from Definition 4.4 fulfils a central limit theorem. Here, we will work by using Assumption 4.6(b). Particularly, this means that our estimator depends on both $0 < \varepsilon < \frac{1}{3}$ and $\Delta > 0$. This is why in this case we will write for the estimator $\hat{\sigma}^2_{\Delta,\varepsilon}(\cdot)$ instead of $\hat{\sigma}^2_{\Delta}(\cdot)$.

Theorem 4.13
Grant Assumptions 1.13, 1.14, 1.18, 1.19, 2.1 and 4.6(b). Then, for every $0 < \varepsilon < \frac{1}{3}$ and $x \in [0,1]$ it holds

$$\sqrt{\Delta^{-\frac{2}{3}}} \cdot \Delta^{\varepsilon} \cdot \left(\frac{\hat{\sigma}^2_{\Delta,\varepsilon}(x)}{\sigma^2(x)} - 1 \right) \xrightarrow{\Delta \to 0} Z \quad \textit{in} \quad \mathbf{P}_m\textit{-distribution,}$$

where Z is a standard normal distributed random variable.

Proof
As in the previous theorem, we set $\Delta > 0$ small in the way that for every $1 \le \ell \le h'^{-1}_{\Delta}$ and $1 \le j \le \lfloor M_{\Delta,\varepsilon} \rfloor$ the stopping times T^{ℓ}_j from (4.7) resp. the time horizon

$$T_{\Delta,\varepsilon} := \sup_{1 \le \ell \le h'^{-1}_{\Delta}} T^{\ell}_{\lfloor M_{\Delta,\varepsilon} \rfloor} + \Delta$$

fulfil

$$\mathbf{P}_m\Big(\big\{ T^{\ell}_j < \infty \big\} \Big) = \mathbf{P}_m\Big(\big\{ T_{\Delta,\varepsilon} < \infty \big\} \Big) = 1,$$

see Theorem 4.10. Thus, we observe the trajectory of the BDI at discrete points in time $i\Delta$, where $0 \le i \le \lfloor T_{\Delta,\varepsilon}/\Delta \rfloor - 1$.

Similarly, as in (4.20) (here without the absolute value) it holds

$$\begin{aligned} \hat{\sigma}^2_{\Delta,\varepsilon}(x) - \sigma^2(x) &= \sum_{\ell=1}^{h'^{-1}_{\Delta}} \mathbb{1}_{C^{\ell}_{\Delta}}(x) \cdot \frac{1}{\lfloor M_{\Delta,\varepsilon} \rfloor} \sum_{j=1}^{\lfloor M_{\Delta,\varepsilon} \rfloor} \Big(\big(Y^*_{T^{\ell}_j}\big)^2 \cdot \mathbb{1}_G\big((\eta_{T^{\ell}_j}, \eta_{T^{\ell}_j+\Delta})\big) - \sigma^2(x) \Big) \\ &\quad + \sum_{\ell=1}^{h'^{-1}_{\Delta}} \mathbb{1}_{C^{\ell}_{\Delta}}(x) \cdot \frac{1}{\lfloor M_{\Delta,\varepsilon} \rfloor} \sum_{j=1}^{\lfloor M_{\Delta,\varepsilon} \rfloor} \big(Y^*_{T^{\ell}_j}\big)^2 \cdot \mathbb{1}_{\widetilde{G} \setminus G}\big((\eta_{T^{\ell}_j}, \eta_{T^{\ell}_j+\Delta})\big) \\ &=: (\mathrm{I}) + (\mathrm{II}). \end{aligned}$$

Consider (I): Taking (4.24) (drop the absolute value in the last line), together with (4.25) it holds

$$(\mathrm{I}) = \mathcal{O}_{\mathbf{P}_m}\big(\sqrt{\Delta}\big) + \mathcal{O}(h_{\Delta}) + \mathcal{O}_{\mathbf{P}_m}(h_{\Delta}) + \sigma^2(x) \cdot \frac{1}{\lfloor M_{\Delta,\varepsilon} \rfloor} \sum_{j=1}^{\lfloor M_{\Delta,\varepsilon} \rfloor} U^{\ell'}_j, \tag{4.36}$$

as $\Delta \to 0$. Defining

$$Z_{\Delta,\varepsilon} := \frac{\frac{1}{\lfloor M_{\Delta,\varepsilon} \rfloor} \cdot \sum_{j=1}^{\lfloor M_{\Delta,\varepsilon} \rfloor} U_j^{\ell'}}{\sqrt{\frac{2}{M_{\Delta,\varepsilon}}}},$$

by (4.12) we can rewrite (4.36) to

$$\text{(I)} = \mathcal{O}_{\mathbf{P}_m}(h_\Delta) + \sqrt{\frac{2}{M_{\Delta,\varepsilon}}} \cdot \sigma^2(x) \cdot Z_{\Delta,\varepsilon}, \tag{4.37}$$

as $\Delta \to 0$.

Consider (II): Using the same calculations as in the previous theorem (notice that (II) is non-negative), it holds

$$\text{(II)} = o_{\mathbf{P}_m}(h_\Delta), \tag{4.38}$$

as $\Delta \to 0$.
Now, combining (4.37) resp. (4.38) it holds

$$\begin{aligned} \hat{\sigma}^2_{\Delta,\varepsilon}(x) - \sigma^2(x) &= \mathcal{O}_{\mathbf{P}_m}(h_\Delta) + \sqrt{\frac{2}{M_{\Delta,\varepsilon}}} \cdot \sigma^2(x) \cdot Z_{\Delta,\varepsilon} + o_{\mathbf{P}_m}(h_\Delta), \\ &= \mathcal{O}_{\mathbf{P}_m}(h_\Delta) + \sqrt{\frac{2}{M_{\Delta,\varepsilon}}} \cdot \sigma^2(x) \cdot Z_{\Delta,\varepsilon}, \end{aligned} \tag{4.39}$$

as $\Delta \to 0$. As we have mentioned before, we make use of Assumption 4.6(b) (instead of Assumption 4.6(a)), i.e., $M_\Delta = M_{\Delta,\varepsilon}$ depends both on $0 < \varepsilon < \frac{1}{3}$ and $\Delta > 0$. Therefore, by using the identity

$$\mathcal{O}_{\mathbf{P}_m}(h_\Delta) = h_\Delta \cdot \Delta^{-\varepsilon} \cdot o_{\mathbf{P}_m}(1),$$

as $\Delta \to 0$, and (4.14), we receive for $0 < \varepsilon < \frac{1}{3}$ in (4.39)

$$\begin{aligned} \hat{\sigma}^2_{\Delta,\varepsilon}(x) - \sigma^2(x) &= h_\Delta \cdot \Delta^{-\varepsilon} \cdot o_{\mathbf{P}_m}(1) + \sqrt{\frac{2}{M_{\Delta,\varepsilon}}} \cdot \sigma^2(x) \cdot Z_{\Delta,\varepsilon} \\ &= h_\Delta \cdot \Delta^{-\varepsilon} \cdot \sigma^2(x) \cdot o_{\mathbf{P}_m}(1) + h_\Delta \cdot \Delta^{-\varepsilon} \cdot \sigma^2(x) \cdot Z_{\Delta,\varepsilon} \\ &= h_\Delta \cdot \Delta^{-\varepsilon} \cdot \sigma^2(x) \cdot \big(o_{\mathbf{P}_m}(1) + Z_{\Delta,\varepsilon}\big), \end{aligned}$$

as $\Delta \to 0$. As $\sigma(\cdot)$ is bounded away from zero, this means

$$\sqrt{\Delta^{-\frac{2}{3}}} \cdot \Delta^{\varepsilon} \cdot \left(\frac{\hat{\sigma}^2_{\Delta,\varepsilon}(x)}{\sigma^2(x)} - 1 \right) = o_{\mathbf{P}_m}(1) + Z_{\Delta,\varepsilon}, \tag{4.40}$$

as $\Delta \to 0$. Because of (4.26), the Central Limit Theorem and Slutsky's lemma (see for instance [24, p. 260]) give

$$Z_{\Delta,\varepsilon} = \frac{\frac{1}{\lfloor M_{\Delta,\varepsilon} \rfloor} \sum_{j=1}^{\lfloor M_{\Delta,\varepsilon} \rfloor} U_j^{\ell'}}{\sqrt{\frac{2}{M_{\Delta,\varepsilon}}}} \xrightarrow{\Delta \to 0} Z \quad \text{in} \quad \mathbf{P}_m\text{-distribution}, \tag{4.41}$$

where Z is a standard normal distributed random variable. Combining (4.40) and (4.41), we obtain by Slutsky's lemma for every $0 < \varepsilon < \frac{1}{3}$

$$\sqrt{\Delta^{-\frac{2}{3}}} \cdot \Delta^{\varepsilon} \cdot \left(\frac{\hat{\sigma}^2_{\Delta,\varepsilon}(x)}{\sigma^2(x)} - 1 \right) \xrightarrow{\Delta \to 0} Z \quad \text{in} \quad \mathbf{P}_m\text{-distribution.}$$

□

Remark 4.14

1. We want to explain why Assumption 4.6(b) (instead of Assumption 4.6(a)) is used in the previous theorem. Because of (4.39), there is a term which is merely stochastically bounded by h_Δ. This is slightly too weak since we need a term which goes to zero in probability in order to achieve a central limit theorem (see lines between (4.39) and (4.40)). This is why we use the identity

$$\mathcal{O}_{\mathbf{P}_m}(h_\Delta) = h_\Delta \cdot \Delta^{-\varepsilon} \cdot o_{\mathbf{P}_m}(1),$$

as $\Delta \to 0$, for sufficiently small $0 < \varepsilon < \frac{1}{3}$. This contributes to the fact that our estimator does not only depend on $\Delta > 0$ but also on ε and that the rates for our central limit theorem are slightly weaker than h_Δ. However, the factor $\Delta^{-\varepsilon}$ is not a severe loss if we consider the following: We remember that according to the first remark in Remark 4.2, Florens-Zmirou's estimator is optimal in the minimax sense under square-error loss for the class of non-negative Lipschitz continuous diffusion coefficients $\sigma(\cdot)$ being bounded and bounded away from zero. Nevertheless, for consistency and a central limit theorem $\sigma(\cdot)$ is assumed to have regularity $r = 3$ (see assumptions in Theorem 4.1), i.e., for technical reasons a greater regularity than $r = 1$ has to be set. In our framework, we merely demand that $\sigma(\cdot)$ is bounded, bounded away from zero and Lipschitz continuous. So, in comparison to Florens-Zmirou's regularity assumption $r = 3$, the factor $\Delta^{-\varepsilon}$ seems an acceptable loss.
Furthermore, the factor 2 in $M_{\Delta,\varepsilon}$ is needed for extinguishing the standard deviation part

$$\sqrt{\mathrm{Var}_m\left(2 \cdot \int_0^1 W_s \, dW_s\right)} = \sqrt{2}$$

which occurs by applying the Central Limit Theorem (see (4.41)).

2. For a central limit theorem, we could also consider the class of $r \in \mathbb{N}\backslash\{1\}$-times continuously differentiable diffusion coefficients $\sigma(\cdot)$, where $\sigma(\cdot)$ is bounded and bounded away from zero and its r derivatives are each bounded. Indeed, the previous proof works analogically by setting

$$\frac{3+2r}{4+8r} < \lambda < \frac{1}{2} \quad \text{resp.} \quad h_\Delta := \Delta^{\frac{1}{1+2r}}$$

and for given $0 < \varepsilon < \frac{1}{1+2r}$

$$M_{\Delta,\varepsilon} := 2 \cdot \Delta^{-\frac{2}{1+2r}+2\varepsilon}.$$

Then, it holds

$$\sqrt{\Delta^{-\frac{2}{1+2r}}} \cdot \Delta^{\varepsilon} \cdot \left(\frac{\hat{\sigma}^2_{\Delta,\varepsilon}(x)}{\sigma^2(x)} - 1 \right) \xrightarrow{\Delta \to 0} Z \quad \text{in} \quad \mathbf{P}_m\text{-distribution},$$

where Z is a standard normal distributed random variable. In this case, in comparison to M_Δ from (4.35), $M_{\Delta,\varepsilon}$ does not even have the power r, i.e., the gap to classical regression framework is greater. Because of this and on account of the same considerations as in the second remark of Remark 4.12, we again restrict ourselves to the case $r = 1$.

$\diamondsuit$

Bibliography

[1] AZEMA, J.; DUFLO, M.; REVUZ, D. (1969): Mesure invariante sur les classes récurrentes de Markov récurrents. Sem. Prob. Strasbourg. 3: 24-33.

[2] BRANDT, C. (2005): Partial reconstruction of the trajectories of a discretely observed branching diffusion with immigration and an application to inference. http://ubm.opus.hbz-nrw.de/volltexte/2005/756/pdf/diss.pdf on 06-01-2012.

[3] DACUNHA-CASTELLE, D.; FLORENS-ZMIROU, D. (1986): Estimation of the Coefficients of a Diffusion from Discrete Observations. An International Journal of Probability and Stochastic Processes. 19(4): 263-284.

[4] COMTE, F.; GENON-CATALOT, V.; ROZENHOLC, Y. (2007): Penalized nonparametric mean square estimation of the coefficients of diffusion processes. Bernoulli. 13(2): 514-543.

[5] DONOHO, D.; JOHNSTONE, I. (1998): Minimax estimation via wavelet shrinkage. The Annals of Statistics. 21: 879-921.

[6] DYNKIN, E. (1965): Markov Processes, Vol II. Grundlehren der mathematischen Wissenschaften 121-122. Berlin: Springer.

[7] FLORENS-ZMIROU, D. (1993): On estimating the diffusion coefficient from discrete observations. J. Appl. Prob. 30(4): 790-804.

[8] GENON-CATALOT, V.; JACOD, J. (1993): On the estimation of the diffusion coefficient for multi-dimensional diffusion processes. Annales de l'institut Henri Poincaré (B) Probabilités et Statistiques. 29: 119-151.

[9] GENON-CATALOT, V.; JACOD, J. (1994): Estimation of the diffusion coefficient for diffusion processes: Random sampling. Scandinavian Journal of Statistics. 21(3): 193-221.

[10] HAMMER, M. (2012): Ergodicity and Regularity of Invariant Measure for Branching Markov Processes with Immigration. http://ubm.opus.hbz-nrw.de/volltexte/2012/3306/pdf/doc.pdf on 01-04-2013.

[11] HÄRDLE, W.; KERKYACHARIAN, G.; PICARD, D.; TSYABKOV, A. (1998): Wavelets, Approximations and statistical Applications. New York: Springer.

[12] HÖPFNER, R.; HOFFMANN, M.; LÖCHERBACH, E. (2002): Non-parametric Estimation of the Death Rate in Branching Diffusions. Scandinavian Journal of Statistics. 29: 665-692.

[13] HÖPFNER, R. (2004): Strange shape of invariant density in branching diffusions with immigration. Unpublished preprint. Johannes Gutenberg-Universität Mainz 2004.

[14] HÖPFNER, R.; LÖCHERBACH, E. (1998): Birth and Death on a Flow: Local Time and Estimation of a Position-Dependent Death Rate. Statist. Inference Stoch. Process. 1: 225-243.

[15] HÖPFNER, R.; LÖCHERBACH, E. (2005): Remarks on ergodicity and invariant occupation measure in branching diffusions with immigration. Annales de l'Institut Henri Poincare (B) Probability and Statistics. 41: 1025-1047.

[16] HÖPFNER, R.; KUTOYANTS, Y. (2010): Estimating discontinuous periodic signals in a time inhomogeneous diffusion. Statist. Inference Stoch. Process. 13: 193-230.

[17] HOFFMANN, M. (1999): L_p estimation of the diffusion coefficient. Bernoulli. 5(3): 447-481.

[18] HOFFMANN, M. (1999): Adaptive estimation in diffusion processes. Stochastic processes and their Applications. 79: 135-163.

[19] HOFFMANN, M. (2001): On estimating the diffusion coefficient: parametric versus nonparametric. Annales de l'Institut Henri Poincare (B) Probability and Statistics. 37(3): 339-372.

[20] IKEDA, N.; NAGASAWA, M; WATANABE, M. (1968): Branching Markov Processes II. J. Math. Kyoto Univ. 8: 365-410.

[21] IKEDA, N.; WATANABE, S. (1989): Stochastic differential equations and diffusion processes. second edition. North-Holland Publ. Co.

[22] JACOD, J.; SHIRYAEV, A.N. (1987): Limit theorems for stochastic processes. Berlin Heidelberg New York: Springer.

[23] JACOD, J. (2000): Non-parametric Kernel Estimation of the Coefficient of a Diffusion. Scandinavian Journal of Statistics. 27: 83-96.

[24] KLENKE, A. (2012): Wahrscheinlichkeitstheorie. 3., überarbeitete und ergänzte Auflage. Berlin Heidelberg: Springer.

[25] KESSLER, M.; LINDNER, A.; SØRENSEN, M. (2012): Statistical Methods for Stochastic Differential Equation. CRC Press.

[26] KUSUOKA, S.; STROOCK, D. (1985): Applications of the Malliavin Calculus II. J. Fac. Sci. Univ. Tokyo. Sect. IA. Math. 32: 1-76.

[27] KUTOYANTS, Y. (1997): Some problems of nonparametric estimation by observations of ergodic diffusion process. Statistics & Probability Letters. 32(3): 311-320.

[28] LÖCHERBACH, E. (2002): Likelihood ratio processes for Markovian particle systems with killing and jumps. Statist. Inference Stoch. Process. 5: 153-177.

[29] LÖCHERBACH, E. (2004): Smoothness of the intensity measure density for interacting branching diffusions with immigrations. Journal of Functional Analysis. 215: 130-177.

[30] NAGASAWA, M. (1977): Basic Models of Branching Processes. Bull. Int. Stat. Inst. XLVII. 2: 423-445.

[31] REVUZ, D.; YOR, M. (2001): Continuous Martingales and Brownian Motion. Corr. 2. print. of the 3. ed. Berlin Heidelberg: Springer.

[32] ROGERS, L.; WILLIAMS, D. (1987): Diffusions, Markov Processes and Martingales. vol 2: Ito calculus. Wiley.

[33] ROGERS, L.; WILLIAMS, D. (1993): Diffusions, Markov Processes and Martingales. vol 1: Foundations. Wiley.

[34] TSYBAKOV, A. (2009): Introduction to Nonparametric Estimation. New York: Springer.