Partial reconstruction of the trajectories of a discretely observed branching diffusion with immigration and an application to inference

Dissertation

zur Erlangung des Grades
"Doktor der Naturwissenschaften"
der Johannes Gutenberg-Universität
in Mainz
am Fachbereich
17 Mathematik und Informatik

Christian Brandt

geboren in Trier

Mainz, 2005

Summary

In this treatise we consider finite systems of branching particles where the particles move independently of each other according to d-dimensional diffusions. Particles are killed at a position dependent rate κ, leaving at their death position a random number of descendants according to a position dependent reproduction law $p(\cdot)$ on $\mathbb{N}_0/\{1\}$. In addition particles immigrate at constant rate c (one immigrant per immigration time). A process φ with above properties is called a branching diffusion with immigration (BDI).

In the first part we present the model in detail and discuss the properties of the BDI under our basic assumptions.

In the second part we consider the problem of reconstruction of the trajectory of φ from discrete observations. We observe the positions of the particles of φ at discrete times $t_i = i\Delta$ for a small step width $\Delta > 0$; in particular we assume that we have no information about the pedigree of the particles.
A natural question arises if we want to apply statistical procedures on the discrete observations: How can we find couples of particle positions which belong to the same particle? We give an easy to implement 'reconstruction scheme' which allows us to redraw or 'reconstruct' parts of the trajectory of φ with high accuracy. Moreover asymptotically the whole path can be reconstructed. Further we present simulations which show that our partial reconstruction rule is tractable in practice.

In the third part we study how the partial reconstruction rule fits into statistical applications. As an extensive example we present a nonparametric estimator for the diffusion coefficient of a BDI where the particles of φ move according to one-dimensional diffusions. This estimator is based on the Nadaraya-Watson estimator for the diffusion coefficient of one-dimensional diffusions and it uses the partial reconstruction rule developed in the second part above. We are able to prove a rate of convergence of this estimator and finally we present simulations which show that the estimator works well even if we leave our set of assumptions.

Contents

Introduction

Spatial branching particle systems have made important contributions to biology and medicine in the areas of population genetics, epidemics and molecular biology (see e.g. the books of Jagers [Jag75] and Yakovlev/Yanev [YY89] and the papers of Sawyer [Saw76] and Iwasa/Teramoto [IT84]). From mathematical point of view this kind of particle systems have been widely investigated, see e.g. the papers of Wakolbinger [Wak95] and Gorostiza and Wakolbinger [GW94] for an intense study of the long time behavior of infinite systems of spatially branching diffusions and the papers of Löcherbach [Löc02b] and Höpfner et al. [HHL02] for results on statistical inference of finite systems of branching diffusions with immigration. Nevertheless there are still many open problems concerning statistical considerations of such particle systems; in particular discretely observed branching diffusions with immigration have not been considered yet.

In this treatise we consider finite systems of branching diffusions with immigration and random branching of particles. Our model can be described as follows:

Each particle of a finite system of particles moves in $\mathbb{R}^d$ independently of the other particles according to a d-dimensional diffusion

$$d\eta_t = b(\eta_t)dt + \sigma(\eta_t)dW_t, \quad t \geq 0,$$

with d-dimensional Brownian motion W and Lipschitz continuous coefficients b and σ.
Independently of each other particles die at position dependent rate $\kappa : \mathbb{R}^d \to \mathbb{R}_+$, i.e. a particle located at time t at position $y \in \mathbb{R}^d$ will die in a short time interval $(t, t+\Delta]$ with probability

$$\kappa(y)\Delta + o\,(\Delta)\,, \qquad \text{for } \Delta \downarrow 0.$$

At its death position the particle gives rise to a random number of offspring according to a position dependent reproduction law $p(\cdot) = (p_k(\cdot))_{k\neq 1}$. The newborn particles move and branch according to the same mechanism as the parent particles did.
Additionally particles immigrate at a constant rate c (one immigrant per immigration event) and choose their position in space according to a probability law π on $\mathbb{R}^d$.

The resulting process $\varphi = (\varphi_t)_{t\geq 0}$ can be constructed as a strong Markov process having càdlàg paths and values in the space S of finite configurations $x = (x^1, \ldots, x^{l(x)})$ with arbitrary length $l(x) \in \mathbb{N}$ of x, $x^i \in \mathbb{R}^d$.

This model is a special case of the model considered in [Löc00] where there is in addition interaction between the particles allowed. We also want to mention the papers of Ikeda et al. ([INW68a],[INW68b] and [INW69]) where general branching Markov processes were introduced the first time. Note that our model includes standard models like binary branching Brownian motions.

Inspired by a paper of Florens-Zmirou [FZ93], where a nonparametric estimator for the diffusion coefficient of an one-dimensional, ergodic diffusion based on discrete observations was presented the first time, we started studying discretely observed BDI's with the aim to construct an estimator for the diffusion coefficient of a BDI by similar methods as in [FZ93]. We realized very soon that there is one crucial difference between discretely observed diffusions and discretely observed BDI's. If we assume that we are able to observe only the positions of the particles of a BDI, we do not know which positions belong to which particle. Trivially this problem does not occur if we observe diffusion processes discretely in time.

We continue with a more detailed description of this problem. We consider a branching diffusion with immigration φ on a fixed time interval $[0, T]$ and we assume that we observe the process φ at discrete times $t_i = i\Delta$ of $[0, T]$, where $\Delta > 0$ is a small step width. Furthermore we assume that we are able to observe only the positions of the particles, i.e. an observation $\beta_{i\Delta}$ of the process φ at time $i\Delta$ is any arrangement of the support of the point measure $\sum_{k=1}^{l(\varphi_{i\Delta})} \epsilon_{\varphi^k_{i\Delta}}$ associated to the configuration $\varphi_{i\Delta} = (\varphi^1_{i\Delta}, \ldots, \varphi^{l(\varphi_{i\Delta})}_{i\Delta})$. In that case we have no information about the pedigree of the particles and, minding that we want apply statistical procedures on the data, one question arises in a very natural way: how shall we redraw or 'reconstruct' the trajectory of the process φ given the discrete observations $\beta_{i\Delta}$ in order to get a good approximation of the true trajectory? In other words: we want to have an easy to implement scheme which allows us to approximate at least parts of the true trajectory of φ with high accuracy.

A second problem we are interested in is to study how above 'reconstruction scheme' fits into statistical applications. As an (extensive) example we develop a nonparametric estimator for the diffusion coefficient of a branching diffusion with immigration in the case that the particles move according to one-dimensional diffusions.

It follows a brief description of the single sections of the present text.

In section 1 we introduce our basic notations and assumptions. Furthermore we give a short review on the construction of branching diffusions via 'elementary processes' (as proposed by Löcherbach in [Lö99]) and we discuss the properties

of our model under the basic assumptions. In particular we have under the assumption of spatial subcriticality ergodicity of the process φ, i.e. the process is Harris recurrent with the void configuration as a recurrence atom.
We close the section with some remarks on the existence of densities of the invariant measure and the invariant occupation measure.

Section 2 is dedicated to the problem of the partial reconstruction of the trajectory of a branching diffusion with immigration from discrete observations. In the beginning of section 2 we give a detailed description of the problem and propose a heuristic how we can solve this problem. The idea is very simple: we will assign to each component $\beta^k_{i\Delta}$ of an observation $\beta_{i\Delta} = (\beta^1_{i\Delta}, \ldots, \beta^{l(\beta_{i\Delta})}_{i\Delta}) \in S$ all components of the successive observation $\beta_{(i+1)\Delta}$ which are contained in a small neighborhood of $\beta^k_{i\Delta}$. Since particles move according to diffusion paths, which provide in particular continuous paths, this assignment seems to be reasonable (at least for a small step width Δ).
In this context two problems arise. First we have to insure that a particle does not fluctuate too much between two successive observation times, i.e. we have to control the probability that a single particle of the process φ (or at least one of its descendants) leaves a small neighborhood between two successive observation times. Secondly there are cases where above scheme does not assign the candidates for the descendants uniquely. This can happen whenever the small neighborhoods corresponding to the components $\beta^1_{i\Delta}, \ldots, \beta^{l(\beta_{i\Delta})}_{i\Delta}$ have a mutually nonempty intersection. In that case we just can guess which assignment is the right one and we prefer not to use observations with intersecting neighborhoods for the reconstruction of the trajectory of φ. Thus we have to show, that we do not loose too much data if we ignore these observations for the reconstruction.

Before we study the fluctuations of a branching diffusion with immigration we first consider in section 2.2 d-dimensional diffusions and prove an exponentially inequality for the probability that a d-dimensional diffusion leaves a small neighborhood of its starting position during a small time interval (lemma 2.2.4). Moreover, under the additional assumption of a bounded drift function b, we achieve an upper bound which is independent of the initial position of the diffusion (lemma 2.2.5) and hence we are able to apply the inequality globally on the single particles of φ.
Thereafter we consider in section 2.3.1 branching diffusions without immigration which start with one single particle. We construct a set which describes the event, that at least one particle (or at least one of its descendants) leaves its neighborhood between two successive observation times. In the main theorem of section 2.3 (theorem 2.3.1) we show that the probability of this event has an

exponentially bound, hence on the complement of this set all particles stay in their neighborhoods with 'high' probability (relative to the step width Δ).
Finally we generalize the result of theorem 2.3.1 for branching diffusions with immigrations in section 2.3.4.

A partial reconstruction rule is developed in section 2.4 where we also study the asymptotic properties of this rule (theorem 2.4.4). There we show that we are able to reconstruct parts of the trajectory of φ with high accuracy. Moreover, under the assumption that the invariant measure of φ admits a Lebesgue density, we are able to reconstruct asymptotically the whole trajectory.
Finally we present in section 2.5 simulations which show that our reconstruction scheme works well for reasonable sets of parameters.

Section 3 engages in the problem of estimating the diffusion coefficient of a BDI from discrete observations. We consider the case where the particles move according to one-dimensional diffusions.
We start our considerations with a brief overview on estimation of the diffusion coefficient of an one-dimensional diffusion (section 3.1). Subsequent we define in section 3.2 an estimator $\hat{\sigma}^2_\Delta(\cdot,\beta)$ for the diffusion coefficient σ^2 of a BDI φ based on discrete observations of the finite time interval $[0,T]$ which uses the partial reconstruction rule from section 2.4.
In section 3.3 we show the consistency of the estimator $\hat{\sigma}^2_\Delta(a,\beta)$ in three steps. First we consider the (hypothetical) case that we observe the whole trajectory of the process φ. In that case there is no need to use the partial reconstruction rule (since we have in particular all information about the pedigree of the particles) and we define a simpler estimator $\bar{\sigma}^2_\Delta(a,\varphi)$ which uses all available discrete data $\varphi_{i\Delta}, i=0,\dots,[T/\Delta]$. We show in theorem 3.3.4 that $\bar{\sigma}^2_\Delta(a,\varphi)$ attains the rate of convergence $\sqrt{\frac{\Delta}{h_\Delta}}$ (which trivially implies the consistency of this estimator).
Thereafter we maintain the assumption that we have all information about the trajectory of φ, but we consider an estimator $\hat{\sigma}^2(a,\varphi)$ which uses only the 'good observations' which are also used by the partial reconstruction rule from section 2.4. We show consistency of $\hat{\sigma}^2(a,\varphi)$ by comparing $\hat{\sigma}^2(a,\varphi)$ with the estimator $\bar{\sigma}^2_\Delta(a,\varphi)$ from the previous section 3.3.2.
In a last comparison between $\hat{\sigma}^2(a,\varphi)$ and $\hat{\sigma}^2_\Delta(a,\beta)$ we are able to show the consistency of our original estimator $\hat{\sigma}^2_\Delta(a,\beta)$ defined in section 3.2.
Finally we present in section 3.4 the main result of section 3 (theorem 3.4.2) which asserts that our estimator $\hat{\sigma}^2_\Delta(a,\beta)$ inherits the rate of convergence of the simpler estimator $\bar{\sigma}^2_\Delta(a,\varphi)$.

Last but not least we present in section 3.5 simulations of the estimator which show that our estimator produces good results already for few observations. Moreover the simulations suggest that the estimator is robust with respect to the size of the neighborhoods considered in the reconstruction rule 2.4.1 and the estimator also works in cases where our assumptions are not fulfilled.

1 Model

1.1 Branching diffusion with immigration

We consider a particle process where finitely many particles living in $\mathbf{R} := \mathbb{R}^d$ and traveling independently of each other according to solutions of the stochastic differential equation

$$d\eta_t = b(\eta_t)dt + \sigma(\eta_t)dW_t \tag{1}$$

with d-dimensional Brownian motions $W = (W_t)_{t\geq 0}$.

Independently of each other, particles are killed at a position-dependent rate

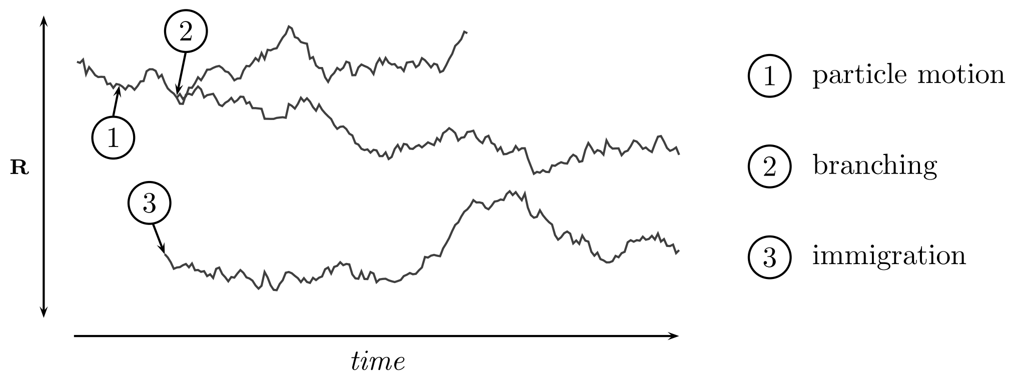
$$\kappa : \mathbf{R} \to (0, \infty),$$

and leave at their death position a random number of descendants according to a family of position dependent reproduction laws

$$p(\cdot) = (p_k(\cdot))_{k\in\mathbb{N}_0\setminus\{1\}}.$$

In other words: a parent particle in position $y \in \mathbf{R}$ at time $t \geq 0$ will die in a short time interval $]t, t + \Delta]$ with probability $\kappa(y)\Delta + o(\Delta), \Delta \downarrow 0$, leaving k descendants with probability $p_k(y)$ at position y.[1]

Finally new particles (one immigrant per immigration time) immigrate at constant rate $c > 0$ and choose their position in space according to a fixed probability law π on $\mathbf{R}$.



A process $\varphi = (\varphi_t)_{t\geq 0}$ with these properties is called a *branching diffusion with immigration (BDI)*.

[1] Since we are interested in statistical applications, we require reproduction laws which avoid the case $k = 1$, because we can't distinguish between a branching with one offspring and the event that no branching occurred.

This kind of process is a special case of branching markov processes considered already in the late 60s by Ikeda et al. ([INW68a], [INW68b] and [INW69]). They studied spatial branching particle systems where the particles move according to general markov processes. More recent work on finite branching particles systems can be found in the papers of Löcherbach and Höpfner (see e.g. [HHL02], [Löc02a], [Löc02b] and [Löc04]). They considered interacting branching diffusions with immigration. For infinite systems of branching diffusion we refer to the papers of Gorostiza and Wakolbinger (see [GW94] and [Wak95]).

1.2 Basic assumptions and properties of the model

The state space S of the BDI φ,

$$S = \bigcup_{\ell=0}^{\infty} \mathbf{R}^{\ell},$$

consists of all (ordered) configurations $x = (x^1, \ldots, x^{\ell}), x^i \in \mathbf{R}, i = 1, \ldots, \ell, \ell \geq 0$, with $\mathbf{R}^0 := \{\delta\}$, where δ denotes the void configuration.[2] Note that in particular S endowed with its Borel σ-algebra $\mathcal{S}$ is a Polish space (see [Lö99, 3.7, p. 15]).

Let $l(x)$ denote the length of the configuration $x \in S$, i.e. $l(x) = \ell \Leftrightarrow x \in \mathbf{R}^{\ell}$. Sometimes we write a configuration $x \in S$ as a point measure on $\mathbf{R}$:

$$x(A) := \begin{cases} \sum_{i=1}^{l(x)} \epsilon_{x^i}(A) & \text{if } l(x) \geq 1, \\ 0 & \text{if } x = \delta \text{ the void configuration.} \end{cases}$$

for $A \in \mathcal{B}(\mathbf{R})$.
By convention we set $f(\delta) := 0$ for any function f on S.

Our basic assumptions are the following:

Assumption A1. *The drift function $b : \mathbf{R} \to \mathbf{R}$ and the diffusion coefficient $\sigma : \mathbf{R} \to \mathbb{R}^{d \times d}$ are globally Lipschitz continuous with Lipschitz constant L.*

Assumption A2. *The branching rate $\kappa : \mathbf{R} \to (0, \infty)$ is supposed to be continuous, bounded and bounded away from 0 by a constant $\underline{\kappa} > 0$ and the reproduction mean $\rho : \mathbf{R} \to (0, \infty)$ given by*

$$\rho(x) := \sum_{k \neq 1} k \; p_k(x), \quad x \in \mathbf{R},$$

is a continuous function, bounded away from 1 by a constant $\bar{\rho} < 1$.

Remark 1.2.1. Assumption A1 ensures the existence and uniqueness of (strong) solutions of the stochastic differential equation (1) (see [IW89, theorem IV.3.1]). Assumption A2 will be needed for the ergodicity of the process. We present the details in section 1.2.3. In particular the assumptions on the branching rate κ and on the reproduction mean ρ are sufficient for the spacial subcriticality of the process φ (we refer to [HL03a, lemma 1.4.b]).

[2] It is also possible to choose the space of all point measures as state space. For the details we refer to appendix A in [Lö99].

1.2.1 The motion and the jump mechanism

Now we want to describe the mechanism of the process in a more detailed way. Write ζ for the (random) lifetime of the process φ and denote by $(T_n)_{n\geq 0}$ the sequence of successive branching or immigration times, $T_0 \equiv 0$ and $T_n \uparrow \zeta$. Then the process φ can be described by the following assertions **I** and **II**:

I) In the random interval $[\![0, T_{n+1} - T_n[\![$ the process $(\varphi_{T_n+s})_{0\leq s\leq T_{n+1}-T_n}$ reduces to the motion of ℓ independent particles according to (1), for suitable $\ell \in \mathbb{N}_0$, stopped at configuration dependent rate

$$\begin{aligned} \alpha(x^1,\ldots,x^\ell) &:= c + \sum_{i=1}^{\ell} \kappa(x^i) \quad \text{for } (x^1,\ldots,x^\ell) \in \mathbf{R}^\ell \text{ with } \ell \geq 1, \\ \alpha(\delta) &:= c, \end{aligned} \tag{2}$$

where δ denotes the void configuration.
In the case $x = \delta$, the process $(\varphi_{T_n+s})_{0\leq s\leq T_{n+1}-T_n}$ is constant equal to δ on the whole interval $[\![0, T_{n+1} - T_n[\![$.

II) The jumps from φ_{T_n-} to φ_{T_n}, $n \geq 1$, are determined by a transition kernel $K(\cdot,\cdot)$ on the configuration space S given by

$$K(x,\cdot) := \sum_{i=1}^{\ell} \frac{\kappa(x^i)}{\alpha(x)} \left(\sum_{k\neq 1} p_k(x^i)\ \epsilon_{C(x,i,k)} \right) + \frac{c}{\alpha(x)} \int_{\mathbf{R}} \pi(dy)\ \epsilon_{(x^1,\ldots,x^\ell,y)} \tag{3}$$

where $C(x,i,k)$ denotes the configuration

$$(x^1,\ldots,x^{i-1},x^{i+1},\ldots,x^\ell,\underbrace{x^i,\ldots,x^i}_{k \text{ times}}),$$

which arises if the i-th particle dies and leaves k offspring at its death position.
In the case $x = \delta$ the first term vanishes and the probability measure $K(\delta, dx)$ simplifies to the measure $\int_{\mathbf{R}} \pi(dy)\ \epsilon_y(dx)$ on S which coincides with π on $\mathbf{R} \subset S$.

1.2.2 Construction of the process

To give a deeper insight into the matter we give a short summary of the construction of BDI's via 'elementary processes' as proposed in [Löc02b]. Note that we are dealing with a simpler model than [Löc02b] since we have no interaction between the particles.

i) We first define an 'elementary process' $\tilde{\varphi}$ with initial configuration $x \in S$ on the (random) interval $[\![0, \tilde{S}[\![$ as follows (the stopping time $\tilde{S}$ is the lifetime of $\tilde{\varphi}$ and will be defined below):

Let $(C, \mathcal{C})$ denote the canonical path space of diffusions (1), i.e. $C = C(\mathbb{R}_+, \mathbf{R})$ is the space of all continuous functions from $\mathbb{R}_+$ to $\mathbf{R}$ and $\mathcal{C}$ denotes the canonical σ-field.
For $\ell := l(x)$ we start with a probability measure on $(C^\ell, \mathcal{C}^\ell)$ given by

$$\tilde{Q}_x^{(\ell)} := \bigotimes_{i=1}^{\ell} \mathcal{L}_{x^i}(\eta^i),$$

where $\eta^1, \ldots, \eta^\ell$ are independent solutions of the stochastic differential equation (1) with starting positions $x^1, \ldots, x^\ell$.
Let $(M, \mathcal{M})$ be the canonical path space for counting processes (see [Bré81]). Define a transition kernel $\tilde{K}^{(\ell)}(\cdot, \cdot)$ from $(C^\ell, \mathcal{C}^\ell)$ to $(M, \mathcal{M})$ by

$$\tilde{K}^{(\ell)}(f, \cdot) := \mathcal{L}\left(\text{Poisson process with intensity } t \mapsto \alpha(f(t))\right), \quad \forall f \in C^\ell,$$

where α is given by (2). With this notation we may define a probability measure

$$\tilde{Q}_x^{(\ell)} \tilde{K}^{(\ell)}(df, dg) := \tilde{Q}_x^{(\ell)}(df) \tilde{K}^{(\ell)}(f, dg)$$

on the filtered probability space $(C^\ell \times M, \mathcal{C}^\ell \otimes \mathcal{M}, \mathbb{G}^\ell \otimes \mathbb{M})$, where $\mathbb{G}^\ell \otimes \mathbb{M} := \left(\mathcal{G}_t^\ell \otimes \mathcal{M}_t\right)_{t \geq 0}$, $\mathbb{G}^\ell = (\mathcal{G}_t^\ell)_{t \geq 0}$ is the canonical filtration on $(C^\ell, \mathcal{C}^\ell)$ and $\mathbb{M} = (\mathcal{M}_t)_{t \geq 0}$ is the canonical filtration on $(M, \mathcal{M})$.
Write $(\tilde{\eta}, \tilde{\xi})$ for the canonical process on $(C^\ell \times M, \mathcal{C}^\ell \otimes \mathcal{M}, \mathbb{G}^\ell \otimes \mathbb{M})$ and set

$$\tilde{S} := \inf\left\{t > 0 : \tilde{\xi}_t \geq 1\right\}.$$

Then $\tilde{S}$ is a $\mathbb{G}^\ell \otimes \mathbb{M}$ stopping time and for $(t, \omega) \in [\![0, \tilde{S}[\![$ we define the elementary process $\tilde{\varphi}$ by

$$\tilde{\varphi}_t(\omega) := \tilde{\eta}_t(\omega).$$

Now generate the post jump configuration $\tilde{\zeta}$ according to the probability measure $K(\tilde{\varphi}_{S-}, \cdot)$ which was introduced in (3). Finally we can define the elementary process on the whole interval $[\![0, \tilde{S}]\!]$:

$$\tilde{\varphi} \mathbb{1}_{[\![0, \tilde{S}]\!]} := \tilde{\eta} \mathbb{1}_{[\![0, \tilde{S}[\![} + \tilde{\zeta} \mathbb{1}_{[\![\tilde{S}]\!]}.$$

Note that

$$\tilde{Q}_x^{(\ell)} \tilde{K}^{(\ell)}(\tilde{S} > t \mid \mathcal{G}_t^\ell) = e^{-\int_0^t \alpha(\tilde{\varphi}_s) ds}. \tag{4}$$

Interprete $\tilde{S}$ as an immigration time if

$$\varphi_{\tilde{S}} = (\varphi^1_{\tilde{S}-}, \ldots, \varphi^\ell_{\tilde{S}-}, y) \quad \text{for some } y \in \mathbf{R}$$

and interprete it as branching time otherwise.

ii) Assume that we have already defined the desired process φ (on a suitable probability space $(\tilde{\Omega}, \tilde{\mathcal{A}}, \tilde{\mathbb{F}})$) on the stochastic interval $[\![0, T_n]\!]$ as a chain of n elementary processes such that for $0 \le m < n$

$$\mathcal{L}\left(\varphi_{T_m+\bullet}\mathbb{1}_{[\![0,T_{m+1}-T_m]\!]}(\bullet)\middle|\,\tilde{\mathcal{F}}_{T_m}\right) = \mathcal{L}\left(\tilde{\eta}\mathbb{1}_{[\![0,\tilde{S}[\![} + \tilde{\zeta}\mathbb{1}_{[\![\tilde{S}]\!]}\middle|\,\tilde{Q}^{(\ell_m)}_{\eta_{T_m}}\tilde{K}^{(\ell_m)}\right) \tag{5}$$

where $\ell_m := l(\eta_{T_m})$ and $\eta_0 \equiv x$ for some $x \in S$.
Now consider another (independent) elementary process $\tilde{\tilde{\varphi}}\mathbb{1}_{[\![0,\tilde{\tilde{S}}]\!]}$ under the measure $\tilde{Q}^{(\ell_n)}_{\varphi_{T_n}}\tilde{K}^{(\ell_n)}$ and append it to the so far defined process:

$$\varphi_{T_n+\bullet}\mathbb{1}_{[\![0,T_{n+1}-T_n]\!]}(\bullet) := \tilde{\tilde{\varphi}}\mathbb{1}_{[\![0,\tilde{\tilde{S}}]\!]} \quad \text{,where} \quad T_{n+1} := T_n + \tilde{\tilde{S}}.$$

iii) Iteration of ii) gives us a process $\varphi\mathbb{1}_{[\![0,\zeta[\![}$ with lifetime $\zeta := \sup_{n\in\mathbb{N}} T_n$.

1.2.3 Ergodicity and invariant measure

Since we are interested in statistical inference we wish to establish the BDI φ to have the following properties **P1** to **P3**:

P1: No accumulation of jumps in finite time intervals. Then in particular the lifetime ζ of φ is a.s. equal to $+\infty$.

P2: Ergodicity, i.e. we wish the process φ to be recurrent in the sense of Harris, having δ as recurrent atom, and such that the invariant measure m on the configuration space S

$$m(F) = \frac{1}{E_\delta(R)} E_\delta\left(\int_0^R dt\ \mathbb{1}_F(\varphi_t)\right), \quad F \in \mathcal{S}, \tag{6}$$

is a finite measure. Here $R := \inf\{T_n : n \ge 1, \varphi_{T_n} = \delta\}$ denotes the time of first return to the void configuration δ.

P3: Finite expected configuration length unter m: associating to m the occupation measure

$$\overline{m}(A) := \int_S m(dx)x(A) = \frac{1}{E_\delta(R)} E_\delta\left(\int_0^R dt\ \varphi_t(A)\right), \quad A \in \mathcal{B}(\mathbf{R}), \tag{7}$$

we wish to have

$$\overline{m}(\mathbf{R}) = \int_S m(dx)\ l(x) < \infty.$$

Theorem 1.2.2. *Under Assumption A1 and A2 the properties* **P1**, **P2** *and* **P3** *hold.*

Proof. The assertion is a direct consequence of theorem 1.6.c in [HL03b]. □

Remark 1.2.3. We will present more results on the invariant measure m and the occupation measure $\overline{m}$ is section 1.3 below. In particular we will see that under additional assumptions the measures m and $\overline{m}$ have densities with respect to the Lebesgue measure (on S, respectively on $\mathbf{R}$).

1.2.4 A canonical path space

We work on a canonical path space $(\Omega, \mathcal{A}, \mathbb{F})$ for branching diffusions with immigration (more details can be found in [Lö99, section 5.1]) with the following properties:

- Ω is a closed subset of the Skorokhod space $D(\mathbb{R}_+, S)$ such that all trajectories $\omega \in \Omega$ have jumps only if the length of the configuration changes. This happens either if a single particle immigrates or if a particle is being killed. Between two jumps w is a continuous function on $\mathbf{R}^\ell$ for a suitable $\ell \in \mathbb{N}_0$ and there is no accumulation of jumps in finite time. In particular the lifetime ζ is a.s. equal to $+\infty$.
- $\mathcal{A}$ is the Borel σ-field of Ω. In particular $\mathcal{A}$ is also the σ-field generated by the canonical process φ, i.e. $\mathcal{A}$ is the σ-field generated by $\varphi_t(\omega) = \omega(t)$, $\omega \in \Omega$, $t \geq 0$.
- Ω is endowed with the canonical filtration $\mathbb{F} = (\mathcal{F}_t)_{t\geq 0}$, where $\mathcal{F}_t := \bigcap_{r>t} \sigma(\varphi_s : 0 \leq s \leq r)$.

We have the following theorem:

Theorem 1.2.4. *Under assumptions A1 and A2 the BDI φ can be constructed as canonical process on the canonical path space $(\Omega, \mathcal{A}, \mathbb{F})$ such that the process is strongly Markov.*

This result is due to Ikeda et al. (see [INW68a]). They considered a more general model where the particle motion is given by a general Markov process. A detailed proof for the case of a BDI can be found in [Lö99].

1.3 Auxiliary results and further remarks

In section 1.2.3 we have seen that assumptions A1 and A2 are sufficient for the finiteness of the invariant occupation measure $\overline{m}$ on $\mathbf{R}$ (property **P3** in section 1.2.3). Since immigrations occur at a constant rate (independently of all other events) we directly conclude that the total expected occupation time of the progeny of one single particle has to be finite as well. Thanks to the work of Höpfner and Löcherbach ([HL99a]) we are able to calculate the total expected occupation time of the progeny of one single particle explicitly.

Let $\varphi^{(s,y)} = (\varphi_t^{(s,y)})_{t \geq s}$ denote a branching diffusion without immigration which starts at time $s \geq 0$ with one single particle located in $y \in \mathbf{R}$. Then by proposition 2.2 in [HL03b] we have the following result.

Lemma 1.3.1. *Under assumptions A1 and A2 the total expected occupation time of a subprocess $\varphi^{(s,y)}, y \in \mathbf{R}$ and $s \geq 0$, is given by*

$$E\left(\int_s^\infty l(\varphi_t^{(s,y)})\ dt\right) = E_y\left(\int_0^\infty e^{-\int_0^t [\kappa(1-\rho)](\eta_u)\ du}\ dt\right),$$

where $\eta = (\eta_t)_{t \geq 0}$ denotes a solution of (1) with initial position $y \in \mathbf{R}$.
In particular we have the upper bound

$$E\left(\int_s^\infty l(\varphi_t^{(s,y)})\ dt\right) \leq \frac{1}{\underline{\kappa}(1-\bar{\rho})} < \infty$$

which is independent of the starting time $s \geq 0$ and the initial position $y \in \mathbf{R}$.

We can interprete the assertion of this lemma as follows. The total expected occupation time of $\varphi^{(s,y)}$ is equal to the expected lifetime of a diffusion (1) killed at rate $\kappa(1-\rho)$. Hence $\kappa(1-\rho)$ can be interpreted as a mass reduction function and in particular assumption A2 is such that the mass reduction $\kappa(1-\rho)$ is bounded away from zero.

We finish section 1 with two remarks on Lebesgue densities of the invariant occupation measure $\overline{m}$ on $\mathbf{R}$ and the invariant measure m on S.

1.3.1 On the invariant occupation density

Now consider the invariant occupation time measure $\overline{m}$ on $\mathbf{R}$. This measure is a well studied object (see [HL99a] and [HL03b] for the case of branching diffusions without interaction and [Löc04] for interacting branching diffusions with immigration) and we present a result which gives us the existence of a bounded and continuous Lebesgue density of $\overline{m}$.

In the sequel $C^k_{(b)}(\mathbb{R}^m, \mathbb{R}^n)$ denotes the space of C^k-functions $\mathbb{R}^n \to \mathbb{R}^m$ for which all partial derivatives of orders $1, \ldots, k$ are bounded and C^k_b denotes the subspace of bounded functions in $C^k_{(b)}$. For $C^k_{(b)}(\mathbb{R}^n, \mathbb{R})$ we write briefly $C^k_{(b)}(\mathbb{R}^n)$.

We need the following set of assumptions:

Assumption Inv: *Suppose that* $\sigma \in C^4_b(\mathbf{R}, \mathbb{R}^{d\times d})$, $b \in C^3_{(b)}(\mathbf{R}, \mathbf{R})$ *and* $\kappa, \rho \in C^2_b(\mathbf{R})$. *Furthermore assume that* b, σ, κ *and* ρ *are such that*

$$[\kappa(1-\rho)](y) + \sum_{i=1}^{d} \frac{\partial b_i}{\partial y_i}(y) - \frac{1}{2} \sum_{i,k=1}^{d} \frac{\partial^2 \sigma^2_{i,k}}{\partial y_i \partial y_k}(y)$$

is bounded away from zero uniformly in $y \in \mathbf{R}$.

Then by theorem 3.5 and lemma 3.6 in [HL03b] we have the following theorem:

Theorem 1.3.2. *Under assumptions A1, A2 and* **Inv** *the invariant occupation measure* $\overline{m}$ *on* $\mathbf{R}$ *has a bounded and continuous density with respect to the Lebesgue measure. Moreover this density is given explicitly.*

We abstain from presenting the explicit form of the density of $\overline{m}$, the existence of a density suffices for our considerations. A detailed derivation of the invariant occupation density can be found in [HL99a] (see also [Löc04] for the case of interacting BDI's).

1.3.2 On the invariant measure on S

In contrast to the occupation measure $\overline{m}$ on $\mathbf{R}$ there is only few known about the invariant measure m on S. Höpfner ([Hö4]) considered the Lebesgue density of a branching diffusion with immigration in a more simple model, where the killing rate κ is assumed to be constant and the reproduction law p does not depend on the position of the particles.

The main result of [Hö4] is the following. Under strong smoothness conditions on the one-particle motion (1) the Lebesgue density of m exists and the density is an increasing limit of 'strange shaped' densities. In this context a density has a 'strange shape' if it has the following properties:

On the open subset $S_0 := \{x \in S : x^k \neq x^{k'} \text{ for } k \neq k'\}$ of S where all particles occupy different positions, the density is smooth, unbounded and takes the value $+\infty$ on a specified collection of 'hyperplans' in the Lebesgue nullset $S \backslash S_0$.

Thus we are dealing with an invariant density which has not common 'nice' properties like global continuity or boundedness. In particular (besides the very rough approximation $m(F) \leq m(S)$) there are no trivial bounds for $m(F)$, $F \in \mathcal{S}$.

2 Partial reconstruction of a BDI from time discrete observations

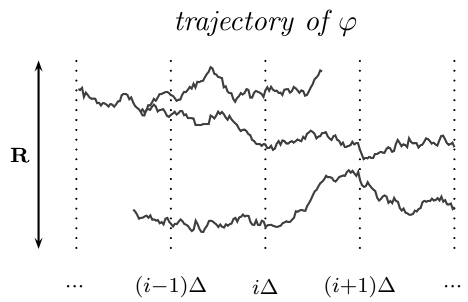
2.1 Presentation of the problem

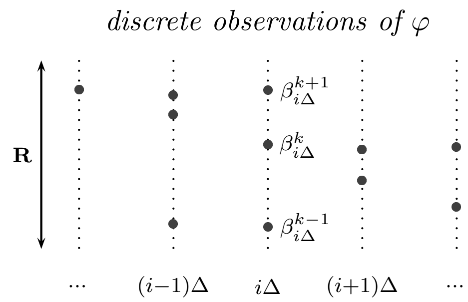
We observe a branching diffusion with immigration $\varphi = (\varphi_t)_{t\geq 0}$ at discrete observation times $t_1, \ldots, t_n$ of the finite time interval $[0, T]$. Without loss of generality, we assume equidistant time steps $t_i = i\Delta$ with step width $\Delta > 0$, $i = 0, \ldots, [T/\Delta]$.[3]

Furthermore we assume that we are able to observe only the positions of the particles, but not their pedigree. In other words: at each time $i\Delta$ the state $\varphi_{i\Delta} = (\varphi^1_{i\Delta}, \ldots, \varphi^{l(\varphi_{i\Delta})}_{i\Delta})$ of the process φ is only observed in form of the point measure $\sum_{k=1}^{l(\varphi_{i\Delta})} \epsilon_{\varphi^k_{i\Delta}}$.

For $i \in \{0, \ldots, [T/\Delta]\}$ let $\beta_{i\Delta} = (\beta^1_{i\Delta}, \ldots, \beta^{l(\varphi_{i\Delta})}_{i\Delta}) \in S$ denote an observation of φ at time $i\Delta$, i.e. $\beta_{i\Delta}$ is any arrangement of the support of the point measure $\sum_{k=1}^{l(\varphi_{i\Delta})} \epsilon_{\varphi^k_{i\Delta}}$. Note that $\beta_{i\Delta}$ and $\varphi_{i\Delta}$ coincide up to permutation of the components.

In the following sketch we see on the left hand side a (part of a) typical trajectory of a BDI φ. On the right hand side we inscribed the particle positions which we actually observe.



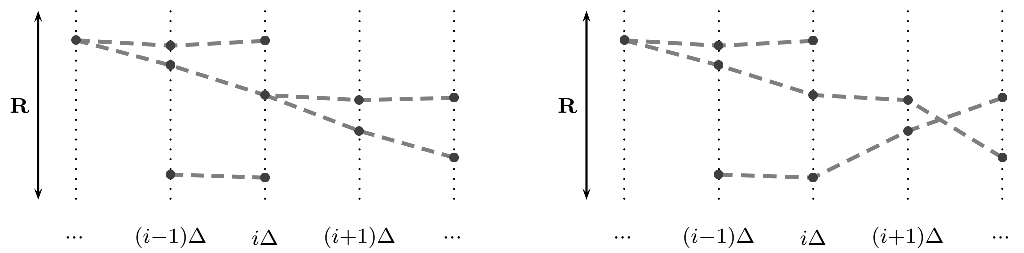


Since we are interested in statistical inference of these processes – in section 3 we will present a nonparametric estimator for the diffusion coefficient σ^2 – one question arises in a very natural way: Given the observations $\beta_{i\Delta}$, how can we 'redraw' or 'reconstruct' (at least parts) of the trajectory of the process φ?
Roughly speaking: how shall we connect the dots in the sketch on the right hand side in order to get a good approximation of the trajectory of φ?

[3]Here $[T/\Delta]$ denotes the largest number $i \in \mathbb{Z}$ such that $i\Delta < T$,
i.e. $[T/\Delta] := \sup\{i \in \mathbb{Z} : i\Delta < T\}$.

Continuing above example, possible 'reconstructions' of the trajectory could look like this:

one possible assignment *another assignment*

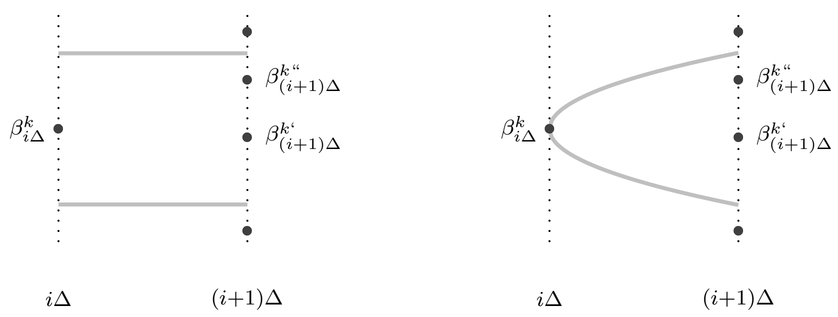


As we can see in this example there are many possibilities to redraw a possible path of φ given the observations $\beta_{i\Delta}$ and the problem we are interested in, is to find an approximation which fits the true assignment (corresponding to the full trajectory of φ) at least on a set with 'high probability'.

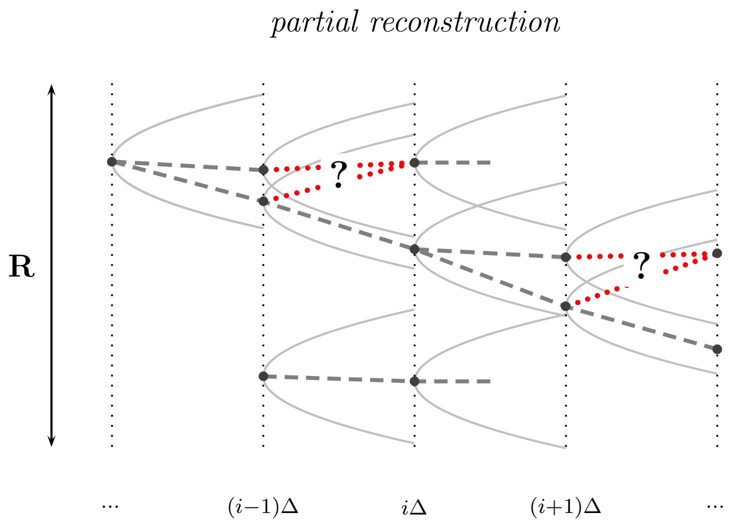
Now the idea is the following. If the step width Δ is small, the trajectories of the particles shouldn't fluctuate too much between two successive observation times (since particles move according to solutions of (1), which provide in particular continuous paths) and we will assign components of two successive observations to the same subtree if the distance in space between them is not too big. More precisely: at time $i\Delta$ we choose for each particle position $\beta^k_{i\Delta}$ a neighborhood of $\beta^k_{i\Delta}$ and assign all components of the observation $\beta_{(i+1)\Delta}$ which are contained in the neighborhood of $\beta^k_{i\Delta}$ to the subtree spanned by a single particle corresponding to the observed position $\beta^k_{i\Delta}$.

Remark 2.1.1. Instead of indicating the neighborhoods of the components $\beta^k_{i\Delta}$ by rectangles (picture 1 below), we draw 'bells' (picture 2 below) in order to improve the legibility of the sketches (the peak of each 'bell' points at the corresponding particle position at time $i\Delta$).

picture 1 *picture 2*



With this notation our reconstruction scheme would give us in our example the following constellation:



In this context two problems arise.

- First we have to precise the expressions 'fluctuate too much' and 'not too big', that means we have to control the probability that a particle leaves a small neighborhood between two successive observation times.

- Secondly we must handle the cases where the observed particle positions are so close that we can't assign those observations uniquely to the particles (question marks in above sketch). This is the case if the neighborhoods belonging to the observed particle positions intersect and in the intersection is at least one particle.

Here is the outline of the following sections:

In section 2.2 we derive an exponential inequality for the probability that a general diffusion (1) leaves a small neighborhood of its initial position during a short time period. Moreover we prove (under additional assumption on the drift term) an exponential bound which is independent of the starting position, so that we are able to apply this inequality globally to our discrete scheme.
In section 2.3 we construct a set which describes the event that at least one particle of a BDI leaves its neighborhood between successive observation times and we show, that the probability of this event has an exponential rate for $\Delta \downarrow 0$. Then, on the complement of this set, all particles will stay in their neighborhoods

with a ‘high probability‘.
After the definition of a ‘partial reconstruction rule‘ we study the asymptotic properties of this reconstruction scheme in section 2.4.
Finally we present simulations in section 2.5, which show that our reconstruction scheme works well in practice.

2.2 Diffusions in a small time interval

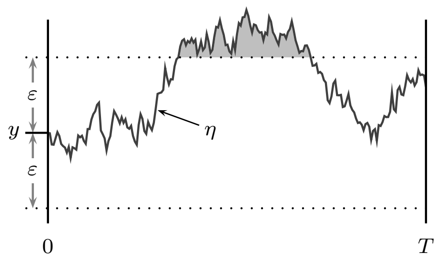
In this section we want to study how we can control the fluctuation of a diffusion (corresponding to the motion of single particles in our model) in a small time interval. For this purpose we consider a strong solution of the d-dimensional stochastic differential equation (1)

$$dX_t = b(X_t)dt + \sigma(X_t)dW_t \quad , X_0 = y \quad , t \geq 0,$$

where $y \in \mathbf{R}$ and $W = (W^1, \ldots, W^d)$ is a d-dimensional Brownian motion. In order to have such a solution we assume A1 (again we refer to [IW89] for general results on the existence and uniqueness of stochastic differential equations).

Let $\eta = (\eta_t)_{t \geq 0}$ be the canonical process on $C(\mathbb{R}_+, \mathbf{R})$. Then $\mathcal{C} := \sigma(\eta_t : t \geq 0)$ is the canonical σ-field and $\mathbb{G} := (\mathcal{G}_t)_{t \geq 0}$ defined by $\mathcal{G}_t := \bigcap_{T>t} \sigma(\eta_s : s \leq T)$ is the canonical filtration relative η. Denote by P_y the uniquely determined distribution of η on $(C, \mathcal{C}, \mathbb{G})$ with $\eta_0 = y \in \mathbf{R}$.

Now we should explain precisely what we intend by saying that we want to control the fluctuation of the diffusion. We want to find an upper bound for the probability that η leaves a ε-neighborhood of the initial position $\eta_0 = y$ in a time interval $[0, T]$, i.e. we want to control the probability of the event $\left\{\sup_{t \in [0,T]} |\eta_t - y| > \varepsilon\right\}$.



In the next section we will first restrict ourself to the case where the drift term of the diffusion vanishes, i.e. where the diffusion process is a local martingale. Then a classical result on continuous local martingales will give us the desired upper bound. In section 2.2.2 we allow arbitrary drift functions and we will show, that this class of drift functions is too big to obtain an inequality which is independent of the initial position. Finally in section 2.2.3 we prove a global exponentially inequality under boundedness assumptions on the drift function.

2.2.1 Exponential inequality for continuous local martingales

We first start with a rather classical result on continuous local martingales (for the one dimensional case see for instance (1.5) in [DvZ01] or exercise (3.16), chapter IV.3 in [RY99]). Let $M = (M_t^1, \dots, M_t^d)_{t\geq 0}$ be a $\mathbf{R}$-valued continuous local martingale with $M_0 = 0$ on a suitable probability space $(\Omega, \mathbb{F}, P)$ and define the maximum process $M^* = (M_t^*)_{t\geq 0}$ by $M_t^* := \sup_{s\leq t} |M_s|$.

Lemma 2.2.1. *Assume that M has a quadratic variation $\langle M \rangle$ such that*

$$\sum_{i,j=1}^{d} \langle M^i, M^j \rangle_t \leq c \cdot t \tag{8}$$

for $t \geq 0$, where c is a positive constant. Then

$$P(M_t^* \geq \varepsilon) \leq 2d \cdot e^{-\frac{\varepsilon^2}{2d\cdot c\cdot t}}.$$

Proof. Since

$$\{\omega \in \Omega : \left|(M_t^1(\omega), \dots, M_t^d(\omega))\right| \geq r\} \ \subseteq \ \bigcup_{i=1}^{d} \{\omega \in \Omega : \left|M_t^i(\omega)\right| \geq r/\sqrt{d}\}$$

for $r > 0$ it suffices to show the assertion in the case $d = 1$.
For $\alpha > 0$ define the Doleans exponential

$$M_t^\alpha := e^{\alpha M_t - \frac{\alpha^2}{2}\langle M \rangle_t}.$$

Then by [JS87, theorem I.4.61] $M^\alpha = (M_t^\alpha)_{t\geq 0}$ is a continuous local martingale. Since the exponential function is convex and increasing we get

$$e^{\sup_{s\leq t}(\alpha M_s - \frac{\alpha^2}{2}\langle M \rangle_s)} \leq \sup_{s\leq t} M_s^\alpha$$

and since $\langle M \rangle_t \leq ct$ by assumption (8) we also have

$$e^{\sup_{s\leq t} \alpha M_s - \frac{\alpha^2}{2} ct} \leq e^{\sup_{s\leq t}(\alpha M_s - \frac{\alpha^2}{2}\langle M \rangle_s)}.$$

Now the maximal inequality for local martingales (Doob's L^1-inequality, [Shi96, theorem VII.3.3]) yields

$$P(\sup_{s\leq t} M_s \geq \varepsilon) \leq P(\sup_{s\leq t} M_s^\alpha \geq e^{\alpha\varepsilon - \frac{\alpha^2}{2}ct}) \leq e^{-\alpha\varepsilon + \frac{\alpha^2}{2}ct} E(|M_t^\alpha|) = e^{-\alpha\varepsilon + \frac{\alpha^2}{2}ct}. \tag{9}$$

The last equation holds since by definition M^α is a nonnegative local martingale with $M_0^\alpha = 1$, which implies $E(|M_t^\alpha|) = E(M_t^\alpha) = E(M_0^\alpha) = 1$.
Inequality (9) holds for all $\alpha > 0$, hence

$$P(\sup_{s\le t} M_s \ge \varepsilon) \le \inf_{\alpha>0} e^{-\alpha\varepsilon + \frac{\alpha^2}{2}ct} = e^{-\frac{\varepsilon^2}{2ct}}. \tag{10}$$

Since $-M$ is also a continuous local martingale with $\langle -M\rangle_t = \langle M\rangle_t \le c\cdot t$ (by assumption (8)) inequality (10) is also valid for $-M$ and we finally conclude

$$P(\sup_{s\le t}|M_s| \ge \varepsilon) \le P(\sup_{s\le t} M_s \ge \varepsilon) + P(\sup_{s\le t}(-M_s) \ge \varepsilon) \le 2e^{-\frac{\varepsilon^2}{2ct}}.$$

□

As a direct consequence of above result we get (assuming A3 below) an upper bound for the probability that a driftless diffusion leaves an ε-neighborhood in $[0,T]$.

Assumption A3. *The matrices $a(x) := \sigma(x)\sigma(x)^T$ satisfy*

$$\sum_{i,j=1}^d a_{ij}(x)y^iy^j \le \bar\sigma^2\sum_{j=1}^d (y^j)^2 \quad , \quad (y^1,\dots,y^d)\in\mathbf{R}$$

uniformly in $x\in\mathbf{R}$ for some constant $0<\bar\sigma^2<\infty$.

Lemma 2.2.2. *Let $b\equiv 0$. Under conditions A1 and A3 for each $\varepsilon>0$ and $T>0$ the inequality*

$$P_y\left(\sup_{t\in[0,T]}|\eta_t - y| > \varepsilon\right) \le c_1\cdot\exp\left[-c_2\frac{\varepsilon^2}{T}\right]$$

holds uniformly in $y\in\mathbf{R}$, where c_1 and c_2 are positive constants independent of the initial position y.

Proof. Since $b\equiv 0$ the process η is a local martingale. Moreover its quadratic variation $\langle\eta\rangle$ is given by $(\langle\eta^i,\eta^j\rangle_t)_{1\le i,j\le d} = \int_0^t a(\eta_s)\,ds$ and by assumption A3

$$\sum_{i,j=1}^d \langle\eta^i,\eta^j\rangle_t = \sum_{i,j=1}^d \int_0^t a_{ij}(\eta_s)\,ds \le \bar\sigma^2\cdot d\cdot t.$$

Hence all conditions of lemma 2.2.1 are fulfilled and the lemma is proven. □

Remark 2.2.3. We are interested in the special case $T = \Delta$. With respect to our motivation in section 2.1 we want to choose $\varepsilon > 0$ (depending on Δ) as big as possible under the condition that the right hand side of the exponential inequality in Lemma 2.2.2 still converges to zero for $\Delta \downarrow 0$. For example choose $\varepsilon = \Delta^\lambda$ for some $\lambda \in (0, \frac{1}{2})$. Then lemma 2.2.2 gives us the following result

$$P_y\left(\sup_{t\in[0,\Delta]} |\eta_t - y| > \Delta^\lambda\right) \leq c_1 \cdot \exp\left[-c_2\left(\frac{1}{\Delta}\right)^{1-2\lambda}\right].$$

Since the upper bound on the right hand side does not depend on the initial position y, the Markov property of η gives

$$P_\bullet\left(\sup_{u\in[0,\Delta]} |\eta_{t+u} - \eta_t| > \Delta^\lambda\right) \leq c_1 \cdot \exp\left[-c_2\left(\frac{1}{\Delta}\right)^{1-2\lambda}\right], \qquad (11)$$

for all $t \geq 0$.

2.2.2 Exponential inequality for general diffusions

Now we will study a diffusion with arbitrary drift function (of course b is assumed to be globally Lipschitz, but nothing else). In this case $\eta = (\eta_t)_{t\geq 0}$ is not longer a martingale so that the local behavior of a particle can strongly depend on the drift term. For example consider the transient Ornstein-Uhlenbeck process $X_t = X_0\ e^{\alpha t} + \sigma \int_0^t e^{\alpha(t-s)}\ dW_s$ where α and σ are positive constants and W is a brownian motion. Here we can always find an initial position X_0 such that the process leaves a small neighborhood in a short time with high probability.
Thus there is no hope to get an exponential inequality independent from the initial point y in the general setting. But we still can state the following result.

Lemma 2.2.4. *Let A3 hold and let $\lambda \in (0, \frac{1}{2})$. Then for $y \in \mathbf{R}$ and $\Delta > 0$*

$$\begin{aligned} &P_y\left(\sup_{t\in[0,\Delta]} |\eta_t - y| > \Delta^\lambda\right) \\ \leq\ & c_1 \cdot \exp\left[-c_2\left(\left[\left(\frac{1}{\Delta}\right)^{1-\lambda} e^{-L\Delta} - |b(y)|\right]^+\right)^2 \Delta\right], \end{aligned}$$

where c_1 and c_2 are positive constants independent of the initial position y.

Proof. Let $t \leq \Delta$. Then using the Lipschitz continuity of b we write

$$\begin{aligned}
|\eta_t - y| &\leq \left|\int_0^t \sigma(\eta_s) dW_s\right| + \int_0^t |b(\eta_s)|\, ds \\
&\leq \sup_{r\in[0,\Delta]} \left|\int_0^r \sigma(\eta_s) dW_s\right| + |b(y)|\, t + \int_0^t |b(\eta_s) - b(y)|\, ds \\
&\leq \sup_{r\in[0,\Delta]} \left|\int_0^r \sigma(\eta_s) dW_s\right| + |b(y)|\, t + L \cdot \int_0^t (|\eta_s - y|) ds \\
&\leq \sup_{r\in[0,\Delta]} \left|\int_0^r \sigma(\eta_s) dW_s\right| + |b(y)|\, \Delta + L \cdot \int_0^t (|\eta_s - y|) ds.
\end{aligned}$$

By a Gronwall lemma (see for instance [Bas98, lemma I.3.3])

$$|\eta_t - y| \leq \left(\sup_{r\in[0,\Delta]} \left|\int_0^r \sigma(\eta_s) dW_s\right| + |b(y)|\, \Delta \right) \cdot e^{Lt}$$

for all $t \leq \Delta$, thus

$$\sup_{t\in[0,\Delta]} |\eta_t - y| \leq \left(\sup_{r\in[0,\Delta]} \left|\int_0^r \sigma(\eta_s) dW_s\right| + |b(y)|\, \Delta \right) \cdot e^{L\Delta}.$$

The last result directly implies

$$P_y \left(\sup_{t\in[0,\Delta]} |\eta_t - y| > \Delta^\lambda \right) \leq P_y \left(\sup_{r\in[0,\Delta]} \left|\int_0^r \sigma(\eta_s) dW_s\right| > \left(\Delta^\lambda e^{-L\Delta} - |b(y)|\, \Delta\right) \right).$$

Now the process $M = (M_t)_{t\geq 0} := (\int_0^t \sigma(\eta_s) dW_s)_t$ is the martingale part of the diffusion η and its quadratic variation process is given by $\langle M \rangle_t = \int_0^t \sigma\sigma^T(\eta_s) ds$. By assumption A3 the quadratic variation fulfills

$$\sum_{i,j=1}^d \langle M^i, M^j \rangle_t \ \leq\ \bar{\sigma}^2 \cdot d \cdot t,$$

for all $t \geq 0$. Thus for the local martingale M the conditions of lemma 2.2.1 are fulfilled and with $\varepsilon = \Delta^\lambda e^{-L\Delta} - |b(y)|\, \Delta$ and $T = \Delta$ we get

$$\begin{aligned}
& P_y \left(\sup_{t\in[0,\Delta]} |\eta_t - y| > \Delta^\lambda \right) \\
\leq\ & c_1 \exp \left(-c_2 \left[\frac{\left([\Delta^\lambda e^{-L\Delta} - |b(y)|\, \Delta]^+\right)^2}{\Delta} \right] \right) \\
=\ & c_1 \cdot \exp \left[-c_2 \left(\left[\left(\frac{1}{\Delta}\right)^{1-\lambda} e^{-L\Delta} - |b(y)| \right]^+ \right)^2 \Delta \right],
\end{aligned}$$

for some suitable constants $c_1, c_2 > 0$. This was the last step of our proof and the assertion is shown. $\square$

2.2.3 Exponential inequality independent of the initial position

The result of lemma 2.2.4 is not strong enough for our purposes. We need independency of the upper bound from the initial point y in order to apply an exponentially inequality to each couple of discrete observations.
As mentioned in the beginning of section 2.2.2 we have to restrict the class of drift functions. We assume

Assumption A4. *The drift function b is bounded, i.e. $|b(y)| \leq \bar{b} \ \forall\, y \in \mathbf{R}$, where $\bar{b}$ is a positive constant.*

For all diffusions meeting A3 and A4 we are now able to prove the following result.

Lemma 2.2.5. *Let A3 and A4 hold. Then for $t \geq 0, \Delta > 0$ and $\lambda \in (0, \frac{1}{2})$*

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

where c_1, c_2 are positive constants independent of t, y.

Proof. First note that we may assume $t = 0$ without restriction since the diffusion is strongly Markov and we want to achieve an upper bound which does not depend on the initial value of the process.
By A4 the drift function b is bounded by a constant $\bar{b} > 0$. Thus

$$\begin{aligned}
& P_y\left(\sup_{t\in[0,\Delta]} |\eta_t - y| > \Delta^\lambda\right) \\
\leq\ & P_y\left(\sup_{t\in[0,\Delta]} \left[\left|\int_0^t \sigma(\eta_s)dW_s\right| + \int_0^t |b(\eta_s)|\, ds\right] > \Delta^\lambda\right) \\
\leq\ & P_y\left(\sup_{t\in[0,\Delta]} \left|\int_0^t \sigma(\eta_s)dW_s\right| > \Delta^\lambda - \bar{b}\Delta\right)
\end{aligned}$$

and an application of lemma 2.2.1 gives us the upper bound

$$c_3 \cdot \exp\left[-c_4\left(\left[\left(\frac{1}{\Delta}\right)^{1-\lambda} - \bar{b}\right]^+\right)^2 \Delta\right]$$

for some positive constants c_3, c_4. Now it is possible to choose $c_1, c_2 > 0$ such that for all $0 < \Delta < \infty$

$$c_3 \cdot \exp\left[-c_4\left(\left[\left(\frac{1}{\Delta}\right)^{1-\lambda} - \bar{b}\right]^+\right)^2 \Delta\right] \leq c_1 \cdot \exp\left(-c_2\left(\frac{1}{\Delta}\right)^{1-2\lambda}\right)$$

and the assertion is shown. □

2.3 On the fluctuation of the particles of a BDI

In section 2.2 we have seen, that under relatively mild conditions (bounded drift function and bounded diffusion coefficient) a single diffusion leaves a Δ^λ-neighborhood ($\lambda \in (0, \frac{1}{2})$) of its starting position in a small time interval $[t, t+\Delta]$ only with small probability (lemma 2.2.5). Now we want to construct a set which describes the event that at least one particle of a branching diffusion leaves its Δ^λ-neighborhood between successive observation times and afterwards we will find an upper bound for the probability of this set. Its complement will be the set of 'high probability' mentioned in section 2.1.

2.3.1 Subprocesses without immigration

In a first step we consider subprocesses of a branching diffusion which start with one single particle.

Throughout this chapter $\varphi^{(s,y)} = (\varphi_t^{(s,y)})_{t \geq s}$ denotes a branching diffusion without immigration which starts at time $s \geq 0$ with a single particle located in $y \in \mathbf{R}$.[4] At each time $t \geq s$ $\varphi_t^{(s,y)}$ is an S-valued random variable and we will write $\varphi_t^{(s,y),k}$ for the $k-th$ component of $\varphi_t^{(s,y)}$, $k = 1, \ldots, l(\varphi_t^{(s,y)})$.

Note that this indication does not allow to decide whether components of $\varphi^{(s,y)}$ at two different times $t \neq t'$ belong to the same subtree if they have equal indices.

We define the desired set for subprocesses $\varphi^{(s,y)}$ in three steps:

i) For $s \geq 0$ and $y \in \mathbf{R}$ we first define the event that a subprocess without immigration, starting at time s with a single particle located in y, leaves a Δ^λ-neighborhood of y during a short time period $[s, s+\Delta]$:

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

Here $B_{\Delta^\lambda}(y) := \{z \in \mathbf{R} : |z-y| \leq \Delta^\lambda\}$ is the (closed) Δ^λ-neighborhood of y and $B^c_{\Delta^\lambda}(y)$ denotes its complement. Furthermore recall the notation $x(A) := \sum_{i=1}^{l(x)} \mathbb{1}_A(x^i)$ for configurations $x \in S$ and $A \subseteq \mathbf{R}$.

ii) Next we consider a configuration $x \in S$ and we define the event that there exist at least one component x^k of x such that the subprocess $\varphi^{(s,x^k)}$ which starts at time s with one particle located in x^k leaves a Δ^λ-neighborhood of

[4] We can think of this process as a BDI with immigration rate $c = 0$.

its starting position up to time $s+\Delta$:

$$A^{\Delta}(s,x) := \begin{cases} \bigcup_{k=1}^{l(x)} a^{\Delta}(s,x^k) & \text{if } x=(x^1,\ldots,x^{\ell}) \text{ for a suitable } \ell \geq 1, \\ \emptyset & \text{if } x \text{ is the void configuration } \delta. \end{cases}$$

iii) Finally we consider the process $\varphi^{(s,y)}$ on a time interval $[s,t]$. For each couple of (successive) discrete observation times $i\Delta, (i+1)\Delta$ in $[s,t]$ and for each component $\varphi_{i\Delta}^{(s,y),k}$ of $\varphi_{i\Delta}^{(s,y)}$ we consider subtrees of $\varphi^{(s,y)}$ which start at time $i\Delta$ in $\varphi_{i\Delta}^{(s,y),k}$ and define the event that at least one particle of these subtrees leaves a Δ^{λ}-neighborhood of $\varphi_{i\Delta}^{(s,y),k}$ during the time interval $[i\Delta,(i+1)\Delta]$:

$$\mathbb{A}^{\Delta}([s,t],y) := \begin{cases} \bigcup_{i=[s/\Delta]+1}^{[t/\Delta]-1} A^{\Delta}(i\Delta, \varphi_{i\Delta}^{(s,y)}) & \text{if } s \leq t-\Delta, \\ \emptyset & \text{if } s > t-\Delta. \end{cases}$$

We are now able to state the main result for branching diffusions $\varphi^{(s,y)}$ without immigration which start at time s with one single particle located in y.

Theorem 2.3.1. *Let A3 and A4 hold. Then for $s,t \geq 0$ and $y \in \mathbf{R}$*

$$Q\left(\mathbb{A}^{\Delta}([s,t],y)\right) \leq \frac{1}{1-e^{-(1-\bar{\rho})\underline{\kappa}\Delta}}\, g_{\lambda}(\Delta) =: \tilde{g}_{\lambda}(\Delta),$$

where the function g_{λ} is defined in lemma 2.2.5.

We will divide the proof of this theorem into two steps. In lemma 1.3.1 we have already seen that the total expected occupation time of a subprocess without immigration $\varphi^{(s,y)}$ is given by

$$E_y\left(\int_0^{\infty} e^{-\int_0^t [\kappa(1-\rho)](\eta_u)\ du}\ dt\right)$$

where η is a d-dimensional diffusion (1). Moreover this expression is finite by assumption A2. Now we will show even more. In the following lemma 2.3.2 we calculate the expected occupation time of the process $\varphi^{(s,y)}(f)$ up to a fixed time t, where $f \in C_b^2(\mathbf{R})$. As a byproduct we get that the expected number of particles of a subprocess $\varphi^{(s,y)}$ living at time $t \geq s$ has the upper bound $e^{-\underline{\kappa}(1-\bar{\rho})t}$.
In a second step we state a theorem which is the key result for the proof of theorem 2.3.1: we calculate an upper bound for the probability, that a subtree of $\varphi^{(s,y)}$ spanned by a single particle leaves a Δ^{λ}-neighborhood of its starting position between two successive observation times.

Lemma 2.3.2. *Let assumptions A1, A2 and A3 hold. Then for $y \in \mathbf{R}$, $f \in C_b^2(\mathbf{R})$ and $t \geq s$*

$$E\left(\varphi_t^{(s,y)}(f)\right) = E_y\left(f(\eta_t)e^{-\int_s^t [\kappa(1-\rho)](\eta_u)du}\right),$$

where $\eta = (\eta_t)_{t\geq s}$ is a diffusion (1) which starts at time s in y.
In particular we obtain for $f \equiv 1$ the upper bound

$$E\left(l(\varphi_t^{(s,y)})\right) \leq e^{-\underline{\kappa}(1-\bar{\rho})(t-s)}$$

for the expected number of particles at time $t \geq s$, independently of the initial position $y \in \mathbf{R}$.

Proof. Since the process $\varphi^{(s,y)}$ is a strong Markov process, there is no loss of generality in assuming $s = 0$.

First recall that the infinitesimal generator A of the diffusion (1) is given by

$$Af(y) := \frac{1}{2}\sum_{i,j=1}^{d} a_{ij}(y)\frac{\partial^2}{\partial y^i \partial y^j} f(y) + \sum_{i=1}^{d} b_i(y)\frac{\partial}{\partial y^i} f(y), \quad y \in \mathbf{R}.$$

(See for instance [IW89, theorem IV.6.1].)

For any function $f \in C_b^2(\mathbf{R})$ we consider a function $\overline{f} \in C_b^2(S)$ defined by

$$\overline{f}(x) := \sum_{i=1}^{l(x)} f(x^i) \text{ for } x = (x^1, \ldots, x^{l(x)}) \in S\backslash\{\delta\} \text{ and } \overline{f}(\delta) := 0.$$

With this notation $\varphi_t^{(0,x)}(f) = \overline{f}(\varphi_t^{(0,x)})$ for any initial configuration $x \in S$.

The Ito formula for $\overline{f}(\varphi^{(0,x)})$ is given by[5]

$$\overline{f}(\varphi_t^{(0,x)}) - \overline{f}(\varphi_0^{(0,x)}) = \int_0^t L^{(1)}\overline{f}(\varphi_s^{(0,x)})\, ds + M_t^c + M_t^d \tag{12}$$

where $L^{(1)}$ is the infinitesimal generator of the branching diffusion (without immigration) $\varphi^{(0,x)}$

$$L^{(1)}\overline{f}(x) = \overline{(Af - \kappa(1-\rho)f)}(x), \qquad \text{for } x \in S, \tag{13}$$

M_t^c is a continuous square integrable local martingale with angle bracket

$$\langle M^c\rangle_t = \int_0^t \overline{\nabla^T f \cdot \sigma\sigma^T \cdot \nabla f}(\varphi_s^{(0,x)})\, ds$$

[5]Apply the Ito formula for continuous semimartingales (see for example theorem I.4.57 in [JS87]) to the process between branching events and compensate the jumps, i.e. the branching events, by the expected number of jumps.

and M_t^d is the purely discontinuous local martingale

$$M_t^d = \sum_{n=1}^{\infty} \mathbb{1}_{\{T_n \le t\}} \left(\overline{f}(\varphi_{T_n}^{(0,x)}) - \overline{f}(\varphi_{T_n -}^{(0,x)}) \right) + \int_0^t \overline{\kappa(1-\rho) f}(\varphi_s^{(0,x)}) \, ds. \tag{14}$$

Here $(T_n)_{n\ge 1}$ denotes the sequence of successive branching times of $\varphi^{(0,x)}$.

By assumption A3 and by lemma 1.3.1 $E(\langle M^c \rangle_\infty) < \infty$ and thus M^c is a martingale (see [RY99, corollary 1.25, p. 130]).
On the other hand by assumption A2 and by lemma 1.3.1 the integrand in equation (14) is integrable and therefore $M^d = (M_t^d)_{t\ge 0}$ is a martingale too.
Hence

$$M_t := M_t^c + M_t^d = \overline{f}(\varphi_t^{(0,x)}) - \overline{f}(\varphi_0^{(0,x)}) - \int_0^t L^{(1)} \overline{f}(\varphi_s^{(0,x)}) \, ds$$

defines a martingale $M = (M_t)_{t\ge 0}$ with $M_0 = 0$.

Taking expectation on both sides of (12) we obtain

$$E\left(\overline{f}(\varphi_t^{(0,x)}) \right) - \overline{f}(\varphi_0^{(0,x)}) = \int_0^t E\left(L^{(1)} \overline{f}(\varphi_s^{(0,x)}) \right) \, ds. \tag{15}$$

Denoting by $(P_t^{(1)})_{t\ge 0}$ the semi-group corresponding to $L^{(1)}$ formula (15) can be rewritten in differential form

$$\frac{\partial}{\partial t} P_t^{(1)} \overline{f}(x) = P_t^{(1)} L^{(1)} \overline{f}(x) = L^{(1)} P_t^{(1)} \overline{f}(x). \tag{15'}$$

In the last equation we used the fact, that we may swap generator and semi-group (see proposition VII.1.2 in [RY99]).

Next consider $(P_t^{(2)})_{t\ge 0}$ defined by $P_t^{(2)} f(y) := E(\overline{f}(\varphi_t^{(0,y)}))$ for $f \in C_b^2(\mathbf{R}), t \ge 0$ and $y \in \mathbf{R}$. This defines a semi-group on $C_b^2(\mathbf{R})$, since by the Markov property and by definition of $P_t^{(2)}$ we have

$$\begin{aligned}
P_{s+t}^{(2)} f(y) &= E\left(\overline{f}(\varphi_{s+t}^{(0,y)}) \right) \\
&= E\left(\sum_{k=1}^{l(\varphi_s^{(0,y)})} E\left(\overline{f}\left(\varphi_{s+t}^{(s,\varphi_s^{(0,y),k})} \right) \Big| \mathcal{F}_s \right) \right) \\
&= E\left(\sum_{k=1}^{l(\varphi_s^{(0,y)})} E\left(\overline{f}\left(\varphi_t^{(0,\varphi_s^{(0,y),k})} \right) \right) \right) \\
&= E\left(\overline{P_t^{(2)} f}\left(\varphi_s^{(0,y)} \right) \right) \\
&= P_s^{(2)} (P_t^{(2)} f)(y).
\end{aligned}$$

Moreover $(P_t^{(2)})_{t\geq 0}$ inherits the continuity in zero from $(P_t^{(1)})_{t\geq 0}$ since by definition $P_t^{(2)} f = P_t^{(1)}\overline{f}\Big|_{\mathbf{R}}$ for $t \geq 0$.

Let $L^{(2)}$ denote the infinitesimal generator on $C_b^2(\mathbf{R})$ corresponding to $(P_t^{(2)})_{t\geq 0}$. Then by definition for $y \in \mathbf{R}$ and $f \in C_b^2(\mathbf{R})$

$$L^{(2)} f(y) = \lim_{t\downarrow 0} \frac{1}{t}\left(P_t^{(2)} f(y) - f(y)\right) = \lim_{t\downarrow 0} \frac{1}{t}\left(P_t^{(1)}\overline{f}(y) - \overline{f}(y)\right) = L^{(1)}\overline{f}(y).$$

Hence (15') implies

$$\frac{\partial}{\partial t}(P_t^{(2)} f) = L^{(2)}(P_t^{(2)} f) = (A - \kappa(1-\rho))\,(P_t^{(2)} f), \tag{15''}$$

where $f \in C_b^2(\mathbf{R})$.

By the theorem of Feynman and Kac (see for instance [Kal02, theorem 24.1]) the solution of this partial differential equation has the representation

$$P_t^{(2)} f(y) = E_y\left(f(\eta_t)\; e^{-\int_0^t [\kappa(1-\rho)](\eta_s)\; ds}\right),$$

where η has the Markov generator A under Q_y. This is precisely the first assertion of the lemma.

Taking again advantage of assumption A2 we get in the special case $f \equiv 1$

$$E(\overline{f}(\varphi_t^{(0,y)})) = E\left(l(\varphi_t^{(0,y)})\right) = E_y\left(e^{-\int_0^t [\kappa(1-\rho)](\xi_s)\; ds}\right) \leq e^{-\underline{\kappa}(1-\bar{\rho})t}$$

and the proof is complete. □

Theorem 2.3.3. *Let A3 and A4 hold. Then for $s \geq 0$ and $y \in \mathbf{R}$*

$$Q(a^{\Delta}(s, y)) \leq g_{\lambda}(\Delta),$$

independently of the initial position y.

Proof. Since $\varphi^{(s,y)}$ is a Markov process we only need to consider the case $s = 0$. In a first step we consider a diffusion η given by (1) which has y as starting position under Q and endow η with an 'alarm clock' giving alarm at position dependent rate κ, i.e. we consider an ascending sequence of stopping times $0 < S_1 < S_2 < \cdots$ such that for $n \in \mathbb{N}$

$$Q(S_{n+1} - S_n > s|\mathcal{F}_{S_n+s}) = e^{-\int_0^s \kappa(\eta_{S_n+u})du}, \quad s \geq 0, \tag{16}$$

(in other words: we set marks on η with position depending rate κ; see also section 1.2).

Next define for $n \in \mathbb{N}$ events that the diffusion η leaves a Δ^λ-neighborhood of its initial position y before $\Delta \wedge S_n$:

$$A_n := \left\{ \sup_{0 \le u \le \Delta \wedge S_n} |\eta_u - y| > \Delta^\lambda \right\}.$$

By definition the laws of η and $\varphi^{(0,y)}$ coincide on $[\![0, S_1[\![$ and furthermore the first branching time T_1 of $\varphi^{(0,y)}$ and the stopping time S_1 are identically distributed. We decompose $a^\Delta(0,y)$ into the events that the initial particle leaves the neighborhood before S_1 or afterwards

$$\begin{aligned} Q(a^\Delta(0,y)) &= Q(A_1 \cap a^\Delta(0,y)) + Q(A_1^c \cap a^\Delta(0,y)) \\ &= Q(A_1) + Q\left(A_1^c, S_1 \le \Delta, \int_{S_1}^{\Delta} \varphi_u^{(0,y)} \left(B^c_{\Delta^\lambda}(y)\right) \, du > 0 \right) \\ &=: (I) \; + \; (II). \end{aligned}$$

Now we will focus on the second term (II). At time S_1 the process $\varphi^{(0,y)}$ branches and releases a random number of descendants (according to the probability law $p(\eta_{S_1}) = (p_m(\eta_{S_1}))_{m \ne 1}$) corresponding to subprocesses

$${}^1\varphi^{(S_1,\eta_{S_1})}, {}^2\varphi^{(S_1,\eta_{S_1})}, {}^3\varphi^{(S_1,\eta_{S_1})}, \ldots,$$

which are independent copies of the subprocess $\varphi^{(S_1,\eta_{S_1})}$.
Thus (II) equals to

$$E\left[\mathbb{1}_{\tilde{A}_1^c} \mathbb{1}_{\{S_1 \le \Delta\}} \sum_{m \ne 1} p_m(\eta_{S_1}) \; Q\left(\bigcup_{i=1}^{m} \left\{ \int_{S_1}^{\Delta} {}^i\varphi_u^{(S_1,\eta_{S_1})}(B^c_{\Delta^\lambda}(y)) \, du > 0 \right\} \right) \right].$$

Since ${}^i\varphi^{(S_1,\eta_{S_1})}$ are identically distributed and independent we get under expectation

$$\begin{aligned} &\sum_{m \ne 1} p_m(\eta_{S_1}) \; Q\left(\bigcup_{i=1}^{m} \left\{ \int_{S_1}^{\Delta} {}^i\varphi_u^{(S_1,\eta_{S_1})}(B^c_{\Delta^\lambda}(y)) \, du > 0 \right\} \right) \\ \le \; &\sum_{m \ne 1} p_m(\eta_{S_1}) \sum_{i=1}^{m} Q\left(\int_{S_1}^{\Delta} {}^i\varphi_u^{(S_1,\eta_{S_1})}(B^c_{\Delta^\lambda}(y)) \, du > 0 \right) \\ = \; &\rho(\eta_{S_1}) \; Q\left(\int_{S_1}^{\Delta} \varphi_u^{(S_1,\eta_{S_1})}(B^c_{\Delta^\lambda}(y)) \, du > 0 \right). \end{aligned} \tag{17}$$

By assumption A2 the mean number of descendants $\rho(\cdot)$ is bounded by 1 and we obtain

$$(II) \le Q\left(A_1^c, S_1 \le \Delta, \int_{S_1}^{\Delta} \varphi_u^{(S_1,\eta_{S_1})}(B^c_{\Delta^\lambda}(y)) \, du > 0 \right).$$

Summarizing above results we have found that

$$Q(a^{\Delta}(0,y)) \leq Q(A_1) + Q\left(A_1^c, S_1 \leq \Delta, \int_{S_1}^{\Delta} \varphi_u^{(S_1,\eta_{S_1})}(B^c_{\Delta^\lambda}(y))\, du > 0\right). \quad (18)$$

Now we want to iterate these steps. Assume that we have shown (18) already for sets A_{n-1} and corresponding stopping times S_{n-1}, $n \geq 2$, i.e.

$$\begin{aligned} &Q(a^{\Delta}(0,y)) \\ &\leq Q(A_{n-1}) + Q\left(A_{n-1}^c, S_{n-1} \leq \Delta, \int_{S_{n-1}}^{\Delta} \varphi_u^{(S_{n-1},\eta_{S_{n-1}})}(B^c_{\Delta^\lambda}(y))\, du > 0\right). \end{aligned}$$

Then we have to consider a branching diffusion $\varphi^{(S_{n-1},\eta_{S_{n-1}})}$ which starts a time S_{n-1} with a single particle located in $\eta_{S_{n-1}}$. As before we study the cases where this subprocess leaves $B_{\Delta^\lambda}(y)$ before or after the first branching event of $\varphi^{(S_{n-1},\eta_{S_{n-1}})}$. By construction this branching time has the same distribution as the stopping time S_n and by the same arguments as above (see (17)) we obtain

$$\begin{aligned} &Q\left(A_{n-1}^c, S_{n-1} \leq \Delta, \int_{S_{n-1}}^{\Delta} \varphi_u^{(S_{n-1},\eta_{S_{n-1}})}(B^c_{\Delta^\lambda}(y))\, du > 0\right) \\ \leq\ & Q(A_n \cap A_{n-1}^c) + Q\left(A_n^c \cap A_{n-1}^c, S_n \leq \Delta, \int_{S_n}^{\Delta} \varphi_u^{(S_n,\eta_{S_n})}(B^c_{\Delta^\lambda}(y))\, du > 0\right) \end{aligned}$$

by replacing at time S_n the descendants of $\varphi^{(S_{n-1},\eta_{S_{n-1}})}$ by a single particle weighted with the mean number of descendants $\rho(\eta_{S_n})$ which is in fact smaller than 1 by assumption A2.

Since $Q(A_{n-1}) + Q(A_n \cap A_{n-1}^c) = Q(A_n)$ and $A_n^c \cap A_{n-1}^c = A_n^c$ by definition of the sets A_n (in fact we have $A_n \subseteq A_{n+1}$ for $n \in \mathbb{N}$) we conclude

$$\begin{aligned} Q(a^{\Delta}(0,y)) &\leq Q(A_n) + Q\left(A_n^c, S_n \leq \Delta, \int_{S_n}^{\Delta} \varphi_u^{(S_n,\eta_{S_n})}(B^c_{\Delta^\lambda}(y))\, du > 0\right) \\ &\leq Q(A_n) + Q(S_n \leq \Delta) \end{aligned} \quad (19)$$

for $n \in \mathbb{N}$.

The boundedness assumption on κ in A2 directly implies $S_n \uparrow \infty$ (recall the definition of the stopping times S_n in (16)) and therefore

$$A_n \uparrow \left\{ \sup_{0 \leq u \leq \Delta} |\eta_u - y| > \Delta^\lambda \right\} \quad \text{and} \quad Q(S_n \leq \Delta) \to 0 \quad \text{for } n \to \infty.$$

Taking in (19) the limit $n \to \infty$ we finally get with lemma 2.2.5

$$Q(a^{\Delta}(0,y)) \leq Q\left(\sup_{0 \leq u \leq \Delta} |\eta_u - y| > \Delta^\lambda\right) \leq g_\lambda(\Delta)$$

which is precisely the assertion of the theorem. □

Combining previous results, we can easily show the assertion of theorem 2.3.1:

Proof of theorem 2.3.1. As in the proofs of lemma 2.3.1 and theorem 2.3.2 we may assume without restriction $s = 0$. By definition

$$\begin{aligned} Q\left(\mathbb{A}^{\Delta}([0,t],y)\right) &= Q\left(\bigcup_{i=[s/\Delta]+1}^{[t/\Delta]-1} A^{\Delta}(i\Delta, \varphi_{i\Delta}^{(s,y)})\right) \\ &\leq \sum_{i=[s/\Delta]+1}^{[t/\Delta]-1} Q\left(\bigcup_{k=1}^{l(\varphi_{i\Delta}^{(0,y)})} a^{\Delta}(i\Delta, \varphi_{i\Delta}^{(0,y),k})\right) \\ &\leq \sum_{i=[s/\Delta]+1}^{[t/\Delta]-1} E_y\left[\sum_{k=1}^{l(\varphi_{i\Delta}^{(0,y)})} Q\left(a^{\Delta}(i\Delta, \varphi_{i\Delta}^{(0,y),k})\right)\right]. \end{aligned}$$

Now we may apply theorem 2.3.3 and get

$$Q\left(\mathbb{A}^{\Delta}([0,t],y)\right) \leq \sum_{i=[s/\Delta]+1}^{[t/\Delta]-1} E\left(l(\varphi_{i\Delta}^{(0,y)})\right) g_{\lambda}(\Delta) \leq \sum_{i=0}^{\infty} E\left(l(\varphi_{i\Delta}^{(0,y)})\right) g_{\lambda}(\Delta).$$

By lemma 2.3.2 this is smaller or equal to

$$\sum_{i=0}^{\infty} e^{-(1-\bar{\rho})\underline{\kappa} i\Delta} g_{\lambda}(\Delta) \leq \frac{1}{1-e^{-(1-\bar{\rho})\underline{\kappa}\Delta}} g_{\lambda}(\Delta) = \tilde{g}_{\lambda}(\Delta)$$

and the theorem is proven. □

2.3.2 Branching diffusion with immigration

Now we consider the whole process, i.e. a branching diffusion with immigration φ. Let (τ_j^I) denote the sequence of successive immigration times and let (ζ_j^I) be the corresponding sequence of immigration positions. Then by construction of the process (see section 1.2.2) (τ_j^I, ζ_j^I) builds a multivariate point process

$$\mu^I(ds, dy) = \sum_{j\geq 1} \mathbb{1}_{\{\tau_j^I < \infty\}} \varepsilon_{(\tau_j^I, \zeta_j^I)}(ds, dy).$$

By definition μ^I is a Poisson random measure on $(0,\infty) \times \mathbb{R}$ under Q_x with intensity $\nu^I(ds, dy) = c\, ds\pi(dy)$ (we refer again to section 1.2.2).

We are interested in the probability, that at least one particle of φ leaves a Δ^{λ}-neighborhood between successive observation times (as described in section 2.1).

In analogy to $\mathbb{A}^{\Delta}([s,t],y)$ defined in the last section we define the desired event corresponding to a branching diffusion with immigration φ by

$$\mathbb{A}_T^{\Delta}(x) := \bigcup_{i=1}^{l(x)} \mathbb{A}^{\Delta}([0,T],x^i) \cup \bigcup_{j\geq 1} \mathbb{A}^{\Delta}([\tau_j^I,T],\zeta_j^I), \tag{20}$$

where $x = (x^1,\ldots,x^{l(x)}) \in S$ is the initial configuration of φ.[6]

Theorem 2.3.4. *Consider a branching diffusion with immigration φ with initial configuration $x \in S$. Under assumptions A3 and A4 the inequality*

$$Q\left(\mathbb{A}_T^{\Delta}(x)\right) \leq \Big(l(x) + c\cdot T\Big)\cdot \tilde{g}_\lambda(\Delta)$$

holds for $\Delta > 0$, where the function $\tilde{g}_\lambda$ is defined in theorem 2.3.1.

Proof. We first separate the sets depending on the immigrating particles from the other sets

$$Q\left(\mathbb{A}_T^{\Delta}(x)\right) \leq \sum_{k=1}^{l(x)} Q\left(\mathbb{A}^{\Delta}([0,T],x^k)\right) + Q\left(\bigcup_{j\geq 1} \mathbb{A}^{\Delta}([\tau_j^I,T],\zeta_j^I)\right). \tag{21}$$

The first term of the sum on the right hand side of (21) consists only of sets depending on branching diffusions without immigration, though we can apply theorem 2.3.1 and get

$$\sum_{k=1}^{l(x)} Q\left(\mathbb{A}^{\Delta}([0,T],x^k)\right) \leq l(x)\cdot \tilde{g}_\lambda(\Delta). \tag{22}$$

Next we consider the second term on the right hand side of (21).

$$\begin{aligned} Q\left(\bigcup_{i\in\mathbb{N}} \mathbb{A}^{\Delta}([\tau_i^I,T],\zeta_i^I)\right) &\leq \sum_{i\in\mathbb{N}} Q\left(\mathbb{A}^{\Delta}([\tau_i^I,T],\zeta_i^I\right) \\ &= E\left(\int_{\mathbb{R}}\int_0^T Q\left(\mathbb{A}^{\Delta}([t,T],z)\right)\mu^I(dt,dz)\right). \end{aligned}$$

Now μ^I is compensated by ν^I, hence $\int_{\mathbb{R}}\int_0^T Q\left(\mathbb{A}^{\Delta}([t,T],z)\right)(\mu^I-\nu^I)(dt,dz)$ has expectation zero and we have

$$\begin{aligned} &E\left(\int_{\mathbb{R}}\int_0^T Q\left(\mathbb{A}^{\Delta}([t,T],z)\right)\mu^I(dt,dz)\right) \\ =&E\left(\int_{\mathbb{R}}\int_0^T Q\left(\mathbb{A}^{\Delta}([t,T],z)\right) c\; ds\; \pi(dz)\right) \end{aligned}$$

[6]Note that $\mathbb{A}^{\Delta}([\tau_j^I,T],\zeta_j^I) \cap \{\tau_j^I \geq T\} = \emptyset$ by definition of the sets $\mathbb{A}^{\Delta}([s,t],y)$.

with $\nu^I(ds, dz) = c\ ds\pi(dz)$.

An application of theorem 2.3.1 gives us the upper bound

$$E\left(\int_{\mathbb{R}}\int_0^T Q\left(\mathbb{A}^{\Delta}([t,T],z)\right) c\ ds\ \pi(dz)\right) \leq cT \cdot \pi(\mathbb{R}) \cdot \tilde{g}_\lambda(\Delta) = cT \cdot \tilde{g}_\lambda(\Delta). \quad (23)$$

Combining (22) and (23) we obtain

$$Q\left(\mathbb{A}_T^{\Delta}(x)\right) \leq \Big(l(x) + c \cdot T\Big)\tilde{g}_\lambda(\Delta),$$

which is our assertion. $\square$

2.4 Partial reconstruction

As we have suggested in section 2.1 we want to reconstruct (at least parts of) the trajectory of a discretely observed branching diffusion with immigration by assigning to each component of an observation certain candidates of descendants in the succeeding observation.

In section 2.4.1 below we define a partial reconstruction rule which is easy to implement (simulations will be presented later in sections 2.5 and 3.5) and which formalizes the heuristics of section 2.1.
In section 2.4.2 we study the asymptotics of the error of this rule.

2.4.1 Notation and definition of the partial reconstruction rule

For each configuration $\varphi_{(i+1)\Delta}$, $i \in \mathbb{N}_0$, we consider the decomposition

$$\varphi_{(i+1)\Delta} \stackrel{d}{=} \bigcup_{k=1}^{l(\varphi_{i\Delta})} \varphi_{(i+1)\Delta}^{(i\Delta,\varphi_{i\Delta}^k)} \cup \varphi_{(i+1)\Delta}^I, \tag{24}$$

where $\varphi_{(i+1)\Delta}^{(i\Delta,\varphi_{i\Delta}^k)}$ has the same law as the subconfiguration of $\varphi_{(i+1)\Delta}$ containing the progeny of a component y of $\varphi_{i\Delta}$ and in $\varphi_{(i+1)\Delta}^I$ we collect all particles of $\varphi_{(i+1)\Delta}$ descending from immigrants born during the time interval $]i\Delta,(i+1)\Delta]$.[7] In the case that there were no immigrants we set $\varphi_{i\Delta}^I := \emptyset$.

Furthermore for $\varepsilon > 0$ consider sets

$$\begin{aligned} S_\varepsilon &:= \left\{x \in S : l(x) \geq 2, \exists\, i \neq j \text{ in } \{1,\dots,l(x)\} \text{ such that } \left|x^i - x^j\right| < \varepsilon\right\} \\ D_\varepsilon &:= S \setminus S_\varepsilon \end{aligned}$$

and

$$S_0 := \lim_{\varepsilon \to 0} S_\varepsilon = \left\{x \in S : l(x) \geq 2, \exists\, i \neq j \text{ in } \{1,\dots,l(x)\} \text{ such that } x^i = x^j\right\}$$

Then S_ε contains all configurations where at least one particle has a neighbor at distance of less than Δ^λ, D_ε is the complement of S_ε and S_0 is the subset of configurations where at least two particles occupy the same position in space.

[7] Of course we could also decompose $\varphi_{i\Delta}^I$ into the progeny of single particles immigrated in $]i\Delta,(i+1)\Delta]$ (there can be more than one immigrant in $]i\Delta,(i+1)\Delta]$), but for our purposes it is sufficient to consider only $\varphi_{i\Delta}^I$.

As mentioned in section 2.1 we assume that we are able to observe $\varphi_{i\Delta}$ only in form of the corresponding point measure $\sum_{k=1}^{l(\varphi_{i\Delta})} \epsilon_{\varphi_{i\Delta}^k}$. In particular we have no information about the pedigree of the particles and hence the subconfigurations corresponding to $\varphi_{(i+1)\Delta}^{(i\Delta,\varphi_{i\Delta}^k)}$ are not observable.

Recall that we denote by $\beta_{i\Delta} \in S$ the observations of the branching diffusion with immigration φ, more precisely $\beta_{i\Delta}$ is an arbitrary arrangement of the support of the point measure $\sum_{k=1}^{l(\varphi_{i\Delta})} \epsilon_{\varphi_{i\Delta}^k}$, $i = 0, \ldots, [T/\Delta]$.
For convenience we define for $x, y \in S$ the equivalence relation

$$x =_p y :\Leftrightarrow \begin{cases} l(x) = l(y) \text{ and there exists a permutation } \pi \\ \text{of } \{1, \ldots, l(x)\} \text{ such that } \pi(x) = y. \end{cases}$$

With this notation $\beta_{i\Delta} =_p \varphi_{i\Delta}$ for all $i \in \{0, \ldots, [T/\Delta]\}$.

The problem will be to find an approximation for $\varphi_{(i+1)\Delta}^{(i\Delta,\beta_{i\Delta}^k)}$ based on the observation $\beta_{(i+1)\Delta}$. Our proposal for such an approximation is given by the following 'partial reconstruction rule'.

Partial reconstruction rule 2.4.1.
Consider observations $\beta_{i\Delta}$ of a BDI φ, $i \in \{0, \ldots, [T/\Delta] - 1\}$.

i) If $\beta_{i\Delta} \in D_{2\Delta^\lambda}$ define

$$\begin{aligned} \beta_{(i+1)\Delta}^{[i\Delta,\beta_{i\Delta}^k]} &:= \left\{\beta_{(i+1)\Delta}^m \in \beta_{(i+1)\Delta} : \left|\beta_{(i+1)\Delta}^m - \beta_{i\Delta}^k\right| \le \Delta^\lambda\right\} \\ &= \beta_{(i+1)\Delta} \cap \bar{B}_{\Delta^\lambda}(\beta_{i\Delta}^k) \end{aligned}$$

for $k = 1, \ldots, l(\beta_{i\Delta})$ and

$$\beta_{(i+1)\Delta}^I := \beta_{(i+1)\Delta} \setminus \left(\bigcup_{k=1}^{l(\beta_{i\Delta})} \beta_{(i+1)\Delta}^{[i\Delta,\beta_{i\Delta}^k]} \right).$$

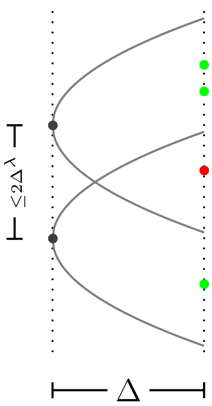
ii) If $\beta_{i\Delta} \in S_{2\Delta^\lambda}$ we don not assign any observation to $\beta_{i\Delta}$.

Loosely speaking rule 2.4.1 works as follows: For every component $\beta_{i\Delta}^k$ of $\beta_{i\Delta}$ we consider a Δ^λ-neighborhood of $\beta_{i\Delta}^k$. In the case that there is at least one couple of neighborhoods which have a nonempty intersection we will not define approximations of the subconfigurations $\varphi_{(i+1)\Delta}^{(i\Delta,\beta_{i\Delta}^k)}$ from (24) in the hope that this case will not occur very often (we will study this topic later in section 2.4.2).
In the case that all these neighborhoods have a mutually empty intersection, we collect for each component $\beta_{i\Delta}^k$ of $\beta_{i\Delta}$ all particles of the succeeding observation

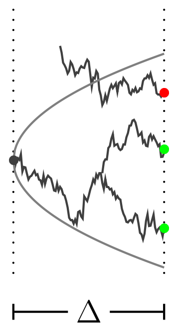
$\beta_{(i+1)\Delta}$ which are contained in the corresponding neighborhood $B_{\Delta^\lambda}(\beta^k_{i\Delta})$ and treat them as possible candidates of descendants of $\beta^k_{i\Delta}$.

Since particles will not leave a Δ^λ-neighborhood on a set with a high probability (which has the lower bound $1 - \tilde{g}_\lambda(\Delta)$ as we have seen in theorem 2.3.4) above rule seems to be a very natural choice.

Remark 2.4.2.



a) First we have to justify why we do not assign observations whenever $\beta_{i\Delta} \in S \backslash D_{2\Delta^\lambda} = S_{2\Delta^\lambda}$. In this case there could be indices $k, k' \in \{1, \dots, l(\beta_{i\Delta})\}, k \neq k'$, such that $\beta_{(i+1)\Delta}$ has a component which is contained in both neighborhoods $B_{\Delta^\lambda}(\beta^k_{i\Delta})$ and $B_{\Delta^\lambda}(\beta^{k'}_{i\Delta})$ (red dot in the sketch on the left hand side). The component in the intersection of the neighborhoods could be an observation of a descendant either of $\beta^k_{i\Delta}$ or of $\beta^{k'}_{i\Delta}$ and even on $(\mathbb{A}^\Delta_T)^c$ we just can guess which assignment could be the right one. For that reason we prefer not to use observations $\beta_{i\Delta} \in S_{2\Delta^\lambda}$ and it remains to show that (asymptotically) we won't loose too much data (see theorem 2.4.4.iii for results on this topic).



b) Secondly we have to calculate the intrinsic error of rule 2.4.1. Of course the subprocess starting in $\beta^k_{i\Delta}$ can leave its Δ^λ-neighborhood. But as we have seen in theorem 2.3.3 this event will occur only with a probability of order $g_\lambda(\Delta)$. There is an other source of error which is worse: new particles may immigrate between $i\Delta$ and $(i+1)\Delta$ and may choose their position that close to $\beta^k_{i\Delta}$, that they add an observation to $\beta^{[i\Delta,k]}_{(i+1)\Delta}$ (trajectory with the red dot in the sketch on the left hand side). This problem is the topic of theorem 2.4.4.i and 2.4.4.ii in the following section 2.4.2.

c) It is possible to define more sophisticated versions of rule 2.4.1 which use more data. For example we could define for observations $\beta_{i\Delta} \in S_{2\Delta^\lambda}$ the subset of 'good components'

$$G(\beta_{i\Delta}) := \left\{ \beta^k_{i\Delta} \in \beta_{i\Delta} : \left| \beta^k_{i\Delta} - \beta^{k'}_{i\Delta} \right| > 2\Delta^\lambda,\ k' \neq k \right\} \subseteq \beta_{i\Delta}, \tag{25}$$

i.e. all components of $\beta_{i\Delta}$ which have no neighbor closer than $2\Delta^\lambda$, and assign to each $\beta^k_{i\Delta} \in G(\beta_{i\Delta})$ an approximation $\beta^{[i\Delta,\beta^k_{i\Delta}]}_{(i+1)\Delta}$ as in rule 2.4.1. But this refinement does not improve the asymptotic behavior of the rule because we improve our rule only for a negligible part of observations (the quota of observations contained in $S_{2\Delta^\lambda}$ converges to zero for $\Delta \downarrow 0$ as we will see in theorem 2.4.4.iii). In our simulations (see sections 2.5 and 3.5) we will use the finer version of the reconstruction rule (which uses the 'good components' as described above) in order to use the capacities of the computer more efficiently.

2.4.2 Asymptotics of the partial reconstruction rule

Now we turn to study the asymptotic error of rule 2.4.1, $\Delta \downarrow 0$. During the whole section we assume that the particle process φ starts in the invariant measure m on S.

Furthermore we need the following strong assumption:

Assumption A5. *The invariant measure m on S admits a density with respect to the Lebesgue measure on S.*[8]

Remark 2.4.3. a) Assumption A5 is indeed a strong assumption, because we are not able to verify the existence of a Lebesgue density by our observations. Nevertheless we already mentioned in section 1.3.2 that in a more simple model the density of the invariant measure m exists (under additional assumptions) and there is no reason why the density should not exist in our general setting.
b) Since S_0 is a countable union of hyperplanes in S which have Lebesgue measure 0 we directly conclude that under assumption A5 S_0 is a m-nullset.

We are interested in the asymptotic behavior of the following objects. First consider the quota of observations where the partial reconstruction rule decided wrong:

$$q_{prr}^{\Delta,T} := \frac{1}{[T/\Delta]} \sum_{i=0}^{[T/\Delta]-1} \mathbb{1}_{D_{2\Delta^\lambda}}(\beta_{i\Delta}) \sum_{k=1}^{l(\beta_{i\Delta})} \mathbb{1}_{\left\{\beta_{(i+1)\Delta}^{[i\Delta,\beta_{i\Delta}^k]} \neq_p \varphi_{(i+1)\Delta}^{(i\Delta,\beta_{i\Delta}^k)}\right\}}$$

and secondly we are interested in the quota of observations with 'close' components, i.e. the quota of observations contained in $S_{2\Delta^\lambda}$:

$$q_{cl}^{\Delta,T} := \frac{1}{[T/\Delta]} \sum_{i=0}^{[T/\Delta]-1} \mathbb{1}_{S_{2\Delta^\lambda}}(\beta_{i\Delta}).$$

We have the following theorem:

[8] The Lebesgue measure λ on S is the usual Lebesgue measure λ on every layer $\mathbf{R}^\ell$ of S, $\ell \geq 1$, and on $\mathbf{R}^0 = \{\delta\}$ it is defined as the Dirac measure ε_δ.

Theorem 2.4.4. *Let assumptions A1 to A5 hold and let* $\lambda \in (0, \frac{1}{2})$.

i) For $i \in \{0, 1, \dots, [T/\Delta] - 1\}$ *and* $k \in \{1, \dots, l(\beta_{i\Delta})\}$ *we have*

$$\begin{aligned} & \mathbb{1}_{D_{2\Delta^\lambda}}(\beta_{i\Delta}) \; Q_m\Big(\Big\{ \beta_{(i+1)\Delta}^{[i\Delta, \beta_{i\Delta}^k]} \neq_p \varphi_{(i+1)\Delta}^{(i\Delta, \beta_{i\Delta}^k)} \Big\} \Big| \, \mathcal{F}_{i\Delta} \Big) \\ \leq \; & \mathbb{1}_{D_{2\Delta^\lambda}}(\beta_{i\Delta}) \left[c \cdot \Delta + l(\beta_{i\Delta}) g_\lambda(\Delta) \right], \end{aligned}$$

where c *is the (constant) immigration rate and the function* $g_\lambda(\cdot)$ *is defined in lemma 2.2.5.*

ii) Assume that $\int_S l^2(x)\ m(dx) < \infty$. *Then the expected quota of observations where the partial reconstruction rule decides wrong is of order* $\mathcal{O}(\Delta)$, *i.e.*

$$E_m(q_{\mathrm{prr}}^{\Delta,T}) = \mathcal{O}(\Delta) \qquad \textit{for } \Delta \downarrow 0.$$

Under the additional assumption that the immigration law π *has a continuous and bounded Lebesgue density we have*

$$E_m(q_{\mathrm{prr}}^{\Delta,T}) = \mathcal{O}\left(\Delta^{1+d\lambda}\right) \qquad \textit{for } \Delta \downarrow 0.$$

where d *is the dimension of* $\mathbf{R}$.

iii) The quota of observations β *with 'close' components up to time* T

$$q_{cl}^{\Delta,T} = \frac{1}{[T/\Delta]} \sum_{i=0}^{[T/\Delta]-1} \mathbb{1}_{S_{2\Delta^\lambda}}(\beta_{i\Delta}).$$

converges in Q_m*-probability to zero for* $\Delta \downarrow 0$.

Remark 2.4.5. a) Theorem 2.4.4.i corresponds to what we have already mentioned in remark 2.4.2.b. The leading part of the error $(c \cdot \Delta)$ comes from immigrating particles whereas the probability that one of the other particles causes an error of the partial reconstruction rule is of order $g_\lambda(\Delta)$. In the (hypothetical) case that we have no immigration, the rate would be of order $\mathcal{O}(g_\lambda(\Delta))$.

b) We have no rate of convergence for the quota of unused observations (theorem 2.4.4.iii). The reason for this problem is the following: even if there exist a density of m with respect to the Lebesgue measure, this density is in general not bounded. The situation is even worse, because the density of m takes in general the value $+\infty$ on $S_0 = \lim_{\varepsilon \downarrow 0} S_\varepsilon$ (see section 1.3.2 for the references). Hence there are no trivial arguments to find a rate of convergence for $m(S_\varepsilon)$ when $\varepsilon \downarrow 0$. Nevertheless we will present simulations in section 2.5 which suggest that the quota decreases fast enough for application.

c) The assertions of theorem 2.4.4 remain true if we use the finer version of the partial reconstruction rule described in remark 2.4.2.c.

Proof. a) We first prove assertion i) of theorem 2.4.4.
Conditioning on the event that a particle of φ leaves a Δ^λ-neighborhood between $i\Delta$ and $(i+1)\Delta$ (this is exactly the event $A^\Delta(i\Delta,\beta_{i\Delta})$ in the notation introduced in section 2.3.1) we obtain

$$\begin{aligned}
& \mathbb{1}_{D_{2\Delta^\lambda}}(\beta_{i\Delta})\cdot Q_m\left(\left\{\beta_{(i+1)\Delta}^{[i\Delta,\beta_{i\Delta}^k]} \neq_p \varphi_{(i+1)\Delta}^{(i\Delta,\beta_{i\Delta}^k)}\right\}\Big|\,\mathcal{F}_{i\Delta}\right)\\
\leq\ & \mathbb{1}_{D_{2\Delta^\lambda}}(\beta_{i\Delta})\cdot Q_m\left(\left\{\beta_{(i+1)\Delta}^{[i\Delta,\beta_{i\Delta}^k]} \neq_p \varphi_{(i+1)\Delta}^{(i\Delta,\beta_{i\Delta}^k)}\right\}\cap (A^\Delta(i\Delta,\beta_{i\Delta}))^c\Big|\,\mathcal{F}_{i\Delta}\right)\\
& +\mathbb{1}_{D_{2\Delta^\lambda}}(\beta_{i\Delta})\cdot Q\left(A^\Delta(i\Delta,\beta_{i\Delta})\right). \qquad (26)
\end{aligned}$$

By theorem 2.3.3

$$Q\left(A^\Delta(i\Delta,\beta_{i\Delta})\right)\leq\sum_{k=1}^{l(\beta_{i\Delta})}Q(a^\Delta(i\Delta,\beta_{i\Delta}^k))\leq l(\beta_{i\Delta})\cdot g_\lambda(\Delta). \qquad (27)$$

Next consider the first summand on the right hand side of (26).

On $(A^\Delta(i\Delta,\beta_{i\Delta}))^c$ all particles of $\varphi^{(i\Delta,\beta_{i\Delta}^k)}$ stay in their Δ^λ-neighborhoods and hence $\varphi_{(i+1)\Delta}^{(i\Delta,\beta_{i\Delta}^k)}\subseteq\beta_{(i+1)\Delta}^{[i\Delta,\beta_{i\Delta}^k]}$ by definition of $\beta_{(i+1)\Delta}^{[i\Delta,\beta_{i\Delta}^k]}$.

Using the decomposition (24) we conclude that for $\beta_{i\Delta}\in D_{2\Delta^\lambda}$

$$\begin{aligned}
& Q_m\left(\left\{\beta_{(i+1)\Delta}^{[i\Delta,\beta_{i\Delta}^k]} \neq_p \varphi_{(i+1)\Delta}^{(i\Delta,\beta_{i\Delta}^k)}\right\}\cap (A^\Delta(i\Delta,\beta_{i\Delta}))^c\Big|\,\mathcal{F}_{i\Delta}\right)\\
=\ & Q_m\left(\left\{\beta_{(i+1)\Delta}^{[i\Delta,\beta_{i\Delta}^k]}\cap\left(\bigcup_{k\neq k'}\varphi_{(i+1)\Delta}^{(i\Delta,\beta_{i\Delta}^{k'})}\cup\varphi_{(i+1)\Delta}^I\right)\neq\emptyset\right\}\cap (A^\Delta(i\Delta,\beta_{i\Delta}))^c\Bigg|\,\mathcal{F}_{i\Delta}\right)\\
=\ & Q_m\left(\left\{\beta_{(i+1)\Delta}^{[i\Delta,\beta_{i\Delta}^k]}\cap\varphi_{(i+1)\Delta}^I\neq\emptyset\right\}\cap (A^\Delta(i\Delta,\beta_{i\Delta}))^c\Big|\,\mathcal{F}_{i\Delta}\right).
\end{aligned}$$

The last step holds since for $\beta_{i\Delta}\in D_{2\Delta^\lambda}$ we have

$$\beta_{(i+1)\Delta}^{[i\Delta,\beta_{i\Delta}^k]}\cap\varphi_{(i+1)\Delta}^{(i\Delta,\beta_{i\Delta}^{k'})}\cap(A^\Delta(i\Delta,\beta_{i\Delta}))^c\subseteq\beta_{(i+1)\Delta}^{[i\Delta,\beta_{i\Delta}^k]}\cap\varphi_{(i+1)\Delta}^{(i\Delta,\beta_{i\Delta}^{k'})}\cap(a^\Delta(i\Delta,\beta_{i\Delta}^{k'}))^c=\emptyset$$

for $k\neq k'$.

Now let $\tau^{i\Delta}$ denote the first immigration time after $i\Delta$ and $\zeta^{i\Delta}$ the corresponding position of the immigrating particle. Then

$$\begin{aligned}
& Q_m\left(\left\{\beta_{(i+1)\Delta}^{[i\Delta,\beta_{i\Delta}^k]}\cap\varphi_{(i+1)\Delta}^I\neq\emptyset\right\}\cap (A^\Delta(i\Delta,\beta_{i\Delta}))^c\Big|\,\mathcal{F}_{i\Delta}\right)\\
\leq\ & Q_m\left(\left\{\beta_{(i+1)\Delta}^{[i\Delta,\beta_{i\Delta}^k]}\cap\varphi_{(i+1)\Delta}^{(\tau^{i\Delta},\zeta^{i\Delta})}\neq\emptyset\right\}\cap\{\tau^{i\Delta}\leq(i+1)\Delta\}\Big|\,\mathcal{F}_{i\Delta}\right). \qquad (28)
\end{aligned}$$

Since immigration occurs with constant rate $c>0$ (independent of all other events)

$$Q_m\left(\left\{\beta_{(i+1)\Delta}^{[i\Delta,\beta_{i\Delta}^k]}\cap\varphi_{(i+1)\Delta}^{(\tau^{i\Delta},\zeta^{i\Delta})}\neq\emptyset\right\}\cap\{\tau^{i\Delta}\leq(i+1)\Delta\}\Big|\,\mathcal{F}_{i\Delta}\right)\ \leq\ c\Delta$$

and the first assertion of the theorem follows with (26) and (27).

b) Now we will show assertion ii). Assume that $\int_S l^2(x)\; m(dx) < \infty$.
Then by (27) and (28)

$$\begin{aligned}
& E_m(q_{\mathrm{prr}}^{\Delta,T}) \\
= \;& \frac{1}{[T/\Delta]} E_m \left(\sum_{i=0}^{[T/\Delta]-1} \mathbb{1}_{D_{2\Delta^\lambda}}(\beta_{i\Delta}) \sum_{k=1}^{l(\beta_{i\Delta})} \mathbb{1}_{\left\{ \beta_{(i+1)\Delta}^{[i\Delta,\beta_{i\Delta}^k]} \neq_p \varphi_{(i+1)\Delta}^{(i\Delta,\beta_{i\Delta}^k)} \right\}} \right) \\
= \;& \frac{1}{[T/\Delta]} E_m \left(\sum_{i=0}^{[T/\Delta]-1} \mathbb{1}_{D_{2\Delta^\lambda}}(\beta_{i\Delta}) \sum_{k=1}^{l(\beta_{i\Delta})} Q_m \left(\beta_{(i+1)\Delta}^{[i\Delta,\beta_{i\Delta}^k]} \neq_p \varphi_{(i+1)\Delta}^{(i\Delta,\beta_{i\Delta}^k)} \middle| \mathcal{F}_{i\Delta} \right) \right) \\
\leq \;& \frac{1}{[T/\Delta]} E_m \left(\sum_{i=0}^{[T/\Delta]-1} \mathbb{1}_{D_{2\Delta^\lambda}}(\beta_{i\Delta}) \sum_{k=1}^{l(\beta_{i\Delta})} \left[Q_m \left(Z_{i\Delta}^k \middle| \mathcal{F}_{i\Delta} \right) + l(\beta_{i\Delta}) g_\lambda(\Delta) \right] \right),
\end{aligned}$$

where

$$Z_{i\Delta}^k := \left\{ \beta_{(i+1)\Delta}^{[i\Delta,\beta_{i\Delta}^k]} \cap \varphi_{(i+1)\Delta}^{(\tau^{i\Delta},\zeta^{i\Delta})} \neq \emptyset \right\} \cap \{\tau^{i\Delta} \leq (i+1)\Delta\}.$$

Consider the cases that the immigrating particle chooses its position in a $2\Delta^\lambda$-neighborhood of $\beta_{i\Delta}^k$ or not. Then

$$\begin{aligned}
& Q_m \left(Z_{i\Delta}^k \middle| \mathcal{F}_{i\Delta} \right) \\
\leq \;& Q_m \left(\left\{ \beta_{(i+1)\Delta}^{[i\Delta,\beta_{i\Delta}^k]} \cap \varphi_{(i+1)\Delta}^{(\tau^{i\Delta},\zeta^{i\Delta})} \neq \emptyset \right\} \cap \left\{ \zeta^{i\Delta} \notin B_{2\Delta^\lambda}(\beta_{i\Delta}^k) \right\} \middle| \mathcal{F}_{i\Delta} \right) \\
& + Q_m \left(\{\tau^{i\Delta} \leq (i+1)\Delta\} \cap \left\{ \zeta^{i\Delta} \in B_{2\Delta^\lambda}(\beta_{i\Delta}^k) \right\} \middle| \mathcal{F}_{i\Delta} \right). \qquad (29)
\end{aligned}$$

Now

$$\begin{aligned}
& \left\{ \beta_{(i+1)\Delta}^{[i\Delta,\beta_{i\Delta}^k]} \cap \varphi_{(i+1)\Delta}^{(\tau^{i\Delta},\zeta^{i\Delta})} \neq \emptyset \right\} \cap \left\{ \zeta^{i\Delta} \notin B_{2\Delta^\lambda}(\beta_{i\Delta}^k) \right\} \\
\subseteq \;& \left\{ \int_{\tau^{i\Delta}}^{(i+1)\Delta} \varphi_s^{(\tau^{i\Delta},\zeta^{i\Delta})}((B_{\Delta^\lambda}(\zeta^{i\Delta}))^c)\; ds > 0 \right\} \\
\subseteq \;& a^\Delta(i\Delta, \zeta^{i\Delta})
\end{aligned}$$

and therefore we get with theorem 2.3.3

$$Q_m \left(\left\{ \beta_{(i+1)\Delta}^{[i\Delta,\beta_{i\Delta}^k]} \cap \varphi_{(i+1)\Delta}^{(\tau^{i\Delta},\zeta^{i\Delta})} \neq \emptyset \right\} \cap \left\{ \zeta^{i\Delta} \notin B_{2\Delta^\lambda}(\beta_{i\Delta}^k) \right\} \middle| \mathcal{F}_{i\Delta} \right) \leq g_\lambda(\Delta). \qquad (30)$$

Since immigrating particles choose their position in space independently of their birth time we get moreover

$$\begin{aligned}
& Q_m \left(\{\tau^{i\Delta} \leq (i+1)\Delta\} \cap \left\{ \zeta^{i\Delta} \in B_{2\Delta^\lambda}(\beta_{i\Delta}^k) \right\} \middle| \mathcal{F}_{i\Delta} \right) \\
= \;& \int_{i\Delta}^{(i+1)\Delta} ce^{-c(s-i\Delta)}\; ds \int_{\mathbf{R}} \mathbb{1}_{B_{2\Delta^\lambda}(\beta_{i\Delta}^k)}(x)\; \pi(dx) \\
\leq \;& c\Delta \cdot \pi(B_{2\Delta^\lambda}(\beta_{i\Delta}^k)).
\end{aligned}$$

Now trivially $\pi(B_{2\Delta^\lambda}(\beta^k_{i\Delta})) \le 1$ (since π is a probability on $\mathbf{R}$), but if we assume additionally that π has a bounded Lebesgue density then $\pi(B_{2\Delta^\lambda}(\beta^k_{i\Delta})) \le K\cdot\Delta^{d\lambda}$, where K is a suitable positive constant.

Hence

$$Q_m\left(\{\tau^{i\Delta} \le (i+1)\Delta\} \cap \left\{\zeta^{i\Delta} \in B_{2\Delta^\lambda}(\beta^k_{i\Delta})\right\} \middle| \mathcal{F}_{i\Delta}\right) \le c\Delta \cdot K^\pi(\Delta), \tag{31}$$

where $K^\pi(\Delta) := 1$ in the case that we pose no further assumptions on π and $K^\pi(\Delta) := K \cdot \Delta^{d\lambda}$ in the case that π has a bounded density.

Combining (31) and (30) we get with (29)[9]

$$Q_m\left(Z^k_{i\Delta}\middle| \mathcal{F}_{i\Delta}\right) = \mathcal{O}\left(\Delta \cdot K^\pi(\Delta)\right) + \mathcal{O}\left(g_\lambda(\Delta)\right) = \mathcal{O}\left(\Delta \cdot K^\pi(\Delta)\right) \quad \text{for } \Delta \downarrow 0.$$

Finally by the invariance property of m

$$\begin{aligned} & E_m(q^{\Delta,T}_{\text{prr}}) \\ = \; & \frac{1}{[T/\Delta]} E_m\left(\sum_{i=0}^{[T/\Delta]-1} \mathbb{1}_{D_{2\Delta^\lambda}}(\beta_{i\Delta}) \sum_{k=1}^{l(\beta_{i\Delta})} \mathbb{1}_{\left\{\beta^{[i\Delta,\beta^k_{i\Delta}]}_{(i+1)\Delta} \neq_p \varphi^{(i\Delta,\beta^k_{i\Delta})}_{(i+1)\Delta}\right\}}\right) \\ \le \; & \frac{1}{[T/\Delta]} E_m\left(\sum_{i=0}^{[T/\Delta]-1} \sum_{k=1}^{l(\beta_{i\Delta})} \left[\mathcal{O}\left(\Delta \cdot K^\pi(\Delta)\right) + l(\beta_{i\Delta}) g_\lambda(\Delta)\right]\right) \\ = \; & \mathcal{O}\left(\Delta \cdot K^\pi(\Delta)\right) \int_S l(x)\; m(dx) + g_\lambda(\Delta) \int_S l^2(x)\; m(dx). \end{aligned}$$

Since $\int_S l(x)\; m(dx) = \overline{m}(\mathbf{R}) < \infty$ by property **P3** (see section 1.2.3) and $\int_S l^2(x)\; m(dx) < \infty$ by assumption

$$\mathcal{O}\left(\Delta \cdot K^\pi(\Delta)\right) \int_S l(x)\; m(dx) + g_\lambda(\Delta) \int_S l^2(x)\; m(dx) = \mathcal{O}\left(\Delta \cdot K^\pi(\Delta)\right)$$

for $\Delta \downarrow 0$ and assertion ii) follows directly with the definition of $K^\pi(\Delta)$.

c) It remains to prove assertion iii). By the invariance property of m we have

$$E_m(q^{\Delta,T}_{cl}) = \frac{1}{[T/\Delta]} \sum_{i=0}^{[T/\Delta]-1} Q_m(\beta_{i\Delta} \in S_{2\Delta^\lambda}) = m(S_{2\Delta^\lambda}).$$

Now $S_{2\Delta^\lambda}$ converges to S_0 for Δ to 0 and S_0 is a union of hyperplanes in S which have Lebesgue measure 0. Hence by assumption A5

$$m(S_{2\Delta^\lambda}) \stackrel{\Delta\downarrow 0}{\longrightarrow} m(S_0) = 0$$

and the proof is finished. $\square$

[9]Note that $g_\lambda(\Delta)$ decreases much faster to zero than any power of Δ.

2.5 Simulations

In this section we present simulations which give a glance on the effectiveness of the partial reconstruction rule 2.4.1 in practice. The simulations were done with the R-package[10] 'BDI' (a link to a download of the package is given in the references; see [Bra04]). This program simulates trajectories of a BDI φ where the particles of φ are living in the one-dimensional space $\mathbb{R}$.

There are two reasons for choosing dimension $d = 1$. First there is less effort (in form of computational capacities) needed to simulate a BDI in the case that the particles move according to one-dimensional diffusions. Secondly, which is in fact the more important reason, the case $d = 1$ is the hardest case for reconstruction since in dimension $d = 1$ the particles have more possibilities to get close to each other (see theorem 2.4.4.ii for the corresponding theoretical result).

We have chosen the following set of parameters for the simulation:[11]

Drift function	$b(x)$	$\equiv$	0
Diffusion coefficient	$\sigma^2(x)$	$=$	$\frac{1}{4}\cos(\pi \cdot x) + \frac{1}{2}$
Killing rate	$\kappa(x)$	$\equiv$	1
Immigration rate	c	$=$	1
Reproduction law	$p_k(x)$	$\equiv$	$\frac{(1.2)^k}{k!}e^{-1.2}, \quad k = 0, 2, 3, \dots$
Simulation step width	sw	$=$	0.00005
Time horizon	T	$=$	5

Remark 2.5.1. There is no loss of generality in choosing a constant drift function b and a constant killing rate κ since these functions are assumed to be bounded functions (assumptions A2 and A4). Even if the diffusion coefficient σ^2 is also assumed to be bounded (assumption A3) we have chosen a smooth function since we want to estimate this function later in section 3 (simulations of the estimator will be presented in section 3.5).

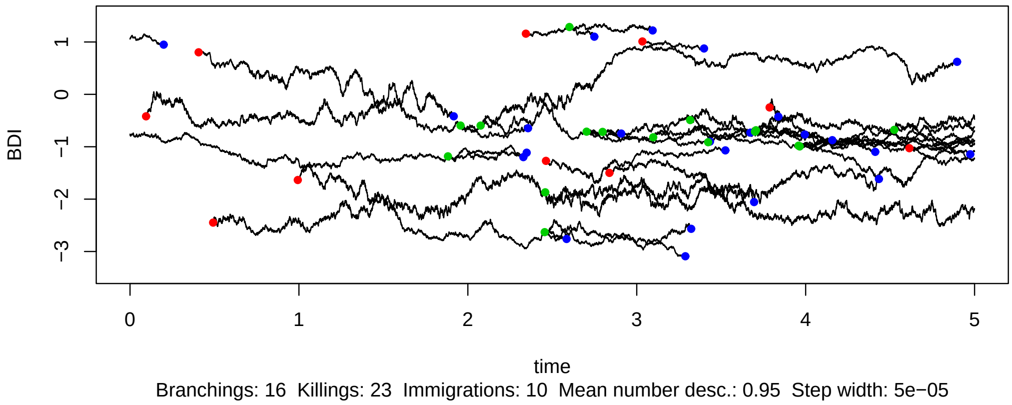
Figure 1 shows the simulated trajectory of the BDI on the interval $[0, 5]$ with above parameters. The green dots denote the branching events (with 2 or more descendants), the red dots denote the immigrations and the blue dots denote

[10] 'R' is an open source project which provides a commandline based statistical software. See 'http://www.r-project.org' for more information.

[11] Note that the reproduction law $(p_k(\cdot))_{k\neq 1}$ defined above has expectation $1.2 - e^{-1.2} < 1$. In particular all parameters are such that assumptions A1 to A5 are fulfilled.

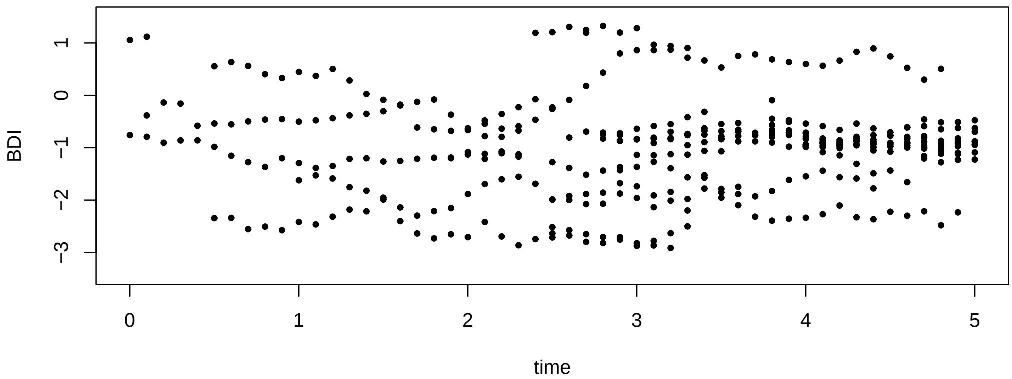
the killings (i.e. branching events where no descendant is born). In this example we had 23 killings, 16 branchings with more than one descendant and 10 immigrations.

Figure 1: Full trajectory of the BDI



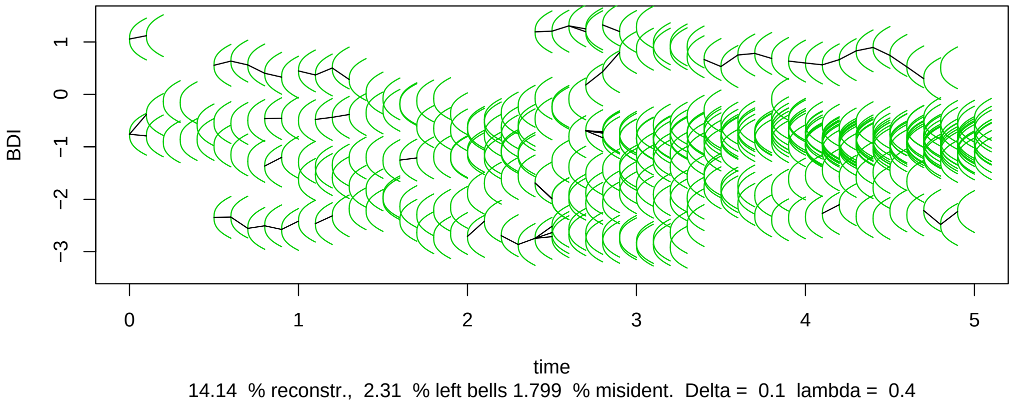
In order to show the mechanism of the reconstruction rule we first have a look at the discrete observations for a big step width $\Delta = 0.1$ (black dots in figure 2).

Figure 2: Discrete Observations for $\Delta = 0.1$



In figure 3 we added for parameter $\lambda = 0.4$ the Δ^λ-neighborhoods corresponding to the discrete observations (green bells) and we draw the reconstructed parts of the trajectory (solid black lines). As mentioned in remark 2.4.2.c we applied the reconstruction rule not only on observations $\beta_{i\Delta} \in D_{2\Delta^\lambda}$, but also on the components $\beta_{i\Delta}^k \in G(\beta_{i\Delta})$ (in the notation of remark 2.4.2.c theses are the 'good' components of $\beta_{i\Delta}$ which have no 'close' neighbor). As we can see, the step width $\Delta = 0.1$ is by far too big to obtain a satisfactory result.

Figure 3: Partial reconstruction for $\Delta = 0.1$ and $\lambda = 0.4$



In the following figures 4 to 15 we plotted the reconstructed trajectories for a finer step width ($\Delta \in \{0.01, 0.005, 0.001\}$) and for different values of the parameter λ (we have chosen $\lambda \in \{0.25, 0.33, 0.40, 0.49\}$).

Tables 1 to 3 recapitulate the results of these reconstructions.

Table 1: Used data in %

		λ			
		0.25	0.33	0.40	0.49
	0.01	19.70	27.79	33.71	42.32
Δ	0.005	23.49	32.58	39.14	49.19
	0.001	31.92	42.50	52.76	64.17

Table 2: Misidentified Observations in %

		λ			
		0.25	0.33	0.40	0.49
	0.01	0.179	0.205	0.538	2.382
Δ	0.005	0.090	0.115	0.294	2.343
	0.001	0.018	0.018	0.084	2.708

In table 1 we calculated the quota of 'good' components $\beta^k_{i\Delta} \in G(\beta_{i\Delta})$ ('used data in %'). This expression corresponds to $1 - q_{\text{cl}}$ in theorem 2.4.4.iii. As we can see this quota increases for decreasing Δ and for every choice of λ. Moreover we observe that the quota of used data increases if we choose smaller bells (or equivalently a bigger value of λ). But for this gain we loose accuracy of the reconstruction as we will see in the following paragraph.

Table 3: Observations which left bells in %

		λ			
		0.25	0.33	0.40	0.49
	0.01	0.00	0.08	0.72	3.92
Δ	0.005	0.00	0.05	0.40	3.53
	0.001	0.00	0.00	0.11	3.17

In table 2 we calculated the quota of components where the reconstruction rule decided wrong (this expression corresponds to q_{prr} in theorem 2.4.4.i and ii). The bigger we choose λ, the more errors occur. On the other hand we observe that for $\lambda = 0.25$, $\lambda = 0.33$ and $\lambda = 0.40$ the quota of misidentified observations decreases for Δ getting smaller. This is exactly what we expect with respect to theorem 2.4.4.ii. But for $\lambda = 0.49$ we cannot affirm the same phenomenon. The reason for that can be explained by the results presented in table 3.

In table 3 we calculated the quota of components $\beta^k_{i\Delta}$ where the corresponding subprocess $\varphi^{(i\Delta,\beta^k_{i\Delta})}$ left the Δ^λ-neighborhood of its starting position in the short time interval $[i\Delta, (i+1)\Delta]$ (these events correspond to the events $a^\Delta(i\Delta, \beta^k_{i\Delta})$ in the notation of section 2.3). We can see that for $\lambda \leq 0.40$ the asymptotics for the quota of observations which left their bell already 'burned in', but for $\lambda = 0.49$ the convergence for $\Delta \downarrow 0$ has not yet started.

Now recall the assertion of theorem 2.4.4.i. There we calculated the intrinsic error of the reconstruction rule 2.4.1 and we have seen that there are two sources of errors which may cause misidentifications: immigrating particles and situations where particles left their Δ^λ-neighborhoods (see also remark 2.4.2.a and b). If we turn our attention again to our example we notice that for $\lambda = 0.49$ the error caused by particles which left their bells is much bigger than the error caused by immigrating particles (which has the upper bound $c \cdot \Delta$ according to theorem 2.4.4.i). This explains why the quota of misidindentified observations does not decrease linearly in Δ in the case $\lambda = 0.49$.

Summarizing above remarks we see that there is a tradeoff between the quota of reconstructed observations and the accuracy of the reconstruction. A good value of λ seems to be $\lambda = 0.33$.

Figure 4: Partial reconstruction for $\Delta = 0.01$ and $\lambda = 0.25$

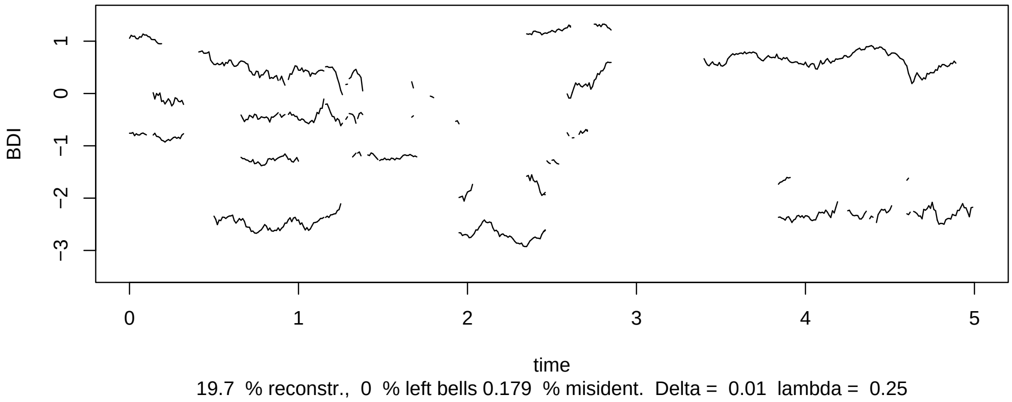


Figure 5: Partial reconstruction for $\Delta = 0.005$ and $\lambda = 0.25$

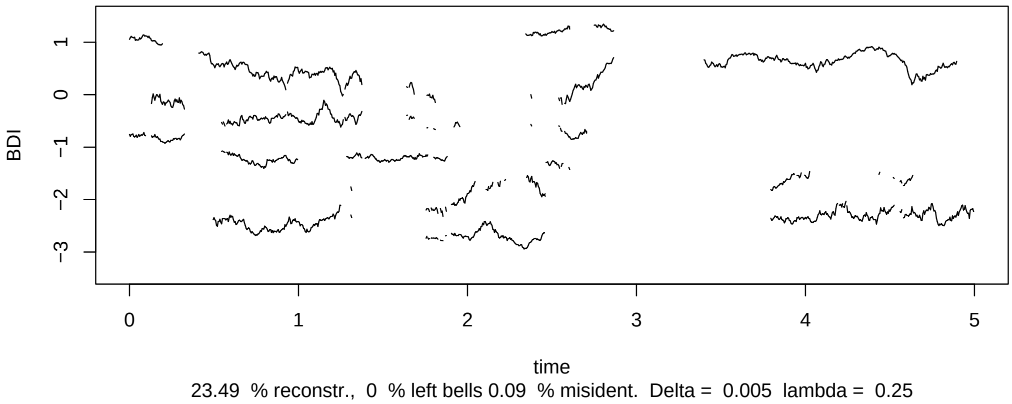


Figure 6: Partial reconstruction for $\Delta = 0.001$ and $\lambda = 0.25$

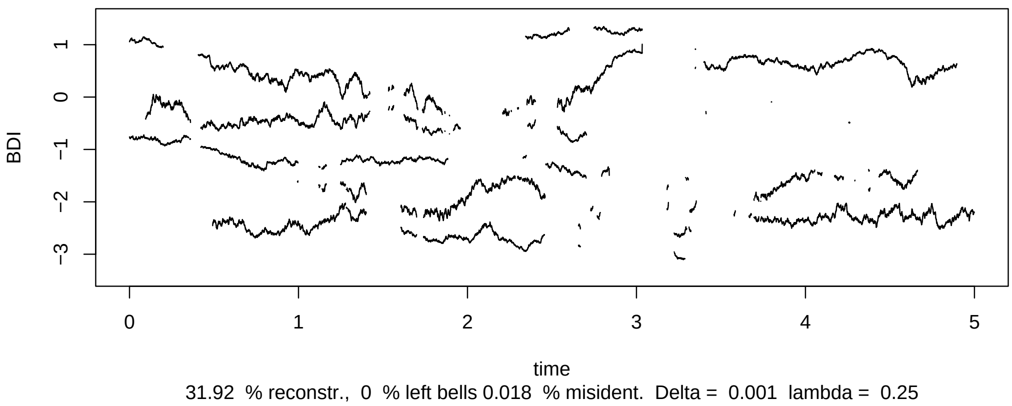


Figure 7: Partial reconstruction for $\Delta = 0.01$ and $\lambda = 0.33$

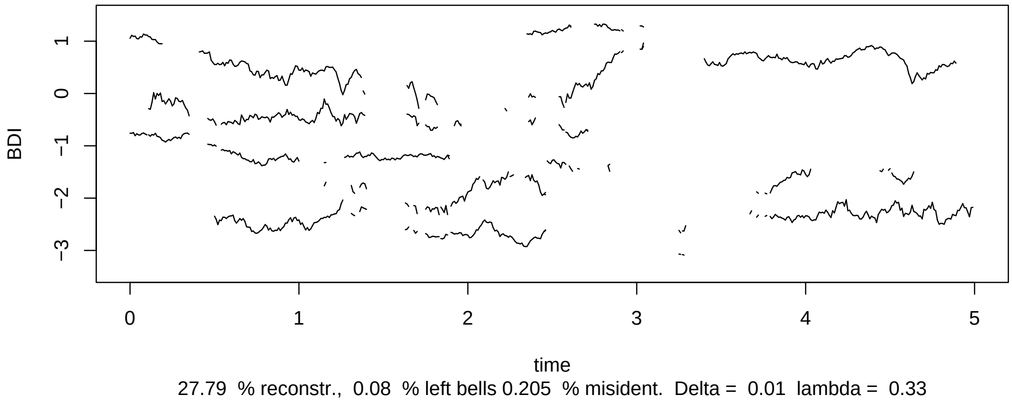


Figure 8: Partial reconstruction for $\Delta = 0.005$ and $\lambda = 0.33$

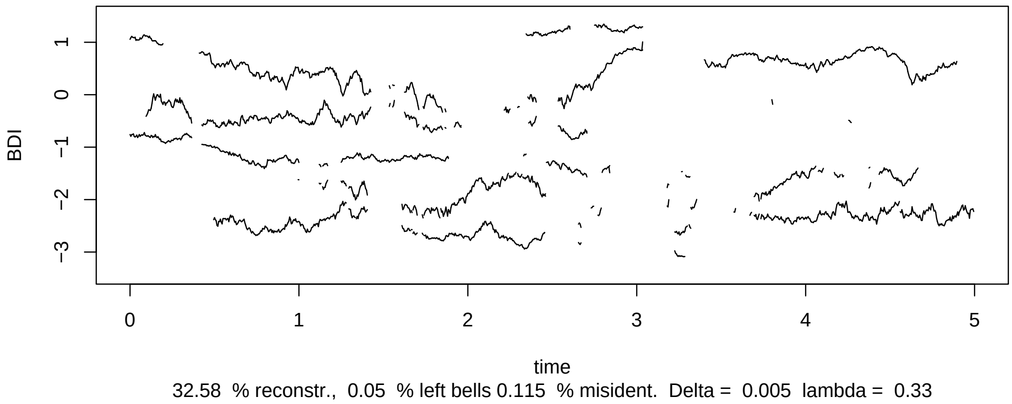


Figure 9: Partial reconstruction for $\Delta = 0.001$ and $\lambda = 0.33$

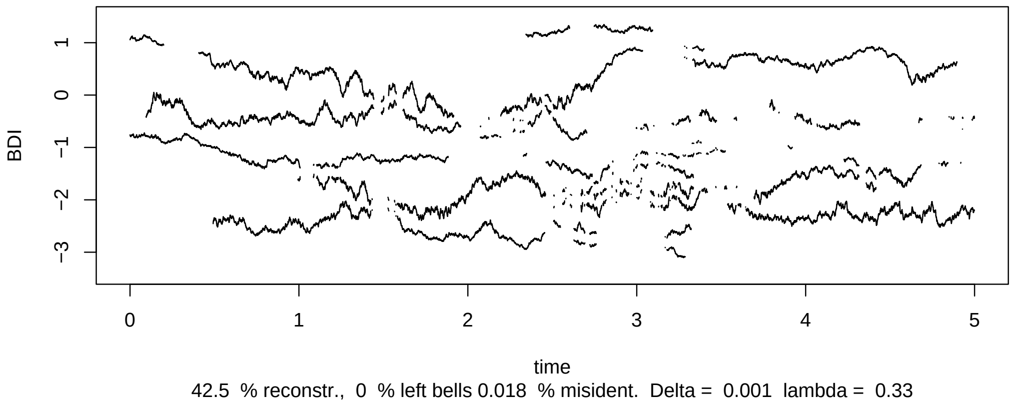


Figure 10: Partial reconstruction for $\Delta = 0.01$ and $\lambda = 0.4$

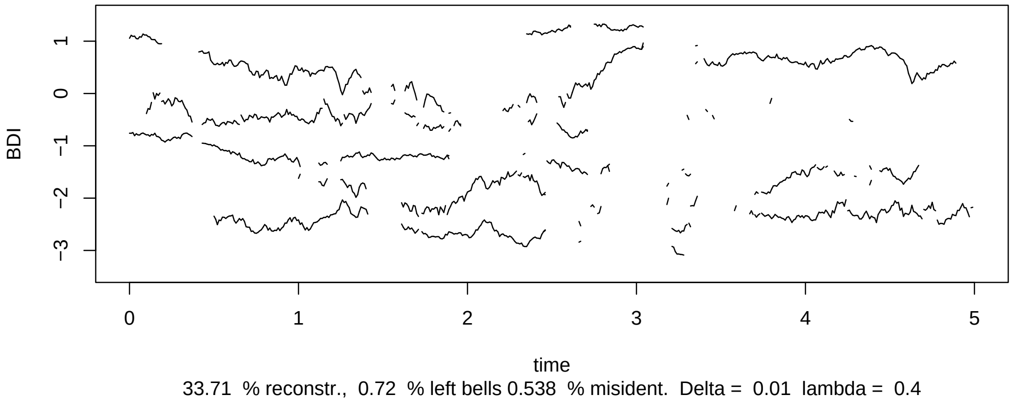


Figure 11: Partial reconstruction for $\Delta = 0.005$ and $\lambda = 0.4$

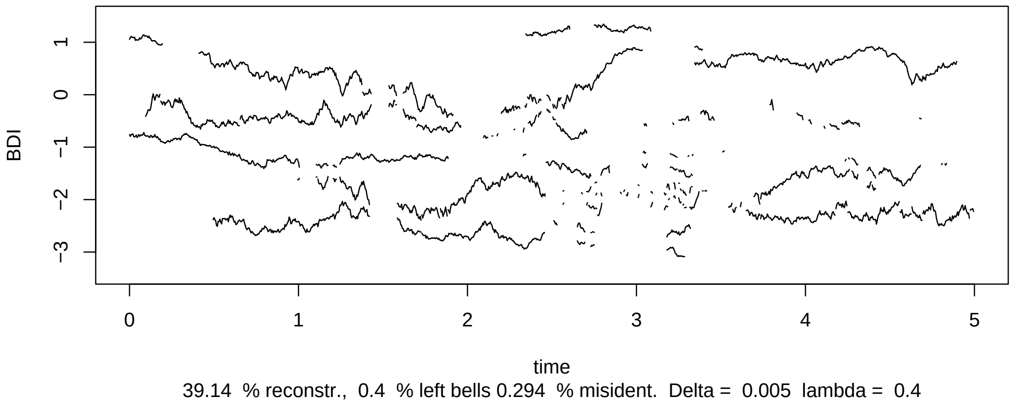


Figure 12: Partial reconstruction for $\Delta = 0.001$ and $\lambda = 0.4$

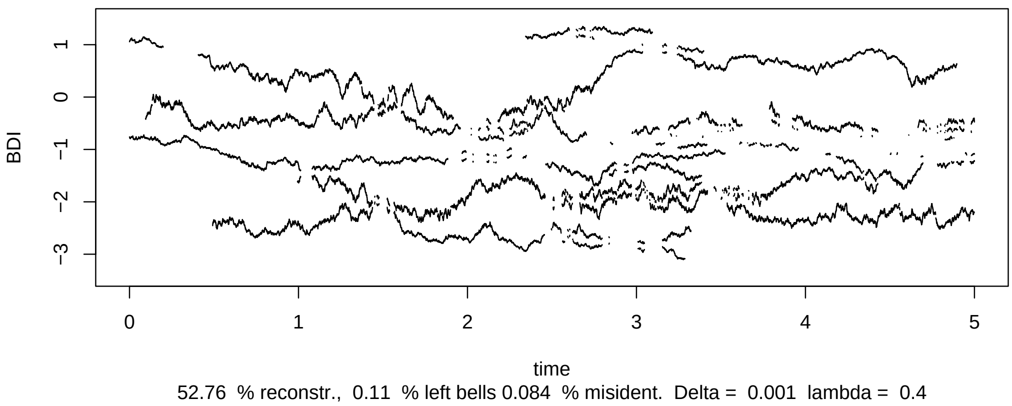


Figure 13: Partial reconstruction for $\Delta = 0.01$ and $\lambda = 0.49$

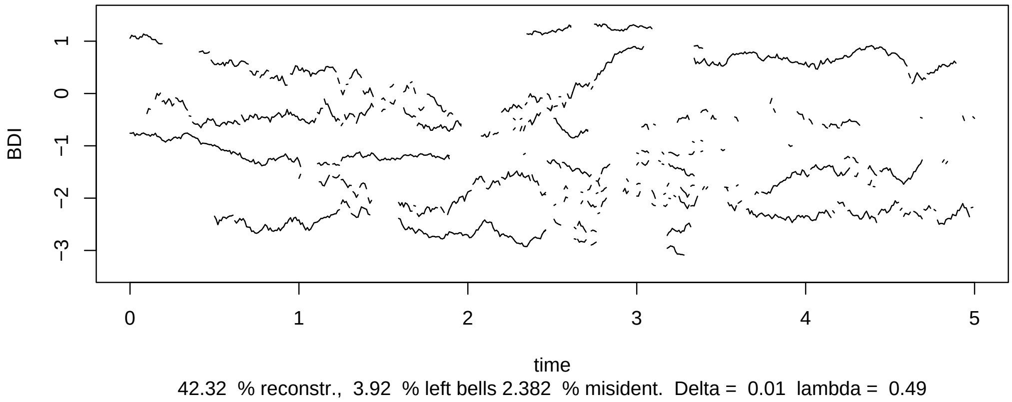


Figure 14: Partial reconstruction for $\Delta = 0.005$ and $\lambda = 0.49$

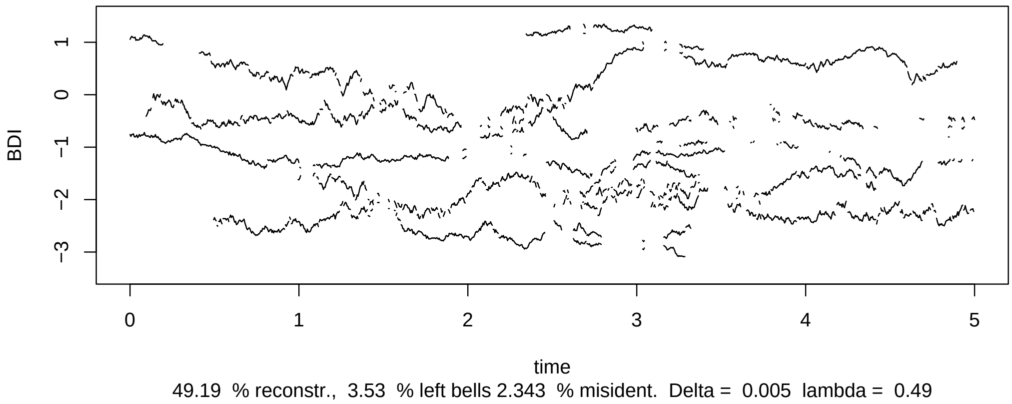
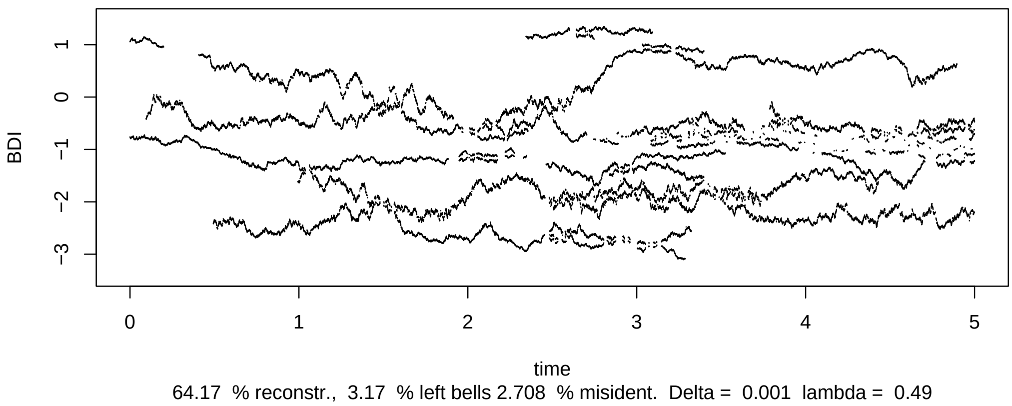


Figure 15: Partial reconstruction for $\Delta = 0.001$ and $\lambda = 0.49$



3 Nonparametric estimation of the diffusion coefficient

In this section we want to show how the partial reconstruction rule 2.4.1 fits into statistical applications. For that we consider the statistical problem of estimating the unknown diffusion coefficient σ^2 of a BDI from discrete data. Section 3.1 provides a short review of results concerning estimation of the diffusion coefficient in the case that we observe an one-dimensional diffusion discretely in time.
In section 3.2 we present a nonparametric estimator for the diffusion coefficient based on discrete observations $\beta_{i\Delta}$ of a branching diffusion with immigration φ on a fixed time interval $[0, T]$.
Consistency of the estimator is shown in section 3.3 and moreover we are able to prove a rate of convergence in section 3.4.
Finally we present simulations of the estimator in section 3.5.

For the whole section we have the following assumptions:

General assumptions for section 3: *We consider the case $d = 1$, i.e. the particles of the BDI move according to one-dimensional diffusions (1).*
Furthermore we assume assumptions A1 to A5 to be true.

Remark 3.0.2. The restriction to $d = 1$ is a natural assumption in statistical problems concerning discretely observed diffusions since in dimensions $d \geq 2$ the occupation time of the diffusion in certain region may be very small or even zero. In that case kernel estimates do not make sense any longer.

3.1 The case of an one-dimensional diffusion

Estimation of the diffusion coefficient from discrete data is a well studied problem in the case that we observe a one-dimensional diffusion. An overview of the different statistical settings (parametric versus nonparametric estimation, constant step widths versus decreasing step widths, etc.) can be found in the preprint of Jacod [Jac01].

For results on the parametric case with constant step width Δ (i.e. observations are taken at times $t_i = i\Delta$, $i = 0, \ldots, n$, $\Delta > 0$ fixed and $n \to \infty$) we refer to the papers of Kessler ([Kes97], [Kes00]) and Kessler/Sørensen [KS99] where the estimation via martingale estimation functions is considered. We also want to mention the paper of Kessler/Paredes [KP02] where computational aspects related to martingale estimation functions are discussed.

For results on the parametric case with decreasing step width (i.e. observations are taken at times $t_i = \frac{i}{n}$, $i = 0, \ldots, n$ and $n \to \infty$) we refer to the papers of Dohnal [Doh87] and Genon-Catalot/Jacod [GCJ94].

Good references for the nonparametric case are the papers of Jacod ([Jac98] and [Jac00]) and the papers of Dacunha/Florens-Zmirou [DCFZ86] and Florens-Zmirou [FZ93].

Here we focus on the nonparametric case. In particular we want to have a closer look at the Nadaraya-Watson estimator for the diffusion coefficient of an one-dimensional ergodic diffusion, introduced in this context the first time by Florens-Zmirou in [FZ93].

For that let $\eta = (\eta_t)_{t \geq 0}$ denote a one-dimensional diffusion with drift $b \in C_b^2(\mathbb{R})$ and diffusion coefficient $\sigma \in C_b^3(\mathbb{R})$. Then an estimator for σ^2 (based on observations at discrete times $i\Delta,\ i = 0, \ldots, [T/\Delta]$) is given by

$$\tilde{\sigma}^2_\Delta(a) := \frac{\sum_{i=0}^{[T/\Delta]-1} \mathbb{1}_{B_{h_\Delta}(a)}(\eta_{i\Delta}) \left[\frac{\eta_{(i+1)\Delta} - \eta_{i\Delta}}{\sqrt{\Delta}}\right]^2}{\sum_{i=0}^{[T/\Delta]-1} \mathbb{1}_{B_{h_\Delta}(a)}(\eta_{i\Delta})} \tag{32}$$

where h_Δ is a sequence of positive numbers with $\lim_{\Delta \downarrow 0} h_\Delta = 0$ and $B_{h_\Delta}(a) := \{y \in \mathbb{R} : |y - a| < h_\Delta\}$ is a h_Δ-neighborhood of $a \in \mathbb{R}$.

Combining proposition 4 and theorem 1' in [FZ93] we have the following result.

Theorem 3.1.1. *If h_Δ is such that $\Delta^{-1} h_\Delta^4$ tends to zero, then $\tilde{\sigma}^2_\Delta(a)$ is a consistent estimator for $\sigma^2(a)$.*
If moreover $\Delta^{-1} h_\Delta^3$ tends to zero, then

$$\sqrt{\frac{h_\Delta}{\Delta}} \left(\frac{\tilde{\sigma}^2_\Delta(a)}{\sigma^2(a)} - 1\right)$$

converges in distribution to $L_T(a)^{-\frac{1}{2}} Z$, where $L_T(a)$ denotes the local time of η and Z is a standard normal variable independent of $L_T(a)$.

Note that this estimator is optimal in the minimax sense under square-error loss for Lipschitz continuous diffusion coefficients (we refer to the paper of Hoffmann [Hof01] where nonparametric estimators for the diffusion coefficient are compared with parametric estimators).

3.2 Definition of the estimator

Thanks to the partial reconstruction rule 2.4.1 we are able to identify coherent parts of the trajectories of the BDI and we will construct an estimator by applying on each part an estimator of type (32). This leads to the following definition.

Definition 3.2.1. *For $a \in \mathbb{R}$, $\lambda \in (0, \frac{1}{2})$ and any sequence h_Δ of positive real numbers with $\lim_{\Delta \downarrow 0} h_\Delta = 0$ we set*

$$\hat{L}_T^\Delta(a) := \frac{\Delta}{2h_\Delta} \sum_{i=0}^{[T/\Delta]-1} \mathbb{1}_{D_{2\Delta^\lambda}}(\beta_{i\Delta})\beta_{i\Delta}(B_{h_\Delta}(a))$$

and

$$\hat{S}_T^\Delta(a, \beta) := \frac{\Delta}{2h_\Delta} \sum_{i=0}^{[T/\Delta]-1} \mathbb{1}_{D_{2\Delta^\lambda}}(\beta_{i\Delta}) \sum_{k=1}^{l(\beta_{i\Delta})} \mathbb{1}_{B_{h_\Delta}(a)}(\beta_{i\Delta}^k)\hat{I}\left(\beta_{i\Delta}^k, \beta_{(i+1)\Delta}^{[i\Delta, \beta_{i\Delta}^k]}\right),$$

where $\hat{I}(\cdot,\cdot)$ is a function from $\mathbb{R} \times S$ to $\mathbb{R}$ defined by

$$\hat{I}(y, x) := \frac{1}{\Delta}\sum_{k=1}^{l(x)}(x^k - y)^2 = x\left(\Delta^{-1}(\cdot - y)^2\right)$$

for $y \in \mathbb{R}$ and $x \in S$ with $l(x) \geq 1$.
If x is the void configuration we set $\hat{I}(y, \delta) := 0$ by convention.

Then our estimator for the diffusion coefficient σ^2 evalued at $a \in \mathbb{R}$ is given by

$$\hat{\sigma}_\Delta^2(a, \beta) := \frac{\hat{S}_T^\Delta(a)}{\hat{L}_T^\Delta(a)} \mathbb{1}_{\left\{\hat{L}_T^\Delta(a) > \varepsilon\right\}}.$$

where $\varepsilon > 0$ can be chosen arbitrary small.

Remark 3.2.2. a) The estimator works as follows: $\hat{L}_T^\Delta(a)$ defines a discrete local time of φ in $a \in \mathbb{R}$ up to time T and $\hat{S}_T^\Delta(a, \beta)$ can be interpreted as quadratic increments of φ integrated with respect to the discrete local time $\hat{L}_T^\Delta(a)$. Hence $\hat{\sigma}_\Delta^2(a, \beta)$ has the same structure as the estimator for the diffusion coefficient of a one-dimensional diffusion presented in section 3.1.

b) Note that in practice one would use a finer version of the estimator defined above, which uses not only observations $\beta_{i\Delta} \in D_{2\Delta^\lambda}$ but also the 'good' components $\beta_{i\Delta}^k \in G(\beta_{i\Delta})$ where $G(\beta_{i\Delta})$ is the subconfiguration of $\beta_{i\Delta}$ where all components have no neighbor in a $2\Delta^\lambda$-neighborhood (see (25) for the definition of $G(\beta_{i\Delta})$ and remark 2.4.2.c for further explanations). As the reader will easily check, all results of the following sections remain true if we use the finer version of the estimator.

c) Note that the estimator $\hat{\sigma}^2_\Delta(a, \beta)$ depends on two parameters: first we have to choose the parameter $\lambda \in (0, \frac{1}{2})$ which determines the size of the 'bells' and secondly we have to choose the bandwidth h_Δ. We will discuss in section 3.5 which choice of λ and h_Δ will lead to good results.

d) There is no restriction in defining the estimator only on the set $\{\hat{L}^\Delta_T(a) > \varepsilon\}$ since on the complement $\{\hat{L}^\Delta_T(a) \leq \varepsilon\}$ the process has visited a small neighborhood of a only for a very short time (depending on the arbitrary small ε) and in this situation the estimator $\hat{\sigma}^2_\Delta(a, \beta)$ does not make sense any longer.

3.3 Consistency of the estimator

We will show the consistency of the estimator in three steps. First we restrict ourself to the simpler case where we are able to observe not only the positions of the particles, but the whole trajectory of the BDI φ. Then we do not need to apply the partial reconstruction rule and we will show that a simpler estimator, which uses all available data, attains a rate of convergence of order $\sqrt{\Delta h_{\Delta}^{-1}}$.
In the second step we compare this estimator with an estimator which uses only the 'good' observations $\beta_{i\Delta} \in D_{2\Delta^{\lambda}}$, still under the assumption that we are able to observe the whole trajectory of our process.
In the last step we abandon the (hypothetical) assumption that we have all information about our process and consider the original problem stated in section 2.1.

Before we are going into detail we introduce in the next section the local time of a branching diffusion with immigration.

3.3.1 Local time of the BDI

Consider a decomposition of a trajectory of a BDI φ into trajectories of the single particles of the process. Let η^p denote the trajectory of the p-th particle of φ and let B_p denote its birth time, D_p its death time, $p \in \mathbb{N}$.
Then (by construction of the branching diffusion) η^p is a one-dimensional diffusion on $[\![B_p, D_p[\![$ with drift b and diffusion coefficient σ.
On the set $]\!]-\infty, B_p]\!] \cup [\![D_p, \infty[\![$ the trajectory of η^p is constant equal to δ (cementary).

Each process η^p has its own local time in $a \in \mathbb{R}$, denoted by $(L_t^p(a))_{t\geq 0}$, which is an increasing process on $[\![B_p, D_p[\![$ (for an introduction to local times see for example [RY99, chapter VI]). On $]\!]-\infty, B_p]\!]$ the local time $L^p(a)$ is constant equal to 0 and on $[\![D_p, \infty[\![$ the local time remains constant at level $L_{D_p-}^p(a)$.

Since local times are additive we may define a local time of φ by summing up the local times of its particles.

Definition 3.3.1. The *local time* of a branching diffusion with immigration φ at position $a \in \mathbb{R}$ is given by

$$L_t^{\mathrm{BDI}}(a) := \sum_{p\in\mathbb{N}} L_t^p(a), \quad t \geq 0.$$

Remark 3.3.2. By theorem 1.2.2 assumptions A1 and A2 are sufficient for property **P1**, i.e. there is no accumulation of jumps in finite time. In particular there are a.s. only finitely many branching or immigration events in finite time and at each branching event there are only a.s. finitely many particles born (the reproduction law has bounded first moments by assumption A2). Hence the sum over p in above definition of L^{BDI} is almost sure finite.

3.3.2 Estimator under knowledge of the whole trajectory

We first consider the case where we are able to observe not only the positions of the particles, but the whole trajectory of φ on the time interval $[0, T]$.
Then there is no need to use a partial reconstruction rule and we may use all available discrete observations.

We define a simpler version of the estimator $\hat{\sigma}^2_\Delta(a, \beta)$ presented in definition 3.2.1 by setting

$$\bar{L}^\Delta_T(a) := \frac{\Delta}{2h_\Delta} \sum_{i=0}^{[T/\Delta]-1} \varphi_{i\Delta}(B_{h_\Delta}(a))$$

and

$$\bar{S}^\Delta_T(a, \varphi) := \frac{\Delta}{2h_\Delta} \sum_{i=0}^{[T/\Delta]-1} \sum_{k=1}^{l(\varphi_{i\Delta})} \mathbb{1}_{B_{h_\Delta}(a)}(\varphi^k_{i\Delta}) \hat{I}(\varphi^k_{i\Delta}, \varphi^{(i\Delta, \varphi^k_{i\Delta})}_{(i+1)\Delta})$$

for $a \in \mathbb{R}$ and h_Δ such that $h_\Delta \downarrow 0$ for $\Delta \downarrow 0$, where $\hat{I}(\cdot, \cdot)$ is given in definition 3.2.1, and finally

$$\overline{\sigma}^2_\Delta(a, \varphi) := \frac{\overline{S}^\Delta_T(a, \varphi)}{\overline{L}^\Delta_T(a)}. \tag{33}$$

By convention we set $\frac{0}{0} := 0$. (Note that $\overline{L}^\Delta_T(a) = 0$ implies $\overline{S}^\Delta_T(a, \varphi) = 0$ by definition of $\overline{L}^\Delta_T(a)$ and $\overline{S}^\Delta_T(a, \varphi)$.)

Before we state the main theorem of this section, we collect some additional assumptions we will need for a proof of this theorem.

Consider the Campbell measure m^c associated to the invariant measure m on S, i.e.

$$m^c(A \times F) := \int_S x(A) \mathbb{1}_F(x)\, m(dx) \quad \text{ for } A \in \mathcal{B}(\mathbb{R}), F \in \mathcal{S}.$$

Since $(S, \mathcal{S})$ is a Polish space (see [Lö99, 3.7, p. 15]) we may disintegrate m^c and obtain

$$m^c(dy, dx) = \overline{m}(dy) m^p(y, dx),$$

where m^p is the so called palm kernel corresponding to m^c (see [Kal02]). Note that the first marginal of m^c coincides with the occupation time measure $\overline{m}$ since by definition $\overline{m}(A) = \int_S x(A)\ m(dx)$.
The palm kernel m^p can be interpreted in the following way: $m^p(y, \cdot)$ is the law of φ_t in a stationary state under the condition that one component of φ_t is located in $y \in \mathbb{R}$.

Our additional assumptions are the following:

Assumption A6.

i) The invariant measure m on S fulfills

$$\int_{\mathbb{R}} l(x)^2\ m(dx) < \infty.$$

ii) The Palm Kernel m^p fulfills

$$\sup_{y \in A} \int_S l(x)\ m^p(y, dx) < \infty$$

for any compact set $A \subset \mathbb{R}$.

iii) Suppose that there is an $\tilde{\varepsilon} > 0$ such that

$$y \mapsto E\left(l(\varphi_1^{(0,y)})^{2+\tilde{\varepsilon}} \right)$$

is a bounded function.

iv) $\overline{m}$ has a continuous and bounded density with respect to the Lebesgue measure on $\mathbb{R}$.

Remark 3.3.3. Assumption A6.i and assumption A6.ii are mainly technical assumptions. In particular assumption A6.ii claims that there is no clustering of particles conditioned on the event that one particle is located in a given compact set A. Furthermore recall that we already have presented sufficient conditions for assumption A6.iv in section 1.3.1 and finally note that assumption A6.iii is trivially fulfilled in the case that $p(\cdot)$ has uniformly finite support, i.e. there is a $N \in \mathbb{N}$ with $p_k(x) = 0$ for all $k \geq N$ and for all $x \in \mathbb{R}$.

Theorem 3.3.4. *Let assumption A6 hold. If h_Δ is such that $\Delta h_\Delta^{-2} \log(\Delta) \to 0$ for $\Delta \downarrow 0$ then $\overline{\sigma}^2_\Delta(a, \varphi)$ is a consistent estimator for $\sigma^2(a)$.*
Moreover we are able to give a rate of convergence. We have

$$E\left(\left| \sqrt{\frac{h_\Delta}{\Delta}} \left(\overline{\sigma}^2_\Delta(a, \varphi) - \sigma^2(a) \right) \right|^2 \middle| \overline{L}_T^\Delta > 0 \right) = \mathcal{O}(1) \qquad \textit{for } \Delta \downarrow 0.$$

We split the proof of this theorem into two parts. First we show in lemma 3.3.5 that $\overline{L}_T^\Delta(a)$ converges Q_y-a.s. to the local time $L_T^{\mathrm{BDI}}(a)$.
After that we propose a technical lemma (lemma 3.3.6) which allows us to show the $\mathbb{L}^2$-boundedness of $\sqrt{\frac{h_\Delta}{\Delta}}\left(\overline{S}_T^\Delta(a,\varphi) - \sigma^2(a)\overline{L}_T^\Delta(a)\right)$ (lemma 3.3.7 below).
Combining lemma 3.3.5 and lemma 3.3.7, theorem 3.3.4 follows immediately.

Lemma 3.3.5. *If $\Delta h_\Delta^{-2}\log\Delta \to 0$ for $\Delta \downarrow 0$ then for $T \geq 0$ and $a \in \mathbb{R}$*

$$\bar{L}_T^\Delta(a) \longrightarrow L_T^{BDI}(a) \quad Q_m\text{-}a.s.$$

for $\Delta \downarrow 0$.

Proof. Consider a decomposition of φ into a union of single particles $(\eta^p)_{p\in\mathbb{N}}$ moving according to (1) as described in section 3.3.1. Rearranging the summands of $\overline{L}_T^\Delta(a)$ gives us the representation

$$\overline{L}_T^\Delta(a) = \frac{\Delta}{2h_\Delta}\sum_{p\in\mathbb{N}}\sum_{i=[B^p/\Delta]}^{[D^p\wedge T/\Delta]} \mathbb{1}_{B_{h_\Delta}(a)}(\eta_{i\Delta}^p).$$

Now by proposition 2 in [FZ93] (under the assumption that $\Delta h_\Delta^{-2}\log\Delta \to 0$) for each p with $B^p < T$

$$\frac{\Delta}{2h_\Delta}\sum_{i=[B^p/\Delta]}^{[D^p\wedge T/\Delta]} \mathbb{1}_{B_{h_\Delta}(a)}(\eta_{i\Delta}^p) \longrightarrow L_T^p(a) \quad Q_m\text{-a.s.}$$

for $\Delta \downarrow 0$.

As mentioned in remark 3.3.2 the sum over p is Q_m-a.s. finite, hence

$$\overline{L}_T^\Delta(a) \longrightarrow \sum_{p\in\mathbb{N}} L_T^p(a) = L_T^{\mathrm{BDI}}(a) \quad Q_m\text{-a.s.}$$

for $\Delta \downarrow 0$, which is the assertion. $\square$

Lemma 3.3.6. *Let assumption A6 hold. Then for $y \in \mathbb{R}$*

i) $E\left(\hat{I}(y,\varphi_\Delta^{(0,y)})\right) = \sigma^2(y) + \mathcal{O}\left(\sqrt{\Delta}\right)$ *for $\Delta \downarrow 0$*

and

ii) $E\left(\left[\hat{I}(y,\varphi_\Delta^{(0,y)})\right]^2\right) = \mathcal{O}(1)$ *for $\Delta \downarrow 0$.*

Proof. a) We first show i).
Recall that by definition of the function $\hat{I}(\cdot,\cdot)$

$$E\left(\hat{I}(y,\varphi_{\Delta}^{(0,y)})\right) = E\left(\varphi_{\Delta}^{(0,y)}\left(\Delta^{-1}(\cdot - y)^2\right)\right).$$

By lemma 2.3.2 the semi-group of $\varphi^{(0,y)}$ coincides on $C_b^2(\mathbb{R})$ with the semi-group of a diffusion η given by (1), killed at position dependent rate $\kappa(\cdot)(1-\rho(\cdot))$. Moreover the 'mass reduction rate' $\kappa(\cdot)(1-\rho(\cdot))$ is a bounded function by assumption A2, thus

$$\begin{aligned} & E\left(\varphi_{\Delta}^{(0,y)}\left(\Delta^{-1}(\cdot - y)^2\right)\right) \\ = \;& E_y\left(\frac{1}{\Delta}(\eta_\Delta - \eta_0)^2 e^{-\int_0^\Delta [\kappa(1-\rho)](\eta_s)\;ds}\right) \\ = \;& E_y\left(\frac{1}{\Delta}(\eta_\Delta - \eta_0)^2\right)\cdot\left(1+\mathcal{O}\left(\Delta\right)\right) \quad \text{for } \Delta \downarrow 0. \end{aligned} \tag{34}$$

Now we will show that $E_y\left(\frac{1}{\Delta}(\eta_\Delta-\eta_0)^2\right) = \sigma^2(\eta_0) + \mathcal{O}\left(\sqrt{\Delta}\right)$ for $\Delta \downarrow 0$.
By (1)

$$\begin{aligned} & \left|E_y\left(\frac{1}{\Delta}(\eta_\Delta - \eta_0)^2\right) - \sigma^2(\eta_0)\right| \\ \le \;& E_y\left(\frac{1}{\Delta}\left(\int_0^\Delta b(\eta_s)\;ds\right)^2\right) + E_y\left(\frac{1}{\Delta}\left|\left(\int_0^\Delta b(\eta_s)\;ds\right)\left(\int_0^\Delta \sigma(\eta_s)\;dW_s\right)\right|\right) \\ & + \left|E_y\left(\frac{1}{\Delta}\left(\int_0^\Delta \sigma(\eta_s)\;dW_s\right)^2\right) - \sigma^2(\eta_0)\right|. \end{aligned}$$

Since the drift function b is assumed to be bounded

$$E_y\left(\frac{1}{\Delta}\left(\int_0^\Delta b(\eta_s)\;ds\right)^2\right) = \mathcal{O}\left(\Delta\right) \quad \text{for } \Delta \downarrow 0 \tag{35}$$

and by the Cauchy-Schwartz inequality

$$\begin{aligned} & E_y\left(\frac{1}{\Delta}\left|\left(\int_0^\Delta b(\eta_s)\;ds\right)\left(\int_0^\Delta \sigma(\eta_s)\;dW_s\right)\right|\right) \\ \le \;& \frac{1}{\Delta}\left(E_y\left[\left(\int_0^\Delta b(\eta_s)\;ds\right)^2\right]\right)^{\frac{1}{2}} \cdot \left(E_y\left[\left(\int_0^\Delta \sigma(\eta_s)\;dW_s\right)^2\right]\right)^{\frac{1}{2}} \\ = \;& \frac{1}{\Delta}\left(E_y\left[\left(\int_0^\Delta b(\eta_s)\;ds\right)^2\right]\right)^{\frac{1}{2}} \cdot \left(E_y\left[\int_0^\Delta \sigma^2(\eta_s)\;ds\right]\right)^{\frac{1}{2}} \\ \le \;& \overline{b}\cdot\overline{\sigma}^2\,\sqrt{\Delta} \end{aligned} \tag{36}$$

where we used again the boundedness of b and σ.

Note that
$$E_y\left[\left(\int_0^\Delta \sigma(\eta_s)\ dW_s\right)^2\right] = E_y\left[\int_0^\Delta \sigma^2(\eta_s)\ ds\right]$$
since $M_t := \int_0^t \sigma(\eta_s)\ dW_s$ defines a martingale with quadratic variation process given by $\langle M\rangle_t = \int_0^t \sigma^2(\eta_s)\ ds$.

It remains to study the asymptotics of the term
$$(*) := \left|E_y\left(\frac{1}{\Delta}\left(\int_0^\Delta \sigma(\eta_s)\ dW_s\right)^2\right) - \sigma^2(\eta_0)\right|$$

By the Lipschitz continuity σ and the boundedness of σ and b
$$\begin{aligned}
(*) &= \left|E_y\left(\frac{1}{\Delta}\int_0^\Delta \sigma^2(\eta_s) - \sigma^2(\eta_0)\ ds\right)\right| \\
&\leq 2\overline{\sigma}L \cdot E_y\left(\sup_{t\in[0,\Delta]} |\eta_t - \eta_0|\right) \\
&\leq 2\overline{\sigma}L \cdot \left[E_y\left(\sup_{t\in[0,\Delta]}\left|\int_0^t b(\eta_s)\ ds\right|\right) + E_y\left(\sup_{t\in[0,\Delta]}\left|\int_0^t \sigma(\eta_t)\ dW_s\right|\right)\right] \\
&\leq 2\overline{\sigma}L\ \overline{b}\cdot\Delta + 2\overline{\sigma}L \cdot E_y\left(\sup_{t\in[0,\Delta]}\left|\int_0^t \sigma(\eta_t)\ dW_s\right|\right).
\end{aligned}$$

An application of a Burkholder-Davis-Gundy inequality (see for instance theorem VII.94 in [DM82]) gives us
$$E_y\left(\sup_{t\in[0,\Delta]}\left|\int_0^t \sigma(\eta_t)\ dW_s\right|\right) \leq K\cdot E_y\left(\left(\int_0^\Delta \sigma^2(\eta_s)\ ds\right)^{\frac{1}{2}}\right) \leq K\overline{\sigma}\cdot\sqrt{\Delta},$$
where K is a suitable constant independent of Δ.

Hence
$$(*) = \mathcal{O}\left(\sqrt{\Delta}\right) \qquad \text{for } \Delta\downarrow 0. \tag{37}$$

Combining (34) to (37) gives us the first assertion.

b) Now we want to show assertion ii), i.e. we want to prove that
$$E\left(\left[\varphi_\Delta^{(0,y)}\left(\Delta^{-1}(\cdot - y)^2\right)\right]^2\right) = \mathcal{O}(1) \qquad \text{for } \Delta\downarrow 0.$$

For $\tilde{\varepsilon} > 0$ choose $p, q > 1$ such that $2p = 2 + \tilde{\varepsilon}$ and $\frac{1}{p} + \frac{1}{q} = 1$. Then we get by an application of the Hölder inequality

$$\begin{aligned}
& E\left(\left[\varphi_{\Delta}^{(0,y)}\left(\Delta^{-1}(\cdot - y)^2\right)\right]^2\right) \\
= \ & \frac{1}{\Delta^2} E\left(\left[l(\varphi_{\Delta}^{(0,y)})\right]^2 \cdot \left[\frac{1}{l(\varphi_{\Delta}^{(0,y)})} \sum_{k=1}^{l(\varphi_{\Delta}^{(0,y)})} \left(\varphi_{\Delta}^{(0,y),k} - y\right)^2\right]^2 \cdot \mathbb{1}_{\left\{l(\varphi_{\Delta}^{(0,y)})>0\right\}}\right) \\
\leq \ & \frac{1}{\Delta^2}\left[E\left(\left[l(\varphi_{\Delta}^{(0,y)})\right]^{2p}\right)\right]^{\frac{1}{p}} \\
& \times \left[E\left(\left[\frac{1}{l(\varphi_{\Delta}^{(0,y)})} \sum_{k=1}^{l(\varphi_{\Delta}^{(0,y)})} \left(\varphi_{\Delta}^{(0,y),k} - y\right)^2\right]^{2q} \cdot \mathbb{1}_{\left\{l(\varphi_{\Delta}^{(0,y)})>0\right\}}\right)\right]^{\frac{1}{q}}.
\end{aligned}$$

By assumption A6.iii

$$E\left(\left[l(\varphi_{\Delta}^{(0,y)})\right]^{2p}\right) = E\left(\left[l(\varphi_{\Delta}^{(0,y)})\right]^{2+\tilde{\varepsilon}}\right) = \mathcal{O}(1) \quad \text{ for } \Delta \downarrow 0$$

and by a Jenssen inequality

$$\begin{aligned}
& \left[E\left(\left[\frac{1}{l(\varphi_{\Delta}^{(0,y)})} \sum_{k=1}^{l(\varphi_{\Delta}^{(0,y)})} \left(\varphi_{\Delta}^{(0,y),k} - y\right)^2\right]^{2q} \cdot \mathbb{1}_{\left\{l(\varphi_{\Delta}^{(0,y)})>0\right\}}\right)\right]^{\frac{1}{q}} \\
\leq \ & \left[E\left(\frac{1}{l(\varphi_{\Delta}^{(0,y)})} \sum_{k=1}^{l(\varphi_{\Delta}^{(0,y)})} \left(\varphi_{\Delta}^{(0,y),k} - y\right)^{4q} \cdot \mathbb{1}_{\left\{l(\varphi_{\Delta}^{(0,y)})>0\right\}}\right)\right]^{\frac{1}{q}} \\
\leq \ & \left[E\left(\varphi_{\Delta}^{(0,y)}\left((\cdot - y)^{4q}\right)\right)\right]^{\frac{1}{q}}.
\end{aligned}$$

Hence it remains to show that for $q > 1$

$$\frac{1}{\Delta^2}\left[E\left(\varphi_{\Delta}^{(0,y)}\left((\cdot - y)^{4q}\right)\right)\right]^{\frac{1}{q}} = \mathcal{O}(1) \quad \text{ for } \Delta \downarrow 0. \tag{38}$$

As in (34)

$$\begin{aligned}
& \left[E\left(\varphi_{\Delta}^{(0,y)}\left((\cdot - y)^{4q}\right)\right)\right]^{\frac{1}{q}} \\
= \ & \left[E_y\left((\eta_\Delta - \eta_0)^{4q} e^{-\int_0^\Delta [\kappa(1-\rho)](\eta_s)\, ds}\right)\right]^{\frac{1}{q}} \\
\leq \ & \left[E_y\left((\eta_\Delta - \eta_0)^{4q}\right)\right]^{\frac{1}{q}},
\end{aligned}$$

where the last step follows with the boundedness of the function $\kappa(\cdot)(1-\rho(\cdot))$.

An application of the Burkholder-Davis-Gundy inequality ([DM82, theorem VII.94]) yields

$$\begin{aligned}
&\left[E_y\left((\eta_\Delta-\eta_0)^{4q}\right)\right]^{\frac{1}{q}}\\
\leq\ &\left[\left(E_y\left[\int_0^\Delta b(\eta_s)\ ds\right]^{4q}\right)^{\frac{1}{4q}}+K_q\cdot\left(E_y\left[\left(\int_0^\Delta \sigma^2(\eta_s)\ ds\right)^{\frac{1}{2}}\right]\right)^{\frac{1}{4q}}\right]^4\\
\leq\ &\left[\overline{b}\ \Delta+K_q\ \overline{\sigma}\ \sqrt{\Delta}\right]^4
\end{aligned}$$

by the boundedness of b and σ, where K_q is a constant which depends only on q.

Thus

$$\frac{1}{\Delta^2}\left[E\left(\varphi_\Delta^{(0,y)}\left((\cdot-y)^{4q}\right)\right)\right]^{\frac{1}{q}}\leq\frac{1}{\Delta^2}\left(\overline{b}\ \Delta+K_q\ \overline{\sigma}\ \sqrt{\Delta}\right)^4=\mathcal{O}(1)\quad \text{for } \Delta\downarrow 0$$

which is exactly (38) and the proof is finished. □

Lemma 3.3.7. *Let assumption A6.ii hold. Then we have for $a\in\mathbb{R}$*

$$E\left(\left[\sqrt{\frac{h_\Delta}{\Delta}}\left(\bar{S}_T^\Delta(a,\varphi)-\sigma^2(a)\bar{L}_T^\Delta(a)\right)\right]^2\right)=\mathcal{O}(1)\quad \textit{for } \Delta\downarrow 0.$$

Proof. We have to consider the following term:

$$\begin{aligned}
&E_m\left(\left|\bar{S}_T^\Delta(a,\varphi)-\sigma^2(a)\bar{L}_T^\Delta(a)\right|^2\right)\\
=\ &E_m\left(\left[\frac{\Delta}{2h_\Delta}\sum_{i=0}^{[T/\Delta]-1}\sum_{k=1}^{l(\varphi_{i\Delta})}\mathbb{1}_{B_{h_\Delta}(a)}(\varphi_{i\Delta}^k)\left(\hat{I}(\varphi_{i\Delta}^k,\varphi_{(i+1)\Delta}^{(i\Delta,\varphi_{i\Delta}^k)})-\sigma^2(a)\right)\right]^2\right)\\
=\ &\frac{\Delta^2}{4h_\Delta^2}\sum_{i=0}^{[T/\Delta]-1}E_m\left(\left[\sum_{k=1}^{l(\varphi_{i\Delta})}\mathbb{1}_{B_{h_\Delta}(a)}(\varphi_{i\Delta}^k)\left(\hat{I}(\varphi_{i\Delta}^k,\varphi_{(i+1)\Delta}^{(i\Delta,\varphi_{i\Delta}^k)})-\sigma^2(a)\right)\right]^2\right)\\
&+\frac{\Delta^2}{4h_\Delta^2}\sum_{i\neq j}E_m\left[\varphi_{i\Delta}\left(\mathbb{1}_{B_{h_\Delta}}(\cdot)\left(\hat{I}(\cdot,\varphi_{(i+1)\Delta}^{(i\Delta,\cdot)})-\sigma^2(a)\right)\right)\right.\\
&\qquad\qquad\left.\times\ \varphi_{j\Delta}\left(\mathbb{1}_{B_{h_\Delta}}(\cdot)\left(\hat{I}(\cdot,\varphi_{(j+1)\Delta}^{(j\Delta,\cdot)})-\sigma^2(a)\right)\right)\right]\\
=:\ &(\mathrm{I})+(\mathrm{II}).
\end{aligned}$$

We first consider for $i\in\{0,\dots,[T/\Delta]-1\}$ and $k\in\{1,\dots,l(\varphi_{i\Delta})\}$ the conditional expectation

$$E_m\left(\left(\hat{I}(\varphi_{i\Delta}^k,\varphi_{(i+1)\Delta}^{(i\Delta,\varphi_{i\Delta}^k)})-\sigma^2(a)\right)\Big|\,\mathcal{F}_{i\Delta}\right)=E\left(\hat{I}(\varphi_{i\Delta}^k,\varphi_{(i+1)\Delta}^{(i\Delta,\varphi_{i\Delta}^k)})\right)-\sigma^2(a).$$

By lemma 3.3.6.i

$$E\left(\hat{I}(\varphi^k_{i\Delta},\varphi^{(i\Delta,\varphi^k_{i\Delta})}_{(i+1)\Delta})\right) = \sigma^2(\varphi^k_{i\Delta}) + \mathcal{O}\left(\sqrt{\Delta}\right) \quad \text{for } \Delta \downarrow 0$$

and by the boundedness and the Lipschitz continuity of σ

$$\mathbb{1}_{B_{h_\Delta}(a)}(\varphi^k_{i\Delta})\left(\sigma^2(\varphi^k_{i\Delta}) - \sigma^2(a)\right) = \mathbb{1}_{B_{h_\Delta}(a)}(\varphi^k_{i\Delta}) \cdot \mathcal{O}\left(h_\Delta\right) \quad \text{for } \Delta \downarrow 0.$$

Hence we have

$$\begin{aligned} & \mathbb{1}_{B_{h_\Delta}(a)}(\varphi^k_{i\Delta})\ E_m\left(\left(\hat{I}(\varphi^k_{i\Delta},\varphi^{(i\Delta,\varphi^k_{i\Delta})}_{(i+1)\Delta}) - \sigma^2(a)\right)\middle| \mathcal{F}_{i\Delta}\right) \\ = \ & \mathbb{1}_{B_{h_\Delta}(a)}(\varphi^k_{i\Delta})\ \mathcal{O}\left(\sqrt{\Delta}+h_\Delta\right) \quad \text{for } \Delta \downarrow 0. \end{aligned} \tag{39}$$

Now we turn our attention to the quadratic term (I):

$$\begin{aligned} & \frac{\Delta^2}{4h^2_\Delta}\sum_{i=0}^{[T/\Delta]-1} E_m\left(\left[\sum_{k=1}^{l(\varphi_{i\Delta})} \mathbb{1}_{B_{h_\Delta}(a)}(\varphi^k_{i\Delta})\left(\hat{I}(\varphi^k_{i\Delta},\varphi^{(i\Delta,\varphi^k_{i\Delta})}_{(i+1)\Delta}) - \sigma^2(a)\right)\right]^2\right) \\ = \ & \frac{\Delta^2}{4h^2_\Delta}\sum_{i=0}^{[T/\Delta]-1} E_m\left[\sum_{k=1}^{l(\varphi_{i\Delta})} \mathbb{1}_{B_{h_\Delta}(a)}(\varphi^k_{i\Delta}) \times (\text{Ia})\right] \\ & + \frac{\Delta^2}{4h^2_\Delta}\sum_{i=0}^{[T/\Delta]-1} E_m\left[\sum_{k\neq k'} \mathbb{1}_{B_{h_\Delta}(a)}(\varphi^k_{i\Delta})\mathbb{1}_{B_{h_\Delta}(a)}(\varphi^{k'}_{i\Delta}) \times (\text{Ib})\right], \end{aligned}$$

where

$$(\text{Ia}) := E_m\left(\left(\hat{I}(\varphi^k_{i\Delta},\varphi^{(i\Delta,\varphi^k_{i\Delta})}_{(i+1)\Delta}) - \sigma^2(a)\right)^2\middle| \mathcal{F}_{i\Delta}\right)$$

and

$$(\text{Ib}) := E_m\left(\left(\hat{I}(\varphi^k_{i\Delta},\varphi^{(i\Delta,\varphi^k_{i\Delta})}_{(i+1)\Delta}) - \sigma^2(a)\right)\left(\hat{I}(\varphi^{k'}_{i\Delta},\varphi^{(i\Delta,\varphi^{k'}_{i\Delta})}_{(i+1)\Delta}) - \sigma^2(a)\right)\middle| \mathcal{F}_{i\Delta}\right).$$

By lemma 3.3.6.i we have $E\left(\hat{I}(\varphi^k_{i\Delta},\varphi^{(i\Delta,\varphi^k_{i\Delta})}_{(i+1)\Delta})\right) = \mathcal{O}\left(1\right)$ for $\Delta \downarrow 0$ and therefore

$$\begin{aligned} (\text{Ia}) &= E\left(\left(\hat{I}(\varphi^k_{i\Delta},\varphi^{(i\Delta,\varphi^k_{i\Delta})}_{(i+1)\Delta}) - \sigma^2(a)\right)^2\right) \\ &= E\left((\hat{I}(\varphi^k_{i\Delta},\varphi^{(i\Delta,\varphi^k_{i\Delta})}_{(i+1)\Delta}))^2\right) + \mathcal{O}\left(1\right) \quad \text{for } \Delta \downarrow 0. \end{aligned}$$

Moreover lemma 3.3.6.ii gives us

$$E\left((\hat{I}(\varphi^k_{i\Delta},\varphi^{(i\Delta,\varphi^k_{i\Delta})}_{(i+1)\Delta}))^2\right) = \mathcal{O}\left(1\right) \quad \text{for } \Delta \downarrow 0$$

and consequently

$$(\mathrm{Ia}) = \mathcal{O}(1) \qquad \text{for } \Delta \downarrow 0.$$

Now we turn to (Ib): Since $\varphi^{(i\Delta,\varphi^k_{i\Delta})}$ and $\varphi^{(i\Delta,\varphi^{k'}_{i\Delta})}$ are independent processes for $k \neq k'$ we get with (39) immediately

$$\begin{aligned}
& \mathbb{1}_{B_{h_\Delta}(a)}(\varphi^k_{i\Delta})\ \mathbb{1}_{B_{h_\Delta}(a)}(\varphi^{k'}_{i\Delta}) \times (\mathrm{Ib}) \\
= \ & \mathbb{1}_{B_{h_\Delta}(a)}(\varphi^k_{i\Delta})\ \mathbb{1}_{B_{h_\Delta}(a)}(\varphi^{k'}_{i\Delta})\ E\left(\hat{I}(\varphi^k_{i\Delta}, \varphi^{(i\Delta,\varphi^k_{i\Delta})}_{(i+1)\Delta}) - \sigma^2(a)\right) \\
& \times E\left(\hat{I}(\varphi^{k'}_{i\Delta}, \varphi^{(i\Delta,\varphi^{k'}_{i\Delta})}_{(i+1)\Delta}) - \sigma^2(a)\right) \\
= \ & \mathbb{1}_{B_{h_\Delta}(a)}(\varphi^k_{i\Delta})\ \mathbb{1}_{B_{h_\Delta}(a)}(\varphi^{k'}_{i\Delta}) \cdot \mathcal{O}\left((\sqrt{\Delta} + h_\Delta)^2\right) \qquad \text{for } \Delta \downarrow 0.
\end{aligned}$$

Putting together the results on (Ia) and (Ib) we get the asymptotics for (I):

$$\begin{aligned}
(\mathrm{I}) \ = \ & \frac{\Delta^2}{4h^2_\Delta} \sum_{i=0}^{[T/\Delta]-1} E_m\left(\varphi_{i\Delta}(B_{h_\Delta}(a))\right) \cdot \mathcal{O}(1) \\
& + \frac{\Delta^2}{4h^2_\Delta} \sum_{i=0}^{[T/\Delta]-1} E_m\left(\sum_{k\neq k'} \mathbb{1}_{B_{h_\Delta}(a)}(\varphi^k_{i\Delta})\mathbb{1}_{B_{h_\Delta}(a)}(\varphi^{k'}_{i\Delta})\right) \cdot \mathcal{O}\left((\sqrt{\Delta} + h_\Delta)^2\right) \\
\leq \ & h^{-1}_\Delta \overline{m}(B_{h_\Delta}(a)) \cdot \mathcal{O}\left(\Delta h^{-1}_\Delta\right) + \int l(x)^2\ m(dx) \cdot \mathcal{O}\left(\Delta h^{-2}_\Delta(\sqrt{\Delta} + h_\Delta)^2\right) \\
= \ & \mathcal{O}\left(\Delta h^{-1}_\Delta\right) \qquad (40)
\end{aligned}$$

for $\Delta \downarrow 0$.
In the last step we used the fact that under assumption A6.iv the invariant occupation measure $\overline{m}$ has a continuous and bounded density with respect to the Lebesgue measure on $\mathbb{R}$ (which implies in particular that $h^{-1}_\Delta \overline{m}(B_{h_\Delta}(a)) = \mathcal{O}(1)$ for $\Delta \downarrow 0$) and that $\int l(x)^2\ m(dx) < \infty$ by assumption A6.i).

It remains to study the asymptotics of the second term (II).
Using (39) twice we directly get

$$(\mathrm{II}) \ = \ \frac{\Delta^2}{h^2_\Delta} \sum_{i\neq j} E_m\Big(\varphi_{i\Delta}(B_{h_\Delta}(a))\varphi_{j\Delta}(B_{h_\Delta}(a))\Big) \cdot \mathcal{O}\left((\sqrt{\Delta} + h_\Delta)^2\right) \qquad \text{for } \Delta \downarrow 0.$$

Now for $i < j$ by lemma 2.3.2[12]

$$\begin{aligned} & E_m\Big(\varphi_{i\Delta}(B_{h_\Delta}(a))\varphi_{j\Delta}(B_{h_\Delta}(a))\Big) \\ = & \int m(dx)x(B_{h_\Delta}(a))E_x\Big(\varphi_{(j-i)\Delta}(B_{h_\Delta}(a))\Big) \\ \leq & \int m(dx)x(B_{h_\Delta}(a))(l(x)+cT) \\ = & \int \overline{m}(dy)\mathbb{1}_{B_{h_\Delta}(a)}(y)\int_S l(x)\ m^p(y,dx) + cT\cdot\overline{m}(B_{h_\Delta}(a)). \end{aligned}$$

By assumption A6.ii $\int_S l(x)\ m^p(\cdot,dx)$ is locally uniformly bounded, hence

$$E_m\Big(\varphi_{i\Delta}(B_{h_\Delta}(a))\varphi_{j\Delta}(B_{h_\Delta}(a))\Big) = \mathcal{O}\left(h_\Delta\right) \quad \text{for } \Delta\downarrow 0.$$

Combining previous two results we get the asymptotics for the second term (II):

$$\text{(II)} = \mathcal{O}\left(h_\Delta^{-1}(\sqrt{\Delta}+h_\Delta)^2\right) = \mathcal{O}\left(\Delta h_\Delta^{-1}\right) \quad \text{for } \Delta\downarrow 0. \tag{41}$$

Finally by (40) and (41)

$$E_m\left(\left[\overline{S}_T^\Delta(a,\varphi)-\sigma^2(a)\overline{L}_T^\Delta(a)\right]^2\right) = \text{(I)}+\text{(II)} = \mathcal{O}\left(\Delta h_\Delta^{-1}\right)$$

for $\Delta\downarrow 0$ which gives us the assertion. □

3.3.3 Partial estimator under knowledge of the pedigree

Now we still assume that we are able to observe the whole path of φ in $[0,T]$ as in section 3.3.2, but we will consider an estimator which uses only the 'good configurations' (contained in $D_{2\Delta^\lambda}$) in order to get closer to the original problem where we observe only positions of particles.

We define an estimator for $\sigma^2(a)$ by setting

$$\hat{L}_T^\Delta(a) := \frac{\Delta}{2h_\Delta}\sum_{i=0}^{[T/\Delta]-1}\mathbb{1}_{D_{2\Delta^\lambda}}(\varphi_{i\Delta})\ \varphi_{i\Delta}\left(B_{h_\Delta}(a)\right),$$

$$\hat{S}_T^\Delta(a,\varphi) := \frac{\Delta}{2h_\Delta}\sum_{i=0}^{[T/\Delta]-1}\mathbb{1}_{D_{2\Delta^\lambda}}(\varphi_{i\Delta})\sum_{k=1}^{l(\varphi_{i\Delta})}\mathbb{1}_{B_{h_\Delta}(a)}(\varphi_{i\Delta}^k)\hat{I}(\varphi_{i\Delta}^k,\varphi_{(i+1)\Delta}^{(i\Delta,\varphi_{i\Delta}^k)})$$

[12] Apply lemma 2.3.2 on the subprocesses of φ which start with one single particle. Note that in expectation we have cT immigration events during the time interval $[0,T]$.

and finally for arbitrary $\varepsilon > 0$

$$\hat{\sigma}^2_\Delta(a,\varphi) := \frac{\hat{S}^\Delta_T(a,\varphi)}{\hat{L}^\Delta_T(a)} \mathbb{1}_{\left\{\hat{L}^\Delta_T(a)>\varepsilon\right\}}. \tag{42}$$

To prove consistency of $\hat{\sigma}^2_\Delta(a,\varphi)$ we need the following assumption.

Assumption A7. The Palm kernel m^p fulfills

$$\lim_{\Delta\to 0} \sup_{y\in A} m^p(y, S_{2\Delta^\lambda}) = 0$$

for any compact set $A \subset \mathbb{R}$.

Remark 3.3.8. Assumption A7 claims that a typical configuration of the stationary process φ conditioned on the event that one component equals a fixed position $y \in \mathbb{R}$ has a.s. no components which occupy the same position in space. Since we already know that $m(S_{2\Delta^\lambda})$ converges to zero for $\Delta \downarrow 0$ (see remark 2.4.3.b) this assumption does not seem to be very strong.

Lemma 3.3.9. *Let assumptions A6 and A7 hold.*
If $\Delta h_\Delta^{-2}\log\Delta \to 0$ then $\hat{\sigma}^2_\Delta(a,\varphi)$ is a consistent estimator for $\sigma^2(a)$.

Proof. Thanks to theorem 3.3.4 we are done if we can show the convergence of $\overline{\sigma}^2_\Delta(a,\varphi) - \hat{\sigma}^2_\Delta(a,\varphi)$ to zero in Q_m-probability.

Consider the decomposition

$$\overline{\sigma}^2_\Delta(a,\varphi) - \hat{\sigma}^2_\Delta(a,\varphi) = \left(\frac{\overline{S}^\Delta_T(a,\varphi)}{\overline{L}^\Delta_T(a)} - \frac{\hat{S}^\Delta_T(a,\varphi)}{\hat{L}^\Delta_T(a)}\right)\cdot \mathbb{1}_{\left\{\hat{L}^\Delta_T(a)>\varepsilon\right\}}$$

$$\begin{aligned} &= \frac{\overline{S}^\Delta_T(a,\varphi) - \hat{S}^\Delta_T(a,\varphi)}{\overline{L}^\Delta_T(a)}\cdot \mathbb{1}_{\left\{\hat{L}^\Delta_T(a)>\varepsilon\right\}} + \frac{\hat{S}^\Delta_T(a,\varphi)}{\hat{L}^\Delta_T(a)}\cdot\frac{\hat{L}^\Delta_T(a) - \overline{L}^\Delta_T(a)}{\overline{L}^\Delta_T(a)}\cdot \mathbb{1}_{\left\{\hat{L}^\Delta_T(a)>\varepsilon\right\}} \\ &=: \quad (\mathrm{I}) \quad + \quad (\mathrm{II}). \end{aligned}$$

In order to prove our assertion it suffices to show the convergence in Q_m-probability of (I) and (II).
We will show $\mathbb{L}^1$ convergence of both expressions.

a) We start with the second term (II).
For $\hat{S}^\Delta_T(a,\varphi)$ consider its compensator

$$\hat{\hat{S}}^\Delta_T(a,\varphi) := \frac{\Delta}{2h_\Delta}\sum_{i=0}^{[T/\Delta]-1} \mathbb{1}_{D_{2\Delta^\lambda}}(\varphi_{i\Delta}) \sum_{k=1}^{(\varphi_{i\Delta})} \mathbb{1}_{B_{h_\Delta}(a)}(\varphi^k_{i\Delta}) E\left(\hat{I}^\Delta_T(\varphi^k_{i\Delta}, \varphi^{(i\Delta,\varphi^k_{i\Delta})}_{(i+1)\Delta})\right).$$

Then trivially

$$E_m\left(\hat{S}_T^\Delta(a,\varphi) - \hat{\hat{S}}_T^\Delta(a,\varphi)\right) = 0 \tag{43}$$

and by lemma 3.3.6.i

$$\hat{\hat{S}}_T^\Delta(a,\varphi) \leq \sup_{y\in B_{h_\Delta}(a)} E\left(\hat{I}(y,\varphi_\Delta^{(0,y)})\right)\cdot \hat{L}_T^\Delta(a) = \mathcal{O}\left(1\right)\cdot \hat{L}_T^\Delta(a) \qquad \text{for } \Delta\downarrow 0.$$

Thus, minding that $\overline{L}_T^\Delta(a) \geq \hat{L}_T^\Delta(a) \geq 0$ and $\hat{S}_T^\Delta(a,\varphi)\geq 0$,

$$\begin{aligned}
& E_m\left(\left|\frac{\hat{S}_T^\Delta(a,\varphi)}{\hat{L}_T^\Delta(a)}\cdot\frac{\hat{L}_T^\Delta(a)-\overline{L}_T^\Delta(a)}{\overline{L}_T^\Delta(a)}\cdot \mathbb{1}_{\left\{\hat{L}_T^\Delta(a)>\varepsilon\right\}}\right|\right) \\
= \; & E_m\left(\frac{\hat{\hat{S}}_T^\Delta(a,\varphi)}{\hat{L}_T^\Delta(a)}\cdot\frac{\overline{L}_T^\Delta(a)-\hat{L}_T^\Delta(a)}{\overline{L}_T^\Delta(a)}\cdot \mathbb{1}_{\left\{\hat{L}_T^\Delta(a)>\varepsilon\right\}}\right) \\
& + E_m\left(\frac{\hat{S}_T^\Delta(a,\varphi)-\hat{\hat{S}}_T^\Delta(a,\varphi)}{\hat{L}_T^\Delta(a)}\cdot\frac{\overline{L}_T^\Delta(a)-\hat{L}_T^\Delta(a)}{\overline{L}_T^\Delta(a)}\cdot \mathbb{1}_{\left\{\hat{L}_T^\Delta(a)>\varepsilon\right\}}\right) \\
\leq \; & E_m\left(\frac{\hat{\hat{S}}_T^\Delta(a,\varphi)}{\hat{L}_T^\Delta(a)}\cdot\frac{\overline{L}_T^\Delta(a)-\hat{L}_T^\Delta(a)}{\varepsilon}\right) + E_m\left(\frac{\hat{S}_T^\Delta(a,\varphi)-\hat{\hat{S}}_T^\Delta(a,\varphi)}{\varepsilon}\cdot 1\right) \\
= \; & \mathcal{O}\left(1\right)\cdot E_m\left(\overline{L}_T^\Delta(a)-\hat{L}_T^\Delta(a)\right) + \varepsilon^{-1}\cdot \underbrace{E_m\left(\hat{S}_T^\Delta(a,\varphi)-\hat{\hat{S}}_T^\Delta(a,\varphi)\right)}_{=0 \text{ by } (43)} \\
= \; & \mathcal{O}\left(1\right)\cdot E_m\left(\overline{L}_T^\Delta(a)-\hat{L}_T^\Delta(a)\right)
\end{aligned}$$

for $\Delta\downarrow 0$.

Furthermore by definition of $\overline{L}_T^\Delta(a)$ and $\hat{L}_T^\Delta(a)$ we obtain

$$\begin{aligned}
& E_m\left(\overline{L}_T^\Delta(a)-\hat{L}_T^\Delta(a)\right) \\
= \; & E_m\left(\frac{\Delta}{h_\Delta}\sum_{i=0}^{[T/\Delta]-1} \mathbb{1}_{S_{2\Delta^\lambda}}(\varphi_{i\Delta})\ \varphi_{i\Delta}(B_{h_\Delta}(a))\right) \\
= \; & \Delta[T/\Delta]\int \mathbb{1}_{S_{2\Delta^\lambda}}(x)\frac{1}{h_\Delta}x(B_{h_\Delta}(a))\ m(dx) \\
= \; & \Delta[T/\Delta]\int_{\mathbb{R}} \frac{1}{2h_\Delta}\mathbb{1}_{B_{h_\Delta}(a)}(y)\ m^p(y,S_{2\Delta^\lambda})\ \overline{m}(dy) \\
\leq \; & \frac{T}{2h_\Delta}\overline{m}(B_{h_\Delta}(a))\sup_{y\in B_{h_\Delta}(a)} m^p(y,S_{2\Delta^\lambda}),
\end{aligned} \tag{44}$$

where m^p is the Palm kernel corresponding to the invariant measure m on S.

Now $\overline{m}$ has a continuous and bounded Lebesgue density (assumption A6.iv), hence $(2h_\Delta)^{-1}\overline{m}(B_{h_\Delta}(a)) \to \frac{d\overline{m}}{d\lambda}(a) < \infty$ for $\Delta \downarrow 0$ and by assumption A7 the $\mathbb{L}^1$-convergence of the second term (II) follows immediately.

b) It remains to show the $\mathbb{L}^1$-convergence of the first term (I).
By lemma 3.3.6.i

$$\begin{aligned} & E_m\left(\left|\frac{\overline{S}_T^\Delta(a,\varphi) - \hat{S}_T^\Delta(a,\varphi)}{\overline{L}_T^\Delta(a)} \cdot \mathbb{1}_{\{\hat{L}_T^\Delta(a)>\varepsilon\}}\right|\right) \\ \leq\ & \varepsilon^{-1} \cdot \sup_{y\in B_{h_\Delta}(a)} E\left(\hat{I}(y,\varphi_\Delta^{(0,y)})\right) \cdot E_m\left(\overline{L}_T^\Delta(a) - \hat{L}_T^\Delta(a)\right) \\ =\ & \mathcal{O}(1) \cdot E_m\left(\overline{L}_T^\Delta(a) - \hat{L}_T^\Delta(a)\right) \qquad \text{for } \Delta \downarrow 0 \end{aligned}$$

and the assertion follows immediately with our previous considerations (see (44)).
□

3.3.4 Estimator using the partial reconstruction rule

Now we consider the case, where we are only able to observe the positions of the particles, but not their pedigree, i.e. we are in the situation of our original problem (see section 2.1).

Theorem 3.3.10. *Let assumptions A6 and A7 hold. If $\Delta h_\Delta^{-2}\log\Delta \to 0$ then $\hat{\sigma}_\Delta^2(a,\beta)$ is a consistent estimator for $\sigma^2(a)$.*

Proof. Since $\hat{\sigma}^2(a,\varphi)$ is a consistent estimator for $\sigma^2(a)$ by lemma 3.3.9 it suffices to show that $\hat{\sigma}^2(a,\varphi) - \hat{\sigma}^2(a,\beta)$ converges in Q_m-probability to zero or equivalently we will show that

$$\hat{S}_T^\Delta(a,\varphi) - \hat{S}_T^\Delta(a,\beta) \xrightarrow{Q_m} 0 \text{ for } \Delta \downarrow 0.$$

By definition

$$\begin{aligned} & E_m\left|\hat{S}_T^\Delta(a,\varphi) - \hat{S}_T^\Delta(a,\beta)\right| \\ =\ & E_m\Bigg|\frac{\Delta}{2h_\Delta}\sum_{i=0}^{[T/\Delta]-1} \mathbb{1}_{D_{2\Delta^\lambda}}(\beta_{i\Delta}) \sum_{k=1}^{l(\beta_{i\Delta})} \mathbb{1}_{B_{h_\Delta}(a)}(\beta_{i\Delta}^k) \\ & \times\left(\hat{I}(\beta_{i\Delta}^k, \varphi_{(i+1)\Delta}^{(i\Delta,\beta_{i\Delta}^k)}) - \hat{I}(\beta_{i\Delta}^k, \beta_{(i+1)\Delta}^{[i\Delta,\beta_{i\Delta}^k]})\right)\Bigg| \end{aligned}$$

which is less equal to

$$\frac{\Delta}{2h_\Delta}\sum_{i=0}^{[T/\Delta]-1} E_m\Bigg(\mathbb{1}_{D_{2\Delta^\lambda}}(\beta_{i\Delta})\sum_{k=1}^{l(\beta_{i\Delta})}\mathbb{1}_{B_{h_\Delta}(a)}(\beta^k_{i\Delta})$$
$$\times\left(\hat{I}(\beta^k_{i\Delta},\varphi^{(i\Delta,\beta^k_{i\Delta})}_{(i+1)\Delta})+\hat{I}(\beta^k_{i\Delta},\beta^{[i\Delta,\beta^k_{i\Delta}]}_{(i+1)\Delta})\right)\mathbb{1}_{\left\{\beta^{[i\Delta,\beta^k_{i\Delta}]}_{(i+1)\Delta}\neq_p\varphi^{(i\Delta,\beta^k_{i\Delta})}_{(i+1)\Delta}\right\}}\Bigg).$$

By definition of $\beta^{[i\Delta,\beta^k_{i\Delta}]}_{(i+1)\Delta}$ in 2.4.1

$$\hat{I}(\beta^k_{i\Delta},\beta^{[i\Delta,\beta^k_{i\Delta}]}_{(i+1)\Delta})\le\Delta^{2\lambda-1}$$

and by lemma 3.3.4.i

$$E_m\left(\hat{I}(\beta^k_{i\Delta},\varphi^{(i\Delta,\beta^k_{i\Delta})}_{(i+1)\Delta})\Big|\,\mathcal{F}_{i\Delta}\right)=\mathcal{O}\left(1\right)\qquad\text{for }\Delta\downarrow 0,$$

thus

$$\begin{aligned}&E_m\left|\hat{S}^\Delta_T(a,\varphi)-\hat{S}^\Delta_T(a,\beta)\right|\\ =\;&\mathcal{O}\left(\Delta^{2\lambda-1}\right)\frac{\Delta}{2h_\Delta}\sum_{i=0}^{[T/\Delta]-1}E_m\Bigg[\mathbb{1}_{D_{2\Delta^\lambda}}(\beta_{i\Delta})\sum_{k=1}^{l(\beta_{i\Delta})}\mathbb{1}_{B_{h_\Delta}(a)}(\beta^k_{i\Delta})\\ &\times Q_m\left(\beta^{[i\Delta,\beta^k_{i\Delta}]}_{(i+1)\Delta}\neq_p\varphi^{(i\Delta,\beta^k_{i\Delta})}_{(i+1)\Delta}\Big|\,\mathcal{F}_{i\Delta}\right)\Bigg]\end{aligned}$$

for $\Delta\downarrow 0$.

Now by theorem 2.4.4.i

$$\mathbb{1}_{D_{2\Delta^\lambda}}(\beta_{i\Delta})Q_m\left(\beta^{[i\Delta,\beta^k_{i\Delta}]}_{(i+1)\Delta}\neq_p\varphi^{(i\Delta,\beta^k_{i\Delta})}_{(i+1)\Delta}\Big|\,\mathcal{F}_{i\Delta}\right)\le\mathbb{1}_{D_{2\Delta^\lambda}}(\beta_{i\Delta})\left[c\cdot\Delta+l(\beta_{i\Delta})g_\lambda(\Delta)\right]$$

and therefore

$$\begin{aligned}&E_m\left|\hat{S}^\Delta_T(a,\varphi)-\hat{S}^\Delta_T(a,\beta)\right|\\ =\;&\mathcal{O}\left(\Delta^{2\lambda-1}\right)\frac{\Delta}{2h_\Delta}\sum_{i=0}^{[T/\Delta]-1}E_m\left(\mathbb{1}_{D_{2\Delta^\lambda}}(\beta_{i\Delta})\beta_{i\Delta}(B_{h_\Delta}(a))\left[c\cdot\Delta+l(\beta_{i\Delta})g_\lambda(\Delta)\right]\right)\\ =\;&\mathcal{O}\left(\Delta^{2\lambda}\right)E_m\left(\hat{L}^\Delta_T(a)\right)+\mathcal{O}\left(\Delta^{2\lambda}+\Delta^{2\lambda-1}g_\lambda(\Delta)h^{-1}_\Delta\right)\int_S l^2(x)\;m(dx)\end{aligned}$$

for $\Delta\downarrow 0$.

Since $E_m\left(\hat{L}^\Delta_T(a)\right)=\Delta\cdot[T/\Delta]\cdot\frac{1}{2h_\Delta}\overline{m}(B_{h_\Delta}(a))$ converges to $T\cdot\frac{dm}{d\mathbb{A}}(a)$, which is a finite value by assumption A6.iv, and since $\int_S l^2(x)\;m(dx)<\infty$ by assumption A6.i the lemma follows with above asymptotics. □

3.4 Rate of convergence of the estimator

Thanks to our results in section 3.3 most of the work is already done. In theorem 3.3.4 we have already calculated the rate of convergence for the estimator $\overline{\sigma}^2_\Delta(a,\varphi)$ and it remains to show that the estimator $\hat{\sigma}^2_\Delta(a,\beta)$, which uses the partial reconstruction rule, inherits the rate of $\overline{\sigma}^2_\Delta(a,\varphi)$.

Nevertheless we need to pose an additional assumption which corresponds to assumptions A6.iv and A7 in section 3.3.

Consider the second order Campbell measure m^{cc} associated to the invariant measure m on S, i.e.

$$m^{cc}(A \times B \times F) := \int_S x(A)x(B)\mathbb{1}_F(x)\ m(dx) \qquad \text{for } A, B \in \mathcal{B}(\mathbb{R}), F \in \mathcal{S}.$$

Desintegration of m^{cc} yields

$$m^{cc}(dy, dy', dx) = \overline{\overline{m}}(dy, dy')m^{pp}(y, y', dx)$$

where $\overline{\overline{m}}$ is the second order measure defined by

$$\overline{\overline{m}}(A \times B) := \int_S x(A)x(B)\ m(dx)$$

and m^{pp} is the Palm kernel of order 2 corresponding to m^{cc}.

Assumption A8. We assume that for any $a \in \mathbb{R}$

$$h_\Delta^{-1} m^{cc}(B_{h_\Delta}(a) \times B_{h_\Delta}(a) \times S_{2\Delta^\lambda}) = \mathcal{O}\left(1\right) \qquad \text{for } \Delta \downarrow 0.$$

Remark 3.4.1. Note that assumption A8 is fulfilled if $\overline{\overline{m}}$ has a continuous and bounded Lebesgue density and if for any compact sets $A, B \in \mathcal{B}(\mathbb{R})$

$$\sup_{(y,y')\in A\times B} m^{pp}(y, y', S_{2\Delta^\lambda}) = \mathcal{O}\left(1\right) \qquad \text{for } \Delta \downarrow 0.$$

We are now able to state the main theorem of section 3.

Theorem 3.4.2. *Assume A6 and A8 and let* $\lambda \in (\frac{1}{4}, \frac{1}{2})$*. Then the estimator* $\hat{\sigma}^2_\Delta(a,\beta)$ *has the rate of convergence* $\sqrt{\Delta h_\Delta^{-1}}$ *for* $\Delta \downarrow 0$*, i.e.*

$$E\left(\left[\sqrt{\frac{h_\Delta}{\Delta}}\left(\hat{\sigma}^2_\Delta(a,\beta) - \sigma^2(a)\right)\right]^2 \middle| \hat{L}^\Delta_T(a) > \varepsilon\right) = \mathcal{O}\left(1\right) \qquad \textit{for } \Delta \downarrow 0.$$

Remark 3.4.3. With our methods we could not achieve a central limit theorem. The missing link is the exact limit of $E\left(\left[\hat{I}(y,\varphi_\Delta^{(0,y)})\right]^2\right)$ for $\Delta \downarrow 0$.
In lemma 3.3.6.ii we could only show that above term is asymptotically bounded for $\Delta \downarrow 0$. As a main technique we used the representation of the semi-group of $l(\varphi)$ presented in lemma 2.3.2. To obtain an exact limit we would need a corresponding result for $l^2(\varphi)$ which requires completely different techniques and considerations.

Proof. Thanks to theorem 3.3.4 the assertion follows if we prove the following two assertions:

$$\text{i)}\quad E_m\left(\left[\sqrt{\frac{h_\Delta}{\Delta}}\left(\hat{L}_T^\Delta(a)-\overline{L}_T^\Delta(a)\right)\right]^2\right) = \mathcal{O}(1) \qquad \text{for } \Delta \downarrow 0$$

and

$$\text{ii)}\quad E_m\left(\left[\sqrt{\frac{h_\Delta}{\Delta}}\left(\hat{S}_T^\Delta(a,\beta)-\overline{S}_T^\Delta(a,\varphi)\right)\right]^2\right) = \mathcal{O}(1) \qquad \text{for } \Delta \downarrow 0.$$

The first assertion follows directly by assumption A8 since we have

$$\begin{aligned}
& E_m\left(\left|\sqrt{\frac{h_\Delta}{\Delta}}\left(\hat{L}_T^\Delta(a)-\overline{L}_T^\Delta(a)\right)\right|^2\right) \\
= \quad & \frac{\Delta}{h_\Delta}E_m\left(\sum_{i=0}^{[T/\Delta]-1}\mathbb{1}_{S_{2\Delta^\lambda}}(\varphi_{i\Delta})\left[\varphi_{i\Delta}(B_{h_\Delta}(a))\right]^2\right) \\
= \quad & h_\Delta^{-1}\int_{S_{2\Delta^\lambda}} x^2(B_{h_\Delta}(a))\; m(dx) \\
= \quad & h_\Delta^{-1}m^{cc}(B_{h_\Delta}(a)\times B_{h_\Delta}(a)\times S_{2\Delta^\lambda}).
\end{aligned}$$

Now we turn to the second assertion ii.

$$\begin{aligned}
& E_m\left(\left|\sqrt{\frac{h_\Delta}{\Delta}}\left(\hat{S}_T^\Delta(a,\beta)-\overline{S}_T^\Delta(a,\varphi)\right)\right|^2\right) \\
\leq \quad & E_m\left(\left|\sqrt{\frac{h_\Delta}{\Delta}}\left(\hat{S}_T^\Delta(a,\beta)-\hat{S}_T^\Delta(a,\varphi)\right)\right|^2\right) \\
& + E_m\left(\left|\sqrt{\frac{h_\Delta}{\Delta}}\left(\hat{S}_T^\Delta(a,\varphi)-\overline{S}_T^\Delta(a,\varphi)\right)\right|^2\right) \\
=: \quad & \text{(I)} \quad + \quad \text{(II)}.
\end{aligned}$$

We first consider term (I). Since by definition $\mathbb{1}_{B_{h_\Delta}(a)}(x^k)\mathbb{1}_{B_{h_\Delta}(a)}(x^{k'}) = 0$ for all $x \in D_{2\Delta^\lambda}$ and $k \neq k'$ we have

$$\begin{aligned}
(\mathrm{I}) &= \frac{\Delta}{h_\Delta} E_m \left(\sum_{i=0}^{[T/\Delta]-1} \mathbb{1}_{D_{2\Delta^\lambda}}(\varphi_{i\Delta}) \left[\sum_{k=1}^{l(\varphi_{i\Delta})} \mathbb{1}_{B_{h_\Delta}(a)}(\beta^k_{i\Delta}) \times (\mathrm{Ia}) \right]^2 \right) \\
&= \frac{\Delta}{h_\Delta} E_m \left(\sum_{i=0}^{[T/\Delta]-1} \mathbb{1}_{D_{2\Delta^\lambda}}(\varphi_{i\Delta}) \sum_{k=1}^{l(\varphi_{i\Delta})} \mathbb{1}_{B_{h_\Delta}(a)}(\beta^k_{i\Delta}) \times (\mathrm{Ia})^2 \right),
\end{aligned}$$

where

$$(\mathrm{Ia}) := \hat{I}(\beta^k_{i\Delta}, \beta^{[i\Delta,\beta^k_{i\Delta}]}_{(i+1)\Delta}) - \hat{I}(\beta^k_{i\Delta}, \varphi^{(i\Delta,\beta^k_{i\Delta})}_{(i+1)\Delta}).$$

Now

$$\begin{aligned}
E\left((\mathrm{Ia})^2 \middle| \mathcal{F}_{i\Delta} \right) &= E\left((\mathrm{Ia})^2 \mathbb{1}_{\left\{ \beta^{[i\Delta,\beta^k_{i\Delta}]}_{(i+1)\Delta} \neq_p \varphi^{(i\Delta,\beta^k_{i\Delta})}_{(i+1)\Delta} \right\}} \right) \\
&\leq E\left(\left[\hat{I}(\beta^k_{i\Delta}, \beta^{[i\Delta,\beta^k_{i\Delta}]}_{(i+1)\Delta}) \right]^2 \mathbb{1}_{\left\{ \beta^{[i\Delta,\beta^k_{i\Delta}]}_{(i+1)\Delta} \neq_p \varphi^{(i\Delta,\beta^k_{i\Delta})}_{(i+1)\Delta} \right\}} \right) \\
&\quad + E\left(\left[\hat{I}(\beta^k_{i\Delta}, \varphi^{(i\Delta,\beta^k_{i\Delta})}_{(i+1)\Delta}) \right]^2 \right)
\end{aligned}$$

and since $\hat{I}(\beta^k_{i\Delta}, \beta^{[i\Delta,\beta^k_{i\Delta}]}_{(i+1)\Delta}) \leq \Delta^{2\lambda-1}$ by definition of $\beta^{[i\Delta,\beta^k_{i\Delta}]}_{(i+1)\Delta}$ we conclude with theorem 2.4.4.i that for $\beta_{i\Delta} \in D_{2\Delta^\lambda}$

$$E\left(\left[\hat{I}(\beta^k_{i\Delta}, \beta^{[i\Delta,\beta^k_{i\Delta}]}_{(i+1)\Delta}) \right]^2 \mathbb{1}_{\left\{ \beta^{[i\Delta,\beta^k_{i\Delta}]}_{(i+1)\Delta} \neq_p \varphi^{(i\Delta,\beta^k_{i\Delta})}_{(i+1)\Delta} \right\}} \right) \leq \Delta^{4\lambda-2}(c\Delta + l(\beta_{i\Delta}) g_\lambda(\Delta)).$$

Moreover by lemma 3.3.6.ii

$$E\left(\left[\hat{I}(\beta^k_{i\Delta}, \varphi^{(i\Delta,\beta^k_{i\Delta})}_{(i+1)\Delta}) \right]^2 \right) = \mathcal{O}(1) \quad \text{for } \Delta \downarrow 0.$$

Hence

$$\begin{aligned}
&(\mathrm{I}) \\
&= \frac{\Delta}{h_\Delta} E_m \left(\sum_{i=0}^{[T/\Delta]-1} \mathbb{1}_{D_{2\Delta^\lambda}}(\varphi_{i\Delta})\ \varphi_{i\Delta}\left(B_{h_\Delta}(a)\right) \right) \cdot \mathcal{O}\left(\Delta^{2\lambda} + \Delta^{2\lambda-1} g_\lambda(\Delta) l(\varphi_{i\Delta}) \right) \\
&= h_\Delta^{-1} \overline{m}(B_{h_\Delta}(a)) \mathcal{O}\left(\Delta^{4\lambda-1}\right) + \mathcal{O}\left(h_\Delta^{-1} \Delta^{4\lambda-2} g_\lambda(\Delta) \right) \int_S l(x)^2\ m(dx) \\
&= \mathcal{O}(1) \quad \text{for } \Delta \downarrow 0.
\end{aligned}$$

It remains to study term (II).

$$
\begin{aligned}
\text{(II)} &= \frac{\Delta}{h_\Delta} E_m \left(\sum_{i=0}^{[T/\Delta]-1} \mathbb{1}_{S_{2\Delta^\lambda}}(\varphi_{i\Delta}) \left[\sum_{k=1}^{l(\varphi_{i\Delta})} \mathbb{1}_{B_{h_\Delta}(a)}(\varphi_{i\Delta}^k) \hat{I}(\varphi_{i\Delta}^k, \varphi_{(i+1)\Delta}^{(i\Delta,\varphi_{i\Delta}^k)}) \right]^2 \right) \\
&= \frac{\Delta}{h_\Delta} E_m \left(\sum_{i=0}^{[T/\Delta]-1} \mathbb{1}_{S_{2\Delta^\lambda}}(\varphi_{i\Delta}) \sum_{k=1}^{l(\varphi_{i\Delta})} \mathbb{1}_{B_{h_\Delta}(a)}(\varphi_{i\Delta}^k) \left[\hat{I}_k \right]^2 \right) \\
&\quad + \frac{\Delta}{h_\Delta} E_m \left(\sum_{i=0}^{[T/\Delta]-1} \mathbb{1}_{S_{2\Delta^\lambda}}(\varphi_{i\Delta}) \sum_{k \neq k'} \mathbb{1}_{B_{h_\Delta}(a)}(\varphi_{i\Delta}^k) \mathbb{1}_{B_{h_\Delta}(a)}(\varphi_{i\Delta}^{k'}) \hat{I}_k \hat{I}_{k'} \right),
\end{aligned}
$$

where

$$\hat{I}_k := \hat{I}(\varphi_{i\Delta}^k, \varphi_{(i+1)\Delta}^{(i\Delta,\varphi_{i\Delta}^k)}).$$

Now by lemma 3.3.6.ii $E\left(\left[\hat{I}_k \right]^2 \middle| \mathcal{F}_{i\Delta} \right) = \mathcal{O}(1)$ for $\Delta \downarrow 0$ and by the independence of $\hat{I}_k$ and $\hat{I}_{k'}$ we have with lemma 3.3.6.i

$$E\left(\hat{I}_k \hat{I}_{k'} \middle| \mathcal{F}_{i\Delta} \right) = E\left(\hat{I}_k \middle| \mathcal{F}_{i\Delta} \right) E\left(\hat{I}_{k'} \middle| \mathcal{F}_{i\Delta} \right) = \mathcal{O}(1)$$

for $\Delta \downarrow 0$.

Hence

$$
\begin{aligned}
\text{(II)} &= \frac{\Delta}{h_\Delta} E_m \left(\sum_{i=0}^{[T/\Delta]-1} \mathbb{1}_{S_{2\Delta^\lambda}}(\varphi_{i\Delta}) \left[\varphi_{i\Delta}(B_{h_\Delta}(a)) \right]^2 \right) \cdot \mathcal{O}(1) \\
&= h_\Delta^{-1} \int_{S_{2\Delta^\lambda}} x^2(B_{h_\Delta}(a)) \, m(dx) \cdot \mathcal{O}(1) \qquad \text{for } \Delta \downarrow 0
\end{aligned}
$$

and by assumption A8 we finally get (II)$= \mathcal{O}(1)$ for $\Delta \downarrow 0$.

All together we have shown the assertion and the proof is finished. □

3.5 Simulations

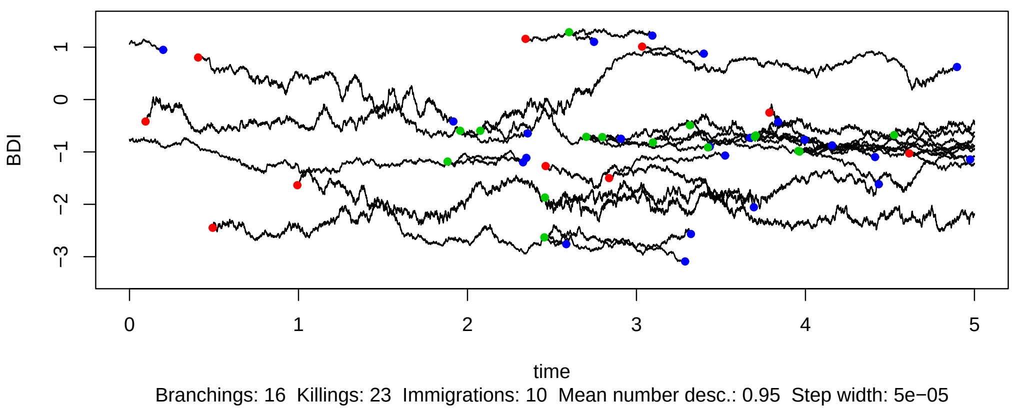
Now we want to apply our estimator $\hat{\sigma}^2_\Delta(\cdot,\beta)$ to the example considered already in the simulations of section 2.5. The computation of the estimator was done with the R-package 'BDI' (see [Bra04] for a link to a download of the package).

Here are the parameters for the simulation:

Drift function	$b(x)$	$\equiv$	0
Diffusion coefficient	$\sigma^2(x)$	$=$	$\frac{1}{4}\cos(\pi\cdot x)+\frac{1}{2}$
Killing rate	$\kappa(x)$	$\equiv$	1
Immigration rate	c	$=$	1
Reproduction law	$p_k(x)$	$\equiv$	$\frac{(1.2)^k}{k!}e^{-1.2}, \quad k=0,2,3,\ldots$
Simulation step width	sw	$=$	0.00005
Time horizon	T	$=$	5

In figure 16 we plotted again the full trajectory of our simulation. Red dots denote immigrations, green dots branching events with two or more descendants and blue dot denote killings (branching events where no particle is born).

Figure 16: Full trajectory of the BDI



In the following figures 17 to 32 we calculated the estimator for different values of Δ and λ (we have chosen $\lambda\in\{0.25,0.33,0.40,0.49\}$ and $\Delta\in\{0.01,0.001\}$). In the figures with odd numbers we plotted the reconstructed trajectories of the BDI φ and the figures with even numbers show the estimator evaluated at points $a_i=0.1\cdot i,\ i=-30,\ldots,10$. Additionally we plotted the discrete local time $\hat{L}^\Delta_5(a_i)$ $(i=-30,\ldots,10)$ normalized to area 1 in the upper half of the coordinate system.

Note that we used for the reconstruction of the trajectories and for the estimation of the diffusion coefficient not only observations $\beta_{i\Delta} \in D_{2\Delta^\lambda}$ (these are the configuration where no component has a neighbor closer than $2\Delta^\lambda$) but also the 'good' components $\beta_{i\Delta}^k \in G(\beta_{i\Delta})$ which have no neighbor in a $2\Delta^\lambda$-neighborhood (see (25) for the definition of $G(\beta_{i\Delta})$ and remarks 2.4.2.c and 3.2.2.b for further details on the finer versions of the reconstruction rule, respectively the estimator). Furthermore note that we have chosen the bandwidth $h_\Delta = \Delta^\lambda$ for all examples since this choice of band width maximizes the number of observations which can be used for estimation by the construction of the estimator.

We can observe different effects if we compare the results for different values of Δ and λ.

First have a look at figures 20, 24 and 28 where we calculated the estimator for $\lambda = 0.25, 0.33, 0.40$ in the case $\Delta = 0.001$. We can see that in all three cases the estimator produces good approximation of the true diffusion coefficient. Thus at a first glance our estimator works well for a reasonable set of parameters and the estimator seems to be robust with respect to the choice of λ.

We now want to examine the effects of the choice of λ in a more detailed way. If we turn our attention to figures 18, 22 and 26, where we calculated the estimator for $\lambda = 0.25, 0.33, 0.40$ in the case of a bigger step width $\Delta = 0.01$, we notice that these plots differ slightly more than the plots in the case $\Delta = 0.001$. It seems that the results get better for increasing values of λ. The reason for that is easy to see. For $\lambda = 0.25$ and $\lambda = 0.33$ we could reconstruct only 19.5%, respectively 27.79%, of the trajectory and in the case $\lambda = 0.40$ we are able to reconstruct 33.71% of the trajectory. Hence bigger values of λ imply a bigger amount of observations which can be used for estimation and thus better approximations of the unknown diffusion coefficient.

But now we have a look at figures 30 and 32. Here we have chosen $\lambda = 0.49$ which is already very close to the upper bound 0.5 for λ. We can see in figure 30 that the estimator seems to underestimate the bigger values of σ^2. This effect stabilizes for the finer step width $\Delta = 0.001$ (see figure 32). The reason for that phenomenon can be explained if we have a look on the quota of observations which left their Δ^λ-neighborhoods. For $\lambda = 0.49$ and $\Delta = 0.001$ there are 3.17% of observations which left their bell. By the construction of the estimator these observations are not used for estimation and therefore the bigger increments are systematicly neglected. It is clear that the estimator cannot work properly if the choice of used observations depends strongly on the parameter we want to estimate (in this case the diffusion coefficient which determines the grade of fluctuation of the particles). We conclude that the estimator is sensible with respect to the accuracy of the reconstruction of the trajectory.

There is also another effect we can observe if we study the results for the smaller step width $\Delta = 0.01$. By construction of the estimator the quality of the estimation improves with the amount of time the process stayed in the region where we want to estimate the diffusion coefficient. Hence the discrete local time is a good indicator for the quality of the estimation. In figure 22 we can see that bigger values of σ^2 are harder to estimate than smaller values (on the local maxima of σ^2 we have less discrete local time cumulated). But we can see that in our example we have also in the 'valleys' of σ^2 only few local time accumulated (especially at position -1 this effect is evident). This happens since particles tend to stay more time in regions where the diffusion coefficient has a local minima and thus the density of particles is much higher than in other regions. In particular there are more observations with intersecting Δ^λ-neighborhoods and thus less observations which can be used for reconstruction, respectively estimation.

Now how shall we choose λ in practice? We are not able to compute the accuracy of the reconstruction from our data and hence we do not know which choice of λ is an optimal choice. But as we have seen from our example we can actually see when the underestimation effect starts. Thus in practice we would start with a small λ and increase λ in small steps. When the underestimation effect becomes visible we stop and a slightly smaller value of λ than the last value of λ should be a good choice.

Simulations suggest that a good choice for the parameter λ are values between 0.33 and 0.40.

Figure 17: Reconstruction for $\Delta = 0.01$ and $\lambda = 0.25$

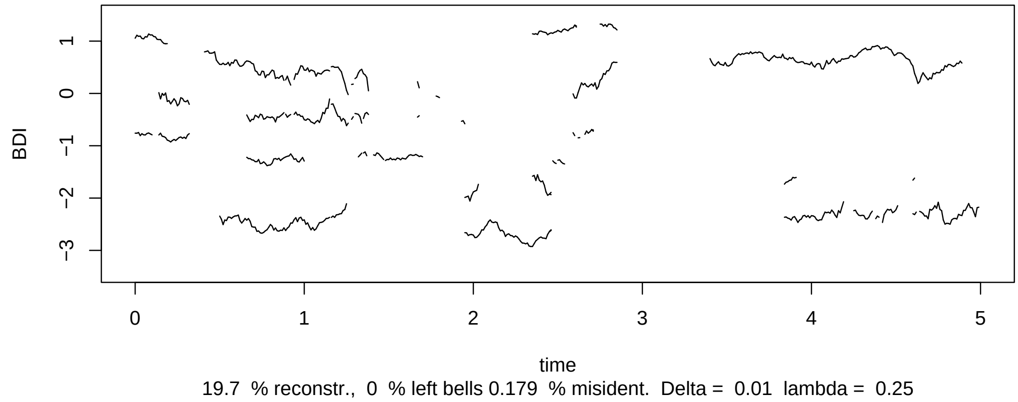


Figure 18: Estimator for σ^2 ($\Delta = 0.01$, $\lambda = 0.25$, $h_\Delta = \Delta^\lambda$)

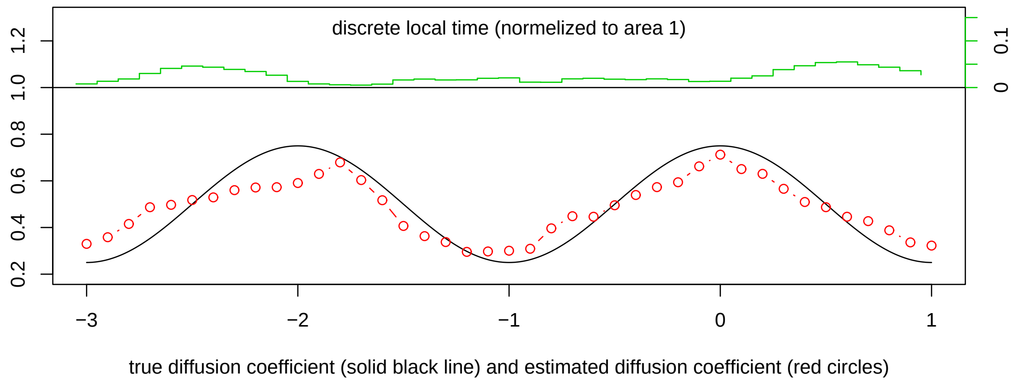


Figure 19: Reconstruction for $\Delta = 0.001$ and $\lambda = 0.25$

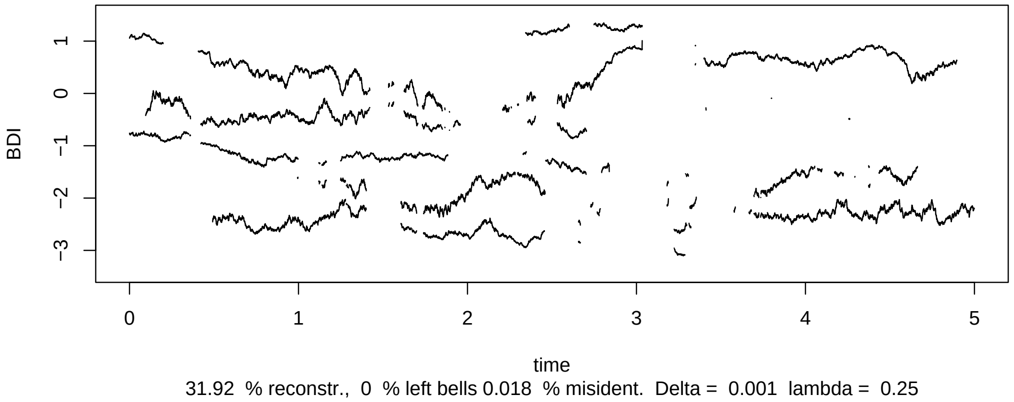


Figure 20: Estimator for σ^2 ($\Delta = 0.001$, $\lambda = 0.25$, $h_\Delta = \Delta^\lambda$)

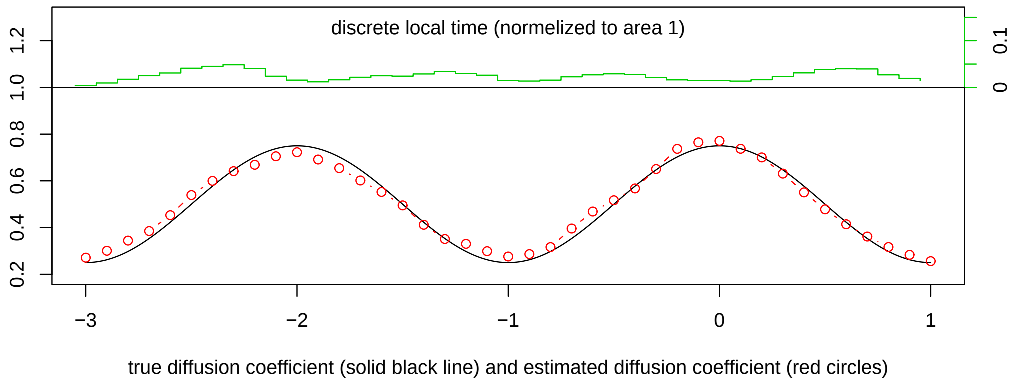


Figure 21: Reconstruction for $\Delta = 0.01$ and $\lambda = 0.33$

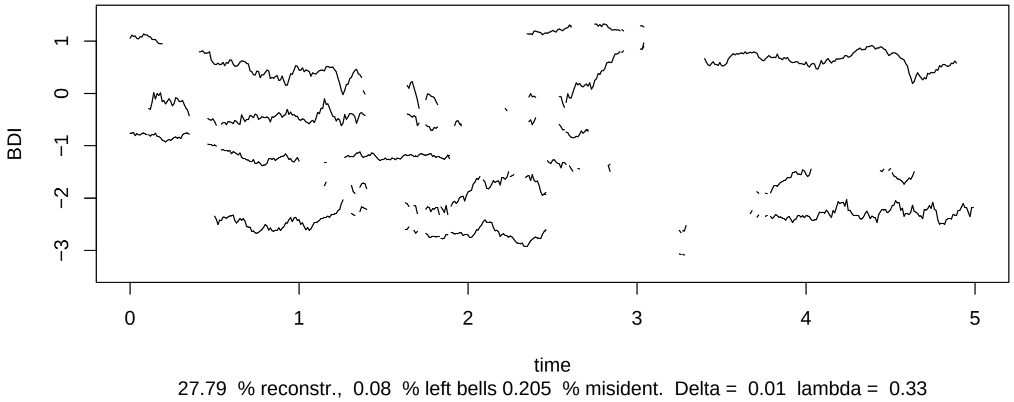


Figure 22: Estimator for σ^2 ($\Delta = 0.01$, $\lambda = 0.33$, $h_\Delta = \Delta^\lambda$)

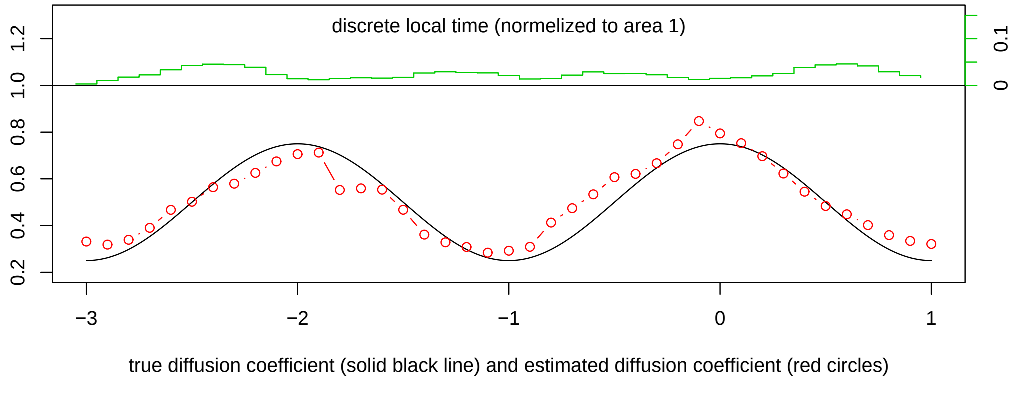


Figure 23: Reconstruction for $\Delta = 0.001$ and $\lambda = 0.33$

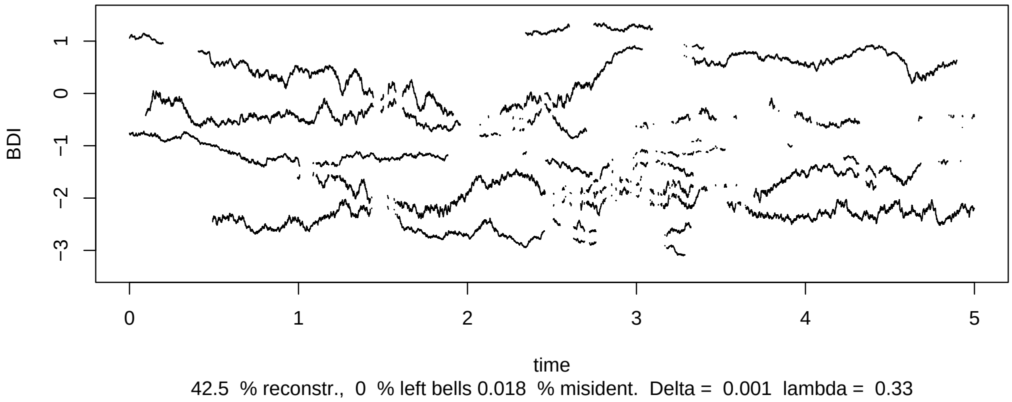


Figure 24: Estimator for σ^2 ($\Delta = 0.001$, $\lambda = 0.33$, $h_\Delta = \Delta^\lambda$)

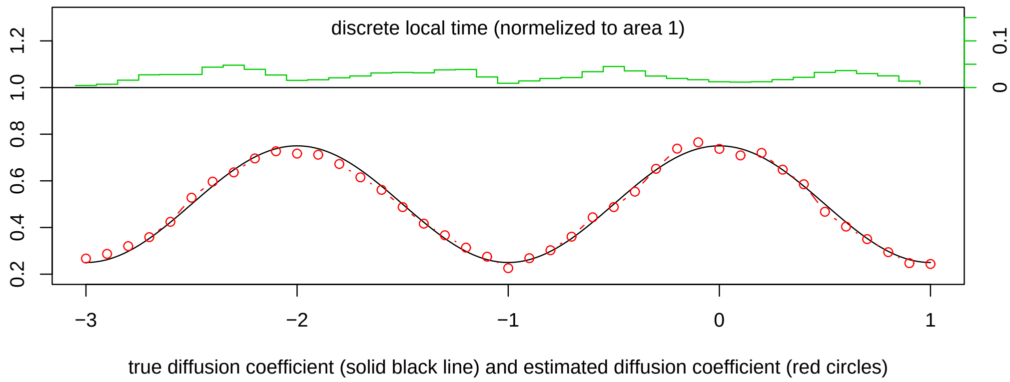


Figure 25: Reconstruction for $\Delta = 0.01$ and $\lambda = 0.40$

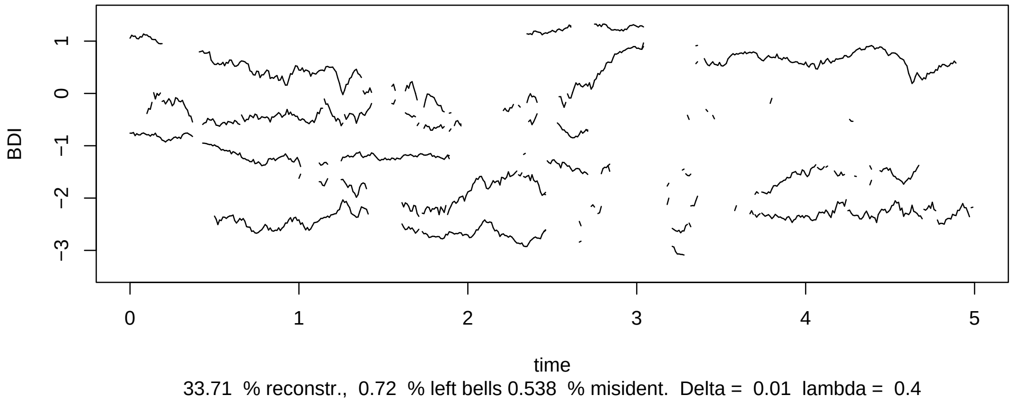


Figure 26: Estimator for σ^2 ($\Delta = 0.01$, $\lambda = 0.40$, $h_\Delta = \Delta^\lambda$)

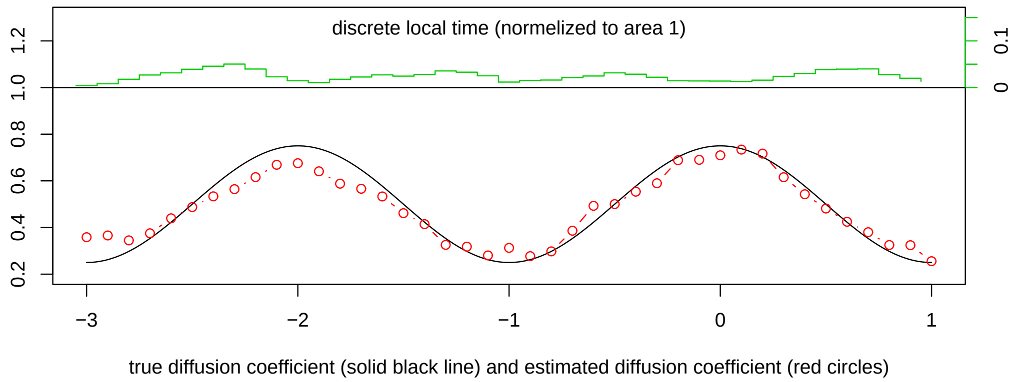


Figure 27: Reconstruction for $\Delta = 0.001$ and $\lambda = 0.40$

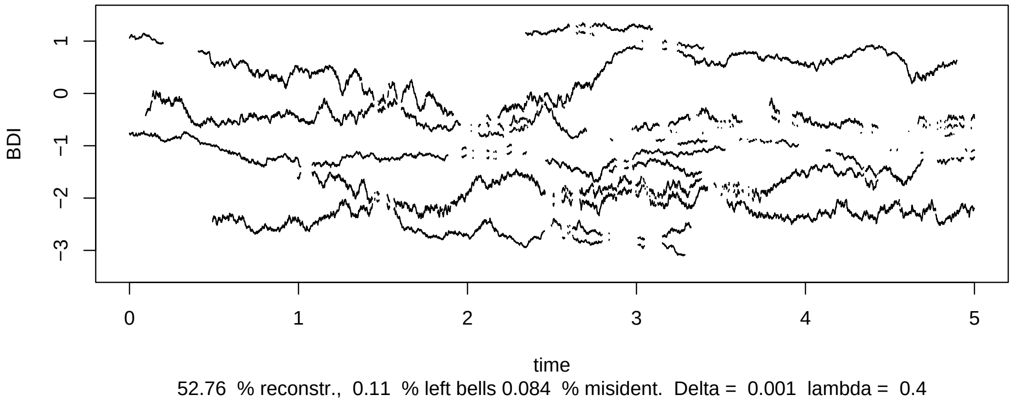


Figure 28: Estimator for σ^2 ($\Delta = 0.001$, $\lambda = 0.40$, $h_\Delta = \Delta^\lambda$)

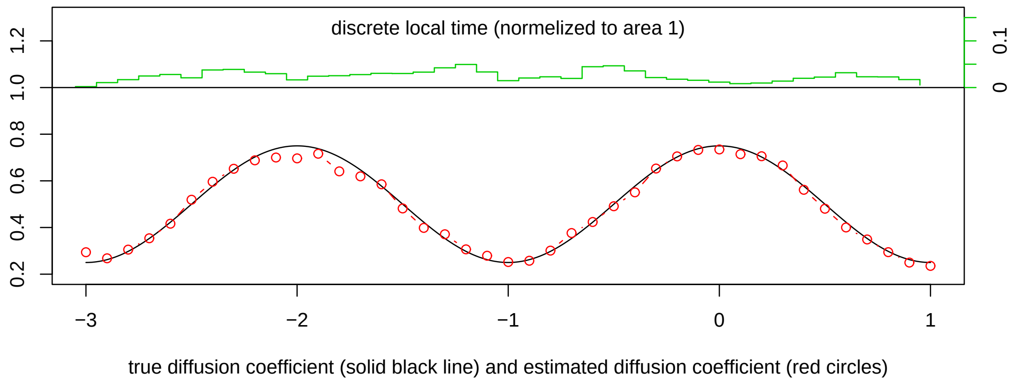


Figure 29: Reconstruction for $\Delta = 0.01$ and $\lambda = 0.49$

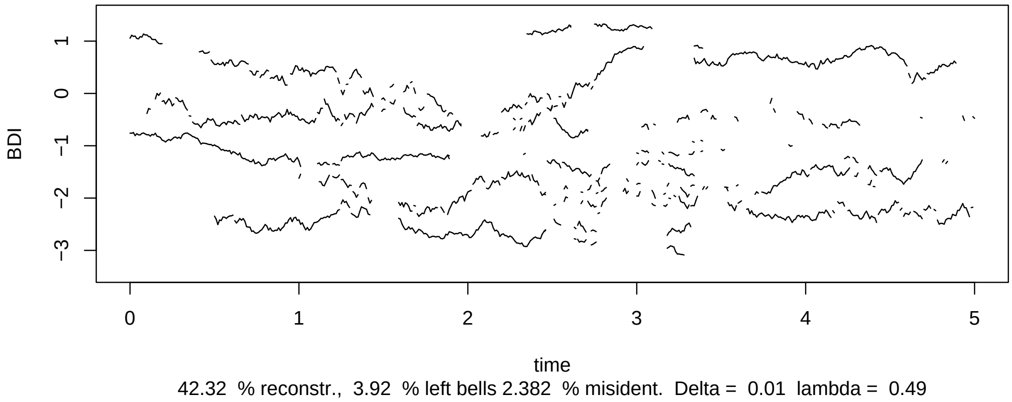


Figure 30: Estimator for σ^2 ($\Delta = 0.01$, $\lambda = 0.49$, $h_\Delta = \Delta^\lambda$)

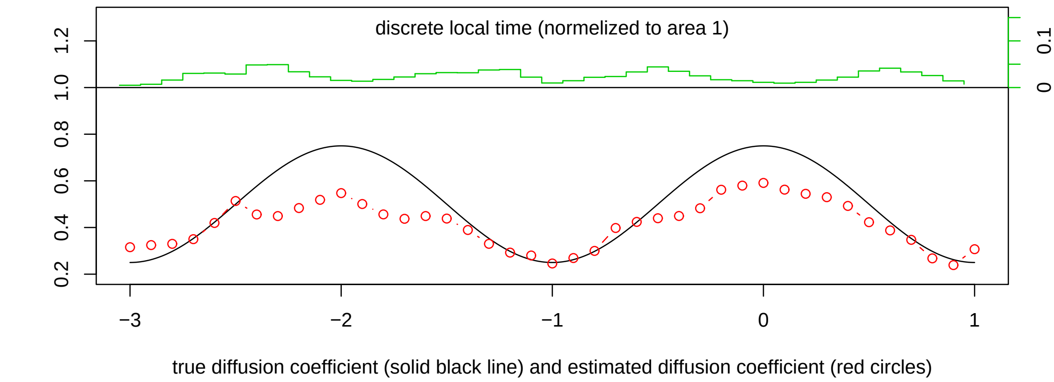


Figure 31: Reconstruction for $\Delta = 0.001$ and $\lambda = 0.49$

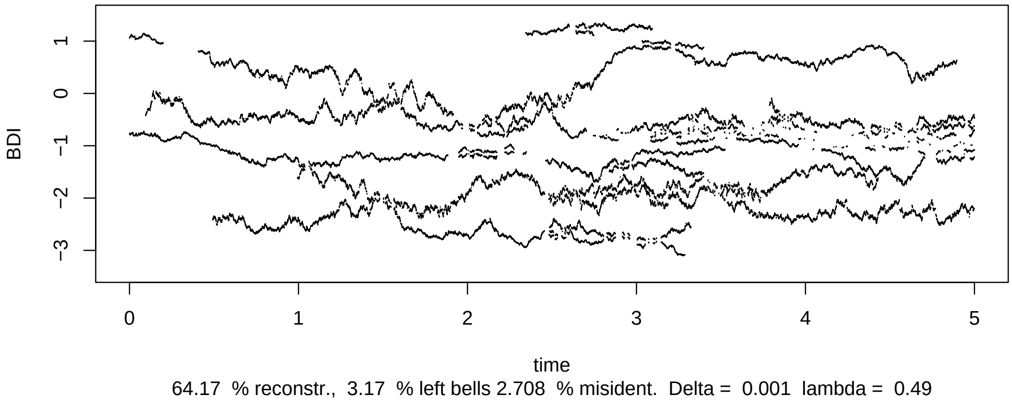
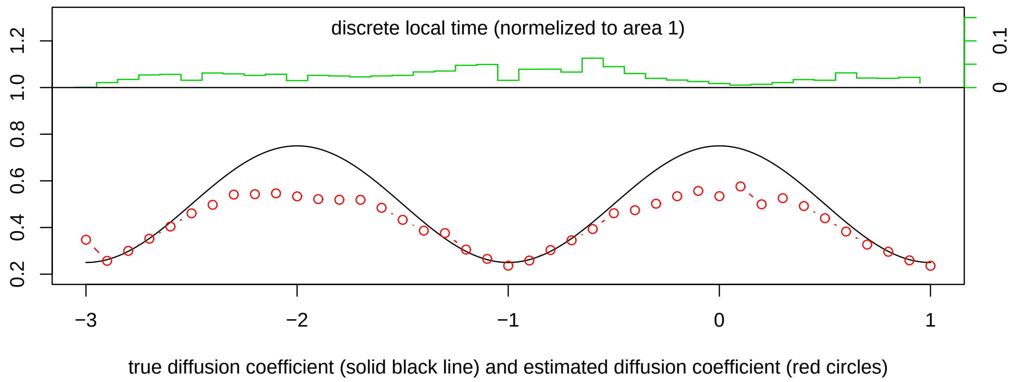


Figure 32: Estimator for σ^2 ($\Delta = 0.001$, $\lambda = 0.49$, $h_\Delta = \Delta^\lambda$)

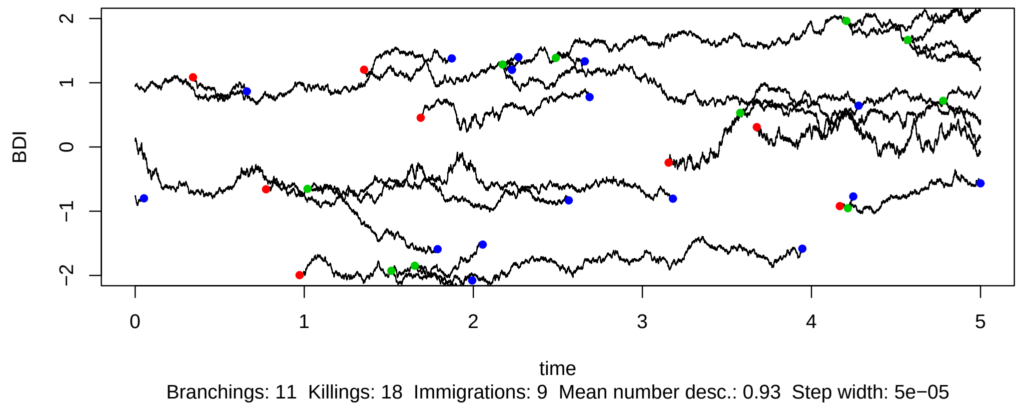


Finally we present another simulation, where we have chosen a step function as diffusion coefficient (all other parameters are the same as in the preceeding example). More precisely we have set

$$\sigma^2(y) := 0.35 \cdot \mathbb{1}_{[-0.5,0.5]}(y) + 0.4, \quad y \in \mathbb{R}.$$

In figure 33 we plotted the simulated trajectory. As before red dots denote immigrations, green dots branchings with two ore more descendants and blue dots denote killings.

Figure 33: Full trajectory of the BDI



Figures 34 and 36 contain the (partially) reconstructed trajectories for $\Delta = 0.001$ and $\lambda = 0.33$, respectively $\lambda = 0.40$. In figures 35 and 37 the estimators for the diffusion coefficient were plotted.

As we can see, the estimator works as good as in the preceeding example, even if the diffusion function is a non-smooth function, which stresses again the robustness of the estimator.

Figure 34: Reconstruction for $\Delta = 0.001$ and $\lambda = 0.33$

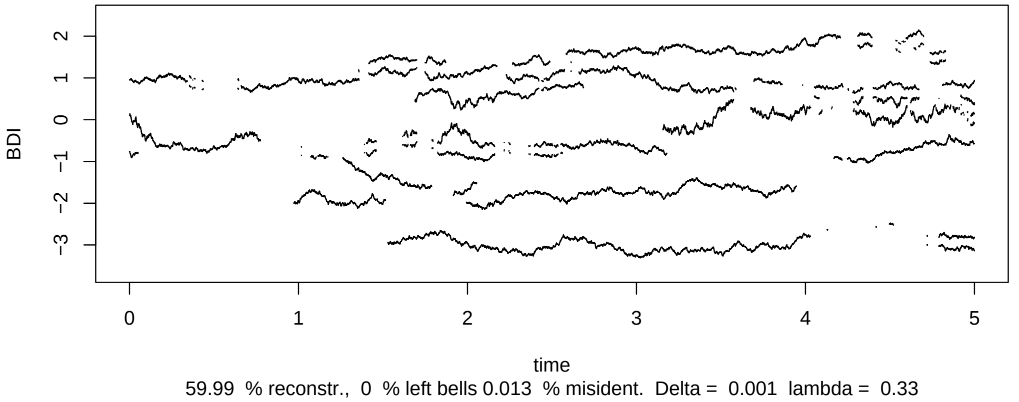


Figure 35: Estimator for σ^2 ($\Delta = 0.001$, $\lambda = 0.33$, $h_\Delta = \Delta^\lambda$)

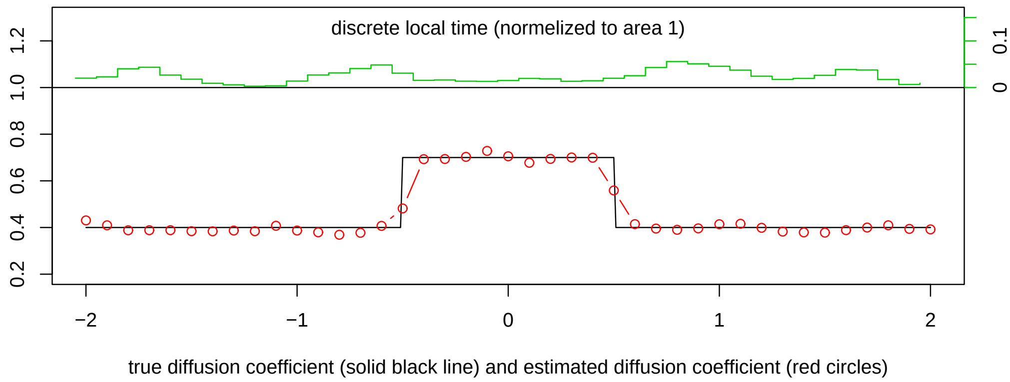


Figure 36: Reconstruction for $\Delta = 0.001$ and $\lambda = 0.40$

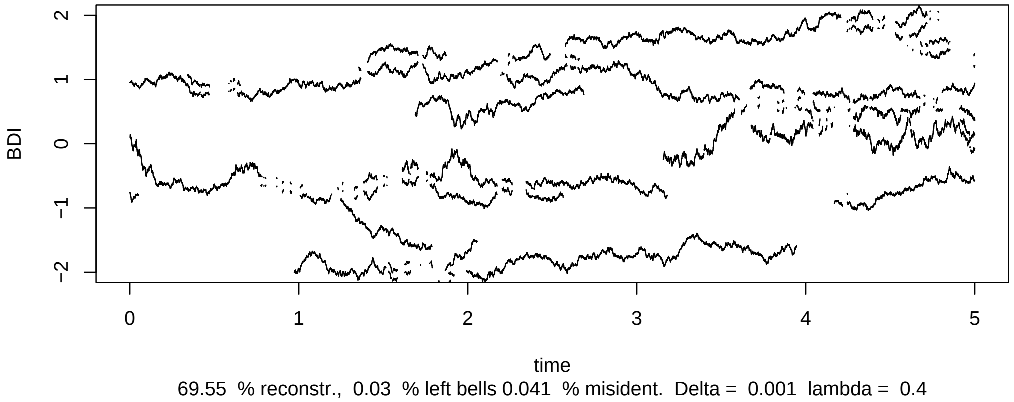
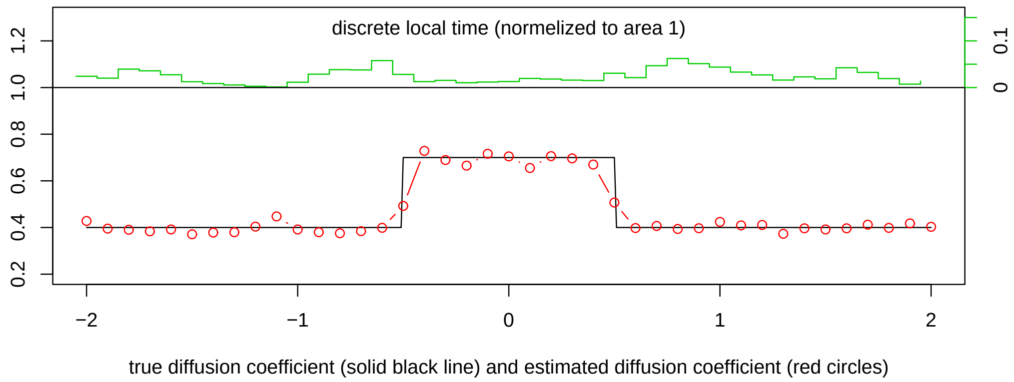


Figure 37: Estimator for σ^2 ($\Delta = 0.001$, $\lambda = 0.40$, $h_\Delta = \Delta^\lambda$)



List of figures

References

[AN04] K. B. Athreya and P. E. Ney. *Branching processes.* Dover Publications Inc., Mineola, NY, 2004. Reprint of the 1972 original [Springer, New York].

[Bas98] R. F. Bass. *Diffusions and Elliptic Operators.* Springer, New York, 1998.

[Bil68] P. Billingsley. *Statistical Inference for Markov Processes.* University of Chicago Press, 1968.

[Bra04] C. Brandt. *BDI, an R-package providing simulations of branching diffusions with immigration.* http://www.informatik.uni-mainz.de/deutsch/institute/arbeitsgr/stochastik/staff/brandt/brandt.html, 2004.

[Bré81] P. Brémaud. *Point processes and queues.* Springer-Verlag, New York, 1981. Martingale dynamics, Springer Series in Statistics.

[DCFZ86] D. Dacunha-Castelle and D. Florens-Zmirou. Estimation of the coefficients of a diffusion from discrete observations. *Stochastics*, 19:263–284, 1986.

[DM82] C. Dellacherie and P.-A. Meyer. *Probabilities and potential. B*, volume 72 of *North-Holland Mathematics Studies.* North-Holland Publishing Co., Amsterdam, 1982. Theory of martingales, Translated from the French by J. P. Wilson.

[Doh87] G. Dohnal. On estimating the diffusion coefficient. *J. Appl. Probab.*, 24:105–114, 1987.

[DVJ03] D.J. Daley and D. Vere-Jones. *An introduction to the theory of point processes. Vol. I: Elementary theory and methods. 2nd ed.* Probability and Its Applications. New York, NY: Springer. xxi, 469 p., 2003.

[DvZ01] K. Dzhaparidze and J. H. van Zanten. On Bernstein-type inequalities for martingales. *Stochastic Process. Appl.*, 93(1):109–117, 2001.

[Eis81] Th. Eisele. Diffusions avec branchement et intéraction. *Ann. Sci. Univ. Clermont-Ferrand II Math.*, (19):21–34, 1981. Saint-Flour Probability Summer Schools (Saint-Flour, 1979/1980).

[Eth93] A. M. Etheridge. Asymptotic behaviour of measure-valued critical branching processes. *Proc. Amer. Math. Soc.*, 118(4):1251–1261, 1993.

[FZ93] D. Florens-Zmirou. On estimating the diffusion coefficient from discrete observations. *J. Appl. Probab.*, 30(4):790–804, 1993.

[GCJ93] V. Genon-Catalot and J. Jacod. On the estimation of the diffusion coefficient for multi-dimensional diffusion processes. *Ann. Inst. H. Poincaré Probab. Statist.*, 29(1):119–151, 1993.

[GCJ94] V. Genon-Catalot and J. Jacod. Estimation of the diffusion coefficient for diffusion processes: Random sampling. *Scand. J. Stat.*, 21(3):193–221, 1994.

[GW94] L. G. Gorostiza and A. Wakolbinger. Long time behavior of critical branching particle systems and applications. In *Measure-valued processes, stochastic partial differential equations, and interacting systems (Montreal, PQ, 1992)*, volume 5 of *CRM Proc. Lecture Notes*, pages 119–137. Amer. Math. Soc., Providence, RI, 1994.

[HÖ4] R. Höpfner. Strange shape of invariant density in branching diffusions with immigration. *Preprint Nr. 05/04, Fachbereich Mathematik und Informatik, Universität Mainz*, 2004.

[Har02] Th. E. Harris. *The theory of branching processes*. Dover Phoenix Editions. Dover Publications Inc., Mineola, NY, 2002. Corrected reprint of the 1963 original [Springer, Berlin].

[Has80] R.Z. Has'minskiĭ. *Stochastic Stability of Differetial Equations*. Sijthoff & Noordhoff, 1980.

[HHL02] R. Höpfner, M. Hoffmann, and E. Löcherbach. Non-parametric estimation of the death rate in branching diffusions. *Scand. J. Statist.*, 29(4):665–692, 2002.

[HJ94a] J. Hoffmann-Jørgensen. *Probability with a view toward statistics. Vol 1: Probabilities.* London: Chapman & Hall. xl, 589 p. , 1994.

[HJ94b] J. Hoffmann-Jørgensen. *Probability with a view toward statistics. Vol. 2: Mathematical statistics.* London: Chapman & Hall. xii, 533 p. , 1994.

[HL99a] R. Höpfner and E. Löcherbach. On invariant measure for branching diffusions. *Unpublished note; see http://www.mathematik.uni-mainz.de/ hoepfner*, 1999.

[HL99b] R. Höpfner and E. Löcherbach. On local asymptotic normality for birth and death on a flow. *Stochastic Process. Appl.*, 83(1):61–77, 1999.

[HL99c] R. Höpfner and E. Löcherbach. Statistical models for birth and death on a flow: local absolute continuity and likelihood ratio processes. *Scand. J. Statist.*, 26(1):107–128, 1999.

[HL03a] R. Höpfner and E. Löcherbach. Limit theorems for null recurrent Markov processes. *Mem. Amer. Math. Soc.*, 161(768):vi+92, 2003.

[HL03b] R. Höpfner and E. Löcherbach. Remarks on ergodicity and invariant occupation measure in branching diffusions with immigration. *(To appear in Ann. Inst. H. Poincar.) Preprint Nr. 12/03, Fachbereich Mathematik und Informatik, Universität Mainz*, 2003.

[Hof99] M. Hoffmann. L_p estimation of the diffusion coefficient. *Bernoulli*, 5(3):447–481, 1999.

[Hof01] M. Hoffmann. On estimating the diffusion coefficient: Parametric versus nonparametric. *Ann. Inst. H. Poincaré Probab. Statist.*, 37(3):339–372, 2001.

[IH81] I. A. Ibragimov and R. Z. Has'minskiĭ. *Statistical Estimation: Asymptotic Theory.* Springer-Verlag, New York, 1981.

[INW68a] N. Ikeda, M. Nagasawa, and S. Watanabe. Branching Markov processes. I. *J. Math. Kyoto Univ.*, 8:233–278, 1968.

[INW68b] N. Ikeda, M. Nagasawa, and S. Watanabe. Branching Markov processes. II. *J. Math. Kyoto Univ.*, 8:365–410, 1968.

[INW69] N. Ikeda, M. Nagasawa, and S. Watanabe. Branching Markov processes. III. *J. Math. Kyoto Univ.*, 9:95–160, 1969.

[IT84] Y. Iwasa and E. Teramoto. Branching-diffusion model for the formation of distributional patterns in populations. *J. Math. Biol.*, 19(1):109–124, 1984.

[Iva80] B. G. Ivanoff. The branching diffusion with immigration. *J. Appl. Probab.*, 17(1):1–15, 1980.

[IW89] N. Ikeda and S. Watanabe. *Stochastic differential equations and diffusion processes.* North-Holland Publ. Co., second edition, 1989.

[Jac98] J. Jacod. Rates of convergence to the local time of a diffusion. *Ann. Inst. Henri Poincare, Probab. Stat.*, 34(4):505–544, 1998.

[Jac00] J. Jacod. Non-parametric Kernel Estimation of the Coefficient of a Diffusion. *Scandinavian Journal of Statistics*, 27:83–96, 2000.

[Jac01] J. Jacod. Inference for stochastic processes. *Prépublication n° 683, Laboratoire de probabilités et modelès aléatoire, Universités de Paris 6 & Paris 7*, 2001.

[Jag75] P. Jagers. *Branching processes with biological applications.* Wiley-Intersci., London, 1975.

[JS87] J. Jacod and A.N. Shiryaev. *Limit theorems for stochastic processes.* Springer, Berlin Heidelberg New York, 1987.

[Kal02] O. Kallenberg. *Foundations of modern probability.* Probability and its Applications (New York). Springer-Verlag, New York, second edition, 2002.

[Kes97] M. Kessler. Estimation of an ergodic diffusion from discrete observations. *Scand. J. Stat.*, 24(2):211–229, 1997.

[Kes00] M. Kessler. Simple and explicit estimating functions for a discretely observed diffusion process. *Scand. J. Stat.*, 27(1):65–82, 2000.

[KP92] P.E. Kloeden and E. Platen. *Numerical Solutions of Stochastic Differential Equations.* Springer, 1992.

[KP02] Mathieu Kessler and Silvestre Paredes. Computational aspects related to martingale estimating functions for a discretely observed diffusion. *Scand. J. Statist.*, 29(3):425–440, 2002.

[KS88] I. Karatzas and S.E. Shreve. *Brownian Motion and Stochastic Calculus.* Springer-Verlag, 1988.

[KS99] M. Kessler and M. Sørensen. Estimating equations based on eigenfunctions for a discretely observed diffusion process. *Bernoulli*, 5(2):299–314, 1999.

[Lö99] E. Löcherbach. *Statistical models and likelihood ratio processes for interacting particle systems with branching and immigration.* PhD thesis, Paderborn: Univ. Paderborn, Fachbereich Mathematik/ Informatik, 100 p., 1999.

[Löc00] E. Löcherbach. Likelihood ratio processes and asymptotic statistics for systems of interacting diffusions with branching and immigration. In *Mathematics and computer science (Versailles, 2000)*, Trends Math., pages 265–274. Birkhäuser, Basel, 2000.

[Löc02a] E. Löcherbach. LAN and LAMN for systems of interacting diffusions with branching and immigration. *Ann. Inst. H. Poincaré Probab. Statist.*, 38(1):59–90, 2002.

[Löc02b] E. Löcherbach. Likelihood ratio processes for Markovian particle systems with killing and jumps. *Stat. Inference Stoch. Process.*, 5(2):153–177, 2002.

[Löc04] E. Löcherbach. Smoothness of the intensity measure density for interacting branching diffusions with immigrations. *J. Funct. Anal.*, 215(1):130–177, 2004.

[LS78] R. S. Liptser and A.N. Shiryayev. *Statistics of Random Processes I and II.* Springer-Verlag, 1977 and 1978.

[Pak75] A. G. Pakes. On Markov branching processes with immigration. *Sankhyā Ser. A*, 37(1):129–138, 1975.

[Rev84] D. Revuz. *Markov Chains.* Elsevier, Amsterdam, 1984.

[RW94] L.C.G. Rogers and D. Williams. *Diffusions, Markov Processes, and Martingales. Vol.I and II.* Wiley, 1987 and 1994.

[RY99] D. Revuz and M. Yor. *Continuous Martingales and Brownian Motion.* Springer, Berlin Heidelberg, second edition, 1999.

[Saw76] S. Sawyer. Branching diffusion processes in population genetics. *Advances in Appl. Probability*, 8(4):659–689, 1976.

[Shi96] A.N. Shiryayev. *Probability.* Springer-Verlag, second edition, 1996.

[SV79] D.W. Stroock and S.R.S. Varadhan. *Multidimensional diffusion processes.* Springer, Berlin Heidelberg New York, 1979.

[Wak95] A. Wakolbinger. Limits of spatial branching populations. *Bernoulli*, 1(1-2):171–189, 1995.

[Yos92] N. Yoshida. Estimation for diffusion processes from discrete observation. *J. Multivariate Anal.*, 41(2):220–242, 1992.

[YY89] A.Y. Yakovlev and N.M. Yanev. *Transient processes in cell proliferation kinetics*. Number 82 in Lecture Notes in Biomath. Springer, 1989.