On the estimation of jumps of continuous-time stochastic processes

Florian Alexander Johann Ueltzhöfer

2012

Lehrstuhl für mathematische Statistik
Technische Universität München
85 748 Garching bei München

Technische Universität München
Lehrstuhl für mathematische Statistik

On the estimation of jumps of continuous-time stochastic processes

Florian Alexander Johann Ueltzhöfer

Vollständiger Abdruck der von der Fakultät für Mathematik der Technischen Universität München zur Erlangung des akademischen Grades eines

Doktors der Naturwissenschaften (Dr. rer. nat.)

genehmigten Dissertation.

Vorsitzende:		Univ.-Prof. Dr. Nina Gantert
Prüfer der Dissertation:	1.	Univ.-Prof. Dr. Claudia Klüppelberg
	2.	Prof. Dr. Jean Jacod, Université Pierre et Marie Curie, Paris, Frankreich
	3.	Univ.-Prof. Dr. Mark Podolskij Ruprecht-Karls-Universität Heidelberg

Die Dissertation wurde am 12. Dezember 2012 bei der Technischen Universität München eingereicht und durch die Fakultät für Mathematik am 12. März 2013 angenommen.

Für 人惠

Zusammenfassung

Diese Doktorarbeit behandelt nichtparametrische Schätzverfahren für das Sprungverhalten zeitstetiger stochastischer Prozesse, die über die Klasse der Lévy-Prozesse hinausgehen. Das Hauptaugenmerk liegt auf der Klasse der rekurrenten Markov-Prozesse welche gleichzeitig ein Itō-Semimartingal sind. Der Lévy-Kern eines solchen Prozesses beschreibt die Verteilung seiner Sprünge. Basierend auf diskreten Beobachtungen des Prozesses konstruieren wir einen Schätzer für die Dichte des Lévy-Kerns. Wir weisen nach, daß unser Schätzer konsistent ist und ein zentraler Grenzwertsatz gilt, wenn sowohl der Zeithorizont als auch die Beobachtungsfrequenz gegen unendlich divergieren. Als Herzstück dieser Arbeit erforschen wir ebenso den Fall, daß stetige Beobachtungen des zugrundeliegenden Prozesses vorliegen. Auf analoge Weise konstruieren wir einen weiteren Schätzer für die Dichte des Lévy-Kerns. Die Konsistenz dieses Schätzers und die Gültigkeit eines zentralen Grenzwertsatzes beweisen wir für eine allgemeinere Klasse von Markow-Prozessen. Praktische Aspekte der Schätzverfahren untersuchen wir in einer Simulationsstudie. Darüber hinaus beschäftigen wir uns in dieser Arbeit mit der Klasse der zeittransformierten Lévy-Prozesse. Wir konstruieren einen analogen Schätzer für die Dichte des Lévy-Maßes eines solchen Prozesses und beweisen dessen Konsistenz und asymptotische Normalität. Zum Abschluß schlagen wir einen bestimmten zeittransformierten Lévy-Prozeß als ein geeignetes Modell für die Intermittenz in Luftturbulenzen vor. Als Teil einer empirischen Studie wenden wir einen verwandten nichtparametrischen Schätzer für die Lévy-Dichte in unserem Modell an.

Summary

In this thesis, non-parametric estimation of the jumps of continuous-time stochastic processes beyond the Lévy case is studied. The main focus lies on recurrent Markov processes which are Itō semi-martingales. The law of their jumps is described by the Lévy kernel. Based on observations on an equidistant time grid, we construct an estimator for the Lévy kernel's density. We prove the consistency of our estimator and a central limit theorem as both, the time horizon and the sampling frequency, tend to infinity. At the core of our study, we also investigate the case where a sample path is continuously observed. Again, we construct an estimator for the Lévy kernel's density. For a more general class of Markov processes than before, we obtain the consistency of our estimator and a central limit theorem. Practical aspects of our estimators are investigated in a simulation study. In addition, we consider time-changed Lévy processes. For an analogous estimator for the Lévy measure's density, we prove its consistency and asymptotic normality. Finally, a particular time-changed Lévy process is advocated as a suitable model for the intermittency in atmospheric turbulence. As a part of an empirical study, a related non-parametric estimator for the Lévy density is applied.

Danksagung

Mein Dank gilt den folgenden Personen, die mich im Zusammenhang mit dieser Arbeit in den vergangen drei Jahren begleitet haben. Die Möglichkeit diese Doktorarbeit zu verwirklichen und die Gelegenheit meine Arbeit auf hochkarätig besetzten Konferenzen vorzustellen, ließ mich viele wertvolle Erfahrungen sammeln.

Um meiner Verlobten Ren Hui meine unendliche Dankbarkeit für Alles auszudrücken, können diese paar Zeilen niemals ausreichen.

Bei meiner Betreuerin Prof. Dr. Claudia Klüppelberg möchte ich mich für das in mich gesetzte Vertrauen, die gute Zusammenarbeit und hilfreiche Diskussionen bedanken.

Particulièrement, je voudrais remercier Jean Jacod pour son assistance qui était toujours encourageante. De même, je sais apprécier tous les séjours à l'Université Paris 6 Pierre et Marie Curie que Jean a rendus possible.

Ich bedanke mich bei Prof. Dr. Mark Podolskij dafür, daß er sich als Prüfer für diese Dissertation zur Verfügung gestellt hat, und unter anderem für die hervorragende Organisation der Konferenz „Dynstoch 2011" in Heidelberg.

Bei Prof. Dr. Reinhard Höpfner möchte ich mich für die Unterstützung im Vorfeld und während meines eineinhalbwöchigen Gastaufenthalts an der Johannes Gutenberg-Universität Mainz zu Beginn meiner Doktorandenzeit bedanken.

Ebenso gilt mein Dank Dr. Eva Löcherbach, die sich auf verschiedenen Konferenzen die Zeit für lohnenswerte Diskussionen nahm. Likewise, I want to thank Dr. Dasha Loukianova for fruitful discussions.

Desweiteren bedanke ich mich bei Prof. Dr. Jürgen Schmiegel für die Zurverfügungstellung des Brookhaven-Datensatzes.

L'eccellente collaborazione con Enzo Ferrazzano terrò sempre in ricordo. Für seine hilfreiche Unterstützung mit dem „Linux-Cluster" gebührt Ulf Schepsmeier die Anerkennung.

Nicht zu unterschätzen für das Gelingen dieser Arbeit ist auch der Wert von guten Freunden mit denen ich nun wunderbare neue gemeinsame Erinnerungen teile: Alexander Bauer, Florian Fuchs, Dr. Martin Moser, Dr. Anastasios Panagiotelis und Christina Steinkohl.

Abschließend gebührt ein besonderes Dankeschön der guten Seele unseres Lehrstuhls: Andrea Grant.

Mit Unterstützung der Technischen Universität München, Institute for Advanced Study, gefördert durch die Exzellenzinitiative des Bundes und der Länder.

Und mit Unterstützung durch die Technische Universität München, International Graduate School of Science and Engineering.

Contents

1 Introduction

In many applied fields, such as neuroscience, geology, computer science, physics, and mathematical finance, continuous-time stochastic processes are used to model specific dynamics. In many cases, it is natural to include sample path discontinuities, also known as jumps, into these models. For instance, when modelling the membrane potential of a neuron, jumps may represent the action potentials – modelling statistically the actual complex biochemical reaction. When modelling tectonic movements (e. g., of the San Andreas Fault in California) jumps may represent earthquakes. And when modelling the CPU load or the main memory usage in supercomputing systems (e. g., at the Leibniz Supercomputing Centre (Leibniz Rechenzentrum)) jumps may represent the commencement and termination of resource intensive jobs.

The statistical inference for continuous-time models with jumps has received significant attention in recent years. Since, usually, only discrete observations are available in practice, one of the main issues encountered is that the jumps are latent. A vast amount of literature has been devoted to the class of processes with stationary and independent increments, called *Lévy processes*. By the Lévy–Khintchine representation (cf. Sato, 1999, Theorem 8.1), the law of their jumps is characterised by their Lévy measure. Early works in the literature on parametric and non-parametric inference for Lévy processes include Rubin and Tucker (1959), Akritas (1982), and Basawa and Brockwell (1982). Numerous non-parametric and semi-parametric approaches for the estimation of the characteristic triplet and, in particular, the Lévy density have been suggested recently. We refer to the special issue Gugushvili, Klaassen, and Spreij (2010) which contains an interesting collection of papers on this topic with ample references to previous literature. Another comprehensive literature review on the Lévy case is presented in the introductory section of Ueltzhöfer and Klüppelberg (2011).

In this thesis, we study the estimation of the jumps of processes beyond the Lévy case. First and foremost, we consider the class of Harris recurrent Markov processes which are Itō semi-martingales. We remark that many important, continuous-time models – at least in finance – are Itō semi-martingales as stochastic integrals and Itō's formula play a prominent role. For such a process $X = (X_t)_{t \geq 0}$, by definition and by the strong Markov property, the law of its jumps $\Delta X_t = X_t - X_{t-}$ is more or less described by a kernel F on the state space, say E. In particular, for every Borel function g on $E \times E$ and $t > 0$, we have

$$\mathbb{E} \sum_{0<s\leq t} g(X_{s-}, \Delta X_s) \mathbb{1}_{\{X_{s-} \neq X_s\}} = \mathbb{E} \int_0^t \mathrm{d}s \int F(X_s, \mathrm{d}y) g(X_s, y), \tag{1.1}$$

where $\mathbb{E}$ denotes the usual expectation. The kernel F is unique (outside an exceptional set). We call it the *Lévy kernel* of X. It is a generalisation of the notion of Lévy measures: Suppose X is a Lévy process with Lévy measure ν; then $F(x, \mathrm{d}y) = \nu(\mathrm{d}y)$ is the Lévy kernel of the Markovian Itō semi-martingale X. We assume that the measures $F(x, \mathrm{d}y)$ on E admit a density $y \mapsto f(x,y)$ and we aim for the non-parametric estimation of the function $(x,y) \mapsto f(x,y)$.

Our main concern is the case where we observe a sample $X_0(\omega), \ldots, X_{n\Delta}(\omega)$ of the process on an equidistant time grid. Such a discrete observation scheme is commonly observed in literature on the estimation of stochastic processes; we refer to the monograph Jacod and Protter (2012) which is entirely devoted to the "discretisation of processes". We study a kernel density estimator for $f(x,y)$ which resembles the Nadaraya–Watson estimator in classical conditional density estimation. For kernel functions g_1, g_2 and a bandwidth $\eta = (\eta_1, \eta_2) > 0$, in particular, our estimator is of the form

$$\hat{f}_n^{\Delta,\eta}(x,y) = \frac{\sum_{k=1}^n g_1(\eta_1^{-1}(X_{(k-1)\Delta} - x)) g_2(\eta_2^{-1}(X_{k\Delta} - X_{(k-1)\Delta} - y))}{\eta_2^d \Delta \sum_{k=1}^n g_1(\eta_1^{-1}(X_{(k-1)\Delta} - x))}. \tag{1.2}$$

As our main results, we show its consistency as $n\Delta \to \infty$, $\Delta \to 0$ and $\eta \to 0$ under a smoothness hypothesis on the estimated density. Also, we prove a central limit theorem: In the positive recurrent case, henceforth also called ergodic case, our estimator is asymptotically normal. In the null recurrent case, we impose an additional assumption which goes back to Darling and Kac (1957). Thereunder, we prove that our estimator is asymptotically mixed normal. As the convergence in our central limit theorem holds stably in law (a notion due to Renyi, 1963), we also obtain a standardised version of our theorem which can be used for the construction of asymptotic confidence intervals.

At the core of this thesis, and en route of the proof of the previously mentioned results, we essentially study the case first, where a continuous sample path $\{X_s(\omega) : s \in [0, t]\}$ is observed and, in particular, all jumps are discerned. In this case, we consider a more general class of strong Markov processes with càdlàg sample paths than Itō semi-martingales only. Benveniste and Jacod (1973) proved that, for a *Hunt process* X, viz. a quasi-left continuous, strong Markov process with càdlàg sample paths, the law of its jumps is more or less described by a pair (F, H), where F is a kernel on the state space and H is a non-decreasing, continuous additive functional of X. A similar equality as in eq. (1.1) holds where, on the right-hand side, the differential '$\mathrm{d}s$' is replaced by '$\mathrm{d}H_s$'. The pair (F, H) is called a *Lévy system*. Apparently, it is a further generalisation of the notions of Lévy measures and Lévy kernels: Suppose X is an Itō semi-martingale with Lévy kernel F as before; then (F, H) with $H_t = t$ is a Lévy system of the Hunt process X.

We emphasise that, in general, neither the Lévy kernel F nor the additive functional H is uniquely defined. With our main results for the Itō semi-martingale case in mind, we restrict ourselves to the case where $H_t = t$. This amounts to assume the existence of a Lévy system for some additive functional $\bar{H}$ which is absolutely continuous. For the – then distinguished – Lévy system (F, t), we have that the Lévy kernel is unique (outside an exceptional set). Again, we call F the *(canonical) Lévy kernel* of X, and assume that the measures $F(x, \mathrm{d}y)$ on E admit a density $y \mapsto f(x, y)$. Based on the additional observed information, we study the following version of our estimator given by eq. (1.2) for the function $(x, y) \mapsto f(x, y)$:

$$\hat{f}_t^{\eta}(x, y) = \frac{\sum_{0<s\le t} g_1(\eta_1^{-1}(X_{s-} - x)) g_2(\eta_2^{-1}(\Delta X_s - y)) \mathbb{1}_{\{X_{s-} \neq X_s\}}}{\eta_2^d \int_0^t g_1(\eta_1^{-1}(X_s - x))}. \tag{1.3}$$

Under slightly weaker assumptions than before, we prove the estimator's consistency and asymptotic (mixed) normality as $t \to \infty$ and $\eta \to 0$.

Along with these main results, we obtain various complementary ones – some of which are of independent interest: Firstly, we prove a triangular array extension of Birkhoff's theorem for additive functionals. The theorem proves useful in the analysis of the asymptotic behaviour of our estimator as, jointly, $t \to \infty$ and $\eta \to 0$. Secondly, we present a new construction of a uniformly ergodic, auxiliary Markov chain which is based on a "splitting" of the sample paths of our Markov process at specific jump times. Although the presentation thereof is tailored specifically for our needs, the technique may provide a helpful alternative to the famous Nummelin splitting. Thirdly, we have thoroughly studied the influence of discretisation on our estimates in the Itō semi-martingale case. In particular, we quantified the difference between our estimator $\hat{f}_n^{\Delta,\eta}(x,y)$ based on discrete observations and the estimator $\hat{f}_{n\Delta}^{\eta}(x,y)$ based on the full observation of a sample path. We use the corresponding results in the following part of this thesis. Lastly, we prove a non-standard limit theorem for a triangular, martingale array scheme.

In the second part of this thesis, we turn our attention to another class of processes which, in general, is non-Markov. We consider time-changed Lévy processes $X = L \circ T$, where L is a Lévy process with Lévy measure F and $\mathrm{d}T_t = Y_t \mathrm{d}t$ for some non-negative, càdlàg process Y which is independent of L. Time-changed Brownian motion, as far as known to us, was first studied by Bochner (1949). In mathematical finance and econometrics, this class of processes received prominent attention (e. g., Clark, 1973; Carr, Geman, Madan, and Yor, 2003; Barndorff-Nielsen and Shephard, 2006). In statistical physics, it may also serve as a building block for the modelling of atmospheric turbulence.

Again, our concern is the case where we observe a sample $X_0(\omega), \dots, X_{n\Delta}(\omega)$ of the process on an equidistant time grid. We assume that the Lévy measure $F(\mathrm{d}x)$ admits a density $x \mapsto f(x)$. We study an estimator which resembles the classical Rosenblatt–Parzen window estimator. For a kernel function g and a bandwidth $\eta > 0$, our estimator specifies to

$$\hat{f}_n^{\Delta,\eta}(x) = \frac{1}{n\Delta\eta^d} \sum_{k=1}^{n} g\left(\frac{X_{k\Delta} - X_{(k-1)\Delta} - x}{\eta}\right). \tag{1.4}$$

Under the assumption that Y is ergodic with finite fourth moment and under a smoothness condition on the estimated density, we prove the consistency of our estimator as $n\Delta \to \infty$, $\Delta \to 0$ and $\eta \to 0$. Under an additional tightness assumption on a functional of the time-change, we also prove that our estimator is asymptotically normal. Since, as in the Markov setting before, the convergence in our central limit theorem holds stably in law, we also obtain a standardised version of our theorem which can be used for the construction of asymptotic confidence intervals.

We remark that the proofs for these results are heavily based on the results and techniques presented in the case of a Markovian Itō semi-martingale. Especially our analysis of the influence of discretisation plays a crucial role. Moreover, we remark that a study which is similar – but distinguished from ours – is presented by Figueroa-López (2009b, 2011). There, estimators of the form $(n\Delta)^{-1}\sum_{k=1}^{n} g(X_{k\Delta} - X_{(k-1)\Delta})$ are investigated. Under the assumption that Y is an ergodic diffusion, Figueroa-López proved the consistency of such an estimator for $\int F(\mathrm{d}x)g(x)$ and a central limit theorem. In our study, in contrast, the fixed function g is replaced by a sequence of functions $(g_n)_{n\in\mathbb{N}}$ (namely, $g_n(z) = \eta_n^{-d} g((z-x)/\eta_n)$) where $\eta_n \to 0$ as $n \to \infty$.

The third part of this thesis is dedicated to the empirical modelling of the intermittency in atmospheric turbulence. This part is joint work with Vincenzo Ferrazzano. Modelling of turbulence is a long-standing problem in physics and mathematics. Since the seminal work of Kolmogorov (1941a,b, 1942) and its refinement Kolmogorov (1962), it is commonly accepted that turbulence can be regarded as a random phenomenon.

In our study, we focus on the modelling of the velocity $V = (V_t)_{t\in\mathbb{R}}$ of a turbulent flow along the main (longitudinal) flow direction at a single, fixed location. Virtually every observed turbulent flow displays several stylised facts. Experimental investigations highlighted that their magnitude depends only on a control parameter called the *Reynolds number*. Our paramount aim is to advocate a statistical model, which is able to reproduce the following essential "intermittent" features of flows with a high Reynolds number, called *fully developed turbulent flows*: Firstly, the velocity increments display a distinctive clustering; the phenomenon originally called *intermittency*. In particular, the squared increments of turbulent flow velocities are significantly correlated; their auto-correlation function is positive and slowly decaying. Secondly, the velocity increments are semi-heavy tailed and display a distinctive scaling: On large time-scales, on the one hand, the distribution of the increments is approximately Gaussian. On small time-scales, on the other hand, the distribution develops exponential tails and is positively skewed.

Barndorff-Nielsen and Schmiegel (2008) proposed a causal continuous-time moving average process

$$V = (V_t)_{t\in\mathbb{R}}, \quad \text{where } V_t = \bar{v} + \int_{-\infty}^{t} g(t-s)\mathrm{d}X_s, \tag{1.5}$$

driven by some normalised random orthogonal martingale measure $\mathrm{d}X$, as a suitable statistical model for a fully developed turbulent flow with mean $\bar{v} > 0$. In such a model, the second-order properties of V depend only on the square-integrable moving-average kernel g. The driving martingale X, henceforth called the *intermittency process*, accounts for all higher-order properties. Deviating from Barndorff-Nielsen and Schmiegel (2008), we advocate that the intermittency process is appropriately modelled by a time-changed Lévy process $X = L \circ \int_0^{\cdot} Y_s \mathrm{d}s$, where L is a purely discontinuous martingale with tempered stable Lévy measure (see Rosiński, 2007) and Y is itself a positive, ergodic, causal continuous-time moving average process – independent of L.

We estimated our model from a data set which consists of measurements taken at the atmospheric boundary layer, about 35m above the ground. Brockwell, Ferrazzano, and Klüppelberg (2012) proposed a method to estimate the kernel g from an observed sample $V_0(\omega), V_\Delta(\omega) \ldots, V_{n\Delta}(\omega)$ of the velocity. Ferrazzano and Fuchs (2012) extended this method to estimate the increments $X_{k\Delta}(\omega) - X_{(k-1)\Delta}(\omega)$ of the intermittency process in addition. Treating these estimated increments as true observations, we estimated the time-change using a method of moment approach (see Kallsen and Muhle-Karbe, 2011). Next, we estimated the Lévy density of the Lévy process L using the projection estimator of Figueroa-López (2009b, 2011) and the penalisation method which Ueltzhöfer and Klüppelberg (2011) studied in the case of Lévy processes. Under a constraint on the moments of the time-changed Lévy process, we also calculated least-squares fits of certain parametric families of tempered stable Lévy densities to our non-parametric estimate. We minimised an information criterion to find an optimal choice of parameters. In a simulation study, we compare a sample of increments from our intermittency model and the data. The fit of the empirical stationary distribution and of the auto-correlation of the squared intermittency increments (that is, the clustering of large increments) is convincing.

General outline

Apart from this introduction, this thesis contains five chapters, some of which are based on papers.

In Chapter 2, we introduce the relevant notions used in this thesis and summarise essential theorems. This includes the notions of semi-martingales and Markov processes and important limit theorems for stochastic processes. Alongside, we develop the relevant notation.

In the first part of this thesis, which consists of Chapters 3 and 4, we study the estimation of the jumps of processes beyond the Lévy case:
Chapter 3 is based on the paper Ueltzhöfer (2012). In this chapter, we study the kernel density estimation of the Lévy kernel of a Markov process. An individual introductory section is provided. In Section 3.2, we study the estimation of the Lévy kernel based on discrete observations; we present the statistical problem, our standing assumptions and our estimator; and we state our main results of this chapter. In Section 3.3, we study the case where continuous-time observations are available. The proofs for the latter section are presented in Section 3.4. This section also contains our extension of Birkhoff's theorem for additive functionals to triangular arrays and the construction of the aforementioned auxiliary Markov chain. The proofs for our main results of Section 3.2 are presented in Section 3.5. Some technical considerations are put off to the supplementary Section 3.6.
In Chapter 4, we study the kernel density estimation of the Lévy measure of a time-changed Lévy process. The chapter is organised analogously to Chapter 3. An individual introductory section is provided. In Section 4.2, we present the statistical problem, our standing assumptions and our estimator; and we state our main results of this chapter. The corresponding proofs are in Section 4.3.

In the second part of this thesis, which consists of Chapters 5 and 6, we study the estimation of jumps in practice:
In Chapter 5, we present a simulation study for the kernel density estimator presented in Chapter 3 and a simulation study for the penalised projection estimation of the Lévy measure of a Lévy process. The latter study is based on Section 4 of the paper Ueltzhöfer and Klüppelberg (2011). In the former study, inter alia, we focus on the influence of discretisation and the importance of suitable bandwidth selection.

Chapter 6 is based on Ferrazzano and Ueltzhöfer (2012). This chapter is joint work with Vincenzo Ferrazzano and is dedicated to the empirical modelling of the intermittency in atmospheric turbulence. An individual introductory section is provided. In Section 6.2 we present the statistical problem and the estimation methods which we apply. In Section 6.3, we perform an empirical study of the so-called Brookhaven wind speed data set. Finally, in Section 6.4, we compare our fitted model and the data set in a short simulation study.

Notational conventions

The following notational conventions are used throughout this thesis without further explanation:

- $\mathbb{R}$ denotes the real numbers; $\mathbb{R}_+$ denotes the non-negative real numbers; $\mathbb{R}_+^*$ denotes the positive real numbers; $\overline{\mathbb{R}} := \mathbb{R} \cup \{-\infty, \infty\}$; $\overline{\mathbb{R}_+} := \mathbb{R}_+ \cup \{\infty\}$
- $\mathbb{Z}$ denotes the integers; $\mathbb{N}$ denotes the non-negative integers; $\mathbb{N}^*$ denotes the natural numbers excluding zero; $\overline{\mathbb{N}} := \mathbb{N} \cup \{\infty\}$
- For $p \in \mathbb{N}^*$, $\mathbb{Z}_p := \mathbb{Z}/(p\mathbb{Z})$ denotes the ring of the residual classes of integers modulo p

2 Semi-martingales, Markov processes, Limit theorems

The theory developed and studied in this thesis is based on various fields of probability theory: Firstly, the general theory of stochastic processes (cf. Jacod, 1979). Secondly, the limit theory for semi-martingales (cf. Jacod and Shiryaev, 2003; Jacod and Protter, 2012). Thirdly, the theory of Markov processes (cf. Getoor, 1975; Sharpe, 1988; Höpfner and Löcherbach, 2003) and their discrete-time analogues – the Markov chains (cf. Revuz, 1984; Meyn and Tweedie, 1993). We emphasise at this point that other fields such as measure theory, topology, and functional analysis are an integral part of the aforementioned. To make this thesis self-contained as far as possible, we dedicate this chapter to summarise the theory applied in subsequent chapters. Alongside, we develop the relevant notation.

2.1 Résumé of the general theory of stochastic processes

The presentation of the general theory of stochastic processes in this section is mainly based on Jacod (1979) and Jacod and Shiryaev (2003). The development of the theory is closely connected to the work of Joseph Leo Doob, Kiyoshi Itō, Paul-André Meyer, Shinzo Watanabe, and Claude Dellacherie, to just name a few.

2.1.1 Random sets; Processes; Optional and predictable σ-field

Let $(\Omega, \mathscr{F}, \mathbb{P})$ be a probability space and let $\mathfrak{F} := (\mathscr{F}_t)_{t\geq 0}$ be a filtration, that is, an increasing sequence of sub-σ-fields of $\mathscr{F}$. For convenience, we suppose $\mathscr{F}_\infty = \mathscr{F}$ and $\mathscr{F}_{\infty-} = \vee_{s\geq 0}\mathscr{F}_s$.

2.1.1 Definition. The filtered probability space $(\Omega, \mathscr{F}, \mathfrak{F}, \mathbb{P})$ is said to *satisfy the usual conditions* if

(i) it is *complete*, that is, $\mathscr{F}$ is $\mathbb{P}$-complete and every $\mathscr{F}_t$ contains all $\mathbb{P}$-null sets of $\mathscr{F}$; and

(ii) the filtration $\mathfrak{F}$ is *right-continuous*, that is, $\mathscr{F}_t = \cap_{s>t}\mathscr{F}_s$ for all $t \geq 0$. ◇

Given a single probability measure $\mathbb{P}$ on $(\Omega, \mathscr{F})$, it is no loss of generality to assume that the usual conditions are satisfied (cf. Jacod and Shiryaev, 2003, I.1.4).

A first notion of interest is that of random sets:

2.1.2 Definition. (i) A subset of $\Omega \times \mathbb{R}_+$ is called a *random set*.

(ii) A random set A is called *evanescent* if $\mathbb{P}(\{\omega : \exists t \geq 0 \text{ s. t. } (\omega, t) \in A\}) = 0$. ◇

An important example of random sets are *stochastic intervals*: Let $S, T : \Omega \to \overline{\mathbb{R}}_+$. Then, we define

$$[\![S, T]\!] := \{(\omega, t) : S(\omega) \leq t \leq T(\omega)\} \tag{2.1.1}$$

and, analogously, $[\![S, T[\![$, $]\!]S, T]\!]$ and $]\!]S, T[\![$. Also, we set $[\![T]\!] := [\![T, T]\!]$.

Having introduced the space on which our main objects of interest, the stochastic processes, will be defined, at this point it remains to introduce the space in which these will take their values: Let $(E, \mathscr{T})$ be a topological space. The σ-field generated by the topology $\mathscr{T}$ is called the *Borel σ-field* on E, denoted $\mathscr{E}^0 := \mathscr{B}(E) := \sigma(\mathscr{T})$. By $\mathcal{M}_{\mathrm{b}}(E)$ we denote the set of all finite (positive) measures on the measurable space $(E, \mathscr{E}^0)$. For every $\mu \in \mathcal{M}_{\mathrm{b}}(E)$, we denote the μ-completion of $\mathscr{E}^0$ by $\mathscr{E}^\mu$. The sets in the family $\mathscr{E}^{\mathrm{u}} := \cap_{\mu \in \mathcal{M}_{\mathrm{b}}(E)} \mathscr{E}^\mu$ are called *universally measurable*; $\mathscr{E}^{\mathrm{u}}$ is called the *universally measurable σ-field* on E. Throughout the remainder of this section, we abbreviate $E = (E, \mathscr{E})$, where $\mathscr{E}$ denotes an "intermediate σ-field" $\mathscr{E}^0 \subseteq \mathscr{E} \subseteq \mathscr{E}^{\mathrm{u}}$.

2.1.3 Definition. (i) A family $X = (X_t)_{t\geq 0}$ of mappings $X_t : \Omega \to E$ is called an *E-valued process*. The process X is called *measurable* if it is $\mathscr{F} \otimes \mathscr{B}(\mathbb{R}_+)$-$\mathscr{E}$-measurable as a mapping $X : \Omega \times \mathbb{R}_+ \to E$. The mappings $t \mapsto X_t(\omega)$ for fixed $\omega \in \Omega$ are called the *trajectories* or *sample paths* of X.

(ii) Let X and Y be two E-valued processes. We call them *indistinguishable* if the random set $\{(\omega, t) : X_t(\omega) \neq Y_t(\omega)\}$ is evanescent. ◇

For $\omega \in \Omega$, we denote the left-limit at time $t > 0$ by $X_{t-}(\omega) := \lim_{s\to t, s<t} X_s(\omega)$ as soon as it exists for the respective trajectory. Also, we agree to set $X_{0-}(\omega) = X_0(\omega)$.

In the case that E is an additive group, we set $\Delta X_t(\omega) := X_t(\omega) - X_{t-}(\omega)$, again, as soon as $X_{t-}(\omega)$ exists. Processes where all trajectories are right-continuous and admit left-limits are of utmost importance to the theory: They are called *càdlàg* for "continu à droite avec des limites à gauche". Likewise, processes which are right-continuous (resp., left-continuous) are called càd (resp., càg). For a càdlàg process X, we denote $X_- := (X_{t-})_{t\geq 0}$ and $\Delta X := (\Delta X_t)_{t\geq 0}$. We say that the trajectory $t \mapsto X_t(\omega)$ has a *jump at time* $t > 0$ if $\Delta X_t(\omega) \neq 0$.

2.1.4 Definition. (i) An E-valued process X is called *adapted (to the filtration* $\mathfrak{F}$*)* if the mappings $X_t : \Omega \to E$ are $\mathscr{F}_t$-$\mathscr{E}$-measurable for all $t \geq 0$.

(ii) The σ-field over $\Omega \times \mathbb{R}_+$ generated by all càdlàg, $\mathfrak{F}$-adapted processes is called the *optional* σ*-field*, denoted $\mathscr{O} = \mathscr{O}(\mathfrak{F})$. The random sets in $\mathscr{O}$ are called *optional*.

(iii) The σ-field over $\Omega \times \mathbb{R}_+$ generated by all càg, $\mathfrak{F}$-adapted processes is called the *predictable* σ*-field*, denoted $\mathscr{P} = \mathscr{P}(\mathfrak{F})$. The random sets in $\mathscr{P}$ are called *predictable*. ⋄

By Proposition I.1.24 of Jacod and Shiryaev (2003), we have $\mathscr{P} \subseteq \mathscr{O}$.

2.1.2 Stopping times; Predictable times; Quasi-left continuity

The notions of optionality and predictability are closely linked to stopping times:

2.1.5 Definition. A mapping $T : \Omega \to \overline{\mathbb{R}}_+$ is called a *stopping time* if $\{T \leq t\} \in \mathscr{F}_t$ for all $t \geq 0$. For a stopping time T, we denote

$$\mathscr{F}_T := \Big\{ A \in \mathscr{F} : A \cap \{T \leq t\} \in \mathscr{F}_t \text{ for all } t \geq 0 \Big\} \tag{2.1.2}$$

and

$$\mathscr{F}_{T-} := \sigma\Big(\mathscr{F}_0 \cup \{ A \cap \{t < T\} : t \in \mathbb{R}_+ \text{ and } A \in \mathscr{F}_t \} \Big). \tag{2.1.3}$$

The process $X^T = (X^T_t)_{t\geq 0}$ given by $X^T(\omega, t) := X(\omega, t \wedge T(\omega))$ is called the process *stopped at time* T. ⋄

By Remark I.1.26 of Jacod and Shiryaev (2003), the optional σ-field is generated by the stochastic intervals $[\![0, T[\![$ where T is a stopping time. By Theorem I.2.2 of

Jacod and Shiryaev (2003), the predictable σ-field is generated by the stochastic intervals $[\![0, T]\!]$ where T is a stopping time and the sets $A \times \{0\}$ where $A \in \mathscr{F}_0$.

The following results give further insight on optional and predictable processes:

2.1.6 Proposition (Jacod and Shiryaev (2003) I.1.21 and I.1.25). *(i) Let X be an optional process and T be a stopping time. Then $X_T \mathbb{1}_{T<\infty}$ is $\mathscr{F}_T$-measurable and the stopped process X^T is optional.*

(ii) Let X be a càdlàg, adapted process. Then the processes X_- and ΔX are optional.

2.1.7 Proposition (Jacod and Shiryaev (2003) I.2.4 and I.2.6).

(i) Let X be a predictable process and T be a stopping time. Then $X_T \mathbb{1}_{T<\infty}$ is $\mathscr{F}_{T-}$-measurable and the stopped process X^T is predictable.

(ii) Let X be a càdlàg, adapted process. Then the process X_- is predictable. If X is predictable in addition, then ΔX is predictable.

Predictable and totally inaccessible (stopping) times play an important role:

2.1.8 Definition. A mapping $T : \Omega \to \overline{\mathbb{R}}_+$ is called a *predictable time* if $[\![0, T[\![\in \mathscr{P}$. A stopping time T is called *totally inaccessible* if $\mathbb{P}(T = S < \infty) = 0$ for all predictable times S. ◇

2.1.9 Definition. A càdlàg process X is called *quasi-left continuous* if, for every predictable time T, $\Delta X_T = 0$ a. s. on $\{T < \infty\}$. ◇

We have:

2.1.10 Proposition (Jacod and Shiryaev (2003) I.2.24 and I.2.26).

(i) Let X be a predictable, càdlàg process. Then there exists a sequence $(T_n)_{n\in\mathbb{N}}$ of predictable times such that $\{\Delta X \neq 0\} = \cup_{n\in\mathbb{N}}[\![T_n]\!]$. Furthermore, $\Delta X_T = 0$ almost surely on $\{T < \infty\}$ for all totally inaccessible stopping times T.

(ii) Let X be a càdlàg, adapted process. Then X is quasi-left continuous if, and only if, there exists a sequence $(T_n)_{n\in\mathbb{N}}$ of totally inaccessible stopping times such that $\{\Delta X \neq 0\} = \cup_{n\in\mathbb{N}}[\![T_n]\!]$ and if, and only if, for every increasing sequence $(S_n)_{n\in\mathbb{N}}$ of stopping times with limit S, we have $X_{S_n} \to X_S$ a. s. on $\{S < \infty\}$ as $n \to \infty$.

2.1.3 Martingales; Increasing processes; Doob–Meyer decomposition

An important class of processes is the class of martingales.

2.1.11 Definition. A process X is called a *martingale* (resp., *sub-martingale*) if it is adapted and a. s. càdlàg such that every X_t is integrable and such that

$$X_s = \mathbb{E}[X_t \mid \mathscr{F}_s] \quad (\text{resp., } X_s \leq \mathbb{E}[X_t \mid \mathscr{F}_s])$$

for every $s \leq t$. A martingale X is called *uniformly integrable* if the family of random variables $(X_t)_{t\geq 0}$ is uniformly integrable, and is called *square-integrable* if $\sup_{t\geq 0} \mathbb{E}\, X_t^2 < \infty$. ⋄

The next theorem is known as Doob's inequality:

2.1.12 Theorem (Jacod and Shiryaev (2003) I.1.43). *Let X be a square-integrable martingale. Then*

$$\mathbb{E} \sup_{t\geq 0} X_t^2 \leq 4 \sup_{t\geq 0} \mathbb{E}\, X_t^2 = 4\, \mathbb{E}\, X_\infty^2. \tag{2.1.4}$$

2.1.13 Definition. A process X is called a *local martingale* if there exists an increasing sequence $(T_n)_{n\in\mathbb{N}}$ of stopping times, called a *localising sequence*, such that $T_n \to \infty$ a. s. and each stopped process X^{T_n} is a martingale. A process X is of *class (D)* if the set of random variables $\{X_T : T \text{ is a finite valued stopping time}\}$ is uniformly integrable. ⋄

The processes of finite variation form a second important class:

2.1.14 Definition. (i) We denote by $\mathscr{V}^+$ (resp., by $\mathscr{V}$) the class of *adapted, increasing processes* (resp., of *processes of finite variation*); that is, of all real-valued, càdlàg, adapted processes such that all its paths are non-decreasing (resp., have finite variation over each finite interval $[0, t]$).

(ii) We denote by $\mathscr{A}^+$ (resp., by $\mathscr{A}$) the class of *integrable, increasing processes* (resp., of *processes of integrable variation*); that is, of all processes $X \in \mathscr{V}^+$ (resp., $X \in \mathscr{V}$) such that $\mathbb{E}\, X_\infty < \infty$ (resp., $\mathbb{E} \operatorname{Var}[X]_\infty < \infty$, where $\operatorname{Var}[X]$ denotes the variation process of X). ⋄

This allows to formulate the Doob–Meyer decomposition for sub-martingales:

2.1.15 Theorem (Jacod and Shiryaev (2003) I.3.15). *Let X be a sub-martingale of class (D). Then there exists a unique (up to indistinguishability) increasing, integrable, predictable process H with $H_0 = 0$ such that $X - H$ is a uniformly integrable martingale.*

There is an important extension of this theorem which we extensively use in this thesis. We remark that the "localising procedure" of Definition 2.1.13 is used analogously for various classes of processes. A process X belongs to the localised class $\mathscr{C}_{\mathrm{loc}}$ of the class $\mathscr{C}$ if there exists a localising sequence $(T_n)_{n\in\mathbb{N}}$ as in Definition 2.1.13 such that the stopped processes X^{T_n} belong to $\mathscr{C}$.

2.1.16 Theorem (Jacod and Shiryaev (2003) I.3.18). *Let $X \in \mathscr{A}_{\mathrm{loc}}$. Then there exists a predictable process X^{p} of locally integrable variation, called the* predictable compensator *of X, which is unique up to an evanescent set, such that $X - X^{\mathrm{p}}$ is a local martingale.*

2.2 Semi-martingales

The presentation of the theory of semi-martingales in this section is mainly based on the monographs Jacod and Protter (2012) and Jacod and Shiryaev (2003). Throughout this section, let $(\Omega, \mathscr{F}, \mathfrak{F}, \mathbb{P})$ be a filtered probability space.

2.2.1 Semi-martingales; Stochastic integrals; Quadratic variation

2.2.1 Definition. A process $X : \Omega \times \mathbb{R}_+ \to \mathbb{R}^d$ is called a d-dimensional *semi-martingale* if each of its components $X^1, \dots, X^d$ is adapted and càdlàg such that $X^i_t - X^i_0$ is the sum of a local martingale and of a process of finite variation. ⋄

Semi-martingales are used as integrators to define stochastic integrals. For a thorough introduction thereof, we refer to Protter (2005) and Section I.4d of Jacod and Shiryaev (2003). For a semi-martingale X and a predictable, locally bounded process H, we denote by $H \cdot X$ the stochastic integral given by

$$(H \cdot X)_t := \int_0^t H_s^{\intercal} \mathrm{d}X_s = \sum_{i=1}^d \int_0^t H^i_s \mathrm{d}X^i_s. \tag{2.2.1}$$

Semi-martingales admit various decompositions. A first decomposition is in terms of the (local) martingale part.

2.2.2 Definition. A local martingale M is called *purely discontinuous* if $M_0 = 0$ and if $MN := (M_t N_t)_{t\geq 0}$ is a local martingale for every continuous local martingale N. $\diamond$

2.2.3 Proposition (Jacod and Shiryaev (2003) I.4.18). *Let M be a local martingale, then there exists a unique (up to indistinguishability) decomposition*

$$M_t = M_0 + M_t^{\mathrm{c}} + M_t^{\mathrm{d}},$$

where $M_0^{\mathrm{c}} = M_0^{\mathrm{d}} = 0$, M^{c} is a continuous martingale, and M^{d} is a purely discontinuous local martingale.

2.2.4 Theorem (Jacod and Shiryaev (2003) I.4.27). *Let X be a semi-martingale. Then there exists a unique (up to indistinguishability) continuous martingale X^{c} with $X_0^{\mathrm{c}} = 0$ such that*

$$X_t = X_0 + A_t + X_t^{\mathrm{c}} + M_t^{\mathrm{d}}, \tag{2.2.2}$$

where $A_0 = M_0^{\mathrm{d}} = 0$, A is an adapted process of finite variation, and M^{d} is a purely discontinuous local martingale.

The process X^{c} is called the *continuous martingale part of* X. We note that a local martingale M with $\mathbb{E}\, M_0^2 < \infty$ such that ΔM is locally bounded is locally square-integrable. By Theorem 2.1.15, thus, there exists an $\mathbb{R}^d \otimes \mathbb{R}^d$-valued process $\langle X^{\mathrm{c}}, X^{\mathrm{c}}\rangle$ such that each of its components $\langle X^{i,\mathrm{c}}, X^{j,\mathrm{c}}\rangle$ is an increasing process and such that each process $X^{i,\mathrm{c}}X^{j,\mathrm{c}} - \langle X^{i,\mathrm{c}}, X^{j,\mathrm{c}}\rangle$ is a local martingale.

2.2.5 Definition. For a semi-martingale X, we call the $\mathbb{R}^d \otimes \mathbb{R}^d$-valued process $[X, X]$ with components given by

$$\left[X^i, X^j\right]_t := \left\langle X^{i,\mathrm{c}}, X^{j,\mathrm{c}}\right\rangle_t + \sum_{0<s\leq t} \Delta X_s^i \Delta X_s^j, \tag{2.2.3}$$

the *quadratic variation* of X. In the case that (each component of) $[X, X]$ is locally integrable, by Theorem 2.1.15, there exists a predictable process $\langle X, X, \rangle$, called the *predictable quadratic variation* of X, such that $[X, X] - \langle X, X\rangle$ is a local martingale. $\diamond$

2.2.2 Random measures; Characteristics

The notion of random measures is essential to the understanding of semi-martingales. A topological space is called a *Polish space* if it is completely metrisable and separable.

2.2.6 Definition. Let $E = (E, \mathscr{E})$ be a Polish space with Borel σ-field $\mathscr{E}$. A family $\mathfrak{m} = (\mathfrak{m}(\omega; \mathrm{d}t, \mathrm{d}x) : \omega \in \Omega)$ of non-negative measures on $(\mathbb{R}_+ \times E, \mathscr{B}_+ \otimes \mathscr{E})$ with $\mathfrak{m}(\omega; \{0\} \times E) = 0$ is called a *random measure*. $\diamond$

We denote $(\tilde{\Omega}, \tilde{\mathscr{F}}) = (\Omega \times \mathbb{R}_+ \times E, \mathscr{F} \otimes \mathscr{B}_+ \otimes \mathscr{E})$, $\tilde{\mathscr{O}} := \mathscr{O} \otimes \mathscr{E}$ and $\tilde{\mathscr{P}} := \mathscr{P} \otimes \mathscr{E}$. A function on $\tilde{\Omega}$ is called *optional* if it is $\tilde{\mathscr{O}}$-measurable. Likewise, it is called *predictable* if it is $\tilde{\mathscr{P}}$-measurable. For a random measure $\mathfrak{m}$ and every optional function g on $\tilde{\Omega}$, we denote by $g \star \mathfrak{m}$ the *stochastic integral* given by

$$g \star \mathfrak{m}_t(\omega) := \begin{cases} \int\limits_{[0,t] \times E} g(\omega, s, x) \mathfrak{m}(\omega; \mathrm{d}s, \mathrm{d}x), & \text{if } |g| \star \mathfrak{m}_t(\omega) < \infty, \\ \infty, & \text{otherwise.} \end{cases} \tag{2.2.4}$$

2.2.7 Definition. (i) A random measure $\mathfrak{m}$ is called *optional* (resp., *predictable*) if the process $g \star \mathfrak{m}$ is optional (resp., predictable) for every optional (resp., predictable) function g on $\tilde{\Omega}$.

(ii) An optional random measure $\mathfrak{m}$ such that $\mathbb{E}\,\mathfrak{m}(\mathbb{R}_+ \times E) < \infty$ is called *integrable*.

(iii) An optional random measure $\mathfrak{m}$ is called *$\tilde{\mathscr{P}}$-σ-finite* if there exists a strictly positive, predictable function h on $\tilde{\Omega}$ such that $h \star \mathfrak{m}_\infty$ is integrable. $\diamond$

Theorem 2.1.16 admits an important generalisation, a "Doob–Meyer decomposition for random measures":

2.2.8 Theorem (Jacod and Shiryaev (2003) II.1.8). *Let $\mathfrak{m}$ be a $\tilde{\mathscr{P}}$-σ-finite random measure. Then there exists a predictable random measure $\mathfrak{m}^{\mathfrak{p}}$, called the* predictable compensator *of $\mathfrak{m}$, which is unique up to a $\mathbb{P}$-null set, such that the following holds: For every predictable function h on $\tilde{\Omega}$ with $|h| \star \mathfrak{m} \in \mathscr{A}^+_{\mathrm{loc}}$, we have $|h| \star \mathfrak{m}^{\mathfrak{p}} \in \mathscr{A}^+_{\mathrm{loc}}$ and $h \star \mathfrak{m} - h \star \mathfrak{m}^{\mathfrak{p}}$ is a local martingale.*

The most important example of a random measure is the so-called *jump-measure* of a semi-martingale; that is, the random measure $\mathfrak{m} = \mathfrak{m}^X$ given by

$$\mathfrak{m}(\omega; \mathrm{d}t, \mathrm{d}x) = \sum_{\{s : \Delta X_s(\omega) \neq 0\}} \epsilon_{(s, \Delta X_s(\omega))}(\mathrm{d}t, \mathrm{d}x), \tag{2.2.5}$$

where ϵ_x denotes the Dirac measure at x. We note that the jump-measure takes its values in $\overline{\mathbb{N}}$, and is optional and $\tilde{\mathscr{P}}$-σ-finite. We denote the predictable compensator of $\mathfrak{m}$, which exists due to Theorem 2.2.8, by $\mathfrak{n} = \mathfrak{n}^X$. By Corollary II.1.19 of Jacod and Shiryaev (2003), we have that X is quasi-left continuous if, and only if, there exists a version of $\mathfrak{n}$ with $\mathfrak{n}(\omega; \{t\} \times E) = 0$ for all $(\omega, t) \in \Omega \times \mathbb{R}_+$.

A construction of an integral $h \star (\mathfrak{m} - \mathfrak{n})$ of some predictable function h on $\tilde{\Omega}$ w. r. t. the compensated jump-measure $\mathfrak{m} - \mathfrak{n}$ is presented, e. g., in Section II.1d of Jacod and Shiryaev (2003). By eqs. (2.1.16) and (2.1.17) of Jacod and Protter (2012), if $(h^2 \wedge |h|) \star \mathfrak{n}_t < \infty$ for all $t \geq 0$, then $h \star (\mathfrak{m} - \mathfrak{n})$ is defined as the unique (up to indistinguishability) purely discontinuous local martingale with jumps given by

$$\Delta(h \star (\mathfrak{m} - \mathfrak{n}))_t = h(t, \Delta X_t) - \int_E h(t, x) \mathfrak{n}(\{t\}, \mathrm{d}x).$$

Eventually, we arrive at a second decomposition of semi-martingales. Jacod and Protter (2012) call it the *isolated big jumps canonical decomposition*:

2.2.9 Theorem (Jacod and Shiryaev (2003) II.2.34). *Let X be a semi-martingale. Then there exists a predictable process of locally finite variation B with $B_0 = 0$ such that*

$$X_t = X_0 + B_t + X_t^c + (x \mathbb{1}_{\|x\| \leq 1}) \star (\mathfrak{m} - \mathfrak{n})_t + (x \mathbb{1}_{\|x\| > 1}) \star \mathfrak{m}_t. \tag{2.2.6}$$

We call $(B, C, \mathfrak{n})$, where $C = \langle X^c, X^c \rangle$, the characteristics *of X.*

2.2.3 Itō semi-martingales; Itō's formula

An important class of semi-martingales are those with absolutely continuous characteristics.

2.2.10 Definition. Let X be a semi-martingale with characteristics $(B, C, \mathfrak{n})$. Then X is called an *Itō semi-martingale* if its characteristics are absolutely continuous with respect to Lebesgue measure; that is,

$$\mathrm{d}B_t = b_t \mathrm{d}t, \quad \mathrm{d}C_t = c_t \mathrm{d}t, \quad \text{and} \quad \mathfrak{n}(\mathrm{d}t, \mathrm{d}x) = \mathrm{d}t \otimes F_t(\mathrm{d}x), \tag{2.2.7}$$

where b (resp., c) is a process with values in $\mathbb{R}^d$ (resp., in $\mathbb{R}^d \otimes \mathbb{R}^d$) and $F_t(\omega; \cdot)$ is a measure on $\mathbb{R}^d$ for each $(\omega, t) \in \Omega \times \mathbb{R}_+$. ⋄

For Itō semi-martingales, Itō's formula reads as follows:

2.2.11 Theorem (Jacod and Protter (2012) eq. 2.1.20). *Let X be a semi-martingale with characteristics $(B, C, \mathfrak{n})$ given by eq. (2.2.7) and let $g \in C^2(\mathbb{R}^d)$. Then $g(X)$ is a semi-martingale and satisfies*

$$\begin{aligned} g(X_t) = g(X_0) &+ \int_0^t \left(b_s^\intercal \nabla g(X_s)\right) \mathrm{d}s + \frac{1}{2} \int_0^t \mathrm{tr}\left(c_s \nabla^2 g(X_s)\right) \mathrm{d}s \\ &+ \left[\left(g(X_- + x) - g(X_-) - x^\intercal \nabla g(X_-)\right) \mathbb{1}_{\|x\| \le 1}\right] \star \mathfrak{n}_t \\ &+ \int_0^t \nabla g(X_{s-})^\intercal \mathrm{d}X_s^c + \left[\left(g(X_- + x) - g(X_-)\right) \mathbb{1}_{\|x\| \le 1}\right] \star (\mathfrak{m} - \mathfrak{n})_t \\ &+ \left[\left(g(X_- + x) - g(X_-)\right) \mathbb{1}_{\|x\| > 1}\right] \star \mathfrak{m}_t, \end{aligned} \tag{2.2.8}$$

where $\mathrm{tr}(\cdot)$ denotes the trace operator on $\mathbb{R}^d \otimes \mathbb{R}^d$ and $\nabla^2 g$ denotes the Hessian of g.

These processes admit another decomposition due to Grigelionis (1971). In the following theorem, let $d' \ge d$ be an integer, and λ be an infinite measure without atom on some arbitrary Polish space E. For the definition of a *very good extension*, see Section 2.1.4 of Jacod and Protter (2012).

2.2.12 Theorem (Jacod and Protter (2012) 2.1.2). *Let X be a d-dimensional Itō semi-martingale with characteristics $(B, C, \mathfrak{n})$ given by eq. (2.2.7). Then there exists a very good extension of the probability space on which a d'-dimensional Brownian motion W and a Poisson random measure $\mathfrak{p}$ with compensator $\mathfrak{q}(\mathrm{d}t, \mathrm{d}x) = \mathrm{d}t \otimes \lambda(\mathrm{d}x)$ are defined such that*

$$X_t = X_0 + \int_0^t b_s \mathrm{d}s + \int_0^t \sigma_s \mathrm{d}W_s + (h \mathbb{1}_{\|h\| \le 1}) \star (\mathfrak{p} - \mathfrak{q})_t + (h \mathbb{1}_{\|h\| > 1}) \star \mathfrak{p}_t, \tag{2.2.9}$$

where σ is an $\mathbb{R}^d \otimes \mathbb{R}^{d'}$-valued, predictable process and h is a predictable function on $\tilde{\Omega} = \Omega \times \mathbb{R}_+ \times E$. Outside a null set, we have $\sigma_t \sigma_t^\intercal = c_t$ and the measure $F_t(\omega; \cdot)$ coincides with the image of the measure λ under the map $x \mapsto h(\omega, t, x)$ restricted to the set $\{x : h(\omega, t, x) \neq 0\}$.

Itō semi-martingales of the form eq. (2.2.9) admit useful estimates.

2.2.13 Proposition (Jacod and Protter (2012) eq. 2.1.43). *Let X be a d-dimensional Itō semi-martingale with characteristics $(B, C, \mathfrak{n})$ given by eq. (2.2.7). Then, for $t > 0$, $p \geq 2$ and every finite stopping time T, there exists a finite constant $\zeta < \infty$ such that*

$$\mathbb{E}\left[\sup_{s \leq t} \|X_{T+s} - X_T\|^p \,\middle|\, \mathscr{F}_T\right] \leq \zeta\, \mathbb{E}\Bigg[\left(\int_T^{T+t} \|b_s\| \mathrm{d}s\right)^p + \left(\int_T^{T+t} \|\sigma_s\|^2 \mathrm{d}s\right)^{p/2}$$
$$+ \int_T^{T+t} \mathrm{d}s \int \lambda(\mathrm{d}z) \|h(s,z)\|^p + \left(\int_T^{T+t} \mathrm{d}s \int_{\{z : \|h(t,z)\| \leq 1\}} \lambda(\mathrm{d}z) \|h(s,z)\|^2\right)^{p/2} \tag{2.2.10}$$
$$+ \left(\int_T^{T+t} \mathrm{d}s \int_{\{z : \|h(t,z)\| > 1\}} \lambda(\mathrm{d}z) \|h(s,z)\|\right)^p \,\Bigg|\, \mathscr{F}_T\Bigg].$$

Moreover, the class of Itō semi-martingales is closed under absolutely continuous time-changes.

2.2.14 Theorem (Jacod (1979) 10.12). *Let X be a d-dimensional Itō semi-martingale with characteristics $(B, C, \mathfrak{n})$ given by eq. (2.2.7). Moreover, let $Y = (Y_t)_{t \geq 0}$ be a positive càdlàg process – independent of X – such that*

$$T_t := \int_0^t Y_s \mathrm{d}s \quad \text{is an } \mathscr{F}_t\text{-stopping time for all } t \geq 0.$$

Then the time-changed process *$X' = (X'_t)_{t \geq 0}$ given by $X'_t := X_{T_t}$ is an Itō semi-martingale w. r. t. the filtration $\mathfrak{F}'$ given by $\mathscr{F}'_t := \mathscr{F}_{T_t}$. Moreover, its characteristics $(B', C', \mathfrak{n}')$ satisfy*

$$\mathrm{d}B'_t = b_t Y_t \mathrm{d}t, \quad \mathrm{d}C'_t = c_t Y_t \mathrm{d}t, \quad \text{and} \quad \mathfrak{n}'(\mathrm{d}t, \mathrm{d}x) = Y_t \mathrm{d}t \otimes F_t(\mathrm{d}x).$$

2.3 Tightness; Convergence of processes; Limit theorems

The definitions and limit theorems in this section are based on Chapters VI and VII of Jacod and Shiryaev (2003). See also Section 2.2 of Jacod and Protter (2012). The notions of almost sure convergence, convergence in probability, and convergence in law for $\mathbb{R}^d$-valued random variables is assumed to be known. A less common notion of convergence (introduced by Renyi, 1963) is stable convergence in law:

2.3.1 Definition. Let $(X_n)_{n\in\mathbb{N}}$ be a sequence of random variables defined on some probability space $(\Omega, \mathscr{F}, \mathbb{P})$. Moreover, let X be a random variable defined on an extension $(\Omega \times \Omega', \mathscr{F} \otimes \mathscr{F}', \tilde{\mathbb{P}})$ where $\tilde{\mathbb{P}}$ is a probability measure whose marginal on Ω is $\mathbb{P}$. Then we say that X_n *converges stably in law* to X if

$$\mathbb{E}[f(X_n)Y] \xrightarrow[n\to\infty]{} \tilde{\mathbb{E}}[f(X)Y]$$

holds for every bounded continuous f on $\mathbb{R}^d$ and every bounded random variable Y on $(\Omega, \mathscr{F})$. We write $X_n \xrightarrow{\mathscr{L}-\mathrm{st}} X$. ⋄

2.3.1 Tightness

Let E be a *Polish space* and let $\mathscr{E}$ denote its Borel σ-field. On the space of probability measures on $(E, \mathscr{E})$, denoted $\mathcal{P}(E)$, we consider the *weak topology* which makes $\mathcal{P}(E)$ another Polish space.

2.3.2 Definition. A subset $\mathcal{A} \subseteq \mathcal{P}(E)$ is called *tight* if, for every $\varepsilon > 0$, there exists a compact $C \subseteq E$ such that $\mu(E \setminus K) \leq \varepsilon$ for every $\mu \in \mathcal{A}$. ⋄

The famous Prokhorov's theorem is essential to the sequel:

2.3.3 Theorem (Jacod and Shiryaev (2003) VI.3.5). *A subset $\mathcal{A} \subseteq \mathcal{P}(E)$ is relatively compact for the weak topology if, and only if, it is tight.*

The importance of this theorem is appreciated in the following corollary:

2.3.4 Corollary. *A tight subset $\mathcal{A} \subseteq \mathcal{P}(E)$ admits at least one limit point in $\mathcal{P}(E)$.*

2.3.2 Skorokhod topology and the convergence of processes

In the following, we introduce various modes of convergence for a sequence of $\mathbb{R}^d$-valued, càdlàg processes. We consider such processes as random variables which take their values in the *Skorokhod space* $\mathcal{D}(\mathbb{R}^d) := \mathcal{D}(\mathbb{R}_+, \mathbb{R}^d)$ of all càdlàg mappings from $\mathbb{R}_+$ to $\mathbb{R}^d$. Skorokhod (1956) introduced a topology on this space which makes it a Polish space. We refer to Sections VII.1 and VII.2 of Jacod and Shiryaev (2003) for further details on this topology and to Billingsley (1995) for details on the space $\mathcal{D}([0,T];\mathbb{R}^d)$ where $T < \infty$.

2.3.5 Definition. Let $(X^n)_{n\in\mathbb{N}}$ be a sequence of càdlàg processes and X be another càdlàg process. We say that

(i) X^n *converges in law* to X if the distributions $\mathscr{L}(X^n)$ of the $\mathcal{D}(\mathbb{R}^d)$-valued random variables X^n converge weakly in $\mathcal{P}(\mathcal{D}(\mathbb{R}^d))$ to the distribution $\mathscr{L}(X)$;

(ii) X^n *converges almost surely* (resp., *in probability*) to X if the $\mathcal{D}(\mathbb{R}^d)$-valued random variables X^n converge almost surely (resp., in probability) to X w. r. t. to the Skorokhod topology;

(iii) X^n *converges uniformly on compacts in probability* (or *in ucp*) to X if, for all $t \geq 0$, we have $\sup_{s\leq t}\|X^n_s - X_s\| \to 0$.

We write $X^n \overset{\mathscr{L}}{\Rightarrow} X$ (resp., $X^n \overset{\text{a.s.}}{\Rightarrow} X$; resp., $X^n \overset{\mathbb{P}}{\Rightarrow} X$; resp., $X^n \overset{\text{ucp}}{\Longrightarrow} X$). ⋄

We note that, as in the random variable case, convergence in law of stochastic processes is equivalent to the convergence of $\mathbb{E}\, f(X^n)$ to $\mathbb{E}\, f(X)$ for every bounded continuous function f; here, f is continuous w. r. t. Skorokhod's topology on the space $\mathcal{D}(\mathbb{R}^d)$. *Stable convergence in law* of processes is defined analogously to Definition 2.3.1.

We summarise important classical results about Skorokhod convergence, which we use in subsequent chapters of this thesis.

2.3.6 Proposition (Höpfner, Jacod, and Ladelli (1990) eqs. 3.2–5).
Let $(A^n)_{n\in\mathbb{N}}$ and $(B^n)_{n\in\mathbb{N}}$ be sequences of non-decreasing, càdlàg processes, let $(X^n)_{n\in\mathbb{N}}$ be a sequence of càdlàg processes, and let A and X be càdlàg processes. In addition, we set $U(n,t) := \inf\{s \geq 0 : A^n_s > t\}$ and $U(t) := \inf\{s \geq 0 : A_s > t\}$.

(i) *If* $(A^n, X^n) \overset{\mathscr{L}}{\Rightarrow} (A, X)$ *where* A *is a.s. continuous such that* $\mathbb{P}(U_{t-} \neq U_t) = 0$ *for all* t*, then* $(A^n, X^n, (U(n,t_i), X^n_{U(n,t_i)})_{i\leq k})) \overset{\mathscr{L}}{\Rightarrow} (A, X, (U(t_i), X_{U(t_i)})_{i\leq k}))$ *for all* $k \in \mathbb{N}^*$ *and* $t_i \geq 0$.

(ii) *Let* $\varepsilon_n \to 0$. *If* $X^n \overset{\mathscr{L}}{\Rightarrow} X$ *where* X *is a.s. continuous, then* $X^n_{t+\varepsilon_n} \overset{\mathscr{L}}{\Rightarrow} X_t$. *If* $A^n \overset{\mathscr{L}}{\Rightarrow} A$ *and* A *is as in* (i)*, then* $U(n, t+\varepsilon_n) \to U(t)$.

(iii) *If* $(A^n, X^n) \overset{\mathscr{L}}{\Rightarrow} (A, X)$ *where* (A, X) *is a.s. continuous, then* $(A^n, X^n_{A^n}) \overset{\mathscr{L}}{\Rightarrow} (A, X_A)$.

(iv) *If* $(A^n, X^n) \overset{\mathscr{L}}{\Rightarrow} (A, X)$ *and* $(A^n, B^n) \overset{\mathscr{L}}{\Rightarrow} (A, A)$*, then* $(A^n, X^n, B^n) \overset{\mathscr{L}}{\Rightarrow} (A, X, A)$. *Moreover, we have* $(A^n, A'^n) \overset{\mathscr{L}}{\Rightarrow} (A, A)$ *where* $A'^n_t := B^n_{t\wedge\sigma(n)} + A^n_t - A^n_{t\wedge\sigma(n)}$ *with* $\sigma(n) = \inf\{t \geq 0 : B^n_t > u\}$ *for some* $u \geq 0$.

Families as in the previous proposition, whose limit points are all laws of continuous processes, play a special role:

2.3.7 Definition. A sequence $(X^n)_{n\in\mathbb{N}}$ of processes is called C-tight if it is tight and if every limit point of the family $\{\mathscr{L}(X^n) : n \in \mathbb{N}\}$ is the law of a continuous process.$_\diamond$

2.3.8 Proposition (Jacod and Shiryaev (2003) VI.3.33). *Let* $(Y^n)_{n\in\mathbb{N}}$ *be a C-tight sequence of* d*-dimensional processes; let* $(Z^n)_{n\in\mathbb{N}}$ *be a tight (resp., C-tight) sequence of* d'*-dimensional processes. Then*

(i) *the sequence* $(Y^n, Z^n)_{n\in\mathbb{N}}$ *of* $(d+d')$*-dimensional processes is tight (resp., C-tight); and*

(ii) *if* $d = d'$*, then the sequence* $(Y^n + Z^n)_{n\in\mathbb{N}}$ *is tight (resp., C-tight).*

2.3.3 Martingale limit theorems

A general, generic scheme to prove limit theorems for stochastic processes was suggested by Prokhorov:

2.3.9 Theorem (Jacod and Shiryaev (2003) VI.3.18).
A sequence $(X^n)_{n\in\mathbb{N}}$ *of càdlàg processes converges to a process* X *if, and only if,*

(i) *the sequence* $(X^n)_{n\in\mathbb{N}}$ *is tight (that is, the family* $\{\mathscr{L}(X^n) : n \in \mathbb{N}\}$ *is tight); and*

(ii) *the law $\mathscr{L}(X)$ is the only possible limit point of the family $\{\mathscr{L}(X^n) : n \in \mathbb{N}\}$.*

For our purposes, the following martingale (central) limit theorem and its finite-dimensional version are the most important:

2.3.10 Theorem (Jacod and Shiryaev (2003) VIII.3.6 and VIII.3.24). *Let X be a continuous Gaussian martingale with characteristics $(0, C, 0)$; let $(X^n)_{n\in\mathbb{N}}$ be a sequence of locally square-integrable martingales with characteristics $(B^n, C^n, \mathfrak{n}^n)$ and $X_0^n = 0$; and let $D \subseteq \mathbb{R}_+$.*

(i) *If $B_t^n \xrightarrow{\mathbb{P}} 0$ and $\sup_{s\le t} |\Delta X_s^n| \xrightarrow{\mathbb{P}} 0$ and*

$$\text{either} \quad [X^n, X^n]_t \xrightarrow{\mathbb{P}} C_t \quad \text{or} \quad \langle X^n, X^n\rangle_t \xrightarrow{\mathbb{P}} C_t \tag{2.3.1}$$

holds for all $t \in D$, then

$$\text{for all } k \in \mathbb{N}^* \text{ and } t_1, \dots, t_k \in D: \quad (X_{t_1}^n, \dots, X_{t_k}^n) \xrightarrow{\mathscr{L}} (X_{t_1}, \dots, X_{t_k}). \tag{2.3.2}$$

(ii) *Suppose that D is dense in $\mathbb{R}_+$ and that the following "Lindeberg condition" holds:*

$$\Big(\|x\|^2 \mathbb{1}_{\{\|x\|>\varepsilon\}}\Big) \star \mathfrak{n}_t^n \xrightarrow[n\to\infty]{\mathbb{P}} 0 \quad \text{for all } t \ge 0, \varepsilon > 0. \tag{2.3.3}$$

Then $X^n \overset{\mathscr{L}}{\Rightarrow} X$ if, and only if, $[X^n, X^n]_t \xrightarrow{\mathbb{P}} C_t$ for all $t \in D$ and also if, and only if, $\langle X^n, X^n\rangle_t \xrightarrow{\mathbb{P}} C_t$ for all $t \in D$.

2.4 Markov chains

Before we come to the class of continuous-time Markov processes, we dedicate this section to give a résumé of their discrete-time counterparts: The Markov chains. The presentation in this section is mostly based on the book of Meyn and Tweedie (1993). Let $E = (E, \mathscr{E})$ be a fixed measurable space.

2.4.1 Definition. (i) A kernel $P : E \times \mathscr{E} \to \mathbb{R}_+$ is called a *transition kernel* if $P(x, E) = 1$ for all $x \in E$. For a kernel P, a σ-finite measure ν, and a measurable function g on E, we denote

$$Pg(x) := \int P(x, \mathrm{d}y)g(y), \quad \nu P(A) := \int \nu(dx)P(x, A), \quad \nu(g) := \int \nu(\mathrm{d}x)g(x).$$

(ii) Let π be a probability measure on E and let P be a transition kernel on E. A family $X = (X_n)_{n\in\mathbb{N}}$ of random variables, defined on some probability space $(\Omega, \mathscr{F}, \mathbb{P}^\pi)$ endowed with a discrete-time filtration $\mathfrak{F} := (\mathscr{F}_n)_{n\in\mathbb{N}}$, is called a *Markov chain with initial law π and transition kernel P* if X is $\mathfrak{F}$-adapted,

a) $\mathbb{P}^\pi(X_0 \in A) = \pi(A)$ for each $A \in \mathscr{E}$; and

b) $\mathbb{E}^\pi[f(X_{n+1}) \mid \mathscr{F}_n] = Pf(X_n)$ for every $n \in \mathbb{N}$ and every bounded $\mathscr{E}$-measurable function f.

(iii) A collection $X = (\Omega, \mathscr{F}, \mathfrak{F}, (X_n)_{n\in\mathbb{N}}, (\mathbb{P}^x)_{x\in E})$ is called a *Markov chain with transition kernel P* if, under every law $\mathbb{P}^x$, $(X_n)_{n\in\mathbb{N}}$ is a Markov chain with initial law ϵ_x and transition kernel P. ⋄

We refer to Section I.2 of Revuz (1984) for a rigorous introduction and the proof of existence of Markov chains. For a probability measure π, we denote the expectation w. r. t. to the law $\mathbb{P}^\pi := \int \mathbb{P}^x\, \pi(\mathrm{d}x)$ by $\mathbb{E}^\pi$. Furthermore, we denote the so-called *n-step transition kernel* by P^n; these are inductively defined by

$$P^0(x, A) = \epsilon_x(A) \quad \text{and} \quad P^n(x, A) := \int P(x, \mathrm{d}y)P^{n-1}(y, A). \tag{2.4.1}$$

The Chapman–Kolmogorov equations (see Theorem 3.4.2 of Meyn and Tweedie, 1993), that is, $P^{m+n} = P^m P^n$ for all $m, n \in \mathbb{N}$, are key to much of the following analysis of Markov chains.

2.4.1 Irreducibility; Small sets; Periodicity; Feller chains

Many of the notions and concepts developed in this and the following subsection will find their counterpart in the continuous-time context.

2.4.2 Definition. Let $A \in \mathscr{E}$. We call

$$\eta_A := \sum_{k=1}^{\infty} \mathbb{1}_A(X_n), \quad \text{and} \quad \tau_A := \min\{n \geq 1 : X_n \in A\},$$

the *sojourn time* of, and the *return time* on A, respectively. ⋄

The concept of irreducibility is best defined in terms of the return times of sets:

2.4.3 Definition. A Markov chain $X = (X_n)_{n \in \mathbb{N}}$ is called *φ-irreducible* if there exists a measure φ on E such that $\varphi(A) > 0$ implies $\mathbb{P}^x(\tau_A < \infty) = 1$ for all x. ⋄

2.4.4 Proposition (Meyn and Tweedie (1993) 4.2.2). *Let X be a φ-irreducible Markov chain. Then there exists a probability ψ on E such that X is ψ-irreducible and, for every φ' for which X is φ'-irreducible, we have $\psi \gg \varphi'$ and, for every ψ-null set $A \in \mathscr{E}$, we have that $\{y : \mathbb{P}^y(\tau_A < \infty) > 0\})$ is ψ-null.*

Such a measure ψ is also called a *maximal irreducibility measure* for X. It is unique up to equivalence. Thus, the following notions are well-defined:

2.4.5 Definition. Let X be a ψ-irreducible Markov chain. A set $A \in \mathscr{E}$ is called *full* if its complement is ψ-null; it is called *absorbing* if $P(x, A) = 1$ for all $x \in A$. Moreover, we set $\mathscr{E}^+ := \{A \in \mathscr{E} : \psi(A) > 0\}$. ⋄

2.4.6 Proposition (Meyn and Tweedie (1993) 4.2.3). *Let X be ψ-irreducible. Then every absorbing set is full, and every full set contains a non-empty, absorbing set.*

An important – at first glance not apparent role – for the asymptotic behaviour of a Markov chain play small sets and petite sets. To prevent confusion, we emphasise that we strictly follow the nomenclature of Meyn and Tweedie (1992).

2.4.7 Definition.

(i) A set $C \in \mathscr{E}$ is called a *small set* if there exists an $m \in \mathbb{N}^*$ and a non-trivial measure ν_m on E such that $P^m(x, A) \geq \nu_m(A)$ for all $x \in C$ and $A \in \mathscr{E}$; C is also called *ν_m-small*.

(ii) A set $C \in \mathscr{E}$ is called a *petite set* if there exits a probability ρ on $\mathbb{N}^*$ and a non-trivial measure ν_ρ on E such that $\sum_{k=1}^{\infty} \rho(k) P^k(x, A) \geq \nu_\rho(A)$ for all $x \in C$ and $A \in \mathscr{E}$. ⋄

Apparently, set $\rho = \epsilon_m$, every small set is petite. The following theorem guarantees the existence of small sets for ψ-irreducible chains.

2.4.8 Theorem (Meyn and Tweedie (1993) 5.2.2 and 5.2.4). *Let X be ψ-irreducible.*

(i) For every $A \in \mathscr{E}^+$, there exists an $m \geq 1$ and a ν_m-small set $C \subseteq A$ such that $C \in \mathscr{E}^+$ and $\nu_m(C) > 0$.

(ii) There exists a countable covering of X by small sets.

(iii) If $C \in \mathscr{E}^+$ is ν_m-small, then there exists an $m' \in \mathbb{N}^*$ such that C is $\nu_{m'}$-small with $\nu_{m'}(C) > 0$.

2.4.9 Definition. Let X be ψ-irreducible.

(i) A collection $D_0, D_1, \dots, D_{p-1} \in \mathscr{E}$ of disjoint sets is called a *p-cycle* if

  a) for every $i \in \mathbb{Z}_p$ and $x \in D_i$, we have $P(x, D_{i+1}) = 1$, and

  b) the complement of $\cup_{i \in \mathbb{Z}_p} D_i$ is ψ-null.

(ii) The largest $p \in \mathbb{N}^*$ for which a p-cycle exists is called the *period* of X.

(iii) We call X *aperiodic* if its period is one. ⋄

2.4.10 Theorem (Meyn and Tweedie (1993) 5.4.4, 5.4.6 and 5.4.7). *Let X be a ψ-irreducible Markov chain.*

(i) Let $C \in \mathscr{E}^+$ be a ν_m-small set with $\nu_m(C) > 0$ and let p denote the greatest common divisor of the set

$$\Big\{ n \in \mathbb{N}^* : C \text{ is } \nu_n\text{-small with } \nu_n = \zeta_n \nu_m \text{ for some } \zeta_n > 0 \Big\}.$$

Then there exists a p-cycle $D_0, \dots, D_{p-1} \in \mathscr{E}$ and, moreover, p is the period of X.

(ii) Furthermore, let X^p denote the so-called sampled Markov chain with transition kernel P^p. Then each set D_i is an absorbing, ψ-irreducible set for X^p and the restriction of X^p to each D_i is aperiodic.

(iii) If X is aperiodic, then every petite set is small.

In general, the identification of petite sets is a tedious task. In an important special case, however, it is straightforward.

2.4.11 Definition. (i) We denote by $\mathcal{C}(E)$ (resp., $\mathcal{C}_b(E)$) the class of *continuous* (resp., *bounded continuous* functions on E.

(ii) A kernel P is called *weak Feller* if $Pg \in \mathcal{C}_b(E)$ for every $g \in \mathcal{C}_b(E)$.

(iii) A kernel P is called *strong Feller* if $Pg \in \mathcal{C}_b(E)$ for every bounded, $\mathscr{E}$-measurable function g. $\diamond$

Obviously, the strong Feller property implies the weak Feller property.

2.4.12 Theorem (Meyn and Tweedie (1993) 6.2.5 (ii) and 6.2.9). *Let X be a ψ-irreducible Markov chain with transition kernel P. If P is weak Feller and if the support of the measure ψ has non-empty interior, then every compact set is petite.*

2.4.2 Recurrence; Invariant measure; Ergodicity

Usually, a presentation of the recurrence properties of Markov chains would also take care of its complementary notion, the transience. Since the transient case is irrelevant in the sequel, we neglect it.

2.4.13 Definition. (i) Let X be ψ-irreducible. We call X *Harris recurrent* if

$$\psi(A) > 0 \implies \mathbb{P}^x(\eta_A = \infty) = 1 \quad \text{for all } x \in E. \tag{2.4.2}$$

(ii) A σ-finite measure μ on E is called an *invariant* measure for X if $\mu P = \mu$. $\diamond$

2.4.14 Proposition (Meyn and Tweedie (1993) 9.1.7 (ii)). *Let X be ψ-irreducible. If there exists a petite set $C \in \mathscr{E}$ such that $\mathbb{P}^x(\eta_C = \infty) = 1$ for all $x \in E$, then X is Harris recurrent.*

2.4.15 Theorem (Meyn and Tweedie (1993) 10.4.4 and 10.4.5). *Let X be Harris recurrent. Then there exists an invariant measure μ which is unique up to constant multiples. If, in addition, X is aperiodic, then a measure is invariant for X if, and only if, it is invariant for every sampled chain X^m with transition kernel P^m, $m \in \mathbb{N}^*$.*

As an essential consequence we obtain the following proposition:

2.4.16 Proposition. *Let X be Harris recurrent with invariant measure μ, period $p > 1$, and p-cycle $D_0, \dots, D_{p-1}$. Then the restriction of the sampled chain X^p with transition kernel P^p to each D_i is aperiodic and Harris recurrent with invariant measure μ_i given by $\mu_i(A) = \mu(A \cap D_i)$.*

Proof. By Theorem 2.4.10 (ii), the sampled chain restricted to each of the sets D_i is aperiodic. Since μ is invariant for P, by definition of the p-cycle we have $\mu_i P = \mu_{i+1}$. By iteration, we obtain $\mu_i P^p = \mu_i$. In other words, μ_i is invariant for P^p. Finally, for each measurable $A \subseteq D_i$ with $\mu_i(A) > 0$, we have $\mathbb{P}^x(\eta_A = \infty) = 1$ for all $x \in D_i$. Thus, we also have Harris recurrence. □

Recurrent processes are further classified in terms of their invariant measure.

2.4.17 Definition. (i) Let X be Harris recurrent with invariant measure μ. We call X *positive (Harris recurrent)* if $\mu(E) < \infty$. Otherwise, we call X *null (Harris recurrent)*. In the positive case, we call the unique invariant probability measure, the *stationary distribution*.

(ii) A Markov chain with transition kernel P and invariant measure μ is called *ergodic* if, for every initial probability π on E, the total variation norm

$$\|\pi P^n - \mu\| := \sup_{\{f:|f|\le 1\}} |\pi P^n f - \mu(f)| \tag{2.4.3}$$

of the signed measures $\pi P^m - \mu$ converges to zero as $n \to \infty$. The chain is called *uniformly ergodic* if

$$\lim_{n\to\infty} \sup_{x\in E} \|\epsilon_x P^n - \mu\| = 0. \tag{2.4.4}$$

2.4.18 Theorem (Meyn and Tweedie (1993) 13.3.3). *An aperiodic, positive Harris recurrent Markov chain is ergodic.*

2.4.19 Theorem (Meyn and Tweedie (1993) 16.2.1 and 16.2.2). *A Markov chain with invariant measure μ is uniformly ergodic if, and only if, there exists a constant $\zeta < 1$ such that $\|\epsilon_x P^n - \mu\| \le \zeta^n$ for every $n \in \mathbb{N}^*$ and $x \in E$. A ψ-irreducible, aperiodic Markov chain with invariant measure μ is uniformly ergodic if, and only if, the state space is petite.*

2.5 Markov processes

In this section, we turn our attention to the class of processes which we study at the core of this thesis: The Markov processes. In the previous section, we have introduced their discrete-time counterparts, the Markov chains, and introduced various notions which we will analogously introduce in the continuous-time case. Due to the abundance of literature on the theory of Markov processes, we only present selected topics. The presentation is based on the monograph Getoor (1975) as well as a couple of individual papers (see also Blumenthal and Getoor, 1968; Sharpe, 1988).

2.5.1 Résumé of the theory of Ray processes; Hunt processes

This subsection is dedicated to briefly touch the embedding of the theory of Markov processes into the general theory of stochastic processes (recall Section 2.1).

The theory of Markov processes, put differently, is the theory of transition semi-groups and resolvents. Let $E = (E, \mathscr{E})$ be a fixed measurable space.

2.5.1 Definition. Let $(R^\alpha)_{\alpha>0}$ be a family of kernels on E. Then $(R^\alpha)_{\alpha>0}$ is called a *sub-Markov resolvent* (resp. *Markov resolvent*) if

i) $\alpha R^\alpha 1 \leq 1$ (resp. $\alpha R^\alpha 1 = 1$) for each $\alpha > 0$,

ii) $R^\alpha - R^\beta = (\beta - \alpha) R^\alpha R^\beta$ for all $\alpha, \beta > 0$. ◇

It is immediate from (i) that each kernel is bounded. Hence, (ii) is well-defined. Additionally, we note that $R^\alpha(\cdot, A)$ is $\mathscr{E}^{\mathrm{u}}$-measurable whenever $A \in \mathscr{E}^{\mathrm{u}}$. Therefore, $(R^\alpha)_{\alpha>0}$ is also a resolvent on $(E, \mathscr{E}^{\mathrm{u}})$. By abuse of notation, in this subsection, we write $f \in \mathscr{E}$ (resp., $f \in \mathrm{p}\mathscr{E}$; resp., $f \in \mathrm{b}\mathscr{E}$) for f being an $\mathscr{E}$-measurable (resp., a positive $\mathscr{E}$-measurable; resp., a bounded $\mathscr{E}$-measurable) function.

2.5.2 Definition. Let $(R^\alpha)_{\alpha>0}$ be a resolvent on $(E, \mathscr{E})$, let $f \in \mathrm{p}\mathscr{E}$, and $\alpha \geq 0$. If $\beta R^{\alpha+\beta} \leq f$ for all $\beta > 0$, then f is called *α-supermedian*. An α-supermedian function is called *α-excessive* if $\lim_{\beta\to\infty} \beta R^{\alpha+\beta} f = f$ pointwise. We denote by $\mathcal{S}^\alpha$ (resp., $\mathcal{E}^\alpha$) the class of all continuous, α-supermedian functions (resp., of all α-excessive functions). In addition, we write $\mathcal{S}^\infty := \bigcup_{\alpha>0} \mathcal{S}^\alpha$. ◇

For the remainder of this subsection, we suppose that E is a compact metric space.

2.5.3 Definition. (i) A sub-Markov resolvent $(R^\alpha)_{\alpha>0}$ is called a *Ray resolvent* if $R^\alpha\mathcal{C} \subseteq \mathcal{C}$ for all $\alpha > 0$ and for each pair $x, y \in E$ with $x \neq y$ there exists a function $f \in \mathcal{S}^\infty$ with $f(x) \neq f(y)$.

(ii) A family $(P_t)_{t\geq 0}$ of sub-Markov kernels on E is called a *sub-Markov semigroup* if $P_{t+s} = P_t P_s$ for all $s, t \geq 0$. ⋄

Remark. We note that P_0 is not assumed to be the identity map.

2.5.4 Theorem (Getoor (1975) 3.6). *Let $(R^\alpha)_{\alpha>0}$ be a Ray resolvent on E. Then there exists a unique sub-Markov semigroup $(P_t)_{t\geq 0}$ such that*

(i) $t \mapsto P_t f(x)$ is right continuous on $\mathbb{R}_+$ for each $x \in E$ and $f \in \mathcal{C}$;

(ii) $R^\alpha f = \int_0^\infty e^{-\alpha t} P_t f \mathrm{d}t$ for all $\alpha > 0$ and $f \in \mathcal{C}$.

In addition,

(iii) a function f is α-supermedian if, and only if, $e^{-\alpha t} P_t f \leq f$ for all $t \geq 0$. Moreover, we have $e^{-\alpha t} P_t f \to P_0 f \in \mathcal{E}^\alpha$ as $t \downarrow 0$.

(iv) Let $D := \{x \in E : \lim_{\alpha\to\infty} \alpha R^\alpha f(x) = f(x)\ \forall f \in \mathcal{C}\}$. Then $D \in \mathscr{E}$ and $P_0(x, \cdot) = \epsilon_x$ if, and only if, $x \in D$. Moreover, $P_t(x, \cdot)$ is carried by D for all $x \in E$ and $t \geq 0$.

(v) $(P_t)_{t\geq 0}$ is Markovian if and only if $(R^\alpha)_{\alpha>0}$ is Markovian.

The set D in (iv) is called the set of *non-branch points* of $(R^\alpha)_{\alpha>0}$ (or of $(P_t)_{t\geq 0}$). Accordingly, $B := E \setminus D$ is called the set of *branch points* of $(R^\alpha)_{\alpha>0}$ (or of $(P_t)_{t\geq 0}$).

Remark. It is standard in literature to reduce the sub-Markovian case to the Markovian case by, firstly, attaching a *cemetery state* to the state space, secondly, constructing a Markov resolvent and, lastly, constructing a Markov semi-group by means of Theorem 2.5.4 (see Getoor, 1975, p. 16).

2.5.5 Definition. (i) Let π be a probability measure on E and let $(P_t)_{t\geq 0}$ be a Markov semi-group. A family $X = (X_t)_{t\geq 0}$ of random variables, defined on some filtered probability space $(\Omega, \mathscr{F}, \mathfrak{F}, \mathbb{P}^\mu)$, is called a *Markov process with initial law π and transition semi-group $(P_t)_{t\geq 0}$* if X is adapted to $\mathfrak{F}$,

a) $\mathbb{P}^\pi(X_0 \in A) = \pi(A)$ for each $A \in \mathscr{E}$; and

b) $\mathbb{E}^\pi[f(X_{t+s}) \mid \mathscr{F}_s] = P_t f(X_s)$ for every $t, s \geq 0$ and bounded measurable function f on E.

(ii) A Markov process $X = (\Omega, \mathscr{F}, \mathfrak{F}, (X_t)_{t\geq 0}, \mathbb{P}^\mu)$ is called *strong Markov* if, for every $\mathfrak{F}$-stopping time, we have

$$\mathbb{E}^\pi[f(X_{T+t})\mathbb{1}_{\{T<\infty\}} \mid \mathscr{F}_T] = P_t f(X_T)\mathbb{1}_{\{T<\infty\}} \tag{2.5.1}$$

for every $t \geq 0$ and bounded measurable function f on E. ⋄

2.5.6 Theorem (Getoor (1975) 5.1). *Let $(R^\alpha)_{\alpha>0}$ be a Markov and a Ray resolvent, and let $(P_t)_{t\geq 0}$ denote the semi-group constructed from $(R^\alpha)_{\alpha>0}$ by means of Theorem 2.5.4. Furthermore, let D denote the set of non-branch points of $(R^\alpha)_{\alpha>0}$ and denote by*

$$\mathcal{W} := \{w \in \mathcal{D}(E) : w(t) \in D \ \forall t \geq 0\}$$

the class of all càd mappings from $\mathbb{R}_+$ to D such that left-limits exist in E on $\mathbb{R}_+^*$. Let X be the canonical process on $\mathcal{W}$ given by $X_t(w) = w(t)$ and set $\mathscr{G}^0 := \sigma(X_t : t \geq 0)$ and $\mathscr{G}_t^0 = \sigma(X_s : s \leq t)$. Then, for every probability measure π on E, there exists a law $\mathbb{P}^\pi$ on $(\mathcal{W}, \mathscr{G}^0)$ such that $X := (\mathcal{W}, \mathscr{G}^0, (\mathscr{G}_t^0)_{t\geq 0}, (X_t)_{t\geq 0}, \mathbb{P}^\pi)$ is a Markov process with initial law μP_0 and transition semi-group $(P_t)_{t\geq 0}$.

By virtue of this theorem, of every Ray resolvent $(R^\alpha)_{\alpha>0}$, there exists a càdlàg realisation. In the following, we summarise some of its interesting properties.

We start with some notation: Let $\theta = (\theta_t)_{t\geq 0}$ denote the *semi-group of shift operators* on $\mathcal{W}$ given by $X_t \circ \theta_s = X_{t+s}$ for all $t, s \geq 0$. For every probability π on $(E, \mathscr{E})$ we denote by $\mathscr{G}^\pi$ the $\mathbb{P}^\pi$-completion of $\mathscr{G}^0$, and by $\mathscr{N}^\pi(\mathscr{G})$ the σ-ideal of $\mathbb{P}^\pi$-null sets in $\mathscr{G}^\pi$. Moreover, $\mathscr{G}_t^\pi := \mathscr{G}_t \vee \mathscr{N}^\mu(\mathscr{G})$. Furthermore, we consider the family $(R^\alpha)_{\alpha>0}$ as a resolvent on $(E,\mathscr{E}^{\mathrm{u}})$ rather than on $(E,\mathscr{E})$. Thus, we change slightly the definition of the α-excessive functions.

2.5.7 Proposition (Getoor (1975) 5.6). *Let $f \in \mathrm{p}\mathscr{E}^{\mathrm{u}}$. Then f is α-excessive if, and only if, $e^{-\alpha t}P_t f \uparrow f$ as $t \downarrow 0$.*

2.5.8 Definition. A numerical function $f \in \mathscr{E}^{\mathrm{u}}$ is called *nearly Borel* (relative to X) if for every initial law π there exist $g, h \in \mathscr{E}$ such that $g \le f \le h$ and the processes $g(X)$ and $h(X)$ are $\mathbb{P}^\pi$-indistinguishable. ⋄

The class $\mathscr{E}^{\mathrm{n}} := \{A \in \mathscr{E}^{\mathrm{u}} : \mathbb{1}_A \text{ is nearly Borel}\}$ forms a σ-algebra, and f is nearly Borel if, and only if, f is $\mathscr{E}^{\mathrm{n}}$-measurable. Also, we have $\mathscr{E} \subseteq \mathscr{E}^{\mathrm{n}} \subseteq \mathscr{E}^{\mathrm{u}}$.

2.5.9 Theorem (Getoor (1975) 5.8 and 5.11). *(i) For every probability π on E, the filtration $\mathfrak{G}^\pi := (\mathscr{G}^\pi_t)_{t\ge 0}$ is right-continuous; that is, $(\mathcal{W}, \mathscr{G}^\pi, \mathfrak{G}^\pi, \mathbb{P}^\pi)$ satisfies the usual hypotheses of the general theory of stochastic processes.*

(ii) The stochastic process X from Theorem 2.5.6 satisfies the strong Markov property relative to $\mathfrak{G}^\pi$.

(iii) Let $\alpha > 0$. Then $\mathcal{E}^\alpha \subseteq \mathscr{E}^{\mathrm{n}}$. In addition, for each $f \in \mathcal{E}^\alpha$, the process $f(X)$ is càdlàg.

(iv) Let π be a probability on E. Let $(T_n)_{n\in\mathbb{N}}$ be an increasing sequence of $\mathfrak{G}^\mu$-stopping times. Set $T := \sup_n T_n$ and $\Lambda := \{w \in \mathcal{W} : T(w) < \infty, T_n(w) < T(w)\ \forall n \in \mathbb{N}\}$. Then, for every bounded, universally measurable f on E,

$$\mathbb{E}^\pi\left[f\circ X_T \mathbb{1}_{\{T<\infty\}} \,\middle|\, \bigvee_n \mathscr{G}^\pi_{T_n} \right] = f\circ X_T \mathbb{1}_{\{T<\infty\}} \mathbb{1}_{\Lambda^c} + P_0 f(X_{T-})\mathbb{1}_\Lambda.$$

2.5.10 Corollary (Getoor (1975) 5.16). *On $\{X_{T-} \in D, T < \infty\}$, we have $X_{T_n} \to X_T$ $\mathbb{P}^\pi$-almost surely. If there are no branch points, that is, if $D = E$, then X is quasi-left continuous.*

2.5.11 Definition. Let $x \in B$. Then x is called a *degenerate branch point* if there exists a $y \in E$ with $\epsilon_x P_0 = \epsilon_y$. The set of degenerate branch points is denoted by B_{d}. ⋄

2.5.12 Theorem (Getoor (1975) 6.2, 6.4, 6.7 and 7.3). *Let π be a probability on E and T be a $\mathfrak{G}^\pi$-stopping time.*

(i) *The set B_{d} is Borel.*

(ii) *Suppose T is predictable. We have $\mathscr{G}_T^\pi = \mathscr{G}_{T-}^\pi$ if, and only if, $\mathbb{P}^\pi(X_{T-} \in B \setminus B_{\mathrm{d}}, 0 < T < \infty) = 0$. Therefore, if $B = B_{\mathrm{d}}$ (in particular, if B is empty), then the filtration $\mathfrak{G}^\pi$ is quasi-left continuous for every π.*

(iii) *If $X_T = X_{T-}$ $\mathbb{P}^\pi$-a. s. on $\{0 < T < \infty\}$, then T is predictable and $\mathscr{G}_T^\pi = \mathscr{G}_{T-}^\pi$.*

(iv) *Let $A = \{0 < T < \infty, X_{T-} \in D, X_T \neq X_{T-}\}$. Then $T_A := T\mathbb{1}_A + \infty\mathbb{1}_{A^c}$ is the totally inaccessible part of T.*

In summary, we have introduced resolvents, semi-groups, and Markov processes. We have presented existence results for an important special case – the Ray resolvents. The corresponding Ray process constructed by Theorems 2.5.6 and 2.5.9 is defined on a probability space which satisfies the usual conditions, has càdlàg sample paths, and is strong Markov. By Corollary 2.5.10, moreover, we are given a criterion for quasi-left continuity. For presentational purposes, we end our presentation of the general theory at this point (although, there would still be much more to say). Throughout the remainder of this thesis, all Markov processes which we deal with are supposed to satisfy the hypothesis (A) of Hunt (1957):

2.5.13 Definition. A Markov process $X = (\Omega, \mathscr{F}, \mathfrak{F}, (X_t)_{t\geq 0}, (\theta_t)_{t\geq 0}, (\mathbb{P}^x)_{x\in E})$ with values in some locally compact, separable space E is called a *Hunt process* if its sample paths are almost surely càdlàg and if it is strong Markov and quasi-left continuous. ⋄

2.5.2 Recurrence; Additive functionals; Ergodic theorems

The notions of recurrence and invariant measure are defined analogously to the discrete-time case (see Definitions 2.4.13 and 2.4.17).

2.5.14 Definition. Let X be a Hunt process with transition semi-group $(P_t)_{t\geq 0}$.

(i) We call X *Harris recurrent* if there exists a σ-finite measure φ on E such that

$$\varphi(A) > 0 \implies \mathbb{P}^x\left(\int_0^\infty \mathbb{1}_A(X_s)\mathrm{d}s = \infty\right) = 1 \quad \text{for all } x \in E. \tag{2.5.2}$$

(ii) A σ-finite measure μ is called *invariant* if $\mu P_t = \mu$ for all $t \geq 0$.

(iii) If X is Harris recurrent with invariant measure μ, then we call X *positive (Harris recurrent)* if $\mu(E) < \infty$; otherwise, we call X *null (Harris recurrent)*. In the positive case, we call the unique invariant probability measure, the *stationary distribution*. ⋄

2.5.15 Theorem (Azéma, Kaplan-Duflo, and Revuz (1967) Théorème I.3).
If X is Harris recurrent, then there exists an invariant measure μ which is unique up to constant multiples.

Next, we introduce the notion of additive functionals:

2.5.16 Definition. A process $H = (H_t)_{t\geq 0}$ is called a *(perfect, homogeneous) additive functional* of X if it is adapted to the filtration $\mathfrak{F}$ and if, for all $s, t \geq 0$, we have $H_{t+s} = H_t \circ \theta_s + H_s$. ⋄

Example. Let g be a measurable function on E and $x \in E$.

(i) The process $H_t := \int_0^t g(X_s)\mathrm{d}s$ is an absolutely continuous additive functional.

(ii) The process $H'_t := \sum_{s\leq t} g(\Delta X_s)\mathbb{1}_{\{\Delta X_s \neq 0\}}$ is a discontinuous additive functional.

(iii) The local time of X at x (see Blumenthal and Getoor, 1964, for details) is a continuous additive functional which is not absolutely continuous.

2.5.18 Definition. Let X be a Hunt process with invariant measure μ and H be an additive functional. We call H *integrable* if $\mu(H) := \mathbb{E}^\mu H_1 < \infty$. ⋄

Then the Chacon–Ornstein theorem for additive functionals reads as follows:

2.5.19 Theorem (Azéma et al. (1967) Théorème II.1). *Let H and H' be integrable additive functionals such that $\mu(H') > 0$. Then*

$$\frac{\mathbb{E}^x H_t}{\mathbb{E}^x H'_t} \xrightarrow[t\to\infty]{} \frac{\mu(H)}{\mu(H')} \quad \mu\text{-almost surely.} \tag{2.5.3}$$

And Birkhoff's theorem for additive functionals reads as follows:

2.5.20 Theorem (Azéma et al. (1967) Théorème II.2). *Let H and H' be given as in Theorem 2.5.19. Then, for every $x \in E$,*

$$\frac{H_t}{H'_t} \xrightarrow[t\to\infty]{} \frac{\mu(H)}{\mu(H')} \quad \mathbb{P}^x\text{-almost surely.} \tag{2.5.4}$$

2.5.3 Deterministic equivalents; Darling–Kac's condition; Mittag–Leffler process

Since $H'_t = t$ is an integrable additive functional if, and only if, $\mu(E) < \infty$, by Theorems 2.5.19 and 2.5.20, we obtain convergence for $t^{-1}H_t$ and $t^{-1}\,\mathbb{E}^x\,H_t$ in the positive recurrent case. In the null recurrent case, we obtain a more differentiated picture.

2.5.21 Definition. A non-decreasing, deterministic function $v : \mathbb{R}_+ \to \mathbb{R}_+$ is called a *deterministic equivalent* of a Markov process X if the families

$$\left\{\mathscr{L}(v(t)^{-1}H_t) \mid \mathbb{P}^\pi : t > 0\right\} \quad \text{and} \quad \left\{\mathscr{L}(v(t)H_t^{-1}) \mid \mathbb{P}^\pi : t > 0\right\} \tag{2.5.5}$$

are tight for every probability π on E and each non-decreasing, additive functional H of X with $0 < \mu(H) < \infty$. ⋄

As seen before, $v(t) = t$ is a deterministic equivalent in the positive recurrent case. Löcherbach and Loukianova (2008) showed that some deterministic equivalent exists whenever X is Harris recurrent. For further details on the generality of their approach, we also refer to Löcherbach and Loukianova (2011).

2.5.22 Definition. Let X be a Harris recurrent Markov process with invariant measure μ and resolvent $(R^\alpha)_{\alpha>0}$. We say that X satisfies *Darling–Kac's condition* if, for some $0 < \delta \leq 1$, there exists a function $v : \mathbb{R}_+ \to \mathbb{R}_+$ which is regularly varying of index δ at infinity such that

$$\frac{1}{v(1/\lambda)} R^\lambda g(x) \xrightarrow[t\to\infty]{} \mu(g) \quad \mu\text{-almost everywhere as } \lambda \downarrow 0 \tag{2.5.6}$$

for every μ-integrable function g on E. ⋄

We note that, in the positive recurrent case, Darling–Kac's condition holds with $\delta = 1$ and $v(t) = t/\mu(E)$.

2.5.23 Definition. For $0 < \alpha < 1$, let $K = (K_t)_{t\geq 0}$ denote the α-stable Lévy subordinator with Laplace transform $\mathbb{E}\,\mathrm{e}^{-\xi K_t} = \mathrm{e}^{-t\xi^\alpha}$ for $\xi, t \geq 0$. Set $L_t := \inf\{s > 0 : K_s > t\}$. We call $L = (L_t)_{t\geq 0}$ the *Mittag-Leffler process of order α*. By abuse of notation, we also call $L_t = t$ the *Mittag-Leffler process of order one*. ⋄

For a brief introduction to the properties of the Mittag-Leffler processes, we refer to Höpfner and Löcherbach (2003).

2.5.24 Theorem (Touati (1987) Théorème 3). *Let X be a Harris recurrent Markov process which satisfies Darling–Kac's condition for some $0 < \delta \leq 1$ and some v. Let $H = (H^1, \dots, H^n)$ be a μ-integrable additive functional of X with (component-wise) non-decreasing paths. Then, under every law $\mathbb{P}^\pi$, we have the following convergence in law in $\mathcal{D}(\mathbb{R}^n)$:*

$$\Big(v(t)^{-1}H^1_{st}, \dots, v(t)^{-1}H^n_{st}\Big)_{s\geq 0} \xrightarrow[t\to\infty]{\mathscr{L}} \Big(\mu(H^1)L_s, \dots, \mu(H^n)L_s\Big)_{s\geq 0}, \tag{2.5.7}$$

where L is the Mittag-Leffler process of order δ.

2.5.4 Jumps of Markov processes; Lévy system; Lévy kernel

2.5.25 Theorem (Benveniste and Jacod (1973) Théorème 1.1). *Let X be a Hunt process on some state space E. Then there exists a kernel F on E with $F(x, \{0\}) = 0$ and a non-decreasing, continuous additive functional H of X such that, for every Borel function $g : E \times E \to \mathbb{R}_+$, every probability π on E, and every $t > 0$,*

$$\mathbb{E}^\pi \sum_{0<s\leq t} g(X_{s-}, \Delta X_s)\mathbb{1}_{\{X_{s-}\neq X_s\}} = \mathbb{E}^\pi \int_0^t \mathrm{d}H_s \int_E F(X_s, \mathrm{d}y)g(X_s, y). \tag{2.5.8}$$

A pair (F, H) satisfying eq. (2.5.8) is called a *Lévy system* (see also Watanabe, 1964). We call F a *Lévy kernel* of the Hunt process X. If there exists a Lévy system (F, H) with $H_t = t$, we call the – then distinguished – kernel F of the Lévy system (F, t) the *(canonical) Lévy kernel* of X.

The proofs in Benveniste and Jacod (1973) show that the continuity of the additive functional H is intimately related to the quasi-left continuity of the process and, hence, to the absence of branch points (recall Corollary 2.5.10). In the terminology of

Section 2.2, the random measure $\mathrm{d}H_t \otimes F(X_t, \mathrm{d}y)$ on $\mathbb{R}_+ \times E$ is the predictable compensator of the jump measure of the càdlàg process X. For conditional expectations w. r. t. to the strict past of jump times, in this thesis we apply the following result.

2.5.26 Theorem (Weil (1971) Théorème 1). *Let X be a Hunt process with Lévy system (F, H) and let $A \subset \{(x, y) \in E \times E : x \neq y\}$. We set $T := \inf\{s > 0 : (X_{s-}, X_s) \in A\}$ and suppose that $(X_{T_-}, X_T) \in A$ almost surely on $\{0 < T < \infty\}$. Then, for every Borel function g on $E \times E$ and every probability π on E, we have*

$$\mathbb{E}^\pi \left[g(X_{T-}, X_T) \mathbb{1}_{0<T<\infty} \mid \mathscr{F}_{T-} \right] = F_A g(X_{T_-}) \mathbb{1}_{0<T<\infty} \quad \mathbb{P}^\pi\text{-almost surely}, \tag{2.5.9}$$

where

$$F_A g(x) := \begin{cases} \frac{\int F(x,\mathrm{d}y) g(x,y) \mathbb{1}_A(x,y)}{\int F(x,\mathrm{d}y) \mathbb{1}_A(x,y)}, & \text{if } 0 < \int F(x, \mathrm{d}y) \mathbb{1}_A(x, y) < \infty, \\ 0, & \text{otherwise.} \end{cases} \tag{2.5.10}$$

The estimation of jumps beyond the Lévy case

3 On non-parametric estimation of the Lévy kernel of Markov processes

This chapter is based on Ueltzhöfer (2012). The individual introduction in Section 3.1 has been edited for presentational purposes in view of the general introduction of this thesis (Chapter 1). Cross-references to the material presented in Chapter 2 have been added. The proof of Lemma 3.4.10 is presented in a more detailed version.

3.1 Introduction

In this chapter, we consider a Harris recurrent Markov process X which is an Itō semi-martingale. In view of Theorem 2.2.12, such a process is a solution of some stochastic differential equation

$$\begin{aligned} dX_t = b(X_t)dt + \sigma(X_t)dW_t + \int \delta(X_{t-},y)\mathbb{1}_{\{\|\delta(X_{t-},y)\|>1\}}\mathfrak{p}(dt,dy) \\ + \int \delta(X_{t-},y)\mathbb{1}_{\{\|\delta(X_{t-},y)\|\le 1\}}(\mathfrak{p}-\mathfrak{q})(dt,dy), \end{aligned} \tag{3.1.1}$$

with coefficients b, σ and δ; the SDE is driven by some Wiener process W and some Poisson random measure $\mathfrak{p}$ (with intensity measure $\mathfrak{q}(dt,dy) = dt \otimes \lambda(dy)$). The law of its jumps is more or less described by the kernel F where, for each x, the measure $F(x,\cdot)$ coincides with the image of the measure λ under the map $y \mapsto \delta(x,y)$ restricted to the set $\{y : \delta(x,y) \neq 0\}$. We call F the *(canonical) Lévy kernel* of X. We assume that the measures $F(x,dy)$ admit a density $y \mapsto f(x,y)$, and we aim for non-parametric estimation of the function $(x,y) \mapsto f(x,y)$.

On an equidistant time grid, we observe a sample $X_0(\omega), X_\Delta(\omega), \dots, X_{n\Delta}(\omega)$ of the process; the jumps are latent. We study a kernel density estimator for $f(x,y)$. We show its consistency as $n\Delta \to \infty$ and $\Delta \to 0$ under a smoothness hypothesis on the estimated density. In the ergodic case, we obtain asymptotic normality. In the null recurrent case, we impose a condition on the resolvent of the process which goes back to Darling and Kac (1957). Thereunder, we prove asymptotic mixed normality. We also provide a standardised version of our central limit theorem for the construction of asymptotic confidence intervals.

Our results are comparable to those in classical non-parametric density estimation. In particular: Our estimator's asymptotic bias and variance resemble those of the Nadaraya–Watson estimator in classical conditional density estimation. Just as in the classical context, moreover, the bandwidth choice is crucial for our estimator's rate of convergence. We conjecture that, for instance, a cross-validation method applies here analogously; see Fan and Yim (2004) and Hall, Racine, and Li (2004). By an optimal choice, if $\Delta \to 0$ fast enough, the rate is $v(n\Delta)^{\alpha_1\alpha_2/[d(\alpha_1+\alpha_2)+2\alpha_1\alpha_2]}$, where $\alpha_1 > 0$ (resp., $\alpha_2 > 0$) stands for the smoothness of f as a function in x (resp., in y), and the function v plays the role of an information rate. In the ergodic case, $v(t) = t$; in the null recurrent case with Darling–Kac's condition imposed (see Definition 2.5.22), $v(t) = t^\delta \ell(t)$ for some $0 < \delta \leq 1$ and some slowly varying function ℓ. We remark that, in the case $\alpha_1 = \alpha_2$, our achieved rate $v(n\Delta)^{\alpha_1/(2\alpha_1+2d)}$ equals the non-parametric minimax rate of smooth density estimation, related to the smoothness of f as a $2d$-dimensional function and with respect to $v(n\Delta)$.

At the core of our statistical problem, we essentially have to study the case first, where the process is observed continuously in time and, in particular, all jumps are discerned. In this case, we can consider a more general class of quasi-left continuous, strong Markov processes with càdlàg sample paths than just Itō semi-martingales. For these, the law of their jumps is again described by their Lévy kernel. We present a version of our estimator which utilises that the sojourn time of certain sets and the jumps are observed. Under slightly weaker assumptions, we prove the estimator's consistency and asymptotic (mixed) normality. As these results are valid for a quite general class of processes, we believe that they are of independent interest, not only as a benchmark for all possible estimators which are based on some discrete observation scheme.

For discrete-time Markov chains, a related result is presented in Karlsen and Tjøstheim (2001). We are aware that our final steps of proof appear to be similar. We emphasise that the main difficulties in our context, however, come in two respects: on the one hand, from establishing an appropriate auxiliary framework where related methods apply; on the other hand, from the discrete observation scheme where our primary objects of interest – the jumps – are latent.

For continuous-time Markov processes, apart from the Lévy process case and as far as known to us, estimation of their Lévy kernel has been confined to the special case of Markov step processes. For these, there exists a one-to-one correspondence between the Lévy kernel and the infinitesimal generator. On the one hand, efficient non-parametric estimation of Markov step process models has been studied by Greenwood and Wefelmeyer (1994). They assume the mean holding times to be bounded, and the transition kernel to be uniformly ergodic. This excludes the null recurrent case. On the other hand, the work on parametric estimation of Markov step processes is more exhaustive. The null recurrent case has been studied, for instance, by Höpfner (1993). There, the process is observed up to a random stopping time such that a deterministic amount of information (or more) has been discerned. Local asymptotic normality is shown in various situations. With a slightly different aim, in contrast, Höpfner et al. (1990) considers Markov step processes observed up to a deterministic time. Accordingly, the observed amount of information is random. Local asymptotic mixed normality (of statistical experiments) is shown under Darling–Kac's condition. Here, we utilise some of their results and methods. We improve upon the restrictions within the aforementioned literature: First and foremost, we do not restrict ourselves to Markov step processes. Secondly, we consider processes, null recurrent in the sense of Harris, in a non-parametric setting. Thirdly, we address the influence of observations on a discrete time grid.

We briefly outline this chapter. In Section 3.2 we study the estimation of the Lévy kernel based on discrete observations. Split into three subsections, we present the statistical problem with our standing assumptions; we give our estimator along with a bias correction; and state our main results – the estimator's consistency and the central limit theorem. In Section 3.3, we study the case where continuous-time observations are available. This section is organised analogously to Section 3.2. The corresponding proofs are in Section 3.4. The proofs for our main results of Section 3.2 are in Section 3.5. Each proofs section comes with its own short outline at its beginning. Since we bring together potential theoretic aspects of Markov processes with functional and martingale limit theory, we put some of our technical considerations off to Section 3.6.

3.2 Density estimation of the Lévy kernel from high-frequency observations

3.2.1 Preliminaries and assumptions

On the filtered probability space(s) $(\Omega, \mathscr{F}, (\mathscr{F}_t)_{t\geq 0}, (\mathbb{P}^x)_{x\in E})$, let $X = (X_t)_{t\geq 0}$ be a Markovian Itō semi-martingale with values in Euclidean space $E = (\mathbb{R}^d, \mathscr{B}^d)$, or a subset thereof, such that $\mathbb{P}^x(X_0 = x) = 1$ for all x. For $n \in \mathbb{N}$ and $\Delta > 0$, we observe $X_0(\omega)$ and the increments

$$\Delta_k^n X(\omega) := X_{k\Delta}(\omega) - X_{(k-1)\Delta}(\omega) \quad k = 1, \dots, n. \tag{3.2.1}$$

We emphasise that the jumps of the process are latent.

Throughout this chapter, we use the notation introduced in Chapter 2 and, moreover: We abbreviate $E^* := E \setminus \{0\}$. For $\alpha \geq 0$ and $A \subseteq E$, in addition, $\mathcal{C}^\alpha_{\mathrm{loc}}(A)$ denotes the class of all continuous functions on A which are $\lfloor \alpha \rfloor$-times continuously differentiable such that every $x \in A$ has a neighbourhood on which the function's (partial) $\lfloor \alpha \rfloor$-derivatives are uniformly Hölder of order $\alpha - \lfloor \alpha \rfloor$.

We recall from Section 2.2.3: The characteristics $(B, C, \mathfrak{n})$ of X are absolutely continuous with respect to Lebesgue measure; there are mappings $b : E \to E$ and $c : E \to E \otimes E$, and a kernel F on E with $F(x, \{0\}) = 0$ such that

$$B_t = \int_0^t b(X_s)\mathrm{d}s, \quad C_t = \int_0^t c(X_s)\mathrm{d}s, \quad \text{and} \quad \mathfrak{n}(\mathrm{d}t, \mathrm{d}y) = \mathrm{d}t \otimes F(X_t, \mathrm{d}y). \tag{3.2.2}$$

The random measure $\mathfrak{n}$ is the predictable compensator of the process's jump measure (see eq. (2.2.5)): For every Borel function $g : E \times E \to \mathbb{R}_+$, (initial) probability π, and $t > 0$, we have

$$\mathbb{E}^\pi \sum_{0<s\leq t} g(X_{s-}, \Delta X_s) \mathbb{1}_{\{X_{s-}\neq X_s\}} = \mathbb{E}^\pi \int_0^t \mathrm{d}s \int_E F(X_s, \mathrm{d}y) g(X_s, y). \tag{3.2.3}$$

We call F the *Lévy kernel*. It is unique outside a set of potential zero. We assume it admits a density $(x, y) \mapsto f(x, y)$ which we want to estimate.

Throughout, we work under the following technical hypothesis on the characteristics:

3.2.1 Assumption. (i) The process X satisfies the following (linear) growth condition: There exists a constant $\zeta < \infty$ and a Lévy measure $\bar{F}$ on E such that

$$\|b(x)\| \leq \zeta(1 + \|x\|), \quad \|c(x)\| \leq \zeta(1 + \|x\|^2), \quad \text{and} \quad F(x, A) \leq (1 + \|x\|)\bar{F}(A)$$

holds for all $x \in E$ and every Borel set $A \subseteq E$. We denote by $\beta \in [0, 2]$ some constant such that $\int \bar{F}(\mathrm{d}w)(\|w\|^\beta \wedge 1)) < \infty$.

(ii) The Lévy measure $\bar{F}$ admits a density $\bar{f}$ which is continuous on E^*.

(iii) There exists a constant $\zeta < \infty$ such that $\sup_{\|z\|>1} \|z\| \bar{f}(z) \leq \zeta$. ⋄

Remark. Apart from the growth condition, there is no assumption on b and c. Whether X is a weak or a strong solution of eq. (3.1.1) is irrelevant to us.

We impose assumptions on the recurrence of X and on the smoothness of f. To obtain consistency for our estimator below, we impose:

3.2.2 Assumption. The process X is Harris recurrent with invariant measure μ (see Definition 2.5.14). ⋄

3.2.3 Assumption. There exists an $\alpha > 0$ for which the Lévy kernel admits a density $f \in \mathcal{C}^\alpha_{\mathrm{loc}}(E \times E^*)$; and the invariant measure from Assumption 3.2.2 admits a continuous density μ'. ⋄

To obtain a central limit theorem, we also impose:

3.2.4 Assumption. For some $0 < \delta \leq 1$, the process X satisfies Darling–Kac's condition with an – at infinity – regularly varying function $v : \mathbb{R}_+ \to \mathbb{R}_+$ of index δ (see Definition 2.5.22). ⋄

Remark. In the positive recurrent case (that is, when μ is finite), Assumption 3.2.4 indeed is satisfied for $\delta = 1$ and with $v(t) = t/\mu(E)$. We also refer to Touati (1987) and to Höpfner and Löcherbach (2003).

3.2.5 Assumption. For some $\alpha_1, \alpha_2 \geq 2$, the Lévy kernel admits a density f which is twice continuously differentiable on $E \times E^*$ such that $x \mapsto f(x,y) \in \mathcal{C}^{\alpha_1}_{\mathrm{loc}}(E)$ for all $y \in E^*$, and $y \mapsto f(x,y) \in \mathcal{C}^{\alpha_2}_{\mathrm{loc}}(E^*)$ for all $x \in E$; and the invariant measure from Assumption 3.2.2 admits a continuous density μ' which is $(\lceil \alpha_1 \rceil - 1)$-times continuously differentiable. ⋄

Example. Suppose that f is bounded and vanishes outside $\{\|x\| \leq 1, \|y\| \leq 1\}$; that is, there are neither jumps with left-limit outside the unit ball nor jumps of size bigger than one. Then our process's recurrence (or transience) is completely determined by drift and volatility. For instance:

(i) If the volatility σ vanishes everywhere and the drift satisfies $b(x) = -x$, then X is positive recurrent.

(ii) If the drift b vanishes everywhere, and the volatility satisfies $\sigma(x) = 1$, then X is not positive. In fact, X has the recurrence (or transience) of Brownian motion: In the univariate case, X is null recurrent and Darling–Kac's condition holds with $\delta = 1/2$; in the bivariate case, X is null recurrent and Darling–Kac's condition fails; and in all other multivariate cases, X is transient.

3.2.2 Kernel density estimator

In principle, we are free to choose our favourite estimation method, for instance, the method of sieves with projection estimators. Here, however, we introduce a kernel density estimator as it allows for a more comprehensible presentation of the proofs. Also, the method is well-understood in the context of classical (conditional) density estimation.

An outline: Firstly, we choose smooth kernels g_1 and g_2 with support $B_1(0)$ (the unit ball centred at zero) which are, at least, of order α_1 and α_2, respectively; that is, for every multi-index $m = (m_1, \ldots, m_d) \in \mathbb{N}^d \setminus \{0\}$ and each $i \in \{1,2\}$, we have

$$|m| := m_1 + \cdots + m_d < \alpha_i \implies \kappa_m(g_i) := \int x_1^{m_1} \cdot \cdots \cdot x_d^{m_d} g_i(x) \mathrm{d}x = 0. \qquad (3.2.4)$$

Secondly, we choose a bandwidth vector $\eta = (\eta_1, \eta_2) > 0$. Lastly, we construct an estimator for $f(x,y)$ using the kernels $g_i^{\eta,x}(z) := \eta_i^{-d} g_i((z-x)/\eta_i)$. If the bandwidth is chosen appropriately, we achieve a consistent estimator which follows a central limit theorem.

3.2.7 Definition. For $\eta = (\eta_1, \eta_2) > 0$, we call $\hat{f}_n^{\Delta,\eta}$ defined by

$$\hat{f}_n^{\Delta,\eta}(x,y) := \begin{cases} \frac{\sum_{k=1}^n g_1^{\eta,x}(X_{(k-1)\Delta}) g_2^{\eta,y}(\Delta_k^n X)}{\Delta \sum_{k=1}^n g_1^{\eta,x}(X_{(k-1)\Delta})} & \text{if } \sum_{k=1}^n g_1^{\eta,x}(X_{(k-1)\Delta}) > 0, \\ 0 & \text{otherwise,} \end{cases} \tag{3.2.5}$$

the *kernel density estimator* of f (w. r. t. bandwidth η based on $X_0, X_\Delta, \dots, X_{n\Delta}$). ⋄

In analogy to classical conditional density estimation, we also introduce a bias correction for our estimator.

3.2.8 Definition. For $\eta = (\eta_1, \eta_2) > 0$, we call $\hat{\gamma}_n^{\Delta,\eta}$ defined by

$$\hat{\gamma}_n^{\Delta,\eta}(x,y) := \begin{cases} \eta_1^{\alpha_1} \sum\limits_{\substack{|m_1+m_2|=\alpha_1 \\ |m_2|\neq 0}} \frac{\kappa_{m_1+m_2}(g_1)}{m_1! m_2!} \frac{\sum_{k=1}^n \frac{\partial^{m_1}}{\partial x^{m_1}} g_1^{\eta,x}(X_{(k-1)\Delta})}{\sum_{k=1}^n g_1^{\eta,x}(X_{(k-1)\Delta})} \frac{\partial^{m_2}}{\partial x^{m_2}} \hat{f}_n^{\Delta,\eta}(x,y) & \\ \quad + \eta_2^{\alpha_2} \sum\limits_{|m|=\alpha_2} \frac{\kappa_m(g_2)}{m!} \frac{\partial^m}{\partial y^m} \hat{f}_n^{\Delta,\eta}(x,y), & \text{if } \begin{array}{c} \sum_{k=1}^n g_1^{\eta,x}(X_{(k-1)\Delta}) > 0 \\ \alpha_1, \alpha_2 \in \mathbb{N}^* \end{array}, \\ 0, & \text{otherwise,} \end{cases}$$

the *bias correction* for $\hat{f}_n^{\Delta,\eta}$. (The sums in the previous equation are over all multi-indices of appropriate length.) ⋄

3.2.3 Consistency and central limit theorem

Here, we present our main results. We agree to the following conventions: Under Assumptions 3.2.2 and 3.2.4, v denotes the regularly varying function given in eq. (2.5.6). Under Assumption 3.2.2 only, v denotes an arbitrary deterministic equivalent of the Markov process X (see Definition 2.5.21). For typographical reasons, we may write v_t for $v(t)$ or $X(t)$ for X_t etc. as convenient.

We utilise the following conditions as $n\Delta \to \infty$ and $\Delta \to 0$, where $0 \le \zeta_1, \zeta_2 < \infty$:

$$v_{n\Delta} \eta_{1,n}^d \eta_{2,n}^d \to \infty, \quad \text{and} \quad \eta_{1,n} \to 0, \quad \eta_{2,n} \to 0; \tag{3.2.6}$$

$$v_{n\Delta}\eta_{1,n}^{d+2\alpha_1}\eta_{2,n}^{d} \to \zeta_1^2, \quad \text{and} \quad v_{n\Delta}\eta_{1,n}^{d}\eta_{2,n}^{d+2\alpha_2} \to \zeta_2^2. \tag{3.2.7}$$

In addition, we also utilise the following conditions due to discretisation, where $\zeta < \infty$ is independent of n:

$$\Delta\eta_{1,n}^{-2-d[(1-2/(\beta+d))\vee 0]} \to 0, \quad \text{and} \quad \Delta\eta_{2,n}^{-2\vee(\beta+d)} \to 0; \tag{3.2.8a}$$

$$n\Delta^2\eta_{1,n}^{d}\eta_{2,n}^{d} \le \zeta, \quad v_{n\Delta}\Delta^2\eta_{1,n}^{d-4-2d[(1-2/(\beta+d)\wedge 0]}\eta_{2,n}^{d} \to 0, \tag{3.2.8b}$$

$$\text{and} \quad v_{n\Delta}\Delta^2\eta_{1,n}^{d}\eta_{2,n}^{d-4\vee 2(\beta+d)} \to 0. \tag{3.2.8c}$$

Remark. If $\Delta \to 0$ fast enough, then eqs. (3.2.6) and (3.2.7) are the crucial conditions.

3.2.9 Theorem. *Grant Assumptions 3.2.1 to 3.2.3. Let $\eta_n = (\eta_{1,n}, \eta_{2,n})$ be such that eqs.* (3.2.6) *and* (3.2.8a) *hold. Moreover, let $(x, y) \in E \times E^*$ be such that $\mu'(x) > 0$ and $F(x, E) > 0$.*

(i) If $n\Delta^2 \to 0$, then, under any law $\mathbb{P}^\pi$, we have the following convergence in probability:

$$\hat{f}_n^{\Delta,\eta_n}(x, y) \xrightarrow[n\to\infty]{\mathbb{P}^\pi} f(x, y). \tag{3.2.9}$$

(ii) Grant Assumption 3.2.4 in addition. If $(n\Delta)^{1-\delta}\Delta \to 0$, then, under any law $\mathbb{P}^\pi$, eq. (3.2.9) *holds as well.*

Remark. By this theorem, our estimator is consistent for every x and $y \neq 0$ if $n\Delta \to \infty$ and $\Delta \to 0$. In practice, however, both n and Δ are given! Then, for instance, if a continuous martingale component is present, our estimator is unreliable for all y close to the origin. To illustrate this important point, suppose that X is a univariate process with constant volatility $\sigma^2 > 0$. Increments with absolute value less than $\zeta\sigma\Delta^{1/2}$, where ζ is quite a large constant (e. g., $\zeta = 5$), are predominantly due to the continuous martingale and not due to jumps. On the set $\{y : |y| \le \zeta\sigma\Delta^{1/2}\}$, therefore, our estimator $\hat{f}_n^{\Delta,\eta}(x, \cdot)$ is unreliable regardless of the chosen bandwidth η. We illustrate this point in a simulation study (see Section 5.1).

For the next theorem, we establish additional notation. On an extension

$$(\tilde{\Omega}, \tilde{\mathscr{F}}, \tilde{\mathbb{P}}) := (\Omega \times \Omega', \mathscr{F} \otimes \mathscr{F}', \mathbb{P}^\pi \otimes \mathbb{P}') \tag{3.2.10}$$

of the probability space, let $V = (V(x,y))_{x\in E, y\in E^*}$ be a standard Gaussian white noise random field (that is, the finite dimensional marginals of V are i. i. d. standard normal) and let $L = (L_t)_{t\geq 0}$ be the Mittag-Leffler process of order δ (see Definition 2.5.23) such that V, L and $\mathscr{F}$ are independent. In the theorem below, convergence holds stably in law (recall Definition 2.3.1).

3.2.10 Theorem. *Grant Assumptions 3.2.1 to 3.2.5. Let $\eta_n = (\eta_{1,n}, \eta_{2,n})$ be such that eqs.* (3.2.6) *and* (3.2.8) *hold, and let $(x_i, y_i)_{i\in I}$ be a finite family of pairwise distinct points in $E \times E^*$ such that $\mu'(x_i) > 0$ and $F(x_i, E) > 0$ for each $i \in I$. If $(n\Delta)^{1-\delta}\Delta \to 0$, then, under any law $\mathbb{P}^\pi$, we have the following stable convergence in law:*

$$\left(\sqrt{v_{n\Delta}\eta_{1,n}^d \eta_{2,n}^d}\left(\hat{f}_n^{\Delta,\eta_n}(x_i,y_i) - \frac{\mu(g_1^{\eta_n,x_i} F g_2^{\eta_n,y_i})}{\mu(g_1^{\eta_n,x_i})}\right)\right)_{i\in I} \xrightarrow[n\to\infty]{\mathscr{L}-\mathrm{st}} \left(\frac{\sigma(x_i,y_i)}{\sqrt{L_1}} V(x_i,y_i)\right)_{i\in I},$$

where the asymptotic variance is given by

$$\sigma(x,y)^2 := \frac{f(x,y)}{\mu'(x)} \int g_1(w)^2 \mathrm{d}w \int g_2(z)^2 \mathrm{d}z. \tag{3.2.11}$$

In addition, let η_n be such that eq. (3.2.7) *holds as well. Suppose either that $\alpha_1, \alpha_2 \in \mathbb{N}^*$ or that $\zeta_1 = \zeta_2 = 0$ in eq.* (3.2.7)*. Then, under any law $\mathbb{P}^\pi$, we have the following stable convergence in law:*

$$\left(\sqrt{v_{n\Delta}\eta_{1,n}^d \eta_{2,n}^d}\left(\hat{f}_n^{\Delta,\eta_n}(x_i,y_i) - f(x_i,y_i)\right)\right)_{i\in I} \xrightarrow[n\to\infty]{\mathscr{L}-\mathrm{st}} \left(\gamma(x_i,y_i) + \frac{\sigma(x_i,y_i)}{\sqrt{L_1}} V(x_i,y_i)\right)_{i\in I},$$

where – in the former case – the asymptotic bias $\gamma(x,y)$ is given by

$$\begin{aligned}\gamma(x,y) = \frac{\zeta_1}{\mu'(x)} &\sum_{\substack{|m_1+m_2|=\alpha_1 \\ |m_2|\neq 0}} \frac{\kappa_{m_1+m_2}(g_1)}{m_1! m_2!} \frac{\partial^{m_1}}{\partial x^{m_1}}\mu'(x) \frac{\partial^{m_2}}{\partial x^{m_2}} f(x,y) \\ &+ \zeta_2 \sum_{|m|=\alpha_2} \frac{\kappa_m(g_2)}{m!} \frac{\partial^m}{\partial y^m} f(x,y),\end{aligned} \tag{3.2.12}$$

and – in the latter case – $\gamma(x,y) = 0$.

Remark. The asymptotic bias and variance of our estimator are analogous to those of the Nadaraya–Watson estimator in classical conditional density estimation (see Hansen, 2009): $\kappa_m(g_i)$ and $\int g_i(z)^2 \mathrm{d}z$ are the relevant *moment* and the *roughness* of the kernel g_i, respectively; and f (resp., μ') plays the role of the conditional (resp., marginal) density.

We recall that v from Darling–Kac's condition (see Definition 2.5.22) satisfies $v_t = t$ in the ergodic case, and $v_t = t^\delta \ell(t)$ for some slowly varying function ℓ in the null recurrent case. If we choose $\eta_{i,n} = v_{n\Delta}^{-\xi_i}$ with

$$\xi_1 = \frac{\alpha_2}{d(\alpha_1+\alpha_2)+2\alpha_1\alpha_2} \quad \text{and} \quad \xi_2 = \frac{\alpha_1}{d(\alpha_1+\alpha_2)+2\alpha_1\alpha_2},$$

then eqs. (3.2.6) and (3.2.7) hold with $\zeta_1 = \zeta_2 = 1$. If $\Delta \to 0$ fast enough such that $n\Delta^{1+[d(\alpha_1+\alpha_2)+2\alpha_1\alpha_2]/\zeta} \to 0$ in addition, where ζ denotes the maximum of

$$(1-\delta)d(\alpha_1+\alpha_2)+2\alpha_1\alpha_2, \quad \delta\alpha_1(\alpha_2+2+d) \quad \text{and} \quad \delta\alpha_2\left(\alpha_1+2+\frac{d^2}{2+d}\right),$$

then our choice of η_n also satisfies eq. (3.2.8) for every $\beta \le 2$. Consequently, our estimator's rate of convergence is

$$v_{n\Delta}^{\alpha_1\alpha_2/[d(\alpha_1+\alpha_2)+2\alpha_1\alpha_2]}. \tag{3.2.13}$$

In the case $\alpha_1 = \alpha_2$, the achieved rate $v_{n\Delta}^{\alpha/(2\alpha+2d)}$ equals the non-parametric minimax rate of smooth density estimation, related to the smoothness of f as a $2d$-dimensional function and w. r. t. $v_{n\Delta}$.

Remark. Bandwidth selection has always been a crucial issue in these kind of studies. Although orders of magnitude are crucial from an asymptotic point of view and $\eta_{i,n} = (n\Delta)^{-\xi_i}$ for some $\xi_i > 0$ may be a good choice , we note that, in practice, $\eta_{i,n} = \zeta(n\Delta)^{-\xi_i}$ with leading constant $\zeta \neq 1$ could be a better one. A detailed analysis would go beyond the scope of this chapter. We briefly comment on two problems: How to choose the bandwidths manually such that conditions (3.2.6–3.2.8) are satisfied for the unknown $v_{n\Delta}$, α_1, α_2 and β? What needs to be considered when employing data-driven methods for selecting optimal bandwidths?

(i) Let $\alpha_0 \geq 2$ and $0 < \delta_0 \leq 1$ such that $\delta_0 > d/(d+\alpha_0)$. If we choose $\eta_{i,n} = (n\Delta)^{-1/(2d+2\alpha_0)}$, then eqs. (3.2.6) and (3.2.7) hold for all processes X such that Assumptions 3.2.4 and 3.2.5 hold for some $\alpha_1, \alpha_2 \geq \alpha_0$ and $\delta_0 < \delta \leq 1$. If $\Delta \to 0$ fast enough such that $n\Delta^{1+2[\alpha_0+d]/[\alpha_0+(2+d)\vee\alpha_0]} \to 0$ in addition, then our chosen bandwidth also satisfies eq. (3.2.8).

(ii) The asymptotic bias and variance are proportional to the value of f and its derivatives at the point of interest. The optimal bandwidth choice in terms of the asymptotic mean squared error, therefore, may depend heavily on x and y. Especially for processes with infinite activity – where $y \mapsto f(x,y)$ has a pole at zero – this is an important issue in practice. In a future study on data-driven bandwidth selection methods like cross-validation, this distinction from estimating a bounded probability density has to be addressed carefully.

Theorem 3.2.10 does not allow for a direct construction of confidence intervals. For this purpose, we also obtain the following standardised version.

3.2.11 Corollary. *Grant Assumptions 3.2.1 to 3.2.5. Let $\eta_n = (\eta_{1,n}, \eta_{2,n})$ be such that eqs.* (3.2.6) *to* (3.2.8) *hold. Suppose either that $\alpha_1, \alpha_2 \in \mathbb{N}^*$ or that $\zeta_1 = \zeta_2 = 0$ in eq.* (3.2.7). *Then under any law $\mathbb{P}^\pi$, we have the following stable convergence in law:*

$$\left(\sqrt{\frac{\eta_{1,n}^d \eta_{2,n}^d \Delta \sum_{k=1}^n g_1^{\eta_n, x_i}(X_{(k-1)\Delta})}{\xi_g^2 \hat{f}_n^{\Delta,\eta_n}(x_i,y_i)}} \left(\hat{f}_n^{\Delta,\eta_n}(x_i,y_i) - \hat{\gamma}_n^{\eta_n}(x_i,y_i) - f(x_i,y_i)\right)\right)_{i\in I} \xrightarrow[n\to\infty]{\mathscr{L}-\mathrm{st}} \left(V(x_i,y_i)\right)_{i\in I},$$

where $\xi_g^2 = \int g_1(w)^2 \mathrm{d}w \int g_2(z)^2 \mathrm{d}z$.

Remark. In principle, the results of this section are extendible to more general Markov models with Lévy kernel F such that eq. (3.2.3) holds. In view of our proofs, the assumption that X is an Itō semi-martingale is crucial for the analysis of the influence of discretisation (see Section 3.5.1). Suppose that an explicit upper bound for the small-time asymptotic "error"

$$\left|\frac{1}{\Delta}\mathbb{E}^x\left[g_2^{\eta,y}(\Delta_1^n X)\right] - \int F(x,\mathrm{d}w) g_2^{\eta,y}(w)\right|$$

and an explicit sufficient condition which ensures

$$\sup_{s\leq 1} \frac{\xi_n}{v_{n\Delta}\eta_{1,n}^d} \left| \Delta \sum_{k=1}^{\lfloor sn \rfloor} h_n(X_{(k-1)\Delta}) - \int_0^{\lfloor sn \rfloor \Delta} h_n(X_r) \mathrm{d}r \right| \xrightarrow[n\to\infty]{\mathbb{P}^\pi} 0$$

for $\xi_n = 1$ or $\xi_n^2 = v_{n\Delta}\eta_{1,n}^d \eta_{2,n}^d$ are available for some Markov process X. Then it is straightforward (see Lemma 3.5.7 and eq. (3.5.31) — Lemmata 3.5.6, 3.5.9 and 3.5.10, respectively) to come up with sufficient conditions for Theorems 3.2.9 and 3.2.10, which replace eq. (3.2.8).

3.3 Density estimation of the Lévy kernel from continuous-time observations — A benchmark

The Lévy kernel of a Markov process is related with jumps. In fact, our estimator eq. (3.2.5) uses $X_{(k-1)\Delta}$ and $\Delta_k^n X$ as proxies for the pre-jump value X_{t-} and the jump size ΔX_t if, at a time $t \in [(k-1)\Delta, k\Delta]$, there is a jump from a neighbourhood of x and of size close to y. Eventually, such time intervals contain either zero or one such jump; never more. Certainly, the statistical analysis simplifies if we observed the whole path of X; introducing proxies would be useless. So, despite observing the whole path of X is somewhat unrealistic, it is theoretically important to study what happens in this case. This section is devoted to this question and can be viewed as a benchmark for what properties are achievable with a more realistic, discrete observation scheme.

3.3.1 Preliminaries and assumptions

On the filtered probability space(s) $(\Omega, \mathscr{F}, (\mathscr{F}_t)_{t\geq 0}, (\mathbb{P}^x)_{x\in E})$, let $X = (X_t)_{t\geq 0}$ be a strong Markov process with values in Euclidean space $E = (\mathbb{R}^d, \mathscr{B}^d)$, or a subset thereof. Its sample paths are supposed to be càdlàg; and X is supposed to be quasi-left continuous (see Definition 2.1.9). In other words, X is a Hunt process (see Definition 2.5.13). We observe – continuously in time – one sample path $\{X_s(\omega) : s \in [0,t]\}$ for $t > 0$; in particular, we discern all jumps.

Benveniste and Jacod (1973) proved the existence of a Lévy system (F, H) where H is continuous (recall Theorem 2.5.25): For every Borel function $g : E \times E \to \mathbb{R}_+$, probability π on E, and $t > 0$, we have

$$\mathbb{E}^\pi \sum_{0<s\le t} g(X_{s-}, \Delta X_s)\mathbb{1}_{\{X_{s-}\neq X_s\}} = \mathbb{E}^\pi \int_0^t \mathrm{d}H_s \int_E F(X_s, \mathrm{d}y) g(X_s, y). \tag{3.3.1}$$

We remark once more that the disintegration into F and H is by no means unique. For an appropriate reference function g_0 with $Fg_0(x) > 0$, nevertheless, ratios of the form $Fg(x)/Fg_0(x)$ are unique outside a set of potential zero.

Throughout this section, we work under the following hypothesis:

3.3.1 Assumption. There exists a Lévy system (F, H) of X where $H_t = t$. ⋄

Recalling eq. (3.2.3), we observe that all Markovian Itō semi-martingales satisfy Assumption 3.3.1. In analogy to the semi-martingale case, we call this F in Assumption 3.3.1 the *(canonical) Lévy kernel* of X. It is unique outside a set of potential zero. Again, we assume it admits a density $(x, y) \mapsto f(x, y)$ which we want to estimate.

Compared to Section 3.2, we slightly weaken the assumptions imposed on the smoothness of f. To obtain consistency for our estimator below, we impose Assumption 3.2.2 and:

3.3.2 Assumption. The canonical Lévy kernel admits a density f, continuous on $E \times E^*$; and the invariant measure from Assumption 3.2.2 admits a continuous density μ'. ⋄

To obtain a central limit theorem, we also impose Assumption 3.2.4 and:

3.3.3 Assumption. For some $\alpha_1, \alpha_2 > 0$, the canonical Lévy kernel admits a density f such that $x \mapsto f(x, y) \in C^{\alpha_1}_{\mathrm{loc}}(E)$ for all $y \in E^*$, and $y \mapsto f(x, y) \in C^{\alpha_2}_{\mathrm{loc}}(E^*)$ for all $x \in E$; and the invariant measure from Assumption 3.2.2 admits a continuous density μ' which is $(\lceil\alpha_1\rceil - 1)$-times continuously differentiable. ⋄

3.3.2 Kernel density estimator

In Section 3.2.2, we introduced a kernel density estimator and its bias correction based on discrete observations. Here, we present corresponding versions which utilise the continuous-time observation scheme. We recall that g_1 and g_2 are kernels with support $B_1(0)$ which are, at least, of order α_1 and α_2, respectively. Given a bandwidth vector $\eta = (\eta_1, \eta_2) > 0$, we utilise the kernels $g_i^{\eta,x}(z) = \eta_i^{-d} g_i((z-x)/\eta_i)$.

3.3.4 Definition. For $\eta = (\eta_1, \eta_2) > 0$, we call $\hat{f}_t^\eta$ defined by

$$\hat{f}_t^\eta(x,y) := \begin{cases} \frac{\sum_{0<s\leq t} g_1^{\eta,x}(X_{s-}) g_2^{\eta,y}(\Delta X_s) \mathbb{1}_{\{X_{s-} \neq X_s\}}}{\int_0^t g_1^{\eta,x}(X_s)\mathrm{d}s} & \text{if } \int_0^t g_1^{\eta,x}(X_s)\mathrm{d}s > 0, \\ 0 & \text{otherwise,} \end{cases}$$

the *kernel density estimator* of f (w. r. t. bandwidth η up to time t). ⋄

Our estimator in Definition 3.2.7 is the discretised analogue from the one presented here: In the numerator of the former, the jumps ΔX_t and the pre-jump left-limits X_{t-} are replaced by the increments $\Delta_k^n X$ and the pre-increment values $X_{(k-1)\Delta}$, respectively. In the denominator, the sojourn time $\int_0^t g_1^{\eta,x}(X_s)\mathrm{d}s$ is replaced by its Riemann sum approximation $\Delta \sum_{k=1}^n g_1^{\eta,x}(X_{(k-1)\Delta})$. In analogy to Definition 3.2.8, we also introduce a bias correction for our estimator:

3.3.5 Definition. For $\eta = (\eta_1, \eta_2) > 0$, we call $\hat{\gamma}_t^\eta$ defined by

$$\hat{\gamma}_t^\eta(x,y) := \begin{cases} \eta_1^{\alpha_1} \sum_{\substack{|m_1+m_2|=\alpha_1 \\ |m_2|\neq 0}} \frac{\kappa_{m_1+m_2}(g_1)}{m_1! m_2!} \frac{\int_0^t \frac{\partial^{m_1}}{\partial x^{m_1}} g_1^{\eta,x}(X_s)\mathrm{d}s}{\int_0^t g_1^{\eta,x}(X_s)\mathrm{d}s} \frac{\partial^{m_2}}{\partial x^{m_2}} \hat{f}_t^\eta(x,y) & \\ \qquad + \eta_2^{\alpha_2} \sum_{|m|=\alpha_2} \frac{\kappa_m(g_2)}{m!} \frac{\partial^m}{\partial y^m} \hat{f}_t^\eta(x,y), & \text{if } \begin{array}{c} \int_0^t g_1^{\eta,x}(X_s)\mathrm{d}s > 0 \\ \alpha_1, \alpha_2 \in \mathbb{N}^* \end{array}, \\ 0, & \text{otherwise,} \end{cases}$$

the *bias correction* for $\hat{f}_t^\eta$. ⋄

3.3.3 Consistency and central limit theorem

Here, we present our results of this section. We continue to use the notation and conventions from Section 3.2.3.

We utilise the following conditions as $t \to \infty$, where $0 \leq \zeta_1, \zeta_2 < \infty$:

$$v_t \eta_{1,t}^d \eta_{2,t}^d \to \infty, \quad \text{and} \quad \eta_{1,t} \to 0, \eta_{2,t} \to 0; \tag{3.3.2}$$

$$v_t \eta_{1,t}^{d+2\alpha_1} \eta_{2,t}^d \to \zeta_1^2, \quad \text{and} \quad v_t \eta_{1,t}^d \eta_{2,t}^{d+2\alpha_2} \to \zeta_2^2. \tag{3.3.3}$$

3.3.6 Theorem. *Grant Assumptions 3.2.2, 3.3.1 and 3.3.2. Let $\eta_t = (\eta_{1,t}, \eta_{2,t})$ be such that eq. (3.3.2) holds. Moreover, let $(x,y) \in E \times E^*$ be such that $\mu'(x) > 0$ and $F(x,E) > 0$. Then, under any law $\mathbb{P}^\pi$, we have the following convergence in probability:*

$$\hat{f}_t^{\eta_t}(x,y) \xrightarrow[t\to\infty]{\mathbb{P}^\pi} f(x,y).$$

3.3.7 Theorem. *Grant Assumptions 3.2.2, 3.2.4, 3.3.1 and 3.3.2. Let $\eta_t = (\eta_{1,t}, \eta_{2,t})$ be such that eq. (3.3.2) holds. Moreover, let $(x_i, y_i)_{i\in I}$ be a finite family of pairwise distinct points in $E \times E^*$ such that $\mu'(x_i) > 0$ and $F(x_i, E) > 0$ for each $i \in I$. Then, under any law $\mathbb{P}^\pi$, we have the following stable convergence in law:*

$$\left(\sqrt{v_t \eta_{1,t}^d \eta_{2,t}^d}\left(\hat{f}_t^{\eta_t}(x_i,y_i) - \frac{\mu(g_1^{\eta_t,x_i} F g_2^{\eta_t,y_i})}{\mu(g_1^{\eta_t,x_i})}\right)\right)_{i\in I} \xrightarrow[t\to\infty]{\mathscr{L}-\mathrm{st}} \left(\frac{\sigma(x_i,y_i)}{\sqrt{L_1}} V(x_i,y_i)\right)_{i\in I},$$

where the asymptotic variance is given by

$$\sigma(x,y)^2 := \frac{f(x,y)}{\mu'(x)} \int g_1(w)^2 \mathrm{d}w \int g_2(z)^2 \mathrm{d}z. \tag{3.3.4}$$

In addition, grant Assumption 3.3.3 and let η_t be such that eq. (3.3.3) holds as well. Suppose either that $\alpha_1, \alpha_2 \in \mathbb{N}^*$ or that $\zeta_1 = \zeta_2 = 0$ in eq. (3.3.3). Then, under any law $\mathbb{P}^\pi$, we have the following stable convergence in law:

$$\left(\sqrt{v_t \eta_{1,t}^d \eta_{2,t}^d}\left(\hat{f}_t^{\eta_t}(x_i,y_i) - f(x_i,y_i)\right)\right)_{i\in I} \xrightarrow[t\to\infty]{\mathscr{L}-\mathrm{st}} \left(\gamma(x_i,y_i) + \frac{\sigma(x_i,y_i)}{\sqrt{L_1}} V(x_i,y_i)\right)_{i\in I},$$

where – in the former case – the asymptotic bias $\gamma(x,y)$ is given by

$$\begin{aligned}\gamma(x,y) = \frac{\zeta_1}{\mu'(x)} \sum_{\substack{|m_1+m_2|=\alpha_1 \\ |m_2|\neq 0}} \frac{\kappa_{m_1+m_2}(g_1)}{m_1! m_2!} \frac{\partial^{m_1}}{\partial x^{m_1}} \mu'(x) \frac{\partial^{m_2}}{\partial x^{m_2}} f(x,y) \\ + \zeta_2 \sum_{|m|=\alpha_2} \frac{\kappa_m(g_2)}{m!} \frac{\partial^m}{\partial y^m} f(x,y),\end{aligned} \tag{3.3.5}$$

and – in the latter case – $\gamma(x,y) = 0$.

We compare Theorems 3.2.10 and 3.3.7: Firstly, we remark that the asymptotic bias and variance of $\hat{f}_n^{\Delta,\eta}$ are equal to those of our benchmark estimator $\hat{f}_t^{\eta}$. Secondly, if we choose $\eta_{i,t} = v_t^{-\xi_i}$ with

$$\xi_1 = \alpha_2/[d(\alpha_1+\alpha_2)+2\alpha_1\alpha_2] \quad \text{and} \quad \xi_2 = \alpha_1/[d(\alpha_1+\alpha_2)+2\alpha_1\alpha_2]$$

again, then eqs. (3.3.2) and (3.3.3) hold with $\zeta_1 = \zeta_2 = 1$. The rate of convergence in Theorem 3.3.7 is

$$v_t^{\alpha_1\alpha_2/[d(\alpha_1+\alpha_2)+2\alpha_1\alpha_2]}; \tag{3.3.6}$$

the rates in eqs. (3.2.13) and (3.3.6) are equivalent. Thirdly, we observe that our remark on the issue of bandwidth selection holds analogously. Lastly, we note that Theorem 3.3.7 does not allow for a direct construction of confidence intervals just as Theorem 3.2.10. In analogy to Corollary 3.2.11, we also obtain the following standardised version.

3.3.8 Corollary. *Grant Assumptions 3.2.2, 3.2.4 and 3.3.1 to 3.3.3. Let $\eta_t = (\eta_{1,t}, \eta_{2,t})$ be such that eqs.* (3.3.2) *and* (3.3.3) *hold. Suppose either that $\alpha_1, \alpha_2 \in \mathbb{N}^*$ or that $\zeta_1 = \zeta_2 = 0$ in eq.* (3.3.3)*. Then under any law $\mathbb{P}^\pi$, we have the following stable convergence in law:*

$$\left(\sqrt{\frac{\eta_{1,t}^d \eta_{2,t}^d \int_0^t g_1^{\eta_t,x_i}(X_s)\mathrm{d}s}{\xi_g^2 \hat{f}_t^{\eta_t}(x_i,y_i)}}\Big(\hat{f}_t^{\eta_t}(x_i,y_i) - \hat{\gamma}_t^{\eta_t}(x_i,y_i) - f(x_i,y_i)\Big)\right)_{i\in I} \xrightarrow[t\to\infty]{\mathscr{L}-\mathrm{st}} \Big(V(x_i,y_i)\Big)_{i\in I},$$

where $\xi_g^2 = \int g_1(w)^2\mathrm{d}w \int g_2(z)^2\mathrm{d}z$.

3.4 Proofs for results of Section 3.3

The notion of a deterministic equivalent of a Markov process plays a crucial role in the limit theory for our estimator (recall Definition 2.5.21). We emphasise the following consequence of Théorème 3 of Touati (1987) (see Theorem 2.5.24): Under Darling–Kac's condition, the function v in eq. (2.5.6) is a deterministic equivalent of X. For every H as in Definition 2.5.21, furthermore, we have that $(v(t)^{-1}H_{st})_{s\geq 0}$ converges in law to a non-trivial process as $t \to \infty$. For Markov processes violating Darling–Kac's condition, the latter convergence may not hold. Nevertheless, Löcherbach and Loukianova (2008) showed that some deterministic equivalent already exists when X is Harris recurrent.

This section is organised as follows: Firstly, in Section 3.4.1 we prove a triangular array extension of Birkhoff's theorem for additive functionals. Secondly, in Section 3.4.2 we introduce auxiliary Markov chains Z and Z' derived from our Markov process X. We show that our result from Section 3.4.1 applies to these chains. Some technicalities are put off to Section 3.6. Thirdly, in Section 3.4.3 we demonstrate a preliminary version of Theorem 3.3.6 which depends only on Z and Z'; we conclude with the final steps in the proof of consistency. Lastly, in Section 3.4.4 we demonstrate a preliminary central limit theorem which depends only on Z and Z'; we conclude with the final steps in the proof of Theorem 3.3.7 and Corollary 3.3.8.

3.4.1 An extension of Birkhoff's theorem

The theorem presented in this subsection is the underlying key result for our proofs. It is a triangular array extension of Birkhoff's theorem for additive functionals (recall Theorem 2.5.20). We prove a rather general version.

3.4.1 Theorem. *Let $Z = (Z_k)_{k\in\mathbb{N}^*}$ be a Markov chain with values in some state space D, with invariant probability ψ, and with transition kernel Ψ. Assume that the state space is* petite, *that is, there exist a probability ρ on $\mathbb{N}^*$ and a non-trivial measure ν_ρ on D such that, for every Borel set $A \subseteq D$,*

$$\inf_{x\in D} \sum_{k=1}^{\infty} \rho(k)\Psi^k(x, A) \geq \nu_\rho(A).$$

Let $(h_n)_{n\in\mathbb{N}^*}$ be a sequence of functions such that $(\Psi h_n)_{n\in\mathbb{N}^*}$ is uniformly bounded. Let $\xi_n > 0$ be such that

$$n\xi_n \to \infty, \quad \xi_n^{-1}\psi(h_n) \to c < \infty, \quad (n\xi_n^2)^{-1}\psi(|h_n|) \to 0 \quad \text{and} \quad (n\xi_n^2)^{-1}\psi(h_n^2) \to 0$$

as $n \to \infty$. Then, under every law $\mathbb{P}^\pi$ for some probability π on D, the following convergence *holds uniformly on compacts in probability:*

$$G^n_s \xrightarrow[n\to\infty]{\text{ucp}} cs, \qquad \textit{where} \quad G^n_s := \frac{1}{n\xi_n}\sum_{k=1}^{\lfloor sn\rfloor} h_n(Z_k). \tag{3.4.1}$$

Remark. If $(h_n)_{n\in\mathbb{N}^*}$ is non-negative (resp., uniformly bounded), then $n\xi_n \to \infty$ and $\xi_n^{-1}\psi(h_n) \to c < \infty$ already imply $(n\xi_n^2)^{-1}\psi(|h_n|) \to 0$ (resp., $(n\xi_n^2)^{-1}\psi(h_n^2) \to 0$).

Proof (of Theorem 3.4.1). Convergence in probability is equivalent to the property that – given any subsequence – there exists a further subsequence which converges almost surely. By Proposition 17.1.6 of Meyn and Tweedie (1993), therefore, it is sufficient to prove this theorem under the law $\mathbb{P}^\psi$ only.

For each $s \geq 0$ and $n \in \mathbb{N}^*$, we observe $G^n_s = H^n_s + H'^n_s$, where

$$H^n_s = \frac{\lfloor sn\rfloor \psi(h_n)}{n\xi_n} \quad \text{and} \quad H'^n_s = \frac{1}{n\xi_n}\sum_{k=1}^{\lfloor sn\rfloor}\Big(h_n(Z_k) - \psi(h_n)\Big).$$

By assumption, we have $H^n_s \to sc$ uniformly in s as $n \to \infty$. It remains to show that H'^n_s converges to zero uniformly on compacts in probability.

We note $\mathbb{E}^\psi[h_n(Z_k)] = \psi(h_n)$ for every $k, n \in \mathbb{N}^*$; thus, $\mathbb{E}^\psi[H'^n_s] = 0$ for all $s \geq 0$. Moreover, its second moment satisfies $\mathbb{E}^\psi[(H'^n_s)^2] = K^n_s + K'^n_s$, where

$$K^n_s = \frac{1}{n^2\xi_n^2}\sum_{k=1}^{\lfloor sn\rfloor}\Big(\psi(h_n^2) - \psi(h_n)^2\Big)$$

and

$$K'^n_s = \frac{2}{n^2\xi_n^2}\sum_{k=1}^{\lfloor sn\rfloor-1}\int \psi(\mathrm{d}z)h_n(z)\sum_{l=k+1}^{\lfloor sn\rfloor}\Big(\Psi^{l-k}h_n(z) - \psi(h_n)\Big).$$

Firstly, we note

$$|K^n_s| \leq \frac{\lfloor sn\rfloor}{n}\left|\frac{\psi(h_n^2)}{n\xi_n^2} - \frac{\psi(h_n)^2}{n\xi_n^2}\right| \xrightarrow[n\to\infty]{} 0. \tag{3.4.2}$$

Secondly, let $m \in \mathbb{N}^*$ denote the period and $D_0, \ldots, D_{m-1}$ denote an m-cycle of Z (recall Theorem 2.4.10). By Proposition 2.4.16, the restriction of the sampled chain with transition kernel Ψ^m to each set D_i is aperiodic and Harris recurrent with invariant probability $m\psi(\,\cdot\,\cap D_i)$. For every $i \in \{1,\ldots,m\}$ and $z \in D_i$, we denote $j(l,z) := (i+l) \bmod m$, where 'mod' stands for the modulo operator. For every $n_0 \in \mathbb{N}^*$, we observe

$$\begin{aligned}\sum_{l=1}^{n_0}\Big(\Psi^l h_n(z) - \psi(h_n)\Big) = &\sum_{k=0}^{\lfloor \frac{n_0}{m}\rfloor}\sum_{l=1}^{m}\Big(\Psi^{km+l} h_{n|D_{j(l,z)}}(z) - m\psi(h_{n|D_{j(l,z)}})\Big)\\ &+\sum_{l=1}^{n_0 \bmod m}\Big(\Psi^{\lfloor\frac{n_0}{m}\rfloor m+l} h_{n|D_{j(l,z)}}(z) - \psi(h_n)\Big).\end{aligned} \tag{3.4.3}$$

Hence,

$$\left|\sum_{l=1}^{n_0}\Big(\Psi^l h_n(z) - \psi(h_n)\Big)\right| \le \sum_{k=0}^{\infty}\sum_{l=1}^{m}\left|\Psi^{km+l} h_{n|D_{j(l,z)}}(z) - m\psi(h_{n|D_{j(l,z)}})\right| + m|\psi(h_n)|.$$

As the state space D is petite w. r. t. Ψ, so is each D_i w. r. t. Ψ^m. By Theorems 16.2.1 and 16.2.2 of Meyn and Tweedie (1993) (see Theorem 2.4.19), there exists a $\zeta < 1$ such that, for every $l = 1,\ldots,m$ and each $k \in \mathbb{N}$,

$$\sup_{z\in D}\left|\Psi^{km+l} h_{n|D_{j(l,z)}}(z) - m\psi(h_{n|D_{j(l,z)}})\right| \le \zeta^k. \tag{3.4.4}$$

Consequently,

$$|K_s'^n| \le \frac{2\lfloor sn\rfloor m}{n}\left(\frac{\zeta\psi(|h_n|)}{(1-\zeta)n\xi_n^2} + \frac{\psi(|h_n|)|\psi(h_n)|}{n\xi_n^2}\right) \xrightarrow[n\to\infty]{} 0. \tag{3.4.5}$$

By eqs. (3.4.2) and (3.4.5), $\mathbb{E}^\psi[(H_s'^n)^2] \to 0$, hence $H_s'^n \to 0$ in probability as $n \to \infty$. It remains to show the local uniformity in s of this convergence.

By eqs. (3.4.3) and (3.4.4), we have that $h_n - \psi(h_n)$ is in the range of $(\mathrm{I} - \Psi)$. Let $\hat{h}_n$ denote its pre-image under $(\mathrm{I} - \Psi)$ (that is, its *potential*), and define the process M^n by

$$M_s^n := \frac{1}{n\xi_n}\sum_{k=1}^{\lfloor sn\rfloor}\Big(\hat{h}_n(Z_k) - \Psi\hat{h}_n(Z_{k-1})\Big).$$

We note that M^n is a $\mathscr{G}_s^n$-martingale where $\mathscr{G}_s^n := \sigma(Z_k : k \le \lfloor sn\rfloor)$. Since $(\Psi h_n)_{n\in\mathbb{N}^*}$ is uniformly bounded by assumption, so is $(\Psi\hat{h}_n)_{n\in\mathbb{N}^*}$. As $n \to \infty$, therefore, we have

$$|H_s'^n - M_s^n| = (n\xi_n)^{-1}|\Psi\hat{h}_n(Z_0) - \Psi\hat{h}_n(Z_{\lfloor sn \rfloor})| \to 0.$$

Likewise,

$$\mathbb{E}^\psi[(M_s^n)^2] \le 2\,\mathbb{E}^\psi(H_s'^n)^2 + 2\,\mathbb{E}^\psi\,|H_s'^n - M_s^n|^2 \to 0.$$

By Doob's inequality (recall Theorem 2.1.12), therefore, $M^n \Rightarrow 0$ in ucp. Hence, also $H'^n \Rightarrow 0$ uniformly on compacts in probability as $n \to \infty$. □

3.4.2 The auxiliary Markov chains

In this subsection, we construct auxiliary Markov chains Z and Z' to which Theorem 3.4.1 applies. Once and for all, we fix our points of interest, i. e., $\{(x_i, y_i) : i \in I\}$ of Theorem 3.3.7 such that $\mu'(x_i) > 0$ and $F(x_i, E) > 0$ for each i. Moreover, we choose a compact set $C \supset \{x_i : i \in I\}$ and constants $0 < \varepsilon, \varepsilon' < \infty$ such that $\varepsilon < \|y_i\| < \varepsilon'$ for all $i \in I$ and such that

$$\inf_{x \in C} F\Big(x, \{y : \varepsilon < \|y\| < \varepsilon'\}\Big) > 0. \tag{3.4.6}$$

Remark. Under Assumptions 3.2.2 and 3.3.2, such a set C always exists by the choice of the points x_i and the continuity of f on $E \times E^*$.

Let $T_1, T_2, \ldots$ denote the successive times of jumps of size between ε and ε' starting from C; that is,

$$T_1 := \inf\Big\{t > 0 : \varepsilon < \|\Delta X_t\| < \varepsilon', X_{t-} \in C\Big\} \quad \text{and} \quad T_{n+1} := T_1 \circ \theta_{T_n} + T_n.$$

The conditional expectation w. r. t. the strict past of the stopping times T_n plays a key role. We set

$$q(x) := F\Big(x, \{y : \varepsilon < \|y\| < \varepsilon'\}\Big)\mathbb{1}_C(x),$$

$$p(x,y) := \begin{cases} q^{-1}(x) f(x,y), & \text{if } x \in C \text{ and } \varepsilon < \|y\| < \varepsilon', \\ 0, & \text{else.} \end{cases}$$

It is well-known that $T_1 < \infty$ a. s. if, and only if, $\mu(q) > 0$. In our case, this holds by eq. (3.4.6). Therefore, $T_n < \infty$ a. s. for all n as well. For convenience, we abbreviate the kernel with density p by Π; its shifted version with density $(x, y) \mapsto p(x, y - x)$ we denote by $\bar{\Pi}$. By Théorème 1 of Weil (1971) (recall Theorem 2.5.26), Π (resp., $\bar{\Pi}$) is the conditional transition probability kernel of the jumps at the time(s) T_n in the following sense: On the set $\{T_n < \infty\}$, for every random variable Y, measurable function g, and all x, we have

$$\mathbb{E}^x[g(\Delta X_{T_n}) \mid \mathscr{F}_{T_n-}] = \Pi g(X_{T_n-}), \tag{3.4.7}$$

$$\mathbb{E}^x[Y \circ \theta_{T_n} \mid \mathscr{F}_{T_n-}] = \bar{\Pi}\, \mathbb{E}^{\cdot}[Y](X_{T_n-}). \tag{3.4.8}$$

We note $\bar{\Pi}\, \mathbb{E}^{\cdot}[Y](x) = \int p(x, y)\, \mathbb{E}^{x+y}[Y] \mathrm{d}y$.

Let $\mathbf{D} := \mathcal{D}([0, 1[; E) \times \mathbb{R}_+ \times C$. For every $k \in \mathbb{N}^*$, we define the $\mathbf{D}$-valued and C-valued random variables

$$Z_k := \Big(s \mapsto X_{(1-s)T_{k-1}+sT_k}, T_k - T_{k-1}, X_{T_k-}\Big) \quad \text{and} \quad Z'_k := X_{T_k-}.$$

The corresponding filtration $(\mathscr{G}_k)_{k \in \mathbb{N}^*}$ is given by $\mathscr{G}_k := \mathscr{F}_{T_k-}$. We emphasise that we exclude time $k = 0$. From eq. (3.4.8) and $T_1 < \infty$ a. s., we deduce that $Z = (Z_k)_{k \in \mathbb{N}^*}$ and $Z' = (Z'_k)_{k \in \mathbb{N}^*}$ are $\mathscr{G}_k$-Markov chains. We denote their transition probabilities by Ψ and Φ, respectively. We refer to Section 3.6 for technical results on these auxiliary Markov chains.

3.4.2 Lemma. *Let $(g, t, x) \in \mathbf{D}$, let $A \subseteq C$ and $\mathbf{A} \subseteq \mathbf{D}$ be measurable, and let $k \in \mathbb{N}^*$. Then*

$$\Phi(x, A) = \bar{\Pi}\, \mathbb{P}^{\cdot}(Z'_1 \in A)(x), \tag{3.4.9}$$

$$\Psi^{k+1}((g, t, x), \mathbf{A}) = \Phi^k \Psi(x, \mathbf{A}). \tag{3.4.10}$$

Proof. We deduce eqs. (3.4.9) and (3.4.10) directly from eq. (3.4.8) and the Markov property of X, respectively. □

By Lemma 3.4.2, Theorem 3.4.1 applies to Z' and, also, to Z.

3.4.3 Lemma. *Grant Assumptions 3.2.2 and 3.3.2. Then the Markov chain Z' is strong Feller. Its state space C is* petite *with respect to Φ.*

Proof. Let f be a bounded Borel function and $x_0 \in C$. Under Assumption 3.3.2, we deduce from Lebesgue's dominated convergence theorem that q is continuous. By eq. (3.4.6), we have that $x \mapsto p(x,y)$ is also continuous for every y and $\sup\{p(x,y) : x \in C, y \in E\} < \infty$. Again by Lebesgue's dominated convergence theorem, we conclude that

$$\lim_{x\to x_0} \bar{\Pi}g(x) = \lim_{x\to x_0} \int p(x,y)g(x+y)\mathrm{d}y = \int p(x_0,y)g(x+y)\mathrm{d}y = \bar{\Pi}g(x_0).$$

By eq. (3.4.9), consequently, $\Phi = \bar{\Pi}\,\mathbb{P}^{\cdot}(Z_1' \in \cdot)$ is strong Feller on C.

By the same argument as for the equivalence of $T_1 < \infty$ a. s. and $\mu(q) > 0$, we have that the measure with μ-density q is an irreducibility measure of Z'. Under Assumption 3.2.2, it is absolutely continuous. Thus, its support has non-empty interior. By Theorems 6.2.5 and 6.2.9 of Meyn and Tweedie (1993) (see Theorem 2.4.12), therefore, every compact set – hence the state space C of Z' – is petite with respect to Φ. □

3.4.4 Corollary. *Grant Assumptions 3.2.2 and 3.3.2. Then the state space* $\mathbf{D}$ *of Z is petite w. r. t. Ψ.*

Proof. By Lemma 3.4.3, there exists a probability ρ on $\mathbb{N}^*$ and a non-trivial measure ν_ρ on C such that, for every Borel set $A \subseteq C$,

$$\inf_{x\in C} \sum_{k=1}^{\infty} \rho(k)\Phi^k(x,A) \geq \nu_\rho(A).$$

Let $(g,t,x) \in \mathbf{D}$, $\mathbf{A} \subseteq \mathbf{D}$ be measurable, and $\tilde{\rho}$ be the probability on $\mathbb{N}^*$ given by $\tilde{\rho}(1) = 0$ and $\tilde{\rho}(k) = \rho(k-1)$ for $k > 1$. By eq. (3.4.10), then

$$\sum_{k=1}^{\infty} \tilde{\rho}(k)\Psi^k((g,t,x),A) = \sum_{k=1}^{\infty} \rho(k)\Phi^k\Psi(x,A) \geq \nu_\rho\Psi(A) =: \tilde{\nu}_{\tilde{\rho}}(A).$$

Since ν_ρ is non-trivial, so is $\tilde{\nu}_{\tilde{\rho}}$. □

3.4.3 Proof of Theorem 3.3.6

Throughout the remainder of Section 3.4, we work under the law $\mathbb{P}^\pi$ for some initial probability π on E and, for presentational purposes, we suppose w. l. o. g. that $\mu(q) = 1$.

We consider the processes $G^{n,\eta}$, $J^{n,\eta}$ and $S^{n,\eta}$ given by

$$G_s^{n,\eta}(x,y) := \frac{1}{n}\sum_{k=1}^{\lfloor sn\rfloor} g_1^{\eta,x}(X_{T_k-})g_2^{\eta,y}(\Delta X_{T_k}), \tag{3.4.11}$$

$$J_s^{n,\eta}(x) := \frac{1}{n}\sum_{k=1}^{\lfloor sn\rfloor} g_1^{\eta,x}(X_{T_k-}) \quad \text{and} \quad S_s^{n,\eta}(x) := \frac{1}{n}\int_0^{T_{\lfloor sn\rfloor}} g_1^{\eta,x}(X_r)\mathrm{d}r. \tag{3.4.12}$$

We emphasise that these processes are of the form $\sum_{k=1}^{\lfloor sn\rfloor} h_n(Z_k)$ where Z is the auxiliary Markov chain defined in Section 3.4.2. We utilise the following preliminary condition as $n \to \infty$ (cf., eq. (3.3.2)):

$$n\eta_{1,n}^d\eta_{2,n}^d \to \infty, \quad \text{and} \quad \eta_{1,n} \to 0, \eta_{2,n} \to 0. \tag{3.4.13}$$

3.4.5 Lemma. *Grant Assumptions 3.2.2, 3.3.1 and 3.3.2. Let $\eta_n = \eta_{1,n}$ be such that eq.* (3.4.13) *holds. Then the following convergences hold uniformly on compacts in probability:*

$$J_s^{n,\eta_n}(x) \underset{n\to\infty}{\overset{\text{ucp}}{\Longrightarrow}} sq(x)\mu'(x) \quad \textit{and} \quad S_s^{n,\eta_n}(x) \underset{n\to\infty}{\overset{\text{ucp}}{\Longrightarrow}} s\mu'(x).$$

Proof. Let ψ and φ denote the invariant probabilities of Z and Z', respectively. We apply Theorem 3.4.1:

(i) We note that $J^{n,\eta_n}(x)$ is of the form eq. (3.4.1) with $\xi_n = \eta_n^d$ and $h_n : C \to \mathbb{R}$ given by $h_n(z) = g_1((z-x)/\eta_n)$; $(h_n)_{n\in\mathbb{N}^*}$ is uniformly bounded. By Corollary 3.6.6 where $\mu(q) = 1$, q is the μ-density of φ. Also q and μ' are continuous. By Lebesgue's differentiation theorem, thus,

$$\eta_n^{-d}\varphi(h_n) = \eta_n^{-d}\int \mu(\mathrm{d}z)q(z)g_1((z-x)/\eta_n) \xrightarrow[n\to\infty]{} q(x)\mu'(x).$$

Since $n\eta_n^d \to \infty$, likewise, $(n\eta_n^{2d})^{-1}\varphi(|h_n|) \to 0$ as $n \to \infty$. ⋄

(ii) We note that $S^{n,\eta_n}(x)$ is of form eq. (3.4.1) with $\xi_n = \eta_n^d$ and $h_n : \mathbf{D} \to \mathbb{R}$ given by $h_n(g,t,z) = t\int_0^1 g_1((g(s)-x)/\eta_n)\mathrm{d}s$. By Corollary 3.6.6, $\psi = \varphi\Psi$. By Lemmata 3.6.2 and 3.6.5, thus,

$$\eta_n^{-d}\varphi(h_n) = \eta_n^{-d}\int \mu(\mathrm{d}z)g_1((z-x)/\eta_n) \xrightarrow[n\to\infty]{} \mu'(x).$$

Likewise, $(n\eta_n^{2d})^{-1}\varphi(|h_n|) \le (n\eta_n^{2d})^{-1}\int\mu(\mathrm{d}z)|g_1((z-x)/\eta_n)| \to 0$. By Corollary 3.6.4, in addition, we observe

$$\frac{\psi(h_n^2)}{n\eta_n^{2d}} \le \frac{2\|g_1\|_\infty}{\inf_{z\in C} q(z)}\frac{\int\mu(\mathrm{d}z)|g_1((z-x)/\eta_n)|}{n\eta_n^{2d}} \xrightarrow[n\to\infty]{} 0. \qquad \square$$

3.4.6 Lemma. *Grant Assumptions 3.2.2, 3.3.1 and 3.3.2. Let $\eta_n = (\eta_{1,n},\eta_{2,n})$ be such that eq.* (3.4.13) *holds. Then the following convergence holds uniformly on compacts in probability:*

$$G_s^{n,\eta_n}(x,y) \overset{\mathrm{ucp}}{\underset{n\to\infty}{\Longrightarrow}} sf(x,y)\mu'(x).$$

Proof. Let $(\mathscr{H}_s^n)_{s\ge 0}$ be the filtration given by $\mathscr{H}_s^n := \mathscr{F}_{T_{\lfloor sn\rfloor+1}-}$. By eq. (3.4.7), we have

$$\mathbb{E}[\Delta G_s^{n,\eta_n} \mid \mathscr{H}_{s-}^n] = g_1^{\eta_n,x}(Z_k')\Pi g_2^{\eta_n,y}(Z_k') \quad \text{for } s = k/n.$$

Thus, the compensator of G^{n,η_n} w.r.t. $(\mathscr{H}_s^n)_{s\ge0}$ is given by

$$H_s^{n,\eta_n} := n^{-1}\sum_{k=1}^{\lfloor sn\rfloor} g_1^{\eta_n,x}(Z_k')\Pi g_2^{\eta_n,y}(Z_k').$$

Fix $s \ge 0$. In analogy to the proof of Lemma 3.4.3, $\Pi g_2^{\eta_n,y}$ is continuous under Assumption 3.3.2. In analogy to Lemma 3.4.5, $s \mapsto n^{-1}\sum_{k=1}^{\lfloor sn\rfloor}|g_1^{\eta_n,x}(Z_k')|$ converges in ucp to a non-trivial process as $n \to \infty$. Therefore,

$$\left|H_s^{n,\eta_n} - \Pi g_2^{\eta_n,y}(x)J_s^{n,\eta_n}(x)\right| \le \sup_{z\in B_{\eta_n}(x)}\left|\Pi g_2^{\eta_n,y}(z) - \Pi g_2^{\eta_n,y}(x)\right| \cdot \frac{1}{n}\sum_{k=1}^{\lfloor sn\rfloor}\left|g_1^{\eta_n,x}(Z_k')\right|$$
$$\xrightarrow[n\to\infty]{} 0.$$

Since p is continuous under Assumption 3.3.2, $\lim_{n\to\infty} \Pi g_2^{\vartheta_n,y}(x) = p(x,y)$ by Lebesgue's differentiation theorem. We recall $f(x,y) = q(x)p(x,y)$. By Lemma 3.4.5, hence,

$$H_s^n \underset{n\to\infty}{\overset{\text{ucp}}{\Longrightarrow}} sf(x,y)\mu'(x).$$

It remains to prove $M_s^n := G_s^n - H_s^n \Rightarrow 0$ uniformly on compacts in probability. By eq. (3.4.13), we have $\sup_s \|\Delta M_s^n\|_\infty \le (n\eta_n^d \vartheta_n^d)^{-1}\|g_1\|_\infty \|g_2\|_\infty \to 0$. By the martingale limit theorem 2.3.10, thus, it is sufficient to show that the predictable quadratic variation $\langle M^n, M^n\rangle_s$ of M^n converges in probability to zero for all s. We observe

$$\begin{aligned}\Big\langle M^n, M^n \Big\rangle_s &= \frac{1}{n^2}\sum_{k=1}^{\lfloor sn\rfloor} \mathbb{E}^\pi\left[g_1^{\eta_n,x}(Z_k')^2 \Big(g_2^{\eta_n,y}(\Delta X_{T_k}) - \Pi g_2^{\eta_n,y}(Z_k')\Big)^2 \,\Big|\, \mathscr{H}_{k/n}^n \right]\\ &\le \frac{1}{n\eta_{1,n}^d \eta_{2,n}^d} \cdot \frac{1}{n}\sum_{k=1}^{\lfloor sn\rfloor} \eta_{1,n}^d g_1^{\eta_n,x}(Z_k')^2 \int\limits_{B_1(0)} p(Z_k', y+\eta_{2,n}z) g_2(z)^2 \mathrm{d}z.\end{aligned}$$

In analogy to Lemma 3.4.5 again, $s \mapsto n^{-1}\sum_{k=1}^{\lfloor sn\rfloor} \eta_{1,n}^d g_1^{\eta_n,x}(Z_k')^2$ converges in ucp to a non-trivial process as $n\to\infty$. As in the proof of Lemma 3.4.3, moreover, p is bounded on $C \times E$. Consequently, $\langle M^n, M^n\rangle_s \to 0$ in probability as $n \to \infty$. □

Next, we carry Lemmata 3.4.5 and 3.4.6 over to the time-scale of X. Let J be the process given by

$$J_t := \sum_{k=1}^{\infty} \mathbb{1}_{[0,t]}(T_k). \tag{3.4.14}$$

We note that J is a non-decreasing additive functional of X. It is the random clock of Z (and Z') in terms of X. By eq. (3.3.1) – where $H_t = t$ –, we have $\mu(J) = \mu(q) = 1$.

3.4.7 Lemma. *Grant Assumptions 3.2.2, 3.3.1 and 3.3.2. Let $v : \mathbb{R}_+ \to \mathbb{R}_+$ denote a deterministic equivalent of X, and let η_t and $(x,y) \in E \times E^*$ be as in Theorem 3.3.6. Then*

$$\text{the family } \left\{ \mathscr{L}\left(G_{J_t/v_t}^{v_t,\eta_t}(x,y), S_{J_t/v_t}^{v_t,\eta_t}(x) \mid \mathbb{P}^\pi \right) : t > 0 \right\} \text{ is tight.} \tag{3.4.15}$$

Moreover, each limit point of the family in eq. (3.4.15) is the law $\mathscr{L}(f(x,y)\mu'(x)\tilde{L}, \mu'(x)\tilde{L})$ for some positive random variable $\tilde{L}$.

Proof. As J is a non-decreasing additive functional of X, by definition, the families $\{\mathscr{L}(J_t/v_t \mid \mathbb{P}^\pi) : t > 0\}$ and $\{\mathscr{L}(v_t/J_t \mid \mathbb{P}^\pi) : t > 0\}$ are tight. By Proposition 2.3.8 and Lemma 3.4.6, thus,

$$\text{the family } \{\mathscr{L}(G^{v_t,\eta_t}(x,y), S^{v_t,\eta_t}(x), J_t/v_t, v_t/J_t \mid \mathbb{P}^\pi) : t > 0\} \text{ is tight.} \tag{3.4.16}$$

Let $\mathbb{Q}$ denote a limit point of the family in eq. (3.4.16), and let $(t_n)_{n\in\mathbb{N}}$ a sequence such that

$$\mathscr{L}(G^{v_{t_n},\eta_{t_n}}(x,y), S^{v_{t_n},\eta_{t_n}}(x), J_{t_n}/v_{t_n}, v_{t_n}/J_{t_n} \mid \mathbb{P}^\pi) \xrightarrow[n\to\infty]{\mathrm{w}} \mathbb{Q}.$$

On some extension of the probability space, w. l. o. g., there exists a random variable $\tilde{L} > 0$ such that $\mathbb{Q} = \mathscr{L}(s \mapsto sf(x,y)\mu'(x), s \mapsto s\mu'(x), \tilde{L}, 1/\tilde{L})$. Since its first and second marginal are the laws of continuous processes, we have

$$\mathscr{L}\Big(G^{v_{t_n},\eta_{t_n}}_{J_{t_n}/v_{t_n}}(x,y), S^{v_{t_n},\eta_{t_n}}_{J_{t_n}/v_{t_n}}(x) \mid \mathbb{P}^\pi\Big) \xrightarrow[n\to\infty]{\mathrm{w}} \mathscr{L}\Big(f(x,y)\mu'(x)\tilde{L}, \mu'(x)\tilde{L}\Big). \qquad \square$$

Proof (of Theorem 3.3.6). For every $t \geq 0$ and each x and y, we have

$$\hat{f}^{\eta_t}_t(x,y) = \frac{G^{v_t,\eta_t}_{J_t/v_t}(x,y)}{S^{v_t,\eta_t}_{J_t/v_t}(x) + v_t^{-1}\int_{T_{J_t}}^t g_1^{\eta_t,x}(X_s)\mathrm{d}s}.$$

Let $h_n : \mathbf{D} \to \mathbb{R}$ be given by $h_n(g,t,z) := t\int_0^1 |g_1^{\eta_n,x}(g(s))|\mathrm{d}s$. By Lemma 3.6.2 and Corollaries 3.6.4 and 3.6.6, we have $\psi(h_n^2) \leq 2\|g_1\|_\infty \eta_{1,n}^{-d}(\inf_{z\in C} q(z))^{-1}\mu(|g_1^{\eta_n,x}|)$. By Markov's inequality, since $v_t^2\eta_{1,t}^d \to \infty$, therefore,

$$v_t^{-1}\int\limits_{T_{J_t}}^{t} g_1^{\eta_t,x}(X_s)\mathrm{d}s \leq v_t^{-1}h_{v_t}(Z_{J_t+1}) \xrightarrow[t\to\infty]{\mathbb{P}^\psi} 0. \tag{3.4.17}$$

By Proposition 17.1.6 of Meyn and Tweedie (1993), in analogy to the proof of Theorem 3.4.1, this convergence in probability holds under every law $\mathbb{P}^\pi$.

We recall the results from Lemma 3.4.7. Let $\tilde{L} > 0$ be a random variable such that the law $\mathscr{L}(f(x,y)\mu'(x)\tilde{L}, \mu'(x)\tilde{L})$ is a limit point of the family in eq. (3.4.15).

Moreover, let $(t_n)_{n\in\mathbb{N}^*}$ be a sequence such that

$$\left(G^{v_{t_n},\eta_{t_n}}_{J_{t_n}/v_{t_n}}(x,y), S^{v_{t_n},\eta_{t_n}}_{J_{t_n}/v_{t_n}}(x)\right) \xrightarrow[n\to\infty]{\mathscr{L}} \left(f(x,y)\mu'(x)\tilde{L}, \mu'(x)\tilde{L}\right).$$

We recall $\mu'(x) > 0$. Consequently, $\hat{f}^{\eta_{t_n}}_{t_n}(x,y) \to f(x,y)$ in law as $n \to \infty$ by the continuous mapping theorem. As this limit is unique and independent of the particular limit point of the family in eq. (3.4.15), we have that $\hat{f}^{\eta_t}_t(x,y)$ converges to $f(x,y)$ in law, hence, in probability. □

3.4.4 Proofs of Theorem 3.3.7 and Corollary 3.3.8

In this subsection, we work on the extended space eq. (3.2.10), L denotes the Mittag-Leffler process of order $0 < \delta \leq 1$, and $W = (W^i)_{i\in I}$ denotes an I-dimensional standard Wiener process such that L, W and $\mathscr{F}$ are independent.

In addition to the processes $G^{n,\eta}$, $J^{n,\eta}$ and $S^{n,\eta}$ given in eqs. (3.4.11) and (3.4.12), we consider the process $U^{n,\eta}$ given by

$$U^{n,\eta}_s(x,y) := \sqrt{n\eta_1^d\eta_2^d}\left(G^{n,\eta}_s(x,y) - \frac{\mu(g_1^{\eta,x}Fg_2^{\eta,y})}{\mu(g_1^{\eta,x})}S^{n,\eta}_s(x)\right). \tag{3.4.18}$$

We emphasise again that these processes are of the form $\sum_{k=1}^{\lfloor sn\rfloor} h_n(Z_k)$ where Z is the auxiliary Markov chain defined in Section 3.4.2.

3.4.8 Lemma. *Grant Assumptions 3.2.2, 3.2.4, 3.3.1 and 3.3.2. Let $\eta_n = (\eta_{1,n}, \eta_{2,n})$ be such that eq.* (3.4.13) *holds. Then we have the following convergence in law in $\mathcal{D}(\mathbb{R}^I)$:*

$$\left(U^{n,\eta_n}_s(x_i,y_i)\right)_{i\in I} \overset{\mathscr{L}}{\underset{n\to\infty}{\Longrightarrow}} \left(\mu'(x_i)\sigma(x_i,y_i)W^i_s\right)_{i\in I},$$

where $\sigma(x,y)^2$ is given by eq. (3.3.4)*.*

Proof. For $n \in \mathbb{N}^*$, let $M^{n,\eta}$ be the process given by

$$M^{n,\eta}_s(x,y) := \frac{\sqrt{\eta_1^d\eta_2^d}}{\sqrt{n}}\sum_{k=1}^{\lfloor sn\rfloor}\left(g_1^{\eta,x}(Z'_k)g_2^{\eta,y}(\Delta X_{T_k}) - \int_{T_{k-1}}^{T_k} g_1^{\eta,x}(X_s)Fg_2^{\eta,y}(X_s)\mathrm{d}s\right),$$

and let $(\mathscr{H}^n_s)_{s\geq 0}$ be given by $\mathscr{H}^n_s := \mathscr{F}_{T_{\lfloor sn \rfloor}}$. By the martingale limit theorem 2.3.10, it is sufficient to prove (i)–(iv) as follows:

(i) We have $U^{n,\eta_n}_s(x,y) - M^{n,\eta_n}_s(x,y) \Rightarrow 0$ in ucp as $n \to \infty$.

(ii) The process $M^{n,\eta}$ is an $\mathscr{H}^n_s$-martingale for each n.

(iii) For all $i, j \in I$, we have

$$\left\langle M^{n,\eta_n}(x_i,y_i), M^{n,\eta_n}(x_j,y_j) \right\rangle_s \xrightarrow[n\to\infty]{\mathbb{P}^\pi} s[\sigma(x_i,y_i)\mu'(x)]^2\delta_{ij}.$$

(iv) We have the "conditional Lyapunov condition"

$$K^{n,\eta_n}_s(x,y) := \sum_{k=1}^{\lfloor sn \rfloor} \mathbb{E}^\pi \left[\left(\Delta M^{n,\eta_n}_{k/n}(x,y) \right)^4 \,\middle|\, \mathscr{H}^n_{k/n-} \right] \xrightarrow[n\to\infty]{\mathbb{P}^\pi} 0.$$

(i) We note that $U^{n,\eta}(x,y) - M^{n,\eta}(x,y)$ is of form eq. (3.4.1) with $h_n : \mathbf{D} \to \mathbb{R}$ given by

$$h_n(g,t,z) = t \int_0^1 g_1\left(\frac{g(s)-x}{\eta_{1,n}}\right) \left(Fg_2^{\eta_n,y}(g(s)) - \frac{\mu(g_1^{\eta_n,x} Fg_2^{\eta_n,y})}{\mu(g_1^{\eta_n,x})} \right) \mathrm{d}s,$$

and $\xi_n = \eta_{1,n}^{d/2}\eta_{2,n}^{-d/2}n^{-1/2}$. By Lemmata 3.6.2 and 3.6.5 and Corollary 3.6.6, we have

$$\xi_n^{-1}\psi(h_n) = \sqrt{n\eta_{1,n}^d\eta_{2,n}^d} \int \mu(\mathrm{d}z) g_1^{\eta,x}(z) \left(Fg_2^{\eta_n,y}(z) - \frac{\mu(g_1^{\eta_n,x} Fg_2^{\eta_n,y})}{\mu(g_1^{\eta_n,x})} \right) \equiv 0.$$

Since $\eta_{2,n} \to 0$, we also observe

$$\frac{\psi(|h_n|)}{n\xi_n^2} \leq \eta_{2,n}^d \left(\mu(|g_1^{\eta_n,x} Fg_2^{\eta_n,y}|) + \mu(|g_1^{\eta_n,x}|) \cdot \left| \frac{\mu(g_1^{\eta_n,x} Fg_2^{\eta_n,y})}{\mu(g_1^{\eta_n,x})} \right| \right) \xrightarrow[n\to\infty]{} 0.$$

By Corollary 3.6.4, likewise,

$$\frac{\psi(h_n^2)}{n\xi_n^2} \leq \frac{2\eta_{2,n}^d \|g_1\|_\infty \|Fg_2^{\eta_n,y}\|_\infty}{\inf_{z\in C} q(z)} \left(\mu\left(\left|g_1^{\eta_n,x} Fg_2^{\eta_n,y}\right|\right) + \mu\left(\left|g_1^{\eta_n,x}\right|\right) \left| \frac{\mu(g_1^{\eta_n,x} Fg_2^{\eta_n,y})}{\mu(g_1^{\eta_n,x})} \right| \right)$$
$$\xrightarrow[n\to\infty]{} 0.$$

Since $n\xi_n \to \infty$, we deduce from Theorem 3.4.1 that (i) holds. $\diamond$

(ii) By construction, $M^{n,\eta}$ is integrable and adapted to $(\mathscr{H}_s^n)_{s\geq 0}$. For $s = k/n$, we note $\mathscr{H}_{s-}^n = \mathscr{F}_{T_{k-1}}$. By eq. (3.3.1) – where $H_t = t$ – the compensator of our process's jump measure is given by $\mathrm{d}t \otimes F(X_t, \mathrm{d}y)$. By Doob's optional sampling theorem, thus,

$$\mathbb{E}^\pi \left[g_1^{\eta,x}(Z_k') g_2^{\eta,y}(\Delta X_{T_k}) - \int_{T_{k-1}}^{T_k} g_1^{\eta,x}(X_s) F g_2^{\eta,y}(X_s) \mathrm{d}s \,\middle|\, \mathscr{F}_{T_{k-1}} \right] = 0$$

for all $k \in \mathbb{N}^*$. Therefore, $M^{n,\eta}(x,y)$ is an $\mathscr{H}_s^n$-martingale. $\diamond$

(iii) Let $i, j \in I$. In analogy to step (ii), we deduce

$$\left\langle M^{n,\eta_n}(x_i,y_i), M^{n,\eta_n}(x_j,y_j) \right\rangle_s$$
$$= \frac{\eta_{1,n}^d \eta_{2,n}^d}{n} \sum_{k=1}^{\lfloor sn \rfloor} \mathbb{E}^\pi \left[g_1^{\eta_n,x_i} g_1^{\eta_n,x_j}(Z_k') g_2^{\eta_n,y_i} g_2^{\eta_n,y_j}(\Delta X_{T_k}) \,\middle|\, \mathscr{F}_{T_{k-1}} \right].$$

For all n large enough, we have $g_1^{\eta_n,x_i} g_1^{\eta_n,x_j} = 0$ whenever $x_i \neq x_j$, and $g_2^{\eta_n,y_i} g_2^{\eta_n,y_j} = 0$ whenever $y_i \neq y_j$. For all ω, if $i \neq j$, thus, $\langle M^{n,\eta_n}(x_i,y_i), M^{n,\eta_n}(x_j,y_j) \rangle_s \to 0$.

Moreover, let $J_s'^{n,\eta_n}(x) := n^{-1}\eta_{1,n}^d \sum_{k=1}^{\lfloor sn \rfloor} \mathbb{E}^{X_{T_{k-1}}}[g_1^{\eta,x}(Z_1')^2]$. We note that J'^{n,η_n} is of form eq. (3.4.1) with $\xi_n = \eta_{1,n}^d$ and $h_n : \mathbf{D} \to \mathbb{R}$ given by

$$h_n(g,t,z) = \mathbb{E}^{g(0)} \left[g_1\Big((Z_1' - x)/\eta_{1,n} \Big)^2 \right].$$

By Lemma 3.6.5 and Corollary 3.6.6 and under Assumption 3.3.2, we observe

$$\eta_{1,n}^{-d} \psi(h_n) = \int \mu'(x+\eta_{1,n}z) q(x+\eta_{1,n}z) g_1(z)^2 \mathrm{d}z \xrightarrow[n\to\infty]{} \mu'(x)q(x) \int g_1(z)^2 \mathrm{d}z.$$

By Theorem 3.4.1, since h_n is non-negative and uniformly bounded, thus,

$$J_s'^{n,\eta_n}(x) \xrightarrow[n\to\infty]{\text{ucp}} sq(x)\mu'(x) \int g_1(z)^2 \mathrm{d}z. \tag{3.4.19}$$

Hence, we observe

$$\begin{aligned}
&\left| \left\langle M^{n,\eta_n}(x,y), M^{n,\eta_n}(x,y) \right\rangle_s - J_s'^{n,\eta_n}(x) p(x,y) \int g_2(w)^2 \mathrm{d}w \right| \\
&\quad \le J_s'^{n,\eta_n}(x) \int g_2(w)^2 \mathrm{d}w \sup_{z,w\in B_1(0)} \left| p(x+\eta_{1,n}z, y+\eta_{2,n}w) - p(x,y) \right| \\
&\quad \xrightarrow[n\to\infty]{\mathbb{P}^\pi} 0.
\end{aligned}$$

Since $f(x,y) = q(x)p(x,y)$, consequently,

$$\left\langle M^{n,\eta_n}(x,y), M^{n,\eta_n}(x,y) \right\rangle_s \xrightarrow[n\to\infty]{\mathbb{P}^\pi} sf(x,y)\mu'(x) \int g_1(w)^2 \mathrm{d}w \int g_2(z)^2 \mathrm{d}z;$$

that is, (iii) holds. ⋄

(iv) We observe $|K_s^{n,\eta_n}(x,y)| \le K_s'^{n,\eta_n} + K_s''^{n,\eta_n}$, where

$$K_s'^{n,\eta_n} := \frac{4\eta_{1,n}^{2d}\eta_{2,n}^{2d}}{n^2} \sum_{k=1}^{\lfloor sn \rfloor} \mathbb{E}^{X_{T_{k-1}}} \left[\left(g_1^{\eta,x}(Z_1') g_2^{\eta,y}(\Delta X_{T_1}) \right)^4 \right],$$

and

$$K_s''^{n,\eta_n} := \frac{4\eta_{1,n}^{2d}\eta_{2,n}^{2d}}{n^2} \sum_{k=1}^{\lfloor sn \rfloor} \mathbb{E}^{X_{T_{k-1}}} \left[\left(\int_0^{T_1} g_1^{\eta,x} F g_2^{\eta,y}(X_s) \mathrm{d}s \right)^4 \right].$$

We note that K'^{n,η_n} and K''^{n,η_n} are of form eq. (3.4.1) with $\xi_n = n\eta_{1,n}^{2d}\eta_{1,n}^{2d}/4$ and, respectively,

$$h_n(g,t,z) = \mathbb{E}^{g(0)} \left[g_1((Z_1'-x)/\eta_{1,n})^4 g_2((\Delta X_{T_1}-y)/\eta_{2,n})^4 \right],$$

and

$$h_n(g,t,z) = \mathbb{E}^{g(0)} \left[\left(\int_0^{T_1} g_1\left(\frac{X_s - x}{\eta_{1,n}}\right) \int F(X_s, \mathrm{d}w) g_2\left(\frac{w-y}{\eta_{2,n}}\right) \right)^4 \right].$$

By Lemma 3.6.5 and Corollary 3.6.6, for K'^n, we have

$$\frac{\psi(h_n)}{\xi_n} = \frac{4}{n\eta_{1,n}^d\eta_{2,n}^d}\iint \mu'(x+\eta_{1,n}z)g_1(z)^4 f(x+\eta_{1,n}z, y+\eta_{2,n}w)g_2(w)^4\mathrm{d}w\mathrm{d}z$$
$$\xrightarrow[n\to\infty]{} 0.$$

By Corollary 3.6.4 and Lemma 3.6.5, for K''^n moreover, there exists a $\zeta < \infty$ such that

$$\frac{\psi(h_n)}{\xi_n} \leq \frac{4\zeta}{n\eta_{1,n}^d\eta_{2,n}^d}\iint \mu'(x+\eta_{1,n}z)|g_1(z)|f(x+\eta_{1,n}z, y+\eta_{2,n}w)|g_2(w)|\mathrm{d}w\mathrm{d}z$$
$$\xrightarrow[n\to\infty]{} 0.$$

Since, in both cases, h_n is non-negative and uniformly bounded, we deduce from Theorem 3.4.1 that $\left|K_s^{n,\eta_n}(x,y)\right| \leq K_s'^{n,\eta_n} + K_s''^{n,\eta_n} \Rightarrow 0$ in ucp as $n \to \infty$. □

Next, we carry Lemma 3.4.8 over to the time-scale of X. We recall that the additive functional J of X, given in eq. (3.4.14), is the random clock of Z (and Z') in terms of X. In addition, let L^t denote the process given by $L_s^t := v_t^{-1}J_{st}$.

We recall that, under Darling–Kac's condition, we have Touati's theorem 2.5.24 at hand. Recalling Lemma 3.4.5, by Proposition 2.3.6 (iii), we directly obtain:

3.4.9 Lemma. *Grant Assumptions 3.2.2, 3.2.4, 3.3.1 and 3.3.2. Let $\eta_t = \eta_{1,t}$ be such that eq.* (3.3.2) *holds. Then we have the following convergence in law in $\mathcal{D}(\mathbb{R}^{1+I})$:*

$$\left(L^t, \left(S_{L^t}^{v_t,\eta_t}(x_i)\right)_{i\in I}\right) \xRightarrow[t\to\infty]{\mathscr{L}} \left(L, \left(\mu'(x_i)L\right)_{i\in I}\right).$$ □

3.4.10 Lemma. *Grant Assumptions 3.2.2, 3.2.4, 3.3.1 and 3.3.2. Let $\eta_t = (\eta_{1,t}, \eta_{2,t})$ be such that eq.* (3.3.2) *holds. Then we have the following convergence in law in $\mathcal{D}(\mathbb{R}^{1+I})$:*

$$\left(L^t, \left(U^{v_t,\eta_t}(x_i,y_i)\right)_{i\in I}\right) \xRightarrow[t\to\infty]{\mathscr{L}} \left(L, \left(\mu'(x_i)\sigma(x_i,y_i)W^i\right)_{i\in I}\right),$$

where $\sigma(x,y)^2$ is given by eq. (3.3.4).

Proof. From Lemma 3.4.9 and Lemma 3.4.8, we infer

$$L^t \xrightarrow[t\to\infty]{\mathscr{L}} L \quad \text{and} \quad \Big(U^{v_t,\eta_t}(x_i,y_i)\Big)_{i\in I} \xrightarrow[t\to\infty]{\mathscr{L}} \Big(\mu'(x_i)\sigma(x_i,y_i)W^i\Big)_{i\in I}. \tag{3.4.20}$$

Thus, the families

$$\{\mathscr{L}(L^t \mid \mathbb{P}^\pi) : t \geq 0\} \quad \text{and} \quad \Big\{\mathscr{L}\Big((U^{v_t,\eta_t}(x_i,y_i))_{i\in I} \mid \mathbb{P}^\pi\Big) : t \geq 0\Big\}$$

are C-tight. By Proposition 2.3.8, we conclude that

$$\text{the family} \quad \Big\{\mathscr{L}\Big(L^t, (U^{v_t,\eta_t}(x_i,y_i))_{i\in I} \mid \mathbb{P}^\pi\Big) : t \geq 0\Big\} \quad \text{is C-tight.} \tag{3.4.21}$$

In the remainder of this proof, we abbreviate $\boldsymbol{U}^{v_t} := (U^{v_t,\eta_t}(x_i,y_i))_{i\in I}$.

Let $(\bar{\Omega},\bar{\mathscr{F}}) := (\mathcal{D}(\mathbb{R}\times\mathbb{R}^I),\mathscr{D}(\mathbb{R}\times\mathbb{R}^I))$ denote the canonical space, and let $(L,\boldsymbol{W})$ be the canonical process. Moreover, let $\bar{\mathbb{P}}$ be an arbitrary limit point of the family in eq. (3.4.21). We deduce from eq. (3.4.20) that its marginals are given by the Mittag-Leffler law of order δ and the I-dimensional (scaled) Wiener law, respectively. For convenience, we abbreviate $\mathbb{Q}_1 := \mathscr{L}(L \mid \bar{\mathbb{P}})$ and $\mathbb{Q}_2 := \mathscr{L}(\boldsymbol{W} \mid \bar{\mathbb{P}})$. Suppose that L and $\boldsymbol{W}$ are independent processes under $\bar{\mathbb{P}}$. Then $\bar{\mathbb{P}} = \mathbb{Q}_1 \otimes \mathbb{Q}_2$ holds. As $\bar{\mathbb{P}}$ is an arbitrary limit point of the family in eq. (3.4.21), then it has to be unique. Hence, $(\mathscr{L}((L^t,\boldsymbol{U}^{v_t}) \mid \mathbb{P}^\pi) \to \mathbb{Q}_1\otimes\mathbb{Q}_2$ weakly as $t\to\infty$. ⋄

Let K denote the right-inverse of L, i. e., $K_t := \inf\{s : L_s > t\}$, and let $(\mathscr{H}_t)_{t\geq 0}$ be the filtration on $\bar{\Omega}$ which is generated by the process $(K,\boldsymbol{W})$. Suppose that – under $\bar{\mathbb{P}}$ – K and $\boldsymbol{W}$ are processes with independent increments relative to $(\mathscr{H}_t)_{t\geq 0}$. (That is, $K_{t+s} - K_t$ and $\mathscr{H}_t$ are independent for all $s,t > 0$, and $\boldsymbol{W}_{t+s} - \boldsymbol{W}_t$ and $\mathscr{H}_t$ are independent for all $s,t > 0$.) Then, in analogy to Step 6 on p. 122 of Höpfner et al. (1990), we deduce that – under $\bar{\mathbb{P}}$ – the pair $(K,\boldsymbol{W})$ itself is a process with independent increments relative to $(\mathscr{H}_t)_{t\geq 0}$. We recall that K is a δ-stable subordinator, thus, purely discontinuous (resp., deterministic if $\delta = 1$). Since $\boldsymbol{W}$ is continuous, hence, K and $\boldsymbol{W}$ are independent processes – under $\bar{\mathbb{P}}$. Consequently, $\bar{\mathbb{P}} = \mathbb{Q}_1\otimes\mathbb{Q}_2$. ⋄

It remains to show that – under $\bar{\mathbb{P}}$ – K and $\boldsymbol{W}$ are processes with independent increments relative to $(\mathscr{H}_t)_{t\geq 0}$. We closely follow Step 7 on pp. 123f of Höpfner et al. (1990):

Let $0 \leq u_1 < \cdots < u_l = u$ and $r > 0$; let also

$$\boldsymbol{V} := \Big((K,\boldsymbol{W})_{u_1},\ldots,(K,\boldsymbol{W})_{u_l}\Big), \quad V' := K_{u+r} - K_u, \quad \text{and} \quad V'' := \boldsymbol{W}_{u+r} - \boldsymbol{W}_u.$$

In analogy to Step 7 on pp. 123f of Höpfner et al. (1990), it is sufficient to show that

$$\begin{aligned} \bar{\mathbb{E}}[g(\boldsymbol{V})h'(V')] &= \bar{\mathbb{E}}[g(\boldsymbol{V})]\,\mathbb{E}_{\mathbb{Q}_1}[h'(K_r)], \\ \bar{\mathbb{E}}[g(\boldsymbol{V})h''(V'')] &= \bar{\mathbb{E}}[g(\boldsymbol{V})]\,\mathbb{E}_{\mathbb{Q}_2}[h''(\boldsymbol{W}_r)] \end{aligned} \tag{3.4.22}$$

holds for all continuous $g : (\mathbb{R} \times \mathbb{R}^{2I})^l \to [0,1]$, $h' : \mathbb{R} \to [0,1]$, and $h'' : \mathbb{R}^{2I} \to [0,1]$) with compact support.

We abbreviate $c_r' := \mathbb{E}_{\mathbb{Q}_1}[h'(K_r)]$ and $c_r'' := \mathbb{E}_{\mathbb{Q}_2}[h''(\boldsymbol{W}_r)]$. By eq. (3.4.21), there exists a sequence $(t_n)_{n\in\mathbb{N}}$ such that $\mathscr{L}((L^t, \boldsymbol{U}^{v_t} \mid \mathbb{P}^\pi) \to \bar{\mathbb{P}}$. For every $y \in E$, moreover, we have

$$\mathscr{L}(L^t \mid \mathbb{P}^y) \to \mathbb{Q}_1 \quad \text{and} \quad \mathscr{L}(\boldsymbol{U}^{v_t} \mid \mathbb{P}^y) \to \mathbb{Q}_2.$$

Since $\mathbb{Q}_1$ (resp., $\mathbb{Q}_2$) is the law of an a. s. continuous process, and K (resp., $\boldsymbol{W}$) has no fixed time of discontinuity, we deduce from Proposition 2.3.6 (ii) that

$$c_r'(y,t_n,\varepsilon) \to c_r' \quad \text{and} \quad c_r''(y,t_n,\varepsilon) \to c_r'' \quad \text{as } n \to \infty \text{ and } \varepsilon \downarrow 0,$$

where

$$c_r'(y,t_n,\varepsilon) := \mathbb{E}^y[h'(K^{t_n}_{r+\varepsilon})] \quad \text{and} \quad c_r''(y,t_n,\varepsilon) := \mathbb{E}^y[h''(\boldsymbol{U}^{v_{t_n}}_{r+\varepsilon})]$$

with $K^{t_n}_u := \inf\{s : L^{t_n}_s > u\}$. Therefore, there exists a set $A \in \mathscr{E}$, $0 < \mu(A) < \infty$ such that, up to a subsequence,

$$|c_r'(y,t_n,\varepsilon) - c_r'| \leq n^{-1} \quad \text{and} \quad |c_r''(y,t_n,\varepsilon) - c_r''| \leq n^{-1} \quad \forall y \in A, |\varepsilon| \leq 1/n. \tag{3.4.23}$$

Let $H_t := \mu(A)^{-1}\int_0^t \mathbb{1}_A(X_r)\mathrm{d}r$, $H^{t_n}_s = v_{t_n}^{-1} H_{st_n}$, $\kappa(t_n) = \inf\{s : H^{t_n}_s > u\}$, and $\hat{L}^{t_n}_s := L^{t_n}_s + s v_{t_n}^{-1}$. We define

$$L'^{t_n}_s = H^{t_n}_{s\wedge\kappa(t_n)} + \hat{L}^{t_n}_s - \hat{L}^{t_n}_{s\wedge\kappa(t_n)} \quad \text{and} \quad L''^{t_n}_s := L^{t_n}_{s\wedge\kappa(t_n)} + \hat{L}^{t_n}_s - \hat{L}^{t_n}_{s\wedge\kappa(t_n)}.$$

Moreover, we set $\hat{K}(t_n, r) = \inf\{s : \hat{L}_s^{t_n} > r\}$,

$$K'(t_n, r) = \inf\{s : L_s'^{t_n} > r\}, \quad \text{and} \quad \boldsymbol{U}_r'^{t_n} = \boldsymbol{U}^{v_{t_n}} \left(L''^{t_n}_{K'(t_n,r)} \right),$$

where we note that $\kappa(t_n) = K'(t_n, u)$ as $\hat{L}$ is strictly increasing. We set also

$$\boldsymbol{V}^{t_n} = \Big((K'(t_n, u_1), \boldsymbol{U}'^{t_n}_{u_1})), \ldots, (K'(t_n, u_l), \boldsymbol{U}'^{t_n}_{u_l})) \Big),$$
$$V'^{t_n} := K'(t_n, u + r) - K'(t_n, u) \quad \text{and} \quad V''^{t_n} := \boldsymbol{U}'^{t_n}_{u+r} - \boldsymbol{U}'^{t_n}_{u}.$$

Due to the choice of $(t_n)_{n\in\mathbb{N}}$, we then have under $\mathbb{P}^\pi$ for the pre-limiting processes and under $\bar{\mathbb{P}}$ for the limit:

$$(L^{t_n}, H^{t_n}) \to (L, L) \quad \text{by eq. (2.5.7);} \tag{3.4.24}$$
$$(L^{t_n}, L'^{t_n}, L''^{t_n}) \to (L, L, L) \quad \text{by eq. (3.4.24) and P. 2.3.6 (iv);} \tag{3.4.25}$$
$$(L^{t_n}, L'^{t_n}, L''^{t_n}, \boldsymbol{U}^{v_{t_n}}) \to (L, L, L, \boldsymbol{W}) \quad \text{by eq. (3.4.25) and P. 2.3.6 (iv);} \tag{3.4.26}$$
$$(L'^{t_n}, L''^{t_n}, \boldsymbol{U}^{v_{t_n}}_{L''^{t_n}}) \to (L, L, \boldsymbol{W}_L), \quad \text{by eq. (3.4.26) and P. 2.3.6 (iii);} \tag{3.4.27}$$
$$(\boldsymbol{V}^{t_n}, V'^{t_n}, V''^{t_n}, L'^{t_n}_{K'(t_n,u)}) \to (\boldsymbol{V}, V', V'', u) \quad \text{by eq. (3.4.27) and P. 2.3.6 (i).} \tag{3.4.28}$$

By eq. (3.4.28), therefore,

$$\mathbb{E}^\pi g(\boldsymbol{V}^{t_n}) \xrightarrow[n\to\infty]{} \bar{\mathbb{E}} g(\boldsymbol{V})$$
$$\mathbb{E}^\pi g(\boldsymbol{V}^{t_n}) h'(V'^{t_n}) \xrightarrow[n\to\infty]{} \bar{\mathbb{E}} g(\boldsymbol{V}) h'(V'),$$
$$\mathbb{E}^\pi g(\boldsymbol{V}^{t_n}) h''(V''^{t_n}) \xrightarrow[n\to\infty]{} \bar{\mathbb{E}} g(\boldsymbol{V}) h''(V'').$$

In addition, we note

$$V'^{t_n}(\omega) = \hat{K}\left(t_n, u + r - L'^{t_n}_{\kappa(t_n)}(\omega), \theta_{t_n \kappa(t_n)}(\omega) \right), \tag{3.4.29}$$

and

$$\begin{aligned} V''^{t_n}(\omega) &= \boldsymbol{U}^{v_{t_n}} \left(\hat{L}^{t_n}_{\hat{K}(t_n, u+r-L'^{t_n}_{\kappa(t_n)}(\omega))}, \theta_{t_n \kappa(t_n)}(\omega) \right) \\ &= \boldsymbol{U}^{v_{t_n}} \left(u + r - L'^{t_n}_{\kappa(t_n)}(\omega)) + O(1/v_{t_n}), \theta_{t_n \kappa(t_n)}(\omega) \right). \end{aligned} \tag{3.4.30}$$

In combination with the Markov property of X, therefore,

$$\mathbb{E}^\pi g(\boldsymbol{V}^{t_n})h'(V'^{t_n}) = \mathbb{E}^\pi g(\boldsymbol{V}^{t_n})c'_r(X_{t_n\kappa(t_n)}, t_n, u - L'^{t_n}_{\kappa(t_n)}),$$
$$\mathbb{E}^\pi g(\boldsymbol{V}^{t_n})h''(V''^{t_n}) = \mathbb{E}^\pi g(\boldsymbol{V}^{t_n})c''_r(X_{t_n\kappa(t_n)}, t_n, u - L'^{t_n}_{\kappa(t_n)} + O(1/v_{t_n})).$$

By the definition of $\kappa(t_n)$, we observe that $X_{t_n\kappa(t_n)} \in A$. By eq. (3.4.28), we observe $L'^{t_n}_{\kappa(t_n)} \to u$ in law. Up to a further subsequence, thus, we can suppose that $\mathbb{P}^\pi(|L'^{t_n}_{\kappa(t_n)} - u| \geq n^{-1}) \leq n^{-1}$ for all $n \in \mathbb{N}$. Recalling eq. (3.4.23), since g, h' and h'' are bounded by one, we conclude that

$$\Big| \mathbb{E}^\pi[g(\boldsymbol{V}^{t_n})h'(V'^{t_n})] - c'_r\,\mathbb{E}^\pi[g(\boldsymbol{V}^{t_n}))] \Big| \leq 2n^{-1},$$
$$\Big| \mathbb{E}^\pi[g(\boldsymbol{V}^{t_n})h''(V''^{t_n})] - c''_r\,\mathbb{E}^\pi[g(\boldsymbol{V}^{t_n}))] \Big| \leq 2n^{-1}.$$

Since $\mathbb{E}^\pi[g(\boldsymbol{V}^{t_n}))] \to \bar{\mathbb{E}}[g(\boldsymbol{V})]$, consequently, eq. (3.4.22) holds. That is, K and $\boldsymbol{W}$ have independent increments relative to $(\mathscr{H}_u)_{u\geq 0}$. □

Next, we demonstrate that the convergence in Lemma 3.4.10 holds stably in law.

3.4.11 Lemma. *Grant Assumptions 3.2.2, 3.2.4, 3.3.1 and 3.3.2. Let η_t be given as in Lemma 3.4.10. Then, we have the following stable convergence in law in $\mathcal{D}(\mathbb{R}^{1+I})$:*

$$\left(L^t, \left(U^{v_t,\eta_t}_{L^t}(x_i, y_i)\right)_{i\in I}\right) \xrightarrow[t\to\infty]{\mathscr{L}-\mathrm{st}} \left(L, \left(\mu'(x_i)\sigma(x_i,y_i)W^i_L\right)_{i\in I}\right),$$

where $\sigma(x,y)^2$ is given by eq. (3.3.4).

Proof. Let h be a bounded, Lipschitz continuous function on $\mathcal{D}(\mathbb{R}^{1+I})$ and Y be a bounded $\mathscr{F}$-measurable random variable. With $\sigma(x,y)^2$ given by eq. (3.3.4), we abbreviate

$$\boldsymbol{U}^{v_t} := \left(U^{v_t,\eta_t}(x_i, y_i)\right)_{i\in I} \quad \text{and} \quad \boldsymbol{W} := \left(\mu'(x_i)\sigma(x_i,y_i)W^i\right)_{i\in I}.$$

We have to demonstrate

$$\mathbb{E}^\pi\left[h(L^t, \boldsymbol{U}^{v_t}_{L^t})Y\right] \xrightarrow[t\to\infty]{} \tilde{\mathbb{E}}\left[h\left(L, \boldsymbol{W}_L\right)\right]\mathbb{E}^\pi Y. \tag{3.4.31}$$

At first, we suppose that Y is $\mathscr{F}_u$-measurable for some $u \geq 0$. Let a^t be given by $a^t_s = (s - ut^{-1})^+$. Then a^t converges to $a_s = s$ as $t \to \infty$. By Lemma 3.4.10, since a^t is non-random, $\mathscr{L}(a^t, L^t, \boldsymbol{U}^{v_t} \mid \mathbb{P}^\pi) \to \mathscr{L}(a, L, \boldsymbol{W} \mid \tilde{\mathbb{P}})$ weakly as $t \to \infty$. The paths of the limit process are a. s. continuous. By Proposition 2.3.6 (iii), therefore,

$$\mathscr{L}\Big(a^t, L^t_{a^t}, \boldsymbol{U}^{v_t} \circ L^t_{a^t} \mid \mathbb{P}^\pi\Big) \xrightarrow[t\to\infty]{\mathrm{w}} \mathscr{L}\Big(a, L, \boldsymbol{W}_L \mid \tilde{\mathbb{P}}\Big).$$

Since $\mathbb{E}^\pi[h(L^t_{a^t} \circ \theta_u, (\boldsymbol{U}^{v_t} \circ L^t_{a^t}) \circ \theta_u)Y] = \mathbb{E}^\pi[\mathbb{E}^{X_u}[h(L^t_{a^t}, \boldsymbol{U}^{v_t} \circ L^t_{a^t})]Y]$ by the Markov property, and since $\mathbb{E}^\pi[\tilde{\mathbb{E}}[h(L, \boldsymbol{W}_L)]Y] = \tilde{\mathbb{E}}[h(L, \boldsymbol{W}_L)]\,\mathbb{E}^\pi Y$, consequently,

$$\mathbb{E}^\pi\Big[h\Big(L^t_{a^t} \circ \theta_u, (\boldsymbol{U}^{v_t} \circ L^t_{a^t}) \circ \theta_u\Big)Y\Big] \xrightarrow[t\to\infty]{} \tilde{\mathbb{E}}\Big[h(L, \boldsymbol{W}_L)\Big]\,\mathbb{E}^\pi Y.$$

For every $r > 0$, we note

$$\sup_{s\leq r}\Big|L^t_s - L^t_{a^t_s} \circ \theta_u\Big| = \sup_{s\leq r}\Big|v_t^{-1} J_{st\wedge u}\Big| \leq v_t^{-1} J_u \xrightarrow[t\to\infty]{\text{a.s.}} 0,$$

and

$$\sup_{s\leq r}\Big\|(\boldsymbol{U}^{v_t} \circ L^t_{a^t_s}) \circ \theta_u - \boldsymbol{U}^{v_t} \circ L^t_s\Big\|_\infty \leq \frac{\|g_1\|_\infty(\|g_2\|_\infty J_u + \eta^d_{2,t}\|Fg_2^{\eta,y}\|_\infty u)}{\sqrt{v_t \eta^d_{1,t}\eta^d_{2,t}}} \xrightarrow[t\to\infty]{\text{a.s.}} 0.$$

Since h is Lipschitz, therefore,

$$\Big|h(L^t, \boldsymbol{U}^{v_t} \circ L^t) - h(L^t_{a^t} \circ \theta_u, (\boldsymbol{U}^{v_t} \circ L^t_{a^t}) \circ \theta_u)\Big| \xrightarrow[t\to\infty]{\text{a.s.}} 0.$$

Since h and Y are bounded, we deduce from Lebesgue's dominated convergence theorem that eq. (3.4.31) holds for all bounded $\mathscr{F}_u$-measurable random variables Y.

Next, for arbitrary bounded $\mathscr{F}$-measurable Y, we have $\mathbb{E}^\pi[Y|\mathscr{F}_u] \to Y$ in $\mathcal{L}^1$ as $u \to \infty$. Consequently, again by Lebesgue's dominated convergence theorem,

$$\lim_{u\to\infty}\sup_{t>0}\Big|\mathbb{E}^\pi\Big[h(L^t, \boldsymbol{U}^{v_t} \circ L^t, \bar{\boldsymbol{U}}^{v_t} \circ L^t)(\mathbb{E}^\pi[Y|\mathscr{F}_u] - Y)\Big]\Big| = 0.$$

Thus, eq. (3.4.31) holds in general. □

By Lemma 3.4.9 and Proposition 2.3.6 (iv), we obtain the following corollary.

3.4.12 Corollary. *Grant Assumptions 3.2.2, 3.2.4, 3.3.1 and 3.3.2. Let η_t be given as in Lemmata 3.4.10 and 3.4.11. Then we have the following stable convergence in law in $\mathcal{D}(\mathbb{R}^{2I})$:*

$$\Big(S^{v_t,\eta_t}_{L^t}(x_i), U^{v_t,\eta_t}_{L^t}(x_i,y_i)\Big)_{i\in I} \overset{\mathscr{L}-\mathrm{st}}{\underset{t\to\infty}{\Longrightarrow}} \Big(\mu'(x_i)L, \mu'(x_i)\sigma(x_i,y_i)W^i_L\Big)_{i\in I},$$

where $\sigma(x,y)^2$ is given by eq. (3.3.4). □

Proof (of Theorem 3.3.7). For every $t \geq 0$ and each x and y, we have

$$\sqrt{v_t\eta^d_{1,t}\eta^d_{2,t}}\,\Big(\hat{f}^{\eta_t}_t(x,y) - \bar{f}^{\eta_t}(x,y)\Big) = \frac{U^{v_t,\eta_t}_{J_t/v_t}(x,y) - \bar{f}^{\eta_t}(x,y)\sqrt{\eta^d_{1,t}\eta^d_{2,t}/v_t}\int_{T_{J_t}}^t g^{\eta_t,x}_1(X_s)\mathrm{d}s}{S^{v_t,\eta_t}_{J_t/v_t}(x) + v_t^{-1}\int_{T_{J_t}}^t g^{\eta_t,x}_1(X_s)\mathrm{d}s},$$

where $\bar{f}^\eta(x,y) := \mu(g^{\eta,x}_1 F g^{\eta,y}_2)/\mu(g^{\eta,x}_1)$. Let $h_n : \mathbf{D} \to \mathbb{R}$ be as in the proof of Theorem 3.3.6. We recall $\psi(h^2_n) \leq \zeta\eta^{-d}_{1,n}$ for some $\zeta < \infty$. We also note $v_t\eta^{-d}_{2,t} \to \infty$. In analogy to eq. (3.4.17), thus,

$$\sqrt{\frac{\eta^d_{1,t}\eta^d_{2,t}}{v_t}}\int_{T_{J_t}}^t g^{\eta_t,x}_1(X_s)\mathrm{d}s \leq \sqrt{\frac{\eta^d_{1,t}\eta^d_{2,t}}{v_t}} h_{v_t}(Z_{J_t+1}) \xrightarrow[t\to\infty]{\mathbb{P}^\pi} 0.$$

Since L and W are independent, $V(x_i,y_i) := L_1^{-1/2}W^i_{L_1}$ defines an I-dimensional standard Gaussian random vector such that L, V and $\mathscr{F}$ are independent. By the continuous mapping theorem and Corollary 3.4.12, consequently,

$$\bigg(\sqrt{v_t\eta^d_{1,t}\eta^d_{2,t}}\Big(\hat{f}^{\eta_t}_t(x_i,y_i) - \bar{f}^{\eta_t}(x_i,y_i)\Big)\bigg)_{i\in I} \xrightarrow[t\to\infty]{\mathscr{L}-\mathrm{st}} \Big(\sigma(x_i,y_i)V(x_i,y_i)L_1^{-1/2}\Big)_{i\in I},$$

where $\sigma(x,y)^2$ is given by eq. (3.3.4). ⋄

In addition, grant Assumption 3.3.3 and let $\eta_t = (\eta_{1,t},\eta_{2,t})$ be such that eq. (3.3.3) holds as well. We abbreviate $\bar{\gamma}^\eta(x,y) = \bar{f}^\eta(x,y) - f(x,y)$ and note

$$\mu(g^{\eta,x}_1)\bar{\gamma}^\eta(x,y) = \iint \mu'(x+\eta_1 z)\Big(f(x+\eta_1 z, y+\eta_2 w) - f(x,y)\Big)g_1(z)g_2(w)\mathrm{d}w\mathrm{d}z.$$

We apply Taylor's theorem to μ' and f: In x, we expand up to the order $\lceil \alpha_1 \rceil - 1$ and, in y, we expand up to the order $\lceil \alpha_2 \rceil - 1$. We recall from eq. (3.2.4) that g_1 and g_2 are, at least, of order α_1 and α_2, respectively. By a classical approximation argument, therefore, there exists a constant $\zeta < \infty$ such that $|\mu(g_1^{\eta,x})\bar{\gamma}^{\eta_t}(x,y)| \leq \zeta(\eta_{1,t}^{\alpha_1} + \eta_{2,t}^{\alpha_2})$. If $\zeta_1 = \zeta_2 = 0$ in eq. (3.3.3), then it is immediate that $(v_t \eta_{1,t}^d \eta_{2,t}^d)^{1/2} \bar{\gamma}^{\eta_t}(x,y) \to 0$. If $\alpha_1, \alpha_2 \in \mathbb{N}^*$, more explicitly,

$$\mu(g_1^{\eta,x})\bar{\gamma}^{\eta}(x,y) = \eta_{1,t}^{\alpha_1} \sum_{\substack{|m_1+m_2|=\alpha_1 \\ |m_2|\neq 0}} \frac{\kappa_{m_1+m_2}(g_1)}{m_1! m_2!} \frac{\partial^{m_1}}{\partial x^{m_1}} \mu'(x) \frac{\partial^{m_2}}{\partial x^{m_2}} f(x,y) + \eta_{2,t}^{\alpha_2} \sum_{|m|=\alpha_2} \frac{\kappa_m(g_2)}{m!} \mu'(x) \frac{\partial^m}{\partial y^m} f(x,y) + o(\eta_{1,t}^{\alpha_1} + \eta_{2,t}^{\alpha_2}).$$

Since $\mu(g_1^{\eta,x}) \to \mu'(x)$, we have $(v_t \eta_{1,t}^d \eta_{2,t}^d)^{1/2} \bar{\gamma}^{\eta_t}(x,y) \to \gamma(x,y)$ given by eq. (3.3.5). □

Proof (of Corollary 3.3.8). In analogy to the proof of Theorem 3.3.7, by Corollary 3.4.12 it remains to show that $(v_t \eta_{1,t}^d \eta_{2,t}^d)^{1/2} \hat{\gamma}_t^{\eta_t}(x,y)$ is a consistent estimator for $\gamma(x,y)$.

We recall that in classical (conditional) density estimation, the (partial) derivatives of a consistent density estimator – provided they exist – are consistent for the (partial) derivatives of the estimated density. In analogy to Lemma 3.4.7, we observe that this is also true in our context. In particular,

$$\frac{\partial^{m_1+m_2}}{\partial x^{m_1} \partial y^{m_2}} \hat{f}_t^{\eta_t}(x,y) \xrightarrow[t\to\infty]{\mathbb{P}^\pi} \frac{\partial^{m_1+m_2}}{\partial x^{m_1} \partial y^{m_2}} f(x,y) \quad \text{and} \quad \frac{\int_0^t \frac{\partial^m}{\partial x^m} g_1^{\eta_t,x}(X_s) \mathrm{d}s}{\int_0^t g_1^{\eta_t,x}(X_s) \mathrm{d}s} \xrightarrow[t\to\infty]{\mathbb{P}^\pi} \frac{\frac{\partial^m}{\partial x^m} \mu'(x)}{\mu'(x)}.$$

If either $\alpha_1, \alpha_2 \in \mathbb{N}^*$ or $\zeta_1 = \zeta_2 = 0$ in eq. (3.3.3), consequently,

$$(v_t \eta_{1,t}^d \eta_{2,t}^d)^{1/2} \hat{\gamma}_t^{\eta_t}(x,y) \xrightarrow[t\to\infty]{\mathbb{P}^\pi} \gamma(x,y).$$

□

3.5 Proofs for results of Section 3.2

Throughout this section, $\zeta < \infty$ denotes some generic constant which may depend on the variables specified at the beginning of each proof. It may change from line to line.

This section is organised as follows: Firstly, in Section 3.5.1 we study the influence of discretisation on our estimator. We prove results for the small-time asymptotic of Itō semi-martingales and for the sojourn time discretisation error. Secondly, in Section 3.5.2 we prove an auxiliary, non-standard martingale limit theorem. Thirdly, in Section 3.5.3 we prove the consistency of our estimator (Theorem 3.2.9) utilising our results from Sections 3.4.3 and 3.5.1. Lastly, in Section 3.5.4 we apply Theorem 3.5.5 from Section 3.5.2 to our case and conclude with the final steps in the proof of the central limit theorem (Theorem 3.2.10 and Corollary 3.2.11) utilising our results from Sections 3.4.4 and 3.5.1.

3.5.1 Small-time asymptotic and sojourn time discretisation error

In this subsection, we study the influence of discretisation.

We compare our estimators in Definitions 3.2.7 and 3.3.4: In the numerator of the former, the jumps ΔX_t and the pre-jump left-limits X_{t-} are replaced by the increments $\Delta_k^n X$ and the pre-increment values $X_{(k-1)\Delta}$, respectively. Our Itō semi-martingale meets the following small-time asymptotic:

3.5.1 Proposition. *Let A be a compact subset of $E \times E^*$, $\eta_0 < \min\{\|y\| : (x,y) \in A\}$, and let g be a twice continuously differentiable kernel with compact support. Grant Assumptions 3.2.1 and 3.2.3. Then, for every $m \in \mathbb{N}^*$, there exists $\zeta < \infty$ such that*

$$\left| \frac{1}{\Delta} \mathbb{E}^x \left[g^{\eta,y}(\Delta_1^n X) \right] - \int F(x, \mathrm{d}w) g^{\eta,y}(w) \right| \\ \leq \zeta \left[\Delta^{(\alpha \wedge 1)/2} + \frac{\Delta}{\eta^{2 \vee (\beta+d)}} \left(1 + \sum_{k=1}^{m} \frac{\Delta^k}{\eta^{2k}} \right) + \frac{\Delta^m}{\eta^{2(m+1)+d}} \right] \tag{3.5.1}$$

holds for every $(x,y) \in A$, $\eta < \eta_0$ and $\Delta \leq 1$, where $g^{\eta,y}(w) = \eta^{-d} g((w-y)/\eta)$.

Remark. For presentational purposes, we have left a small gap in the finite activity case. For instance, if f is locally bounded on $E \times E$, then we can improve the bound in eq. (3.5.1) replacing $\eta^{2 \vee (\beta+d)}$ by η^2 independently of the dimension d.

In the former estimator's denominator, the sojourn time $\int_0^t g_1^{\eta,x}(X_s) \mathrm{d}s$ is replaced by its Riemann sum approximation $\Delta \sum_{k=1}^{n} g_1^{\eta,x}(X_{(k-1)\Delta})$.

3.5.2 Proposition. *Let $x \in E$, $v : \mathbb{R}_+ \to \mathbb{R}_+$ be a non-decreasing function, $\xi_n > 0$, $\eta_n \to 0$, and $(h_n)_{n\in\mathbb{N}^*}$ be a uniformly bounded family of twice continuously differentiable functions supported on $B_{\eta_n}(x)$ such that $(\eta_n^{|m|}\partial^m h_n)_{n\in\mathbb{N}^*}$ is uniformly bounded for every multi-index m with $|m| \in \{1,2\}$. As $n\Delta \to \infty$ and $\Delta \to 0$, we suppose $v(n\Delta)\eta_n^d \to \infty$ and $\xi_n \Delta \eta_n^{-2-d[(1-2/(\beta+d))\vee 0]} \to 0$.*

(i) Grant Assumptions 3.2.1 to 3.2.3. If $n\Delta^2\xi_n \to 0$ and $v(s) = \bar{v}(st)$ for some deterministic equivalent $\bar{v}$ of X and some $t > 0$, then, under any law $\mathbb{P}^\pi$, we have the following convergence in probability:

$$\sup_{s\le t} \frac{\xi_n}{v(n\Delta)\eta_n^d}\left|\Delta\sum_{k=1}^{\lfloor sn\rfloor} h_n(X_{(k-1)\Delta}) - \int_0^{\lfloor sn\rfloor\Delta} h_n(X_r)\mathrm{d}r\right| \xrightarrow[n\to\infty]{\mathbb{P}^\pi} 0. \tag{3.5.2}$$

(ii) Grant Assumptions 3.2.1 to 3.2.4. If $(n\Delta)^{1-\delta}\Delta\xi_n \to 0$ and v is the regularly varying function from eq. (2.5.6), then, under any law $\mathbb{P}^\pi$, eq. (3.5.2) holds for all $t > 0$.

Before we turn to the proofs of Propositions 3.5.1 and 3.5.2, we present two auxiliary upper bounds for the small-time asymptotic of Itō semi-martingales. Below, we heavily utilise the results and notation presented in Section 2.2.

We recall that our underlying process X is an Itō semi-martingale with absolutely continuous characteristics $(B, C, \mathfrak{n})$ satisfying eq. (3.2.2). By Grigelionis decomposition theorem 2.2.12, we can assume w. l. o. g. that there exists a d-dimensional Wiener process W, defined on $(\Omega, \mathscr{F}, (\mathscr{F}_t)_{t\ge 0}, (\mathbb{P}^x)_{x\in E})$, and an $E\otimes E$-valued function σ with $c = \sigma\sigma^\intercal$ such that

$$X_t = X_0 + \int_0^t b(X_s)\mathrm{d}t + \int_0^t \sigma(X_s)\mathrm{d}W_s + (w\mathbb{1}_{\|w\|\le 1}) \star (\mathfrak{m}-\mathfrak{n})_t + (w\mathbb{1}_{\|w\|>1}) \star \mathfrak{m}_t.$$

For $\xi > 0$, we denote by $T^\xi := \inf\{t > 0 : \|\Delta X_t\| > \xi\}$ the first time of a jump greater than ξ . Also, we introduce the following decomposition of our semi-martingale X:

$$X_t = X_0 + X_t^\xi + X_t'^\xi, \quad \text{where } X_t'^\xi := (w\mathbb{1}_{\|w\|>\xi}) \star \mathfrak{m}_t = \sum_{s\le t} \Delta X_s \mathbb{1}_{\|\Delta X_s\|>\xi}.$$

We note that X^ξ and X'^ξ are again Itō semi-martingales; we denote their characteristics by $(B^\xi, C, \mathfrak{n}^\xi)$ and $(B'^\xi, 0, \mathfrak{n}'^\xi)$, respectively. Furthermore, we decompose X^ξ into drift B^ξ, continuous martingale part M^c, and purely discontinuous martingale part M^ξ. These are given by

$$B_t^\xi = \int_0^t b^\xi(X_s)\mathrm{d}s, \quad M_t^c = \int_0^t \sigma(X_s)\mathrm{d}W_s \quad \text{and} \quad M_t^\xi = (w1_{\|w\|\leq\xi}) \star (\mathfrak{m}-\mathfrak{n})_t,$$

where $b^\xi(x) = b(x) - \int_{\xi<\|w\|\leq 1} F(x,\mathrm{d}w)w$ in the case $\xi < 1$, and $b^\xi(x) = b(x) + \int_{1<\|w\|\leq\xi} F(x,\mathrm{d}w)w$ in the case $\xi \geq 1$. Under Assumption 3.2.1, we derive the following two lemmata.

3.5.3 Lemma. *Let $\xi_0 > 0$ and $p \geq 2$. Grant Assumption 3.2.1. Then, there exists a constant $\zeta < \infty$ such that, for every $0 < \xi \leq \xi_0$, $x \in E$, and $t \leq 1$, we have*

$$\mathbb{E}^x \sup_{s\leq t} \|X^\xi_{s\wedge T^\xi}\|^p \leq \zeta(1+\|x\|^p)t.$$

Proof. In this proof, $\zeta < \infty$ may depend on ξ_0 and p but neither on t, x, ξ nor ζ'.

(i) Let $1 \leq \xi \leq \xi_0$. We emphasise that, in this case,

$$\|b^\xi(x)\| \leq \|b(x)\| + \xi_0^{d+1} F(x, \{1 < \|w\| \leq \xi_0\}). \tag{3.5.3}$$

By eq. (3.2.2), we have $\mathfrak{n}^\xi(\mathrm{d}t, A) = \mathrm{d}tF^\xi(X_t, A) := \mathrm{d}tF(X_t, A\cap B_\xi(0))$ for every Borel set A. By construction, $X_t'^\xi = 0$ on $\{t < T^\xi\}$. By Proposition 2.2.13, thus,

$$\begin{aligned}
&\mathbb{E}^x \sup_{s\leq t}\|X^\xi_{s\wedge T^\xi}\|^p \leq \\
&\quad \zeta\, \mathbb{E}^x \Bigg[t^{p-1}\int_0^t \|b^\xi(X_0 + X^\xi_{s\wedge T^\xi})\|^p \mathrm{d}s + t^{p/2-1}\int_0^t \|c(X_0 + X^\xi_{s\wedge T^\xi})\|^{p/2}\mathrm{d}s \\
&\qquad + \int_0^t \mathrm{d}s \int F^{\xi_0}(X_0 + X_s^\xi, \mathrm{d}w)\|w\|^p \\
&\qquad + t^{p/2-1}\int_0^t \mathrm{d}s \left(\int F^{\xi_0}(X_0 + X_s^\xi, \mathrm{d}w)\|w\|^2\right)^{p/2}\Bigg].
\end{aligned}$$

Under Assumption 3.2.1, for all $t \leq 1$, we observe

$$\mathbb{E}^x \sup_{s \leq t} \|X^{\xi}_{s\wedge T^{\xi}}\|^p \leq \zeta \int_0^t (1 + \mathbb{E}^x \|X_0 + X^{\xi}_{s\wedge T^{\xi}}\|^p) \mathrm{d}s.$$

For $\zeta' > 0$, let $S^{\zeta'} := \inf\{s > 0 : \|X^{\xi}_s\| > \zeta'\}$. Then

$$\mathbb{E}^x \sup_{s \leq t} \|X^{\xi}_{s\wedge T^{\xi}\wedge S^{\zeta'}}\|^p \leq \zeta \int_0^t (1 + \mathbb{E}^x \|X_0 + X^{\xi}_{s\wedge T^{\xi}\wedge S^{\zeta'}}\|^p) \mathrm{d}s,$$

where we note $\sup_{s\leq t} \|X^{\xi}_{s\wedge T^{\xi}\wedge S^{\zeta'}}\| \leq \zeta' + \xi$. By the Grönwall–Bellmann inequality, thus,

$$\mathbb{E}^x \sup_{s \leq t} \|X^{\xi}_{s\wedge T^{\xi}\wedge S^{\zeta'}}\|^p \leq \zeta(1 + \|x\|^p) \left(t + \int_0^t \zeta \mathrm{e}^{\zeta(t-s)} \mathrm{d}s \right) = \zeta(1 + \|x\|^p)(\mathrm{e}^{\zeta t} - 1).$$

Since $S^{\zeta'} \wedge T^{\xi} \to T^{\xi}$ as $\zeta' \to \infty$, therefore, $\mathbb{E}^x \sup_{s\leq t} \|X^{\xi}_{s\wedge T^{\xi}}\|^p \leq \zeta(1 + \|x\|^p)t$. ◇

(ii) Let $0 < \xi < 1$. We note that $X^{\xi}_t \mathbb{1}_{t<T^{\xi}} = (X_t - X_0)\mathbb{1}_{t<T^{\xi}}$ holds, and that X^{ξ} is continuous at T^{ξ} outside the null set $\{\|\Delta X_{T^{\xi}}\| = \xi\}$. As $T^{\xi} \leq T^1$ for all ω, thus,

$$\sup_{s\leq t} \|X^{\xi}_{s\wedge T^{\xi}}\| = \sup_{s\leq t} \|(X_s - X_0)\mathbb{1}_{s<T^{\xi}}\| \leq \sup_{s\leq t} \|(X_s - X_0)\mathbb{1}_{s<T^1}\| = \sup_{s\leq t} \|X^1_{s\wedge T^1}\|$$

almost surely. By case $\xi \geq 1$, consequently, $\mathbb{E}^x \sup_{s\leq t} \|X^{\xi}_{s\wedge T^{\xi}}\|^p \leq \zeta(1 + \|x\|^p)t$. □

3.5.4 Lemma. *Let $y \neq 0$ and $\eta_0 < \|y\|$. Grant Assumption 3.2.1. Then, for every $m \in \mathbb{N}^*$, there exists a constant $\zeta < \infty$ – non-increasing in $\|y\|$ – such that, for every $x \in E$, $\eta < \eta_0$, and $t \leq 1$,*

$$\mathbb{P}^x(X_t \in B_\eta(X_0 + y)) \leq \zeta \left(1 + \|x\|^{2(m+1)} + \|y\|^{2(m+1)}\right) \left[t\eta^d \left(1 + \sum_{k=1}^m \frac{t^k}{\eta^{2\vee(\beta+d)+2(k-1)}} \right) + \frac{t^m}{\eta^{2m}} \right]. \tag{3.5.4}$$

Proof. Let $1 < \zeta' < (\|y\|/\eta_0)^{1/(m+1)}$, $\varepsilon := (\zeta'^{m+1}\eta_0 - \zeta'^{m}\eta_0)/6 > 0$ and $\xi < \varepsilon/2$. In addition, let g be a $\mathcal{C}^2$-kernel such that $\mathbb{1}_{B_1(0)} \le g \le \mathbb{1}_{B_{(\zeta'+1)/2}(0)}$. We set $g_\eta(z) = g((z-x-y)/\eta)$ and abbreviate $h(t,\eta) := \mathbb{P}^x(X_t \in B_\eta(x+y)) \le \mathbb{E}^x\, g_\eta(X_t)$. In this proof, $\zeta < \infty$ may depend on η_0, ζ', β and m, but neither on x, t nor η.

By Itō's formula eq. (2.2.8), we have $h(t,\eta) \le |H_t^\eta| + |H_t'^\eta| + |H_t''^\eta|$, where

$$
\begin{aligned}
H_t^\eta &:= \mathbb{E}^x \int_0^t b(X_s)^\intercal \nabla g_\eta(X_s) ds + \frac{1}{2}\,\mathbb{E}^x \int_0^t \operatorname{tr}\Big(c(X_s)\nabla^2 g_\eta(X_s)\Big)\mathrm{d}s,\\
H_t'^\eta &:= \mathbb{E}^x \int_0^t \mathrm{d}s \mathbb{1}_{B_{\zeta'\eta}(x+y)}(X_s)\\
&\qquad\qquad \cdot \int F(X_s,\mathrm{d}w)\Big\{g_\eta(X_s+w) - g_\eta(X_s) - w^\intercal \nabla g_\eta(X_s)\mathbb{1}_{\|w\|\le 1}\Big\},\\
H_t''^\eta &:= \mathbb{E}^x \int_0^t \mathrm{d}s \mathbb{1}_{B_{\zeta'\eta}(x+y)^c}(X_s) \int F(X_s,\mathrm{d}w) g_\eta(X_s+w).
\end{aligned}
$$

Under Assumption 3.2.1, $b(z)$ and $c(z)$ are bounded in norm by $\zeta(1+\|z\|^2)$. Moreover, the gradient and Hessian of g_η vanish outside $B_{(\zeta'+1)\eta/2}(x+y)$ and satisfy $\|\partial_i g_\eta\| \le \zeta\eta^{-1}$ and $\|\partial_{ij} g_\eta\| \le \zeta\eta^{-2}$. Hence,

$$
\big|H_t^\eta\big| \le \zeta(1+\|x\|^2+\|y\|^2)\eta^{-2}\,\mathbb{E}^x \int_0^t \mathbb{1}_{B_{(\zeta'+1)\eta/2}(x+y)}(X_s)\mathrm{d}s.
$$

For $z \in B_{\zeta'\eta}(x+y)$, furthermore,

$$
\begin{aligned}
\int F(z,\mathrm{d}w)\Big\{g_\eta(z+w) - g_\eta(z) - w^\intercal \nabla g_\eta(z)\mathbb{1}_{\|w\|\le 1}\Big\} &\le \\
&\frac{\zeta(1+\|z\|)}{\eta^2}\int \bar{F}(\mathrm{d}w)(1\wedge\|w\|^2).
\end{aligned}
$$

Therefore,

$$
\big|H_t^\eta\big| + \Big|H_t'^\eta\Big| \le \frac{\zeta(1+\|x\|^2+\|y\|^2)}{\eta^2}\int_0^t h(s,\zeta'\eta)\mathrm{d}s. \tag{3.5.5}
$$

Suppose that $|H_t''^{\eta}| \leq \zeta(1+\|x\|^3+\|y\|^3)(t\eta^d+t^2\eta^{-\beta})$ holds. Then,

$$h(t,\eta) \leq$$

$$\zeta(1+\|x\|^3+\|y\|^3)t\eta^d(1+t\eta^{-(\beta+d)}) + \frac{\zeta(1+\|x\|^2+\|y\|^2)}{\eta^2}\int_0^t h(s,\zeta'\eta)\mathrm{d}s.$$

By iteration, we obtain eq. (3.5.4) after m steps. $\diamond$

It remains to prove $|H_t''^{\eta}| \leq \zeta(1+\|x\|^3+\|y\|^3)(t\eta^d+t^2\eta^{-\beta})$. Under Assumption 3.2.1 (iii), on the one hand, we have

$$\begin{aligned}\int F(z,\mathrm{d}w)g_\eta(z+w) &\leq \zeta(1+\|z\|)\eta^d\int \bar{f}(y+x-z+\eta w)g(w)\mathrm{d}w\\ &\leq \begin{cases}\zeta(1+\|x\|)\eta^d, & \text{if } z\in B_{3\varepsilon}(x),\\ \zeta(1+\|x+y\|)\eta^d & \text{if } z\in B_{1+\zeta'\eta}(x+y)^{\mathrm{c}}.\end{cases}\end{aligned}$$

For $z\in B_{1+\zeta'\eta}(x+y)\setminus B_{\zeta'\eta}(x+y)$, on the other hand, we have

$$\int F(z,\mathrm{d}w)g_\eta(z+w) \leq \frac{\zeta(1+\|z\|)}{((\zeta'-1)\eta/2)^\beta}\int \mathrm{d}w g\left(\frac{w+z-x-y}{\eta}\right)\bar{f}(w)\|w\|^\beta.$$

Since $\eta^d \leq \eta^{-\beta}$ and $\int \bar{F}(\mathrm{d}w)(\|w\|^\beta\wedge 1)<\infty$ by assumption, thus,

$$\int F(z,\mathrm{d}w)g_\eta(z+w) \leq \begin{cases}\zeta(1+\|x+y\|)\eta^{-\beta}, & \text{if } z\in B_{\zeta'\eta}(x+y)^{\mathrm{c}},\\ \zeta(1+\|x\|)\eta^d, & \text{if } z\in B_{3\varepsilon}(x).\end{cases} \tag{3.5.6}$$

Let $S^{\varepsilon,\xi} := \inf\{t>0 : \|X_t^\xi\| > 3\varepsilon\}$, and $\Omega_t^{\varepsilon,\xi} := \{S^{\varepsilon,\xi}\leq T^\xi\wedge t\}$. We split the set $\Omega\times[0,t]$ into

$$\begin{aligned}A_1 &:= \Omega\times[\![0,t\wedge T^\xi\wedge S^{\varepsilon,\xi}[\![,\\ A_2 &:= (\Omega_t^{\varepsilon,\xi})^{\mathrm{c}}\times[\![T^\xi\wedge t,t]\!],\\ A_3 &:= \Omega_t^{\varepsilon,\xi}\times[\![S^{\eta,\xi},t]\!].\end{aligned}$$

Then we obtain the following:

Firstly: Since $\sup_{s\le t}\|X^{\xi}_{s\wedge T^{\xi}\wedge S^{\varepsilon,\xi}} - X_0\| \le 3\varepsilon$, by eq. (3.5.6), we obtain

$$\iint_{A_1} \mathrm{d}\mathbb{P}^x\,\mathrm{d}s\,\mathbb{1}_{B_{\zeta'\eta}(x+y)^c}(X_s)\int F(X_s,\mathrm{d}w)g_\eta(X_s+w) \le \zeta(1+\|x\|)t\eta^d.$$

Secondly: Under Assumption 3.2.1, we have

$$\mathbb{P}^x(T^{\xi} \le t\wedge S^{\varepsilon,\xi}) \le \mathbb{E}^x\int_0^t \mathrm{d}s\,\mathbb{1}_{B_{3\varepsilon}(x)}(X_s)F(X_s,\|w\|>\xi) \le \zeta(1+\|x\|)t.$$

By the Markov property and eq. (3.5.6), therefore,

$$\begin{aligned}
&\iint_{A_2} \mathrm{d}\mathbb{P}^x\,\mathrm{d}s\,\mathbb{1}_{B_{\zeta'\eta}(x+y)^c}(X_s)\int F(X_s,\mathrm{d}w)g_\eta(X_s+w)\\
&\quad\le \mathbb{E}^x\,\mathbb{1}_{\{T^{\xi}\le t\wedge S^{\varepsilon,\xi}\}}\,\mathbb{E}^{X_{T^{\xi}}}\int_0^t \mathrm{d}s\,\mathbb{1}_{B_{\zeta'\eta}(x+y)^c}(X_s)\int F(X_s,\mathrm{d}w)g_\eta(X_s+w)\\
&\quad\le \zeta(1+\|x+y\|)t\eta^{-\beta}\,\mathbb{P}^x(T^{\xi}\le t\wedge S^{\varepsilon,\xi})\\
&\quad\le \zeta(1+\|x\|^2+\|y\|^2)t^2\eta^{-\beta}.
\end{aligned} \tag{3.5.7}$$

Thirdly: By Lemma 3.5.3, we have $\mathbb{P}^x(\Omega_t^{\varepsilon,\xi}) \le \zeta(1+\|x\|^2)t$. By the Markov property and eq. (3.5.6), therefore,

$$\begin{aligned}
&\iint_{A_3} \mathrm{d}\mathbb{P}^x\,\mathrm{d}s\,\mathbb{1}_{B_{\zeta'\eta}(x+y)^c}(X_s)\int F(X_s,\mathrm{d}w)g_\eta(X_s+w)\\
&\qquad\le \zeta(1+\|x+y\|)t\eta^{-\beta}\,\mathbb{P}^x(\Omega_t^{\varepsilon,\xi})\\
&\qquad\le \zeta(1+\|x\|^3+\|y\|^3)t^2\eta^{-\beta}.
\end{aligned} \tag{3.5.8}$$

□

We turn to the proofs of Propositions 3.5.1 and 3.5.2.

Proof (of Proposition 3.5.1). Let $1 < \zeta' < (\min\{\|y\| : (x,y)\in A\}/\eta_0)^{1/(m+2)}$, and $\varepsilon,\xi > 0$ be given as in the proof of Lemma 3.5.4. In this proof, $\zeta < \infty$ may depend on η_0, ζ', β, m and the set A, but neither on x, y, Δ nor η.

Let $\eta \le \eta_0$, and $(x,y)\in A$. W.l.o.g., we assume that g is supported on $B_1(0)$. To avoid cumbersome notation, we abbreviate $h_\eta = g^{\eta,x+y}$. From eq. (3.2.2) and Itō's formula eq. (2.2.8), we obtain $\mathbb{E}^x h_\eta(X_\Delta) = H^\eta_\Delta + H'^\eta_\Delta + H''^\eta_\Delta$, where

$$\begin{aligned} H^\eta_\Delta &= \mathbb{E}^x \int_0^\Delta b(X_t)^\intercal \nabla h_\eta(X_t)\mathrm{d}t + \frac{1}{2}\,\mathbb{E}^x \int_0^\Delta \operatorname{tr}\Big(c(X_t)\nabla^2 h_\eta(X_t)\Big)\mathrm{d}t, \\ H'^\eta_\Delta &= \mathbb{E}^x \int_0^\Delta \mathrm{d}t \mathbb{1}_{B_{\zeta'\eta}(x+y)}(X_t) \\ &\qquad \cdot \int F(X_t,\mathrm{d}w)\Big\{h_\eta(X_t+w) - h_\eta(X_t) - w^\intercal \nabla h_\eta(X_t)\mathbb{1}_{\|w\|\le 1}\Big\}, \\ H''^\eta_\Delta &= \mathbb{E}^x \int_0^\Delta \mathrm{d}t \mathbb{1}_{B_{\zeta'\eta}(x+y)^c}(X_t) \int F(X_t,\mathrm{d}w) h_\eta(X_t+w). \end{aligned}$$

By eq. (3.5.5), we observe

$$\left|H^\eta_\Delta\right| + \left|H'^\eta_\Delta\right| \le \frac{\zeta}{\eta^{d+2}} \int_0^\Delta \mathbb{P}^x(X_t \in B_{\zeta'\eta}(x+y))\mathrm{d}t.$$

By the choice of ζ', Lemma 3.5.4 implies

$$\left|H^\eta_\Delta\right| + \left|H'^\eta_\Delta\right| \le \zeta \left[\frac{\Delta^2}{\eta^2}\left(1 + \sum_{k=1}^m \frac{\Delta^k}{\eta^{2\vee(\beta+d)+2(k-1)}}\right) + \frac{\Delta^{m+1}}{\eta^{2(m+1)+d}}\right]. \tag{3.5.9}$$

Suppose

$$\begin{aligned} \left| H''^\eta_\Delta - \int F(x,\mathrm{d}w)h_\eta(x+w)\int_0^\Delta \mathrm{d}t\, \mathbb{P}^x(X_t \notin B_{\zeta'\eta}(x+y))\right| & \\ \le \zeta(\Delta^{1+(\alpha\wedge 1)/2} + \Delta^2\eta^{-(\beta+d)}). & \end{aligned} \tag{3.5.10}$$

Combining eq. (3.5.9) and eq. (3.5.10), we obtain eq. (3.5.1). ⋄

It remains to prove eq. (3.5.10). By eq. (3.5.6), we observe

$$\int F(z,\mathrm{d}w)h_\eta(z+w) \le \begin{cases} \zeta\eta^{-(\beta+d)}, & \text{if } z \in B_{\zeta'\eta}(x+y)^c, \\ \zeta, & \text{if } z \in B_{3\varepsilon}(x). \end{cases} \tag{3.5.11}$$

Let the stopping time $S^{\varepsilon,\xi}$, and the event $\Omega_\Delta^{\varepsilon,\xi}$ be given as in the proof of Lemma 3.5.4. We split the set $\Omega \times [0,\Delta]$ into $A_1 := \Omega \times [\![0, \Delta \wedge T^\xi \wedge S^{\varepsilon,\xi}[\![$, $A_2 := (\Omega_\Delta^{\varepsilon,\xi})^c \times [\![T^\xi \wedge \Delta, \Delta]\!]$ and $A_3 := \Omega_\Delta^{\varepsilon,\xi} \times [\![S^{\eta,\xi}, \Delta]\!]$. For convenience, we also abbreviate

$$\tilde{f}^\eta_{x,y}(z,w) := f(z, y + x - z + \eta w) - f(x, y + \eta w).$$

Then we obtain, firstly: By the choice of ε, we have that the convex hull of the set

$$\{(z, y + (x - z) + \eta w) : (x,y) \in A, \|z - x\| \leq 3\varepsilon, \|w\| \leq 1\}$$

is a compact subset of $E \times E^*$. By Assumption 3.2.3 and for all $(z,w) \in B_{3\varepsilon}(x) \times B_1(0)$, we have $|\tilde{f}^\eta_{x,y}(z,w)| \leq \zeta \|z - x\|^{\alpha\wedge 1}$. By Lemma 3.5.3, therefore,

$$\iint\limits_{A_1} \mathrm{d}\mathbb{P}^x \,\mathrm{d}t \int \mathrm{d}w g(w) \tilde{f}^\eta_{x,y}(X_t, w) \leq \zeta \Delta\, \mathbb{E}^x \sup_{t\leq\Delta} \|X^\xi_{t\wedge T^\xi \wedge S^{\varepsilon,\xi}}\| \leq \zeta \Delta^{1+(\alpha\wedge 1)/2}.$$

Secondly and thirdly: We compare eqs. (3.5.6) and (3.5.11). In analogy to eqs. (3.5.7) and (3.5.8), respectively, by the Markov property and eq. (3.5.11), therefore,

$$\iint\limits_{A_i} \mathrm{d}\mathbb{P}^x \,\mathrm{d}t \mathbb{1}_{B_{\zeta'\eta}(x+y)^c}(X_t) \int \mathrm{d}w g(w) \tilde{f}^\eta_{x,y}(X_t, w) \leq \zeta \Delta^2 \eta^{-(\beta+d)},$$

for $i \in \{2,3\}$. In summary, we proved eq. (3.5.10). □

Proof (of Proposition 3.5.2). W. l. o. g., we assume $\eta < 1/4$. In this proof, $\zeta < \infty$ may neither depend on n, Δ nor η.

By Itō's formula eq. (2.2.8), we observe

$$\frac{\xi_n}{v_{n\Delta}\eta_n^d} \left| \int\limits_0^{\lfloor sn \rfloor \Delta} h_n(X_r)\mathrm{d}r - \Delta \sum_{k=1}^{\lfloor sn \rfloor} h_n(X_{(k-1)\Delta}) \right| \leq |H^n_s| + |H'^n_s| + |H''^n_s| + |M^n_s|,$$

where

$$H^n_s := \frac{\xi_n}{v_{n\Delta}\eta_n^d} \sum_{k=1}^{\lfloor sn \rfloor} \int\limits_{(k-1)\Delta}^{k\Delta} \mathrm{d}t \int\limits_{(k-1)\Delta}^{t} \left(b(X_r)^\intercal \nabla h_n(X_r) + \frac{1}{2} \operatorname{tr}\left(c(X_r) \nabla^2 h_n(X_r) \right) \right) \mathrm{d}r,$$

$$H_s'^n := \frac{\xi_n}{v_{n\Delta}\eta_n^d} \sum_{k=1}^{\lfloor sn \rfloor} \int_{(k-1)\Delta}^{k\Delta} \mathrm{d}t \int_{(k-1)\Delta}^{t} \mathrm{d}r \cdot \int_{\|w\|\le 1} F(X_r, \mathrm{d}w) \Big\{ h_n(X_r + w) - h_n(X_r) - w^\intercal \nabla h_n(X_r) \Big\},$$

$$H_s''^n := \frac{\xi_n}{v_{n\Delta}\eta_n^d} \sum_{k=1}^{\lfloor sn \rfloor} \int_{(k-1)\Delta}^{k\Delta} \mathrm{d}t \sum_{(k-1)\Delta < r \le t} \mathbb{1}_{\|\Delta X_r\|>1} \Big\{ h_n(X_{r-} + \Delta X_r) - h_n(X_{r-}) \Big\},$$

and

$$M_s^n := \frac{\xi_n}{v_{n\Delta}\eta_n^d} \sum_{k=1}^{\lfloor sn \rfloor} \int_{(k-1)\Delta}^{k\Delta} \mathrm{d}t \left(\int_{(k-1)\Delta}^{t} \nabla h_n(X_r)^\intercal \sigma(X_r) \mathrm{d}W_r + \int_{(k-1)\Delta}^{t} \int_{\|w\|\le 1} \Big\{ h_n(X_{r-} + w) - h_n(X_{r-}) \Big\} (\mathfrak{m} - \mathfrak{n})(\mathrm{d}r, \mathrm{d}w) \right).$$

It remains to show:

(i) Under Assumptions 3.2.1 to 3.2.3, if $v(s) = \bar{v}(st)$ for some deterministic equivalent $\bar{v}$ of X and some $t > 0$, and if $n\Delta^2\xi_n \to 0$, then H_s^n, $H_s'^n$, $H_s''^n$ and M_s^n converge to zero uniformly on $\{0 \le s \le t\}$ in probability.

(ii) Under Assumptions 3.2.1 to 3.2.4, if v is the regularly varying function from eq. (2.5.6), and if $(n\Delta)^{1-\delta}\Delta\xi_n \to 0$, then H_s^n, $H_s'^n$, $H_s''^n$ and M_s^n converge to zero uniformly for $\{0 \le s \le t\}$ in probability for all $t > 0$.

(a) Under Assumption 3.2.1, $b(z)$ and $c(z)$ are bounded in norm by $\zeta(1 + \|z\|^2)$. Moreover, the gradient and Hessian of h_n vanish outside $B_{\eta_n}(x)$ and satisfy $\|\partial_i h_n\| \le \zeta\eta_n^{-1}$ and $\|\partial_{ij} h_n\| \le \zeta\eta^{-2}$, by assumption. Thus,

$$\left| b(z)^\intercal \nabla h_n(z) + \frac{1}{2} \operatorname{tr}\left(c(z) \nabla^2 h_n(z) \right) \right| \le \zeta(1 + \|z\|)\eta^{-2} \mathbb{1}_{B_{\eta_n}(x)}(z).$$

By Fubini's theorem, therefore,

$$\sup_{r \le s} |H_r^n| \le \zeta(1 + \|x\|^2) \frac{\Delta\xi_n}{\eta^2} S_s'^{n,\Delta,\eta_n}, \quad \text{where } S_s'^{n,\Delta,\eta} = \frac{1}{v_{n\Delta}\eta^d} \int_0^{\lfloor sn \rfloor \Delta} \mathbb{1}_{B_\eta(x)}(X_r) \mathrm{d}r.$$

In case (i), we deduce from Lemma 3.4.7 that the family $\{\mathscr{L}(S_t^{\prime n,\Delta,\eta_n} \mid \mathbb{P}^x) : n \in \mathbb{N}^*\}$ is tight under Assumptions 3.2.2 and 3.2.3. As $\Delta\xi_n\eta_n^{-2} \to 0$, $\sup_{s\le t}|H_s^n| \to 0$ in probability. In case (ii), we obtain from Lemma 3.4.9 that $S^{\prime n,\Delta,\eta_n}$ converges stably in law to a non-trivial process. As $\Delta\xi_n\eta_n^{-2} \to 0$, $\sup_{s\le t}|H_s^n| \to 0$ in probability for all $t > 0$. ⋄

(b) Let $\zeta' > 1$ and $\kappa = 1 \wedge 2/(\beta+d)$. Under Assumption 3.2.1, we have

$$\left| \int_{\|w\|\le 1} F(z,\mathrm{d}w)\{h_n(z+w) - h_n(z) - w^\intercal \nabla h_n(z)\} \right| \\ \le \begin{cases} \zeta(1+\|z\|)\eta_n^{-2}\int_{\|w\|\le 1}\bar{F}(\mathrm{d}w)\|w\|^2, & \text{for } \|z-x\| \le \zeta'\eta_n^\kappa, \\ \zeta(1+\|z\|)\eta_n^{-\kappa\beta}\int_{\|w\|\le 1}\bar{F}(\mathrm{d}w)\|w\|^\beta, & \text{for } \zeta'\eta_n^\kappa < \|z-x\| \le 1+\eta_n, \\ 0, & \text{else.} \end{cases} \tag{3.5.12}$$

Again by Fubini's theorem, therefore,

$$\sup_{t\le s}|H_t^{\prime n}| \le \zeta(1+\|x\|)\left(\frac{\Delta\xi_n\eta_n^{\kappa d}}{\eta_n^{d+2}}S_s^{\prime n,\Delta,\zeta'\eta_n^\kappa} + \frac{\Delta\xi_n}{\eta^{d+\kappa\beta}}S_s^{\prime n,\Delta,1+\eta_n}\right).$$

In analogy to step (a), since $\Delta\xi_n\eta_n^{-2-d(1-\kappa)} \to 0$, $H_s^{\prime n} \to 0$ uniformly on $\{0 \le s \le t\}$ in probability in case (i); and for all $t > 0$ in case (ii). ⋄

(c) In analogy to steps (a) and (b), we note

$$\begin{aligned} |H_s^{\prime\prime n}| &\le \xi_n\Delta(v_{n\Delta}\eta_n^d)^{-1}(|h_n(X_- + w)| + |h_n(X_-)|)\mathbb{1}_{\|w\|>1} \star \mathfrak{m}_{\lfloor sn\rfloor\Delta} \\ &\le |K_s^n| + |N_{\lfloor sn\rfloor/n}^n| + |K_s^{\prime n}| + |N_{\lfloor sn\rfloor/n}^{\prime n}|, \end{aligned}$$

where

$$\begin{aligned} K_s^n &:= \xi_n\Delta(v_{n\Delta}\eta_n^d)^{-1}|h_n(X_- + w)|\mathbb{1}_{\|w\|>1} \star \mathfrak{n}_{\lfloor sn\rfloor\Delta}, \\ K_s^{\prime n} &:= \xi_n\Delta(v_{n\Delta}\eta_n^d)^{-1}|h_n(X_-)|\mathbb{1}_{\|w\|>1} \star \mathfrak{n}_{\lfloor sn\rfloor\Delta}, \\ N_s^n &:= \xi_n\Delta(v_{n\Delta}\eta_n^d)^{-1}|h_n(X_- + w)|\mathbb{1}_{\|w\|>1} \star (\mathfrak{m}-\mathfrak{n})_{sn\Delta}, \\ N_s^{\prime n} &:= \xi_n\Delta(v_{n\Delta}\eta_n^d)^{-1}|h_n(X_-)|\mathbb{1}_{\|w\|>1} \star (\mathfrak{m}-\mathfrak{n})_{sn\Delta}. \end{aligned}$$

Under Assumption 3.2.1, since $\int_{\|w\|>1}F(z,\mathrm{d}w)|h_n(z+w)| = 0$ for $z \in B_{1-2\eta_n}(x)$, we have

$$\int_{\|w\|>1} F(z, \mathrm{d}w)|h_n(z+w)| \leq \zeta(1 + \|x\|).$$

In both cases (i) and (ii), therefore,

$$\sup_{s \leq t} |K_s^n| \leq \zeta(1 + \|x\|) \frac{tn\Delta^2 \xi_n}{v_{n\Delta}} \xrightarrow[n\to\infty]{} 0,$$

for all $t > 0$. Furthermore, we observe that N^n is a martingale w. r. t. the filtration $(\mathscr{F}_{sn\Delta})_{s\geq 0}$. Its predictable quadratic variation satisfies

$$\langle N^n, N^n \rangle_s = \frac{\Delta^2 \xi_n^2}{v_{n\Delta}^2} |h_n(X_- + w)|^2 \mathbb{1}_{\|w\|>1} \star \mathfrak{n}_{sn\Delta} \leq \zeta(1 + \|x\|) \frac{sn\Delta^3 \xi_n^2}{v_{n\Delta}^2 \eta_n^d} \xrightarrow[n\to\infty]{} 0.$$

Since $\lfloor sn \rfloor / n \to s$, $N^n_{\lfloor sn \rfloor / n} \to 0$ uniformly on $\{0 \leq s \leq t\}$ in probability for all $t > 0$.

In addition, we recall that $F(z, \{\|w\| > 1\} \leq \zeta(1 + \|z\|)$ under Assumption 3.2.1. Thus,

$$\sup_{s \leq t} |K_s'^n| \leq \zeta(1 + \|x\|) \xi_n \Delta S_t'^{n,\Delta,\eta_n} \xrightarrow[n\to\infty]{\mathbb{P}^\pi} 0$$

in case (i); and for all $t > 0$ in case (ii). Again, we observe that N'^n is a martingale w. r. t. the filtration $(\mathscr{F}_{sn\Delta})_{s\geq 0}$. Its predictable quadratic variation satisfies

$$\langle N'^n, N'^n \rangle_s = \frac{\Delta^2 \xi_n^2}{v_{n\Delta}^2 \eta_n^{2d}} |h_n(X_-)|^2 \mathbb{1}_{\|w\|>1} \star \mathfrak{n}_{sn\Delta} \leq \frac{\zeta(1 + \|x\|) \Delta^2 \xi_n^2}{v_{n\Delta} \eta_n^d} S_s'^{n,\Delta,\eta_n} \xrightarrow[n\to\infty]{} 0.$$

Thus, $N'^n_{\lfloor sn \rfloor / n} \to 0$ uniformly on $\{0 \leq s \leq t\}$ in probability in case (i); and for all $t > 0$ in case (ii). ◇

(d) Let $(M_s'^n)_{s\geq 0}$ and $(M_s''^n)_{s\geq 0}$ denote the $\mathscr{F}_{sn\Delta}$-martingales given by

$$M_s'^n := \frac{\xi_n}{v_{n\Delta} \eta_n^d} \int_0^{sn\Delta} \varphi_\Delta(r) \nabla h_n(X_r)^\intercal \sigma(X_r) \mathrm{d}W_r,$$

$$M_s''^n := \frac{\xi_n}{v_{n\Delta} \eta_n^d} \varphi_\Delta(r)(h_n(X_- + w) - h_n(X_-)) \mathbb{1}_{\|w\|\leq 1} \star (\mathfrak{m} - \mathfrak{n})_{sn\Delta},$$

where $\varphi_\Delta(r) := \Delta - (r - \lfloor r/\Delta \rfloor \Delta)$. The predictable quadratic variation of M'^n satisfies

$$\begin{aligned}\langle M'^n, M'^n\rangle_s &= \frac{\xi_n^2}{v_{n\Delta}^2 \eta_n^{2d}} \int_0^{sn\Delta} \varphi_\Delta(r)^2 \nabla h_n(X_r)^\intercal c(X_r) \nabla h_n(X_r) \mathrm{d}t \\ &\le \frac{\zeta(1+\|x\|^2)\Delta^2 \xi_n^2}{v_{n\Delta}\eta_n^{d+2}} S_s^{\prime n,\Delta,\eta_n}.\end{aligned}$$

As $\Delta\xi_n\eta_n^{-2} \to 0$ and $v_{n\Delta}\eta_n^d \to \infty$, $M_s'^n \to 0$ uniformly on $\{0 \le s \le t\}$ in probability in case (i); and for all $t > 0$ in case (ii).

In addition, the predictable quadratic variation of M''^n satisfies

$$\begin{aligned}\langle M''^n, M''^n\rangle_s &= \frac{\xi_n^2}{v_{n\Delta}^2 \eta_n^{2d}} \varphi_\Delta(r)^2 (h_n(X_- + w) - h_n(X_-))^2 \mathbb{1}_{\|w\|\le 1} \star \mathfrak{n}_{sn\Delta} \\ &\le \frac{\Delta^2 \xi_n^2}{v_{n\Delta}^2 \eta_n^{2d}} \int_0^{sn\Delta} \mathrm{d}r \int_{\|w\|\le 1} F(X_r, \mathrm{d}w)(h_n(X_r + w) - h_n(X_r))^2.\end{aligned}$$

Let $\zeta' > 1$ and $\kappa = 1 \wedge 2/(\beta + d)$ be as in step (b). By eq. (3.5.12),

$$\begin{aligned}&\left| \int_{\|w\|\le 1} F(z, \mathrm{d}w)(h_n(z+w) - h_n(z))^2 \right| \\ &\le \begin{cases} \zeta(1+\|z\|)\eta_n^{-2} \int_{\|w\|\le 1} \bar{F}(\mathrm{d}w)\|w\|^2, & \text{for } \|z - x\| \le \zeta'\eta_n^\kappa, \\ \zeta(1+\|z\|)\eta_n^{-\kappa\beta} \int_{\|w\|\le 1} \bar{F}(\mathrm{d}w)\|w\|^\beta, & \text{for } \zeta'\eta_n^\kappa < \|z - x\| \le 1 + \eta_n, \\ 0, & \text{else.} \end{cases}\end{aligned}$$

Therefore,

$$\langle M''^n, M''^n\rangle_s \le \frac{\zeta(1+\|x\|)\Delta\xi_n}{v_{n\Delta}\eta_n^d} \left(\frac{\Delta\xi_n\eta_n^{\kappa d}}{\eta_n^{d+2}} S_s^{\prime n,\Delta,\zeta'\eta_n^\kappa} + \frac{\Delta\xi_n}{\eta_n^{d+\kappa\beta}} S_s^{\prime n,\Delta,1+\eta_n} \right).$$

Again since $\Delta\xi_n\eta_n^{-2-d(1-\kappa)} \to 0$, $M_s''^n \to 0$ uniformly on $\{0 \le s \le t\}$ in probability in case (i); and for all $t > 0$ in case (ii). □

3.5.2 Auxiliary martingale limit theorem

The theorem presented in this subsection serves as a preliminary result for the proof of our central limit theorem (Theorem 3.2.10 and Corollary 3.2.11). It is a non-standard limit theorem for a triangular, martingale array scheme.

Here, we work on the extension eq. (3.2.10) of the probability space, L denotes the Mittag-Leffler process of order $0 < \delta \leq 1$, and $W = (W^i)_{i \in I}$ denotes an I-dimensional standard Wiener process such that L, W and $\mathscr{F}$ are independent.

3.5.5 Theorem. *For $n \in \mathbb{N}^*$, let $(\mathscr{G}^n_s)_{s>0}$ be the filtration given by $\mathscr{G}^n_s := \mathscr{F}_{\lfloor sn \rfloor \Delta}$, and I be a finite index set. Moreover, let $h_n : E \times E \to \mathbb{R}^I$ be such that $\|h_n\|_\infty \to 0$ as $n \to \infty$. Grant Assumptions 3.2.2 and 3.2.4, and suppose that the process M^n given by*

$$M^n_s := \sum_{k=1}^{\lfloor sn \rfloor} h_n(X_{(k-1)\Delta}, \Delta^n_k X) \tag{3.5.13}$$

is a $\mathscr{G}^n_s$-martingale such that the predictable quadratic co-variation $\langle M^{ni}, M^{nj} \rangle$ is identically zero for every $i \neq j$ and all n large enough. If $(\langle M^{ni}, M^{ni} \rangle)_{i \in I}$ converges stably in law in $\mathcal{D}(\mathbb{R}^I)$ to $(\varsigma_i^2 L)_{i \in I}$, then

$$M^n \overset{\mathscr{L}-\mathrm{st}}{\underset{n\to\infty}{\Longrightarrow}} (\varsigma_i W^i_L)_{i \in I}.$$

Proof. Let $\delta = 1$. Then we have $L_s = s$. Therefore, the convergence of M^n to $(\varsigma_i^2 W^i)_{i \in I}$ follows directly from the martingale limit theorem Theorem 2.3.10.

For the remainder, let $0 < \delta < 1$. We consider the processes L^{ni}, $\bar{L}^n$, K^n and N^n given by

$$L^{ni}_s := \langle M^{ni}, M^{ni} \rangle_s = \sum_{k=1}^{\lfloor sn \rfloor} \mathbb{E}^{X_{(k-1)\Delta}} h^i_n(X_{(k-1)\Delta}, \Delta^n_k X)^2,$$

$$\bar{L}^n_s := \sum_{i \in I} L^{ni}_s, \quad K^n_u := \inf\{s > 0 : \bar{L}^n_s > u\} \quad \text{and} \quad N^n_s := M^n_{K^n_s}.$$

We emphasise that $N^n_{\bar{L}^n_s} = M^n_s + \Delta M^n_{K^n(\bar{L}^n_s)}$ holds for all s. As $\|\Delta M^n\|_\infty \leq \|h_n\|_\infty \to 0$, it is sufficient to prove the following stable convergence in law in $\mathcal{D}(\mathbb{R} \times \mathbb{R}^I)$:

$$\left(\bar{L}^n, N^n\right) \overset{\mathscr{L}-\mathrm{st}}{\underset{n\to\infty}{\Longrightarrow}} \left(\bar{\varsigma}^2 L, \left((\varsigma_i/\bar{\varsigma}) W^i\right)_{i \in I}\right), \quad \text{where } \bar{\varsigma}^2 := \sum_{i \in I} \varsigma_i^2. \tag{3.5.14}$$

By the continuous mapping theorem, we obtain

$$\left(\bar{L}^n, \left(L^{ni}\right)_{i\in I}\right) \overset{\mathscr{L}-\text{st}}{\underset{n\to\infty}{\Longrightarrow}} \left(\bar{\varsigma}^2 L, \left(\varsigma_i^2 L\right)_{i\in I}\right). \tag{3.5.15}$$

In addition, we remark that K_u^n is a predictable $\mathscr{G}_s^n$-stopping time for all $u \geq 0$. Thus, N^n is a martingale w. r. t. the time-changed filtration $\mathscr{H}_s^n := \mathscr{G}_{K_s^n}^n$. Moreover, we observe that its predictable quadratic variation satisfies

$$\langle N^{ni}, N^{ni}\rangle_s = L_{K_s^n}^{ni}.$$

By eq. (3.5.15), we have $|L^{ni} - (\varsigma_i/\bar{\varsigma})^2 \bar{L}^n| \to 0$ uniformly on compacts in probability for all $i \in I$. We note that the (scaled) Mittag-Leffler process $\bar{\varsigma}^2 L$ is a. s. continuous. Its right-inverse K given by $K_u := \inf\{s > 0 : \bar{\varsigma}^2 L_s > u\}$ is a (deterministically time-changed) δ-stable Lévy process, hence, without fixed time of discontinuity. By Proposition 2.3.6 (i), therefore, $L_{K_s^n}^{ni} \to (\varsigma_i/\bar{\varsigma})^2 s$ in law for every $s \geq 0$; hence, in probability. By construction, we have that $\|\Delta N_s^n\|_\infty$ is bounded above by $\|h_n\|_\infty$. This bound converges to zero. By the martingale limit theorem 2.3.10, consequently,

$$N^n \overset{\mathscr{L}-\text{st}}{\underset{n\to\infty}{\Longrightarrow}} \left((\varsigma_i/\bar{\varsigma}) W^i\right)_{i\in I}. \tag{3.5.16}$$

In analogy to the proof of Lemma 3.4.10 and Steps 6 and 7 on pp. 122–124 of Höpfner et al. (1990), we obtain that the pair $(\bar{L}^n, N^n)$ converges in law in $\mathcal{D}(\mathbb{R} \times \mathbb{R}^I)$ to the process $(\bar{\varsigma}^2 L, ((\varsigma_i/\bar{\varsigma})W^i)_{i\in I})$. Finally, the stable convergence in law and the independence from $\mathscr{F}$ follows in analogy to the proof of Lemma 3.4.11. □

3.5.3 Proof of Theorem 3.2.9

Throughout the remainder of Section 3.5, we work under the law $\mathbb{P}^\pi$ for some initial probability π on E, and we denote $E_\oplus := \{x \in E : \mu'(x) > 0, F(x, E) > 0\}$.

We consider the processes $G^{n,\Delta,\eta}$ and $R^{n,\Delta,\eta}$ given by

$$G_s^{n,\Delta,\eta}(x,y) := \frac{1}{v_{n\Delta}} \sum_{k=1}^{\lfloor sn \rfloor} g_1^{\eta,x}(X_{(k-1)\Delta}) g_2^{\eta,y}(\Delta_k^n X), \tag{3.5.17}$$

$$R_s^{n,\Delta,\eta}(x) := \frac{\Delta}{v_{n\Delta}} \sum_{k=1}^{\lfloor sn \rfloor} g_1^{\eta,x}(X_{(k-1)\Delta}). \tag{3.5.18}$$

3.5.6 Lemma. *Grant Assumptions 3.2.1 to 3.2.3. Let $\eta_n = \eta_{1,n}$ be such that eq. (3.2.6) holds, and let $x \in E_\oplus$.*

(i) If $n\Delta^2 \to 0$, then,

$$\text{the family } \left\{ \mathscr{L}\left(R_1^{n,\Delta,\eta_n}(x) \mid \mathbb{P}^\pi \right) : n \in \mathbb{N}^* \right\} \text{ is tight.} \tag{3.5.19}$$

(ii) Grant Assumption 3.2.4 in addition. If $(n\Delta)^{1-\delta}\Delta \to 0$, then, eq. (3.5.19) holds as well.

In both cases, each limit point of the family in eq. (3.5.19) is the law $\mathscr{L}(\mu'(x)\tilde{L})$ for some positive random variable $\tilde{L}$.

Proof. Let $S_s^{t,\eta}(x) := v_t^{-1} \int_0^{st} g_2^{\eta,x}(X_r)\mathrm{d}r$. By Lemma 3.4.7, the family $\{\mathscr{L}(S_1^{n\Delta,\eta_n}(x) \mid \mathbb{P}^\pi) : n \in \mathbb{N}^*\}$ is tight; moreover, each of its limit points is the law $\mathscr{L}(\mu'(x)\tilde{L})$ for some random variable $\tilde{L} > 0$. In both cases (i) and (ii), since η_n is such that eq. (3.2.6) holds, we have

$$\left| S_1^{n\Delta,\eta_n}(x) - R_1^{n,\Delta,\eta_n}(x) \right| \xrightarrow[n\to\infty]{\mathbb{P}^\pi} 0$$

by Proposition 3.5.2. Consequently, the family $\{\mathscr{L}(R_1^{n,\Delta,\eta}(x) \mid \mathbb{P}^\pi) : n \in \mathbb{N}^*\}$ is tight; moreover, each of its limit points is a limit point of $\{\mathscr{L}(S_1^{t,\eta}(x) \mid \mathbb{P}^\pi) : t > 0\}$, hence, the law $\mathscr{L}(\mu'(x)\tilde{L})$ for some random variable $\tilde{L} > 0$. □

3.5.7 Lemma. *Grant Assumptions 3.2.1 and 3.2.3. Let $\eta_n = (\eta_{1,n}, \eta_{2,n})$ be such that $\eta_{1,n} \to 0$, $\eta_{2,n} \to 0$ and $\Delta\eta_{2,n}^{-2\vee(\beta+d)} \to 0$. Moreover, let $(x,y) \in E_\oplus \times E^*$, and let g be a C^2-function with compact support. Then*

$$\lim_{n\to\infty} \sup_{z \in B_{\eta_{1,n}}(x)} \left| \frac{1}{\Delta} \mathbb{E}^z g^{\eta_n,y}(\Delta_1^n X) - f(x,y) \int g(w)\mathrm{d}w \right| = 0. \tag{3.5.20}$$

Proof. Firstly, by Proposition 3.5.1 – where we choose m large enough – we have

$$\lim_{n\to\infty} \sup_{z\in B_{\eta_{1,n}}(x)} \left|\Delta^{-1}\,\mathbb{E}^z\, g^{\eta,x}(\Delta_1^n X) - Fg^{\eta,y}(z)\right| = 0.$$

Secondly, under Assumption 3.2.3, $f \in \mathcal{C}^{\alpha}_{\text{loc}}(E\times E^*)$ for some $\alpha > 0$. Therefore,

$$\lim_{n\to\infty} \sup_{z\in B_{\eta_{1,n}}(x)} \left|Fg^{\eta_n,y}(z) - Fg^{\eta_n,y}(x)\right| \le \lim_{n\to\infty} \zeta\eta_{1,n}^{\alpha\wedge 1} = 0.$$

Lastly, by Lebesgue's differentiation theorem, we observe

$$\lim_{n\to\infty} \left|Fg^{\eta_n,y}(x) - f(x,y)\int g(w)\mathrm{d}w\right| = 0.$$

□

3.5.8 Lemma. *Grant Assumptions 3.2.1 to 3.2.3. Let $\eta_n = (\eta_{1,n},\eta_{2,n})$ be such that eq.* (3.2.6) *holds. Moreover, let $(x,y)\in E_\oplus \times E^*$. Then, in both cases as in Lemma 3.5.6,*

$$\text{the family } \left\{\mathscr{L}\left(G_1^{n,\Delta,\eta_n}(x,y), R_1^{n,\Delta,\eta_n}(x) \mid \mathbb{P}^\pi\right) : n\in\mathbb{N}^*\right\} \text{ is tight.} \tag{3.5.21}$$

Moreover, each limit point of the family in eq. (3.5.21) *is the law $\mathscr{L}(f(x,y)\mu'(x)\tilde{L}, \mu'(x)\tilde{L})$ for some positive random variable $\tilde{L}$.*

Proof. We note that $G_s^{n,\Delta,\eta}(x,y) = f(x,y)R_s^{n,\Delta,\eta}(x) + H_s^{n,\Delta,\eta}(x,y) + M_s^{n,\Delta,\eta}(x,y)$ with

$$H_s^{n,\Delta,\eta}(x,y) = \frac{1}{v_{n\Delta}}\sum_{k=1}^{\lfloor sn\rfloor} g_1^{\eta,x}(X_{(k-1)\Delta})\left(\mathbb{E}^{X_{(k-1)\Delta}}[g_2^{\eta,x}(\Delta_1^n X)] - \Delta f(x,y)\right), \tag{3.5.22}$$

$$M_s^{n,\Delta,\eta}(x,y) = \frac{1}{v_{n\Delta}}\sum_{k=1}^{\lfloor sn\rfloor} g_1^{\eta,x}(X_{(k-1)\Delta})\left(g_2^{\eta,y}(\Delta_k^n X) - \mathbb{E}^{X_{(k-1)\Delta}}[g_2^{\eta,x}(\Delta_1^n X)]\right). \tag{3.5.23}$$

By Lemma 3.5.6, it is sufficient to prove that $H_1^{n,\Delta,\eta_n}(x,y)$ and $M_1^{n,\Delta,\eta_n}(x,y)$ converge to zero in probability as $n\to\infty$.

(H) We observe

$$\left|H_1^{n,\Delta,\eta}(x,y)\right| \le \sup_{z\in B_{\eta_1}(x)} \left|\Delta^{-1}\,\mathbb{E}^z[g_2^{\eta,x}(\Delta_1^n X)] - f(x,y)\right| \sum_{k=1}^{n} \frac{\Delta h^{\eta,x}(X_{(k-1)\Delta})}{v_{n\Delta}}, \tag{3.5.24}$$

where h is a $\mathcal{C}^2$-function dominating $|g_1|$. In analogy to Lemma 3.5.6, the sequence $(v_{n\Delta}^{-1} \sum_{k=1}^{n} \Delta h^{\eta_n,x}(X_{(k-1)\Delta}))_{n\in\mathbb{N}^*}$ is tight. As

$$\sup_{z\in B_{\eta_{1,n}}(x)} \left|\frac{1}{\Delta}\, \mathbb{E}^z[g_2^{\eta_n,x}(\Delta_1^n X)] - f(x,y)\right| \xrightarrow[n\to\infty]{} 0$$

by Lemma 3.5.7, we have $H_1^{n,\Delta,\eta_n}(x,y) \to 0$ in law, hence, in probability. $\diamond$

(M) We observe that $M^{n,\Delta,\eta}$ is an $\mathscr{F}_{\lfloor sn\rfloor\Delta}$-martingale. We note $\sup_{s\le 1}\|\Delta M_s^{n,\Delta,\eta_n}\|_\infty \le (v_{n\Delta}\eta_{1,n}^d\eta_{2,n}^d)^{-1}\|g_1\|_\infty\|g_2\|_\infty \to 0$ by eq. (3.2.6). By the martingale limit theorem 2.3.10, thus, it is sufficient to show that the predictable quadratic variation of M^{n,Δ,η_n} at time one, denoted $\langle M^{n,\Delta,\eta_n}, M^{n,\Delta,\eta_n}\rangle_1$, converges to zero in probability.

We observe

$$\left\langle M^{n,\Delta,\eta}, M^{n,\Delta,\eta}\right\rangle_1 \le \frac{\|g_1\|_\infty}{v_{n\Delta}\eta_1^d\eta_2^d} \sup_{z\in B_{\eta_1}(x)} \left|\frac{\eta_2^d}{\Delta}\, \mathbb{E}^z\, g_2^{\eta,y}(\Delta_1^n X)^2\right| v_{n\Delta}^{-1} \sum_{k=1}^{n} \Delta h^{\eta,x}(X_{(k-1)\Delta}).$$

By Lemma 3.5.7,

$$\sup_{z\in B_{\eta_{1,n}}(x)} |\Delta^{-1}\, \mathbb{E}^z\, \eta_{2,n}^d g_2^{\eta_n,x}(\Delta_1^n X)^2| \xrightarrow[n\to\infty]{} f(x,y) \int g_1(w)^2 \mathrm{d}w.$$

In analogy to step (H), since $v_{n\Delta}\eta_{1.n}^d\eta_{2,n}^d \to \infty$, we have $\langle M^{n,\Delta,\eta_n}, M^{n,\Delta,\eta_n}\rangle_1 \to 0$ in law, hence, in probability. $\square$

Proof (of Theorem 3.2.9). We recall the results from Lemma 3.5.8. Let $\tilde{L} > 0$ be a random variable such that the law $\mathscr{L}(f(x,y)\mu'(x)\tilde{L}, \mu'(x)\tilde{L})$ is a limit point of the family in eq. (3.5.21), and let $(n_k)_{k\in\mathbb{N}^*}$ be a sequence such that

$$\left(G_1^{n_k,\Delta,\eta_{n_k}}(x,y), R_1^{n_k,\Delta,\eta_{n_k}}(x)\right) \xrightarrow[k\to\infty]{\mathscr{L}} \left(f(x,y)\mu'(x)\tilde{L}, \mu'(x)\tilde{L}\right).$$

Since $\mu'(x) > 0$, by the continuous mapping theorem, we conclude

$$\hat{f}_{n_k}^{\Delta,\eta_{n_k}}(x,y) = \frac{G_1^{n_k,\Delta,\eta_{n_k}}(x,y)}{R_1^{n_k,\Delta,\eta_{n_k}}(x)} \xrightarrow[k\to\infty]{\mathscr{L}} f(x,y).$$

As this limit is unique and independent of the particular limit point of the family in eq. (3.5.21), we have that $\hat{f}_n^{\Delta,\eta_n}(x,y)$ converges to $f(x,y)$ in law, hence, in probability.□

3.5.4 Proofs of Theorem 3.2.10 and Corollary 3.2.11

Throughout this subsection, we work on the extension eq. (3.2.10) of the probability space, L denotes the Mittag-Leffler process of order $0 < \delta \leq 1$, and $W = (W^i)_{i\in I}$ denotes an I-dimensional standard Wiener process such that L, W and $\mathscr{F}$ are independent.

We consider the processes $G^{n,\Delta,\eta}$ and $R^{n,\Delta,\eta}$ given by eq. (3.5.17) and eq. (3.5.18), and the processes $U^{n,\Delta,\eta}$ and $R'^{n,\Delta,\eta}$ given by

$$U_s^{n,\Delta,\eta}(x,y) := \sqrt{v_{n\Delta}\eta_1^d\eta_2^d}\left(G_s^{n,\Delta,\eta}(x,y) - \frac{\mu(g_1^{\eta,x}Fg_2^{\eta,y})}{\mu(g_1^{\eta,x})}R_s^{n,\Delta,\eta}(x)\right) \tag{3.5.25}$$

$$R_s'^{n,\Delta,\eta}(x) := \frac{\Delta}{v_{n\Delta}}\sum_{k=1}^{\lfloor sn\rfloor}\eta_1^d g_1^{\eta,x}(X_{(k-1)\Delta})^2. \tag{3.5.26}$$

We recall that, under Darling–Kac's condition, Touati's theorem 2.5.24 at hand. First, we obtain an extension of Lemma 3.5.6.

3.5.9 Lemma. *Grant Assumptions 3.2.1 to 3.2.4. Let $\eta_n = \eta_{1,n}$ be such that eq. (3.2.6) and eq. (3.2.8a) hold, and let $(x_i)_{i\in I}$ be a family of pairwise distinct points in $E_\oplus$. If $(n\Delta)^{1-\delta}\Delta \to 0$, then, under any law $\mathbb{P}^\pi$, we have the following stable convergence in law in $\mathcal{D}(\mathbb{R}^{2I})$:*

$$\left(R^{n,\Delta,\eta_n}(x_i), R'^{n,\Delta,\eta_n}(x_i)\right)_{i\in I} \overset{\mathscr{L}-\text{st}}{\underset{n\to\infty}{\Longrightarrow}} \left(\mu'(x_i)L, \mu'(x_i)\int g_2(w)^2\mathrm{d}wL\right)_{i\in I}. \tag{3.5.27}$$

Proof. Let $S_s^{t,\eta}(x) := v_t^{-1}\int_0^{st} g_1^{\eta,x}(X_r)\mathrm{d}r$ and $S_s'^{t,\eta}(x) := v_t^{-1}\int_0^{st}\eta^d g_1^{\eta,x}(X_r)^2\mathrm{d}r$. We note that $\mu(g_1^{\eta_n,x}) \to \mu'(x)$ and $\mu(\eta_n^d(g_1^{\eta_n,x})^2) \to \mu'(x)\int g_1(w)^2\mathrm{d}w$ for all x. By Theorems 2.5.24 and 3.4.1, we deduce – in analogy to Corollary 3.4.12 – that

$$\left(S^{n\Delta,\eta_n}(x_i), S'^{n\Delta,\eta_n}(x_i)\right)_{i\in I} \overset{\mathscr{L}-\text{st}}{\underset{t\to\infty}{\Longrightarrow}} \left(\mu'(x_i)L, \mu'(x_i)\int g_2(w)^2\mathrm{d}wL\right)_{i\in I}.$$

For every x, moreover, we deduce from Proposition 3.5.2 that

$$\left|R^{n,\Delta,\eta_n}(x) - S^{n\Delta,\eta_n}(x)\right| \overset{\text{ucp}}{\underset{n\to\infty}{\Longrightarrow}} 0 \quad \text{and} \quad \left|R'^{n,\Delta,\eta_n}(x) - S'^{n\Delta,\eta_n}(x)\right| \overset{\text{ucp}}{\underset{n\to\infty}{\Longrightarrow}} 0.$$

Consequently, we obtain eq. (3.5.27). □

In view of Theorem 3.5.5, we obtain the following preliminary result.

3.5.10 Lemma. *Grant Assumptions 3.2.1 to 3.2.5. Let $\eta_n = (\eta_{1,n}, \eta_{2,n})$ be such that eqs.* (3.2.6) *and* (3.2.8) *hold, and let $(x_i, y_i)_{i\in I}$ be a finite family of pairwise distinct points in $E_\oplus \times E^*$. If $(n\Delta)^{1-\delta}\Delta \to 0$, then, under any law $\mathbb{P}^\pi$, we have the following stable convergence in law in $\mathcal{D}(\mathbb{R}^I)$:*

$$\left(R^{n,\Delta,\eta_n}(x_i), U^{n,\Delta,\eta_n}(x_i,y_i)\right)_{i\in I} \overset{\mathscr{L}-\text{st}}{\underset{n\to\infty}{\Longrightarrow}} \left(\mu'(x_i)L, \sigma(x_i,y_i)\mu'(x_i)W_L^i\right)_{i\in I}, \tag{3.5.28}$$

where $\sigma(x,y)^2$ is given by eq. (3.2.11)*.*

Proof. Let $(\mathscr{G}_s^n)_{s\geq 0}$ be given by $\mathscr{G}_s^n = \mathscr{F}_{\lfloor sn\rfloor\Delta}$, and let the process $M^{n,\Delta,\eta}$ be given by

$$M_s^{n,\Delta,\eta}(x,y) := \sqrt{\frac{\eta_1^d\eta_2^d}{v_{n\Delta}}}\sum_{k=1}^{\lfloor sn\rfloor} g_1^{\eta,x}(X_{(k-1)\Delta})\left(g_2^{\eta,y}(\Delta_k^n X) - \mathbb{E}^{X_{(k-1)\Delta}} g_2^{\eta,y}(\Delta_1^n X)\right).$$

We note that $M^{n,\Delta,\eta}$ is a $\mathscr{G}_s^n$-martingale of the form eq. (3.5.13). The proof is divided into four steps: Firstly, we prove

$$\left|U^{n,\Delta,\eta_n}(x,y) - M^{n,\Delta,\eta_n}(x,y)\right| \overset{\text{ucp}}{\underset{n\to\infty}{\Longrightarrow}} 0. \tag{3.5.29}$$

Secondly, we show that the predictable quadratic variation of $M^{n,\Delta,\eta}(x,y)$ satisfies

$$\left(\langle M^{n,\Delta,\eta_n}(x_i,y_i), M^{n,\Delta,\eta_n}(x_i,y_i)\rangle\right)_{i\in I} \overset{\mathscr{L}-\text{st}}{\underset{n\to\infty}{\Longrightarrow}} \left([\sigma(x_i,y_i)\mu'(x_i)]^2 L\right)_{i\in I} \tag{3.5.30}$$

in $\mathcal{D}(\mathbb{R}^I)$. Thirdly, we show that $\langle M^{n,\Delta,\eta_n}(x_i,y_i), M^{n,\Delta,\eta_n}(x_j,y_j)\rangle$ vanishes for all n large enough if $i \neq j$. Lastly, we argue

$$\left(R^{n,\Delta,\eta_n}(x_i), \langle M^{n,\Delta,\eta_n}(x_i,y_i), M^{n,\Delta,\eta_n}(x_i,y_i)\rangle\right)_{i\in I} \overset{\mathscr{L}-\text{st}}{\underset{n\to\infty}{\Longrightarrow}} \left(\mu'(x_i)L, [\sigma(x_i,y_i)\mu'(x_i)]^2 L\right)_{i\in I}$$

in $\mathcal{D}(\mathbb{R}^{2I})$. By Theorem 3.5.5 and Proposition 2.3.6 (iv), we then obtain eq. (3.5.28).

(i) We note $U_s^{n,\Delta,\eta}(x,y) - M_s^{n,\Delta,\eta}(x,y) = H_s^{n,\Delta,\eta}(x,y) + H_s^{\prime n,\Delta,\eta}(x,y)$ with

$$H_s^{n,\Delta,\eta}(x,y) := \sqrt{v_{n\Delta}\eta_1^d\eta_2^d}\frac{\Delta}{v_{n\Delta}}\sum_{k=1}^{\lfloor sn\rfloor} g_1(X_{(k-1)\Delta})\left(Fg_2^{\eta,y}(X_{(k-1)\Delta}) - \frac{g_1^{\eta,x}Fg_2^{\eta,y}}{\mu(g_1^{\eta,x})}\right),$$

and

$$\left|H_s^{\prime n,\Delta,\eta}(x,y)\right| \leq \sqrt{v_{n\Delta}\eta_1^d\eta_2^d}\sup_{z\in B_{\eta_1}(x)}\left|\frac{1}{\Delta}\,\mathbb{E}^z\, g_2^{\eta,y}(\Delta_1^n X) - Fg_2^{\eta,y}(z)\right| R_s^{\prime\prime n,\Delta,\eta}(x),$$

where $R_s^{\prime\prime n,\Delta,\eta}(x) = \Delta v_{n\Delta}^{-1}\sum_{k=1}^{\lfloor sn\rfloor} h^{\eta,x}(X_{(k-1)\Delta})$ for some $\mathcal{C}^2$-function h, dominating $|g_1|$. Under Assumption 3.2.5, $Fg_2^{\eta,y}$ is twice continuously differentiable. Since eq. (3.2.8) holds, by Proposition 3.5.2 and step (i) in the proof of Lemma 3.4.8, $H^{n,\Delta,\eta_n}(x,y) \Rightarrow 0$ in ucp as $n\to\infty$. By Proposition 3.5.1 – where we choose m large enough – we have

$$\sup_{z\in B_{\eta_1}(x)}\left|\frac{1}{\Delta}\,\mathbb{E}^z\, g_2^{\eta,y}(\Delta_1^n X) - Fg_2^{\eta,y}(z)\right| \leq \zeta\left(\sqrt{\Delta} + \Delta\eta_2^{-2\vee(\beta+d)}\right)$$

since eq. (3.2.8a) holds. Since, moreover, eq. (3.2.8) holds, therefore,

$$\sqrt{v_{n\Delta}\eta_{1,n}^d\eta_{2,n}^d}\sup_{z\in B_{\eta_{1,n}}(x)}\left|\frac{1}{\Delta}\,\mathbb{E}^z\, g_2^{\eta_n,y}(\Delta_1^n X) - Fg_2^{\eta_n,y}(z)\right| \xrightarrow[n\to\infty]{} 0. \tag{3.5.31}$$

In analogy to Lemma 3.5.9, $R^{\prime\prime n,\Delta,\eta_n}(x)$ converges stably in law to some non-trivial process. So, $|H_s^{\prime n,\Delta,\eta}(x,y)| \Rightarrow 0$ in ucp as $n\to\infty$. Thus, eq. (3.5.29) holds. ◇

(ii) We note $\langle M^{n,\Delta,\eta}(x,y), M^{n,\Delta,\eta}(x,y)\rangle_s = K_s^{n,\Delta,\eta}(x,y) - K_s^{\prime n,\Delta,\eta}(x,y)$, where

$$K_s^{n,\Delta,\eta}(x,y) = \frac{\eta_1^d\eta_2^d}{v_{n\Delta}}\sum_{k=1}^{\lfloor sn\rfloor} g_1^{\eta,x}(X_{(k-1)\Delta})^2\left(\mathbb{E}^{X_{(k-1)\Delta}}\, g_2^{\eta,y}(\Delta_1^n X)^2\right),$$

and

$$|K_s^{\prime n,\Delta,\eta}(x,y)| \leq \sup_{z\in B_{\eta_1}(x)}\left|\frac{1}{\Delta^2}\left(\mathbb{E}^{X_{(k-1)\Delta}}\, g_2^{\eta,y}(\Delta_1^n X)\right)^2\right|\Delta\eta_2^d R_s^{\prime n,\Delta,\eta}(x).$$

By Lemma 3.5.7 and the continuous mapping theorem,

$$\sup_{z\in B_{\eta_{1,n}}(x)} \left| \frac{1}{\Delta^2} \left(\mathbb{E}^z g_2^{\eta_n,y}(\Delta_1^n X)\right)^2 \right| \xrightarrow[n\to\infty]{} f(x,y)^2.$$

By Lemma 3.5.9, $R_s'^{n,\Delta,\eta_n}(x)$ converges stably in law. Since $\Delta\eta_{2,n}^d \to 0$, we observe that $|K_s'^{n,\Delta,\eta_n}(x,y)|$ converges to zero uniformly on compacts in probability as $n\to\infty$.

Again by Lemma 3.5.7,

$$\sup_{z\in B_{\eta_{1,n}}(x)} \left| \frac{\eta_{2,n}^d}{\Delta} \mathbb{E}^{X_{(k-1)\Delta}} g_2^{\eta_n,y}(\Delta_1^n X)^2 - f(x,y)\int g_2(w)^2 \mathrm{d}w \right| \xrightarrow[n\to\infty]{} 0.$$

In analogy to $K'^{n,\Delta,\eta}(x,y)$, therefore,

$$\left| K^{n,\Delta,\eta_n}(x,y) - f(x,y)\int g_1(w)^2 \mathrm{d}w R'^{n,\Delta,\eta_n}(x) \right| \overset{\text{ucp}}{\underset{n\to\infty}{\Longrightarrow}} 0. \tag{3.5.32}$$

By Lemma 3.5.9, consequently,

$$\left(K^{n,\Delta,\eta_n}(x_i,y_i)\right)_{i\in I} \overset{\mathscr{L}-\text{st}}{\underset{n\to\infty}{\Longrightarrow}} \left(f(x_i,y_i)\int g_1(w)^2 \mathrm{d}w \mu'(x_i)\int g_2(z)^2 \mathrm{d}z L\right)_{i\in I};$$

hence, eq. (3.5.30) holds. ◇

(iii) Let $i,j\in I$. We note that for all n large enough such that $\eta_{1,n},\eta_{2,n}$ are small enough, we have $g_1^{\eta_n,x_i} g_1^{\eta_n,x_j} \equiv 0$ whenever $x_i \neq x_j$, and $g_2^{\eta_n,y_i} g_2^{\eta_n,y_j} \equiv 0$ whenever $y_i \neq y_j$. For all ω and n large enough, thus, $\langle M^{n,\Delta,\eta_n}(x_i,y_i), M^{n,\Delta,\eta_n}(x_j,y_j)\rangle_s \equiv 0$ if $i\neq j$. ◇

(iv) By Lemma 3.5.9 and eq. (3.5.32), we obtain the joint convergence of the processes $(R^{n,\Delta,\eta_n}(x_i))_{i\in I}$ and $\langle M^{n,\Delta,\eta_n}(x_i,y_i), M^{n,\Delta,\eta_n}(x_i,y_i)\rangle)_{i\in I}$ to the required limit. □

Proof (of Theorem 3.2.10). For every n, and $(x,y)\in E_\oplus \times E^*$, we have

$$\sqrt{v_{n\Delta}\eta_{1,n}^d\eta_{2,n}^d}\left(\hat{f}_n^{\Delta,\eta_n}(x,y) - \bar{f}^{\eta_n}(x,y)\right) = \frac{U_1^{n,\Delta,\eta_n}(x,y)}{R_1^{n,\Delta,\eta_n}(x)},$$

where $\bar{f}^\eta(x,y) := \mu(g_1^{\eta,x} F g_2^{\eta,y})/\mu(g_1^{\eta,x})$. Since L and W are independent, $V(x_i,y_i) :=$

$L_1^{-1/2}W^i_{L_1}$ defines an I-dimensional standard Gaussian random vector such that L, V and $\mathscr{F}$ are independent. By the continuous mapping theorem and Lemma 3.5.10, consequently,

$$\sqrt{v_{n\Delta}\eta_{1,n}^d\eta_{2,n}^d}\left(\hat{f}_n^{\Delta,\eta_n}(x_i,y_i)-\bar{f}^{\eta_n}(x_i,y_i)\right)_{i\in I}\xrightarrow[n\to\infty]{\mathscr{L}-\mathrm{st}}\left(\sigma(x_i,y_i)V(x_i,y_i)L_1^{-1/2}\right)_{i\in I},$$

where $\sigma(x,y)^2$ is given by eq. (3.2.11).

In addition, let $\eta_n=(\eta_{1,n},\eta_{2,n})$ be such that eq. (3.2.7) holds as well. It remains to prove $(v_{n\Delta}\eta_{1,n}^d\eta_{2,n}^d)^{1/2}(\bar{f}^{\eta_n}(x,y)-f(x,y))\to\gamma(x,y)$. This, however, follows in analogy to the proof of Theorem 3.3.7. □

Proof (of Corollary 3.2.11). In analogy to the proof of Theorem 3.2.10, by Lemma 3.5.10 it remains to show that $(v_{n\Delta}\eta_{1,n}^d\eta_{2,n}^d)^{1/2}\hat{\gamma}_n^{\eta_n}(x,y)$ is a consistent estimator for $\gamma(x,y)$. This, however, follows in analogy to the proof of Corollary 3.3.8. □

3.6 On the auxiliary Markov chains Z and Z'

In this section, we derive an explicit representation for the transition kernel Φ of the auxiliary process Z', and (in-)equalities for expectations of the form $\mathbb{E}^x(\int_0^{T_1}h(X_s)\mathrm{d}s)^k$. In addition, we derive representations for the stationary probability measures ψ and φ of the processes Z and Z'.

We invoke technical results on resolvents of semi-groups. We recall from Section 2.5 that the *resolvent* $(R_\lambda)_{\lambda>0}$ of a semi-group $(P_t)_{t\geq 0}$ is given by $R_\lambda:=\int_0^\infty\exp(-\lambda t)P_t\mathrm{d}t$. For bounded measurable functions h, the generalised resolvent kernel R_h is given by

$$R_h(x,A):=\mathbb{E}^x\int_0^\infty e^{-\int_0^t h(X_s)\mathrm{d}s}\mathbb{1}_A(X_t)\mathrm{d}t\qquad\forall x\in E,A\in\mathscr{E}.$$

These kernels were first introduced by Neveu (1972). For a comprehensive interpretation, we refer to Section 4 of Down, Meyn, and Tweedie (1995).

3.6.1 Lemma. *Let $(R_\lambda)_{\lambda>0}$ be the resolvent of X, and let $(R^*_\lambda)_{\lambda>0}$ be given by*

$$R^*_\lambda := R_\lambda \sum_{k=0}^{\infty} \Big((\mathrm{I}_q - \mathrm{I}_q \bar{\Pi}) R_\lambda \Big)^k, \quad \textit{where} \quad \mathrm{I}_q h(x) := q(x)h(x). \tag{3.6.1}$$

Then $(R^*_\lambda)_{\lambda>0}$ is the resolvent of a positive contraction semi-group. For its corresponding process X^*, we have that the laws of $X^* \mathbb{1}_{[\![0,T_1[\![}$ and $X \mathbb{1}_{[\![0,T_1[\![}$ are equal.

Proof. Since $\mathrm{I}_q \bar{\Pi}$ is a bounded kernel, $(R^*_\lambda)_{\lambda>0}$ is the resolvent of a positive contraction semi-group by Theorem 4.2 of Bass (1979). It follows from Sawyer (1970) and Chapter 6 of Bass (1979) that, for the process X^* (corresponding to $(R^*_\lambda)_{\lambda>0}$), we have that the laws of $X^* \mathbb{1}_{[\![0,T_1[\![}$ and $X \mathbb{1}_{[\![0,T_1[\![}$ are equal. □

3.6.2 Lemma. *Let h be a measurable function on E. Then*

$$\mathbb{E}^x\, h(Z_1') = R^*_q I_q h(x) \quad \textit{and} \quad \mathbb{E}^x \int_0^{T_1} h(X_s)\mathrm{d}s = R^*_q h(x), \tag{3.6.2}$$

where R^*_q denotes the generalised resolvent kernel associated with the modified resolvent $(R^*_\lambda)_{\lambda>0}$ and the function q. For every $\lambda_q \geq \|q\|_\infty$, we have $R^*_q = \sum_{k=0}^\infty R^*_{\lambda_q} (\mathrm{I}_{\lambda_q - q} R^*_{\lambda_q})^k$.

Proof. We recall that the laws of $X^* \mathbb{1}_{[\![0,T_1[\![}$ and $X \mathbb{1}_{[\![0,T_1[\![}$ are equal. The expectation of $h(Z_1')$ under $\mathbb{P}^x$, therefore, coincides with the expectation of $h(X^*)$ sampled at an independent killing time according to the multiplicative functional $\exp(-\int_0^\cdot q(X^*_s)\mathrm{d}s)$. In formulas, we have

$$\mathbb{E}^x\, h(Z_1') = \mathbb{E}^x \int_0^\infty e^{-\int_0^t q(X^*_s)\mathrm{d}s} q(X^*_t) h(X^*_t) \mathrm{d}t.$$

By eq. (19) of Down et al. (1995), hence, $\mathbb{E}^x\, h(Z_1') = R^*_q \mathrm{I}_q h(x)$, where R^*_q denotes the generalised resolvent kernel associated with the modified resolvent $(R^*_\lambda)_{\lambda>0}$. By Chapter 7 of Neveu (1972), $R^*_q = \sum_{k=0}^\infty R^*_{\lambda_q} (\mathrm{I}_{\lambda_q - q} R^*_{\lambda_q})^k$ holds for every $\lambda_q \geq \|q\|_\infty$.

Similarly, we observe

$$\mathbb{E}^x \int_0^{T_1} h(X_s)\mathrm{d}s = \mathbb{E}^x \int_0^\infty e^{-\int_0^t q(X^*_u)\mathrm{d}u} q(X^*_t) \int_0^t h(X^*_s)\mathrm{d}s\mathrm{d}t. \tag{3.6.3}$$

By Fubini's theorem (cf., eq. (20) of Down et al. (1995)), consequently,

$$\mathbb{E}^x \int_0^{T_1} h(X_s)\mathrm{d}s = \mathbb{E}^x \int_0^{\infty} e^{-\int_0^t q(X_s^*)\mathrm{d}s} h(X_t^*)\mathrm{d}t = R_q^* h(x).$$

□

Remark. It is immediate from Lemma 3.4.2 that $\Phi = \bar{\Pi} R_q^* \mathrm{I}_q$.

We obtain two corollaries:

3.6.3 Corollary. *Let $h_1, \ldots, h_k$ be measurable functions on E. Then*

$$\mathbb{E}^x \prod_{j=1}^{k} \int_0^{T_1} h_j(X_s)\mathrm{d}s = \sum_{j=1}^{k} \mathbb{E}^x \int_0^{\infty} e^{-\int_0^t q(X_u^*)\mathrm{d}u} h_j(X_t^*) \prod_{l \neq j} \int_0^t h_l(X_s^*)\mathrm{d}s\mathrm{d}t. \tag{3.6.4}$$

Proof. In analogy to eq. (3.6.3), we observe

$$\mathbb{E}^x \prod_{j=1}^{k} \int_0^{T_1} h_j(X_s)\mathrm{d}s = \mathbb{E}^x \int_0^{\infty} e^{-\int_0^t q(X_u^*)\mathrm{d}u} q(X_t^*) \prod_{j=1}^{k} \int_0^t h_j(X_s^*)\mathrm{d}s\mathrm{d}t.$$

By the Leibniz rule, moreover,

$$\prod_{j=1}^{k} \int_0^t h_j(X_s^*)\mathrm{d}s = \sum_{j=1}^{k} \int_0^t h_j(X_s^*) \prod_{l \neq j} \int_0^s h_l(X_r^*)\mathrm{d}r\mathrm{d}s.$$

By Fubini's theorem, therefore, we have eq. (3.6.4). □

3.6.4 Corollary. *Let h be a bounded measurable function on E. For all $k \in \mathbb{N}^*$, if $\inf_{x \in \mathrm{supp}(h)} q(x) > 0$, then*

$$\mathbb{E}^x \left(\int_0^{T_1} h(X_s)\mathrm{d}s \right)^k \leq \frac{k! \|h\|_\infty^{k-1}}{(\inf_{x \in \mathrm{supp}(h)} q(x))^{k-1}} R_q^* |h|(x). \tag{3.6.5}$$

Proof (by induction). By Lemma 3.6.2, we immediately have eq. (3.6.5) for $k = 1$. We assume that eq. (3.6.5) holds for some $k \in \mathbb{N}^*$. Then we deduce from Corollary 3.6.3 and $|h| \leq q\|h\|_\infty / (\inf_{x \in \mathrm{supp}(h)} q(x))$ that eq. (3.6.5) holds for $k + 1$. □

3.6.5 Lemma. $\mu \mathrm{I}_q \bar{\Pi} R_q^* = \mu$.

Proof. By Theorem 4.2 of Bass (1979) and Section 7 of Neveu (1972), we have

$$(\mathrm{I}_q \bar{\Pi} - (\mathrm{I} - R_1^{-1}))R_q^* = \mathrm{I},$$

where the formal inverse of R_1 is defined by $R_1^{-1} := \sum_{k=0}^\infty (\mathrm{I} - R_1)^k$. Since μ is invariant w. r. t. $(P_t)_{t \geq 0}$, we also have $\mu R_1 = \mu$ and $\mu = \mu R_1^{-1}$. Hence, $\mu \mathrm{I}_q \bar{\Pi} = \mu(\mathrm{I}_q \bar{\Pi} - (\mathrm{I} - R_1^{-1}))$. Therefore, $\mu \mathrm{I}_q \bar{\Pi} R_q^* = \mu$. □

3.6.6 Corollary. *The measures* $\varphi := (\mu(q))^{-1} \mu \mathrm{I}_q$ *and* $\psi := \varphi \Psi$ *are the invariant probability measures w. r. t.* Φ *and* Ψ*.*

Proof. Since $\Phi = \bar{\Pi} R_q^* \mathrm{I}_q$, we observe $\mu \mathrm{I}_q \Phi = \mu \mathrm{I}_q$. By eq. (3.4.10), $\varphi \Psi^{k+1} = \varphi \Phi^k \Psi = \varphi \Psi$ in addition. □

4 On the estimation of the Lévy measure of time-changed Lévy processes

This chapter is dedicated to the case of a time-changed Lévy process. In general, such a process is no longer Markov; unless, for instance, the time-change is a Lévy subordinator itself. In analogy to Chapter 3, we study a kernel density estimator for the density of the Lévy measure. Our results and methods of proof are adapted from the Markov case.

4.1 Introduction

In this chapter, we consider a process $X_t = L_{T_t}$ where L is a Lévy process with Lévy measure F and T is an absolutely continuous time-change. We assume that F admits a density $x \mapsto f(x)$, which we want to estimate in a non-parametric way.

On an equidistant time grid, we observe a sample $X_0(\omega), X_\Delta(\omega), \ldots, X_{n\Delta}(\omega)$ of the process; the jumps and the time-change are latent. We study a kernel density estimator for $f(x)$. We show its consistency as $n\Delta \to \infty$ and $\Delta \to 0$ under a smoothness hypothesis on the estimated density and an ergodicity assumption on the time-change. In addition, we prove a central limit theorem. A standardised version for the construction of asymptotic confidence intervals is provided.

Our results are comparable to those in the positive recurrent Markov case and to those in classical non-parametric density estimation. By an optimal choice of the bandwidth, if $\Delta \to 0$ fast enough, our estimator's rate is $(n\Delta)^{\alpha/(2\alpha+d)}$, where α stands for the smoothness of the function f. We remark that our achieved rate equals the non-parametric minimax rate of smooth density estimation.

Several non-parametric estimation methods and divers statistical frameworks for the estimation of the Lévy measure of a time-changed Lévy process have been considered in literature. For the low-frequency case, where observations take place on a fixed sampling grid with $\Delta > 0$, we refer, for instance, to Belomestny (2011). Figueroa-López (2009b, 2011) considers the high-frequency setting. Consistency and a central limit theorem are proved for an estimator of the integral $\int F(\mathrm{d}x)g(x)$ of a test function g w. r. t. to the Lévy measure. Estimators for such integrals serve as the main building block for projection estimators. Although Figueroa-López's estimator is related to ours, the results are clearly distinguishable. In our study, we consider a sequence of functions $(g_n)_{n\in\mathbb{N}}$ which satisfy $\int F(\mathrm{d}x)g_n(x) \to f(x)$; this is in contrast to the usage of a fixed function g in Figueroa-López (2009b, 2011).

We briefly outline this chapter. In Section 4.2 we study the estimation of the Lévy measure based on discrete observations. Split into three subsections, we present the statistical problem with our standing assumptions; we give our estimator along with a bias correction; and state our main results – the estimator's consistency and the central limit theorem. The corresponding proofs are in Section 4.3.

4.2 Estimation of the Lévy density from high-frequency observations

Throughout this chapter, we use the notation introduced in Chapters 2 and 3.

4.2.1 Preliminaries and assumptions

On the filtered probability space $(\Omega, \mathscr{F}, (\mathscr{G}_t)_{t\geq 0}, \mathbb{P})$, let $L = (L_t)_{t\geq 0}$ be a Lévy process with values in $E = \mathbb{R}^d$ and characteristic triple (b, c, F). Moreover, let $Y = (Y_t)_{t\geq 0}$ be a positive càdlàg process – independent of L – such that

$$T_t := \int_0^t Y_s \mathrm{d}s \quad \text{is a } \mathscr{G}_t\text{-stopping time for all } t \geq 0.$$

By Corollaire 10.12 of Jacod (1979) (recall Theorem 2.2.14), the *time-changed Lévy process* X given by $X_t := L_{T_t}$ is an Itō semi-martingale w. r. t. the filtration $(\mathscr{F}_t)_{t\geq 0}$ given by $\mathscr{F}_t := \mathscr{G}_{T_t}$. Moreover, its characteristics $(B, C, \mathfrak{n})$ satisfy

$$\mathrm{d}B_t = bY_t\mathrm{d}t, \quad \mathrm{d}C_t = cY_t\mathrm{d}t, \quad \text{and} \quad \mathfrak{n}(\mathrm{d}t, \mathrm{d}x) = Y_t\mathrm{d}t \otimes F(\mathrm{d}x).$$

For $n \in \mathbb{N}^*$ and $\Delta > 0$, we observe the increments

$$\Delta_k^n X(\omega) := X_{k\Delta}(\omega) - X_{(k-1)\Delta} \quad k = 1, \dots, n.$$

We emphasise that the jumps of the process and the time-change are latent.

Throughout, we impose the following assumption on the time-change:

4.2.1 Assumption. The process Y is ergodic with stationary distribution μ on $\mathbb{R}_+^*$ such that $\mathbb{E}^\mu Y_t^4 = \int x^4 \mu(\mathrm{d}x) < \infty$. ⋄

W. l. o. g., we suppose that $\mathbb{E}^\mu Y_t = 1$. Moreover, we assume that the Lévy measure F of L admits a density $x \mapsto f(x)$ which we want to estimate. Also, we impose an assumption on the smoothness of f:

4.2.2 Assumption. There exists an $\alpha > 0$ for which the Lévy measure F of L admits a density $f \in C_{\mathrm{loc}}^\alpha(E^*)$. ⋄

To obtain a central limit theorem, we suppose in addition:

4.2.3 Assumption. The family $\{(\sqrt{t}(t^{-1}T_{st} - s))_{s \geq 0} : t > 0\}$ is tight. ⋄

4.2.2 Kernel density estimator

We briefly outline our method of estimation: Firstly, we choose a smooth kernel g with support $B_1(0)$ which is, at least, of order α; that is, for every multi-index $m = (m_1, \dots, m_d) \in \mathbb{N}^d \setminus \{0\}$, we have

$$|m| := m_1 + \cdots + m_d < \alpha_i \implies \kappa_m(g) := \int x_1^{m_1} \cdot \cdots \cdot x_d^{m_d} g(x)\mathrm{d}x = 0. \qquad (4.2.1)$$

Secondly, we choose a bandwidth $\eta > 0$. Lastly, we construct an estimator for $f(x)$ using the kernel $g^{\eta,x}(z) := \eta^{-d} g(\eta^{-1}(z - x))$. If the bandwidth is chosen appropriately, we achieve a consistent and asymptotically normal estimator.

4.2.4 Definition. For $\eta > 0$, we call $\hat{f}_n^{\Delta,\eta}$ defined by

$$\hat{f}_n^{\Delta,\eta}(x) := \frac{1}{n\Delta} \sum_{k=1}^n g^{\eta,x}(\Delta_k^n X) \qquad (4.2.2)$$

the *kernel density estimator* of f (w. r. t. bandwidth η based on $X_0, X_\Delta, \dots, X_{n\Delta}$). ⋄

In analogy to classical kernel density estimation, we also introduce a bias correction for our estimator.

4.2.5 Definition. For $\eta > 0$, we call $\hat{\gamma}_n^{\Delta,\eta}$ defined by

$$\hat{\gamma}_n^{\Delta,\eta}(x) := \begin{cases} \eta^\alpha \sum_{|m|=\alpha} \frac{\kappa_m(g)}{m!} \frac{\partial^m}{\partial x^m} \hat{f}_n^{\Delta,\eta}(x), & \text{if } \alpha \in \mathbb{N}^*, \\ 0, & \text{otherwise,} \end{cases} \tag{4.2.3}$$

the *bias correction* for $\hat{f}_n^{\Delta,\eta}(x)$. ⋄

4.2.3 Consistency and central limit theorem

In this subsection, we present our main results of this chapter. We utilise the following conditions as $n\Delta \to \infty$ and $\Delta \to 0$, where $0 \leq \zeta < \infty$:

$$n\Delta\eta_n^d \to \infty, \quad \text{and} \quad \eta_n \to 0; \tag{4.2.4}$$

$$n\Delta\eta_n^{d+2\alpha} \to \zeta^2. \tag{4.2.5}$$

In addition, we also utilise the following conditions due to discretisation:

$$\Delta\eta_n^{-2-d} \to 0; \tag{4.2.6a}$$

$$n\Delta^2\eta_n^d \to 0, \quad \text{and} \quad n\Delta^3\eta_n^{-4-d} \to 0. \tag{4.2.6b}$$

Remark. If $\Delta \to 0$ fast enough, then eqs. (4.2.4) and (4.2.5) are the crucial conditions.

4.2.6 Theorem. *Grant Assumptions 4.2.1 and 4.2.2. Let η_n be such that eqs.* (4.2.4) *and* (4.2.6a) *hold. Moreover, let $x \neq 0$. Then we have the following convergence in probability:*

$$\hat{f}_n^{\Delta,\eta_n}(x) \xrightarrow[n\to\infty]{\mathbb{P}} f(x). \tag{4.2.7}$$

For the next theorem, we establish additional notation. On an extension

$$(\tilde{\Omega}, \tilde{\mathscr{F}}, \tilde{\mathbb{P}}) := (\Omega \times \Omega', \mathscr{F} \otimes \mathscr{F}', \mathbb{P} \otimes \mathbb{P}') \tag{4.2.8}$$

of the probability space, let $V = (V(x))_{x \in \mathbb{R}^*}$ be a standard Gaussian white noise random field such that V and $\mathscr{F}$ are independent. In the theorem below, convergence holds stably in law (recall Definition 2.3.1).

4.2.7 Theorem. *Grant Assumptions 4.2.1 to 4.2.3. Let $(x_i)_{i \in I}$ be a finite family of pairwise distinct points in E^*, and let η_n be such that eqs.* (4.2.4) *and* (4.2.6) *hold. Then we have the following stable convergence in law:*

$$\left(\sqrt{n \Delta \eta_n^d} \left(\hat{f}_n^{\Delta, \eta_n}(x_i) - F g^{\eta_n, x_i} \right) \right)_{i \in I} \xrightarrow[n \to \infty]{\mathscr{L}-\mathrm{st}} \left(\sigma(x_i) V(x_i) \right)_{i \in I}, \tag{4.2.9}$$

where the asymptotic variance is given by

$$\sigma(x)^2 := f(x) \int g(z)^2 \mathrm{d}z. \tag{4.2.10}$$

In addition, let η_n be such that eq. (4.2.5) *holds as well. Suppose either that $\alpha \in \mathbb{N}^*$ or that $\zeta = 0$ in eq.* (4.2.5)*. Then we have the following stable convergence in law:*

$$\left(\sqrt{n \Delta \eta_n^d} \left(\hat{f}_n^{\Delta, \eta_n}(x_i) - f(x_i) \right) \right)_{i \in I} \xrightarrow[n \to \infty]{\mathscr{L}-\mathrm{st}} \left(\gamma(x_i) + \sigma(x_i) V_i \right)_{i \in I}, \tag{4.2.11}$$

where – in the former case – the asymptotic bias $\gamma(x)$ is given by

$$\gamma(x) := \zeta \sum_{|m| = \alpha} \frac{\kappa_m(g)}{m!} \frac{\partial^m}{\partial x^m} f(x), \tag{4.2.12}$$

and – in the latter case – $\gamma(x) = 0$.

Remark. The asymptotic bias and variance of our estimator are analogous to those of our estimators in the Markov case (Chapter 3) and, also, analogous to those of the Rosenblatt–Parzen window estimator in classical density estimation.

If we choose $\eta_n = (n\Delta)^{-1/(2\alpha+d)}$, then eqs. (4.2.4) and (4.2.5) hold with $\zeta = 1$. If $\Delta \to 0$ fast enough such that $n\Delta^{1+(2\alpha+d)/(\alpha+d+2)} \to 0$ in addition, then our choice of η_n also satisfies eq. (4.2.6). Consequently, our estimator's rate of convergence is

$$(n\Delta)^{\alpha/(2\alpha+d)}.$$

This equals the non-parametric minimax rate of smooth density estimation.

Theorem 4.2.7 does not allow for a direct construction of confidence intervals. For this purpose, we also obtain the following standardised version.

4.2.8 Corollary. *Grant Assumptions 4.2.1 to 4.2.3. Let η_n be such that eqs.* (4.2.4) *to* (4.2.6) *hold. Suppose either that $\alpha \in \mathbb{N}^*$ or that $\zeta = 0$ in eq.* (4.2.5). *Then we have the following stable convergence in law:*

$$\left(\sqrt{\frac{n\Delta\eta_n^d}{\hat{f}_n^{\Delta,\eta_n}(x_i)\int g(z)^2\mathrm{d}z}}\left(\hat{f}_n^{\Delta,\eta_n}(x_i)-\hat{\gamma}_n^{\Delta,\eta_n}(x_i)-f(x_i)\right)\right)_{i\in I}\xrightarrow[n\to\infty]{\mathscr{L}-\mathrm{st}}\left(V(x_i)\right)_{i\in I}.$$

4.3 Proofs

Throughout this section, we work on the extension of the probability space given in eq. (4.2.8). Given the index family I, we denote by $W=(W^i)_{i\in I}$ an I-dimensional Wiener process such that W and $\mathscr{F}$ are independent.

Proof (of Theorem 4.2.6). For $n\in\mathbb{N}^*$ and $x\neq 0$, let $G^{n,\Delta,\eta}(x)$ be given by

$$G_s^{n,\Delta,\eta}(x)=\frac{1}{n\Delta}\sum_{k=1}^{\lfloor sn\rfloor}g^{\eta,x}(\Delta_k^n X).$$

Moreover let $(\mathscr{H}_s^n)_{s\geq 0}$ be the filtration given by $\mathscr{H}_s^n:=\mathscr{F}_{\lfloor sn\rfloor\Delta}$. For $s=k/n$ and $k\in\mathbb{N}$, we have $\mathscr{H}_{s-}^n=\mathscr{F}_{(k-1)\Delta}$. We decompose the process $G^{n,\Delta,\eta}(x)$ as follows:

$$G_s^{n,\Delta,\eta}(x)=H_s^{n,\Delta,\eta}(x)+H_s'^{n,\Delta,\eta}(x)+M_s^{n,\Delta,\eta}(x),$$

where

$$\begin{aligned}H_s^{n,\Delta,\eta}(x)&:=(n\Delta)^{-1}T_{\lfloor sn\rfloor\Delta}Fg^{\eta,x},\\ H_s'^{n,\Delta,\eta}(x)&:=\frac{1}{n\Delta}\sum_{k=1}^{\lfloor sn\rfloor}\mathbb{E}\left[g^{\eta,x}(\Delta_k^n X)-Fg^{\eta,x}\Delta_k^n T\,\middle|\,\mathscr{F}_{(k-1)\Delta}\right],\end{aligned}$$

and

$$M_s^{n,\Delta,\eta}(x) := \frac{1}{n\Delta}\sum_{k=1}^{\lfloor sn \rfloor}\Big(g^{\eta,x}(\Delta_k^n X) - Fg^{\eta,x}\Delta_k^n T\Big) - H_s'^{n,\Delta,\eta}(x).$$

We prove that $H_s^{n,\Delta,\eta_n}(x) \to sf(x)$, $H_s'^{n,\Delta,\eta_n}(x) \to 0$ and $M_s^{n,\Delta,\eta_n}(x) \to 0$ in probability for every $s \geq 0$ as $n \to \infty$.

(H) Since Y is ergodic (Assumption 4.2.1), we have $(n\Delta)^{-1}T_{\lfloor sn \rfloor \Delta} \to s$ in probability as $n\Delta \to \infty$ for every $s \geq 0$. Since f is continuous at x (Assumption 4.2.2), moreover, we observe

$$Fg^{\eta,x} = \int f(x+\eta z)g(z)\mathrm{d}z \to f(x) \quad \text{as } \eta \to 0.$$

Consequently, $H_s^{n,\Delta,\eta_n}(x) \to sf(x)$ in probability as $n \to \infty$. ⋄

(H') We recall that – conditionally on T – the distribution $\mathscr{L}(\Delta_k^n X \mid \Delta_k^n T)$ is equal to the distribution $\mathscr{L}(L_{\Delta_k^n T} \mid \Delta_k^n T)$. Since L and Y are independent, by Proposition 3.5.1 where we choose $m = 3$, we obtain that there exists a $\zeta < \infty$ such that

$$\begin{aligned}\Big|\mathbb{E}\Big[g^{\eta,x}(\Delta_k^n X) - Fg^{\eta,x}\Delta_k^n T \,\Big|\, \mathscr{F}_{(k-1)\Delta}\Big]\Big| &\leq \\ \zeta\,\mathbb{E}\Bigg[(\Delta_k^n T)^{3/2} + \frac{(\Delta_k^n T)^2}{\eta^{2+d}}\Bigg(1+\sum_{j=1}^{3}\frac{(\Delta_k^n T)^j}{\eta^{2j}} + \frac{(\Delta_k^n T)^2}{\eta^6}\Bigg)\,\Bigg|\,\mathscr{F}_{(k-1)\Delta}\Bigg].\end{aligned} \tag{4.3.1}$$

By Hölder's inequality, we have $(\Delta_k^n T)^p \leq \Delta^{p-1}\int_{(k-1)\Delta}^{k\Delta} Y_s^p \mathrm{d}s$ for $p \geq 1$. Since Y is ergodic and $\mathbb{E}^\mu Y_t^4 < \infty$ (Assumption 4.2.1), moreover, we have

$$\frac{1}{n\Delta}\sum_{k=1}^{\lfloor sn \rfloor}\mathbb{E}\left[\int_{(k-1)\Delta}^{k\Delta} Y_s^p \mathrm{d}s \,\middle|\, \mathscr{F}_{(k-1)\Delta}\right] \xrightarrow[n\to\infty]{\mathbb{P}} s\,\mathbb{E}^\mu Y_t^p \quad \text{for all } p \leq 4.$$

Since $\Delta\eta_n^{-2-d} \to 0$ by eq. (4.2.6a), thus, $H_s'^{n,\Delta,\eta_n}(x) \to 0$ in probability as $n \to \infty$. ⋄

(M) We note that $M^{n,\Delta,\eta_n}(x)$ is an $\mathscr{H}_s^n$-martingale. Since $n\Delta\eta_n^d \to \infty$ by eq. (4.2.4), we observe that $\sup_{r \leq s}|\Delta M_r^{n,\Delta,\eta_n}(x)| \to 0$ in probability as $n \to \infty$. By the martingale limit theorem 2.3.10, therefore, it is sufficient to prove that its predictable quadratic variation at time $s = 1$ converges to zero as $n \to \infty$.

We observe

$$\left\langle M^{n,\Delta,\eta}(x), M^{n,\Delta,\eta}(x)\right\rangle_s \leq \frac{4}{n^2\Delta^2}\sum_{k=1}^{\lfloor sn\rfloor}\mathbb{E}\left[g^{\eta,x}(\Delta_k^n X)^2 + (Fg^{\eta,x})^2\,(\Delta_k^n T)^2\,\middle|\,\mathscr{F}_{(k-1)\Delta}\right].$$

In analogy to the case (H'), we have

$$\frac{\eta_n^d}{n\Delta}\sum_{k=1}^{\lfloor sn\rfloor}\mathbb{E}\left[g^{\eta_n,x}(\Delta_k^n X)^2 - F(g^{\eta_n,x})^2\Delta_k^n T\,\middle|\,\mathscr{F}_{(k-1)\Delta}\right]\xrightarrow[n\to\infty]{\mathbb{P}} 0, \tag{4.3.2}$$

$$\begin{aligned}\left(\frac{Fg^{\eta_n,x}}{n\Delta}\right)^2 &\sum_{k=1}^{\lfloor sn\rfloor}\mathbb{E}[(\Delta_k^n T)^2 \mid \mathscr{F}_{(k-1)\Delta}] \leq \\ &\frac{(Fg^{\eta_n,x})^2}{n}\cdot\frac{1}{n\Delta}\sum_{k=1}^{\lfloor sn\rfloor}\mathbb{E}\left[\int\limits_{(k-1)\Delta}^{k\Delta} Y_s^2\mathrm{d}s\,\middle|\,\mathscr{F}_{(k-1)\Delta}\right]\xrightarrow[n\to\infty]{\mathbb{P}} 0.\end{aligned} \tag{4.3.3}$$

Since $\eta_n^d F(g^{\eta_n,x})^2(n\Delta)^{-1}\sum_{k=1}^{\lfloor sn\rfloor}\mathbb{E}[\Delta_k^n T \mid \mathscr{F}_{(k-1)\Delta}] \to sf(x)\int g(z)^2\mathrm{d}z$ in addition, we obtain $\langle M^{n,\Delta,\eta_n}(x), M^{n,\Delta,\eta_n}(x)\rangle_s \to 0$ in probability as $n \to \infty$. □

Proof (of Theorem 4.2.7). For $n \in \mathbb{N}^*$ and $x \neq 0$, let $U^{n,\Delta,\eta}(x)$ be given by

$$U_s^{n,\Delta,\eta}(x) := \sqrt{n\Delta\eta^d}\Big(G_s^{n,\Delta,\eta}(x) - sf(x)\Big).$$

We decompose the process $U^{n,\Delta,\eta}(x)$ as follows:

$$U_s^{n,\Delta,\eta}(x) = \sqrt{n\Delta\eta^d}\Big(H_s^{n,\Delta,\eta}(x) - sf(x)\Big) + \sqrt{n\Delta\eta^d}H_s'^{n,\Delta,\eta}(x) + \sqrt{n\Delta\eta^d}M_s^{n,\Delta,\eta}(x).$$

We prove that $(n\Delta\eta_n^d)^{1/2}(H_s^{n,\Delta,\eta_n}(x) - sf(x)) \to s\gamma(x)$ and $(n\Delta\eta_n^d)^{1/2}H_s'^{n,\Delta,\eta_n}(x) \to 0$ in probability as well as $((n\Delta\eta_n^d)^{1/2}M^{n,\Delta,\eta_n}(x_i))_{i\in I} \to (\sigma(x_i)W^i)_{i\in I}$ stably in law as $n \to \infty$.

(H) Immediately, we have

$$\sqrt{n\Delta\eta_n^d}\Big(H_s^{n,\Delta,\eta_n}(x) - sf(x)\Big) = s\sqrt{n\Delta\eta_n^d}\Big(Fg^{\eta_n,x} - f(x)\Big) + \sqrt{n\Delta\eta_n^d}\left(\frac{T_{\lfloor sn\rfloor\Delta}}{n\Delta} - s\right).$$

Under Assumption 4.2.2, by Taylor's theorem – in the case $\alpha \in \mathbb{N}^*$ –

$$s\sqrt{n\Delta\eta_n^d}\Big(Fg^{\eta_n,x} - f(x)\Big) = s\sqrt{n\Delta\eta_n^d}\left(\eta_n^\alpha \sum_{|m|=\alpha} \frac{\kappa_m(g)}{m!}\frac{\partial^m}{\partial x^m}f(x) + o(\eta_n^\alpha)\right)$$
$$\xrightarrow[n\to\infty]{} s\gamma(x)$$

holds as g is of order α and – in the case $\zeta = 0$ in eq. (4.2.5) – there is some constant $\zeta' < \infty$ such that

$$\left|s\sqrt{n\Delta\eta_n^d}\Big(Fg^{\eta_n,x} - f(x)\Big)\right| \le \zeta' s\sqrt{n\Delta\eta_n^d}\eta_n^\alpha \xrightarrow[n\to\infty]{} 0.$$

In addition, we directly obtain

$$\sqrt{\eta_n^d}\cdot\sqrt{n\Delta}\,\Big((n\Delta)^{-1}T_{\lfloor sn\rfloor\Delta} - s\Big) \xrightarrow[n\to\infty]{\mathbb{P}} 0$$

under Assumption 4.2.3. ⋄

(H') By step (H') in the proof of Theorem 4.2.6, we directly obtain

$$\sqrt{n\Delta\eta_n^d}H_s'^{n,\Delta,\eta_n}(x) \xrightarrow[n\to\infty]{\mathbb{P}} 0$$

since $n\Delta^2\eta_n^d \to 0$ and $n\Delta^3\eta_n^{-d-4} \to 0$ by eq. (4.2.6b). ⋄

(M) Since $n\Delta\eta_n^d \to \infty$ and $\eta_n \to 0$ by eq. (4.2.4), we have

$$\sup_{r\le s}\sqrt{n\Delta\eta_n^d}\left|\Delta M_r^{n,\Delta,\eta_n}(x_i)\right| \xrightarrow[n\to\infty]{\mathbb{P}} 0$$

for all $i \in I$. Let $N^{n,\Delta,\eta}$ and $N'^{n,\Delta,\eta}$ be the $\mathscr{H}_s^n$-martingales given by

$$N_s^{n,\Delta,\eta}(x) := \sqrt{\frac{\eta^d}{n\Delta}}\sum_{k=1}^{\lfloor sn\rfloor} g^{\eta,x}(\Delta_k^n X) - \mathbb{E}\left[g^{\eta,x}(\Delta_k^n X) \mid \mathscr{F}_{(k-1)\Delta}\right]$$
$$N_s'^{n,\Delta,\eta}(x) := \sqrt{\frac{\eta^d}{n\Delta}}Fg^{\eta,x}\sum_{k=1}^{\lfloor sn\rfloor}\int_{(k-1)\Delta}^{k\Delta} Y_r\mathrm{d}r - \mathbb{E}\left[\int_{(k-1)\Delta}^{k\Delta} Y_r\mathrm{d}r \,\middle|\, \mathscr{F}_{(k-1)\Delta}\right].$$

We note that $(n\Delta\eta_n^d)^{1/2}M_s^{n,\Delta,\eta}(x) = N_s^{n,\Delta,\eta}(x) - N_s'^{n,\Delta,\eta}(x)$ and, for all $s \geq 0$, we prove

$$\left\langle N'^{n,\Delta,\eta_n}(x), N'^{n,\Delta,\eta_n}(x)\right\rangle_s \xrightarrow[n\to\infty]{\mathbb{P}} 0,$$
$$\left\langle N^{n,\Delta,\eta_n}(x_i), N^{n,\Delta,\eta_n}(x_j)\right\rangle_s \xrightarrow[n\to\infty]{\mathbb{P}} s\sigma(x_i)^2\delta_{ij}.$$

In analogy to eq. (4.3.3), we observe

$$\left\langle N'^{n,\Delta,\eta_n}(x), N'^{n,\Delta,\eta_n}(x)\right\rangle_s \leq$$
$$\Delta\eta_n^d(Fg^{\eta_n,x})^2 \cdot \frac{1}{n\Delta}\sum_{k=1}^{\lfloor sn\rfloor}\mathbb{E}\left[\int_{(k-1)\Delta}^{k\Delta} Y_r^2\mathrm{d}r \,\middle|\, \mathscr{F}_{(k-1)\Delta}\right] \xrightarrow[n\to\infty]{\mathbb{P}} 0.$$

This also implies

$$\frac{\eta_n^d}{n\Delta}\sum_{k=1}^{\lfloor sn\rfloor}\mathbb{E}\left[\Delta_k^n T \,\middle|\, \mathscr{F}_{(k-1)\Delta}\right]^2 \xrightarrow[n\to\infty]{\mathbb{P}} 0 \tag{4.3.4}$$

Next, we have

$$\left\langle N^{n,\Delta,\eta_n}(x_i), N^{n,\Delta,\eta_n}(x_j)\right\rangle_s =$$
$$\frac{\eta_n^d}{n\Delta}\sum_{k=1}^{\lfloor sn\rfloor}\mathbb{E}\left[g^{\eta_n,x}(\Delta_k^n X)^2 \,\middle|\, \mathscr{F}_{(k-1)\Delta}\right] - \mathbb{E}\left[g^{\eta_n,x}(\Delta_k^n X) \,\middle|\, \mathscr{F}_{(k-1)\Delta}\right]^2.$$

We observe that

$$\frac{\eta_n^d}{n\Delta}\sum_{k=1}^{\lfloor sn\rfloor}\mathbb{E}\left[g^{\eta_n,x}(\Delta_k^n X) \,\middle|\, \mathscr{F}_{(k-1)\Delta}\right]^2 \leq \tag{4.3.5a}$$
$$\frac{4\eta_n^d}{n\Delta}\sum_{k=1}^{\lfloor sn\rfloor}(Fg^{\eta,x})^2\,\mathbb{E}\left[\Delta_k^n T \,\middle|\, \mathscr{F}_{(k-1)\Delta}\right]^2 \tag{4.3.5b}$$
$$+\frac{4\eta_n^d}{n\Delta}\sum_{k=1}^{\lfloor sn\rfloor}\mathbb{E}\left[g^{\eta,x}(\Delta_k^n X) - Fg^{\eta,x}\Delta_k^n T \,\middle|\, \mathscr{F}_{(k-1)\Delta}\right]^2, \tag{4.3.5c}$$

where the summand in eq. (4.3.5b) goes to zero by eq. (4.3.4), and – recall eq. (4.3.1) – the summand in eq. (4.3.5c) goes to zero in analogy to step (H') in the proof of Theorem 4.2.6.

Finally, we note that $g^{\eta,x}g^{\eta,y} \equiv 0$ for all η small enough if $x \neq y$, and recall that $\eta^d F(g^{\eta,x})^2 \to f(x) \int g(z)^2 \mathrm{d}z$ as $\eta \to 0$. Consequently, we deduce in combination with eq. (4.3.2) that

$$\frac{\eta_n^d}{n\Delta} \sum_{k=1}^{\lfloor sn \rfloor} \mathbb{E}\left[g^{\eta,x_i} g^{\eta,x_j}(\Delta_k^n X) \,\middle|\, \mathscr{F}_{(k-1)\Delta} \right] \xrightarrow[n\to\infty]{\mathbb{P}} \begin{cases} s f(x_i) \int g(z)^2 \mathrm{d}z & \text{if } i = j, \\ 0, & \text{else.} \end{cases} \tag{4.3.6}$$

By the martingale limit theorem 2.3.10, therefore,

$$\left(\sqrt{n\Delta\eta_n^d} M^{n,\Delta,\eta_n}(x_i) \right)_{i\in I} \overset{\mathscr{L}}{\underset{n\to\infty}{\Longrightarrow}} \left(\sigma(x_i) W^i \right)_{i\in I},$$

where $\sigma(x)^2 = f(x) \int g(z)^2 \mathrm{d}z$. It remains to prove that this convergence holds stably in law. This, however, follows in analogy to the proof of Lemma 3.4.11 for instance. □

Proof (of Corollary 4.2.8). It remains to show that $(n\Delta\eta_n^d)^{1/2}\hat{\gamma}_n^{\Delta,\eta_n}(x) - \gamma(x) \to 0$ in probability as $n \to \infty$. This, however, follows in analogy to the proof of Corollary 3.3.8. □

The estimation of jumps in practice: Simulation studies and the empirical modelling of intermittency

5 Simulation studies

The theoretical results developed in the previous chapters are asymptotical. We dedicate this chapter to investigate on the performance of the various estimators in practice. We illustrate, inter alia, the influence of discretisation and the importance of suitable bandwidth selection. In Section 5.1, we study the kernel density estimator for the Lévy kernel of a Markovian Itō semi-martingale for an example process with finite activity, and another example process with infinite activity. Section 5.2 is based on Section 4 of Ueltzhöfer and Klüppelberg (2011). Within each section, all figures and tables are put off to the end.

5.1 Markovian Itō semi-martingales

In this section, we present a simulation study for the kernel density estimators (Definitions 3.2.7 and 3.3.4) of the canonical Lévy kernel of Markovian Itō semi-martingales. We implemented numerical simulation schemes for a process with finite activity (that is, with almost surely finitely many jumps on compact time-intervals) and for a process with infinite activity. In particular, we considered the univariate Itō semi-martingales with characteristics $(B, C, \mathfrak{n})$ given by

$$\mathrm{d}B_t = -bX_t\mathrm{d}t, \quad \mathrm{d}C_t = c\mathrm{d}t, \quad \text{and} \quad \mathfrak{n}(\mathrm{d}t, \mathrm{d}y) = f(X_t, y)\mathrm{d}t\mathrm{d}y, \tag{5.1.1}$$

where $b, c > 0$ and

i) the density of the Lévy kernel of the process with finite activity is a mixture of the normal density $\varphi(\,\cdot\,; 0, \sigma^2)$ with mean zero and variance $\sigma^2 > 0$ and the exponential density $\rho(\,\cdot\,;\lambda)$ with mean $1/\lambda > 0$; in particular,

$$f(x,y) := \begin{cases} \zeta\rho(y;\lambda), & \text{if } x \in \,]-\infty,-\xi], \\ \zeta[m(x)\varphi(y;0,\sigma^2) + (1-m(x))\rho(y;\lambda)] & \text{if } x \in \,]-\xi,0], \\ \zeta[m(x)\varphi(y;0,\sigma^2) + (1-m(x))\rho(-y;\lambda)] & \text{if } x \in \,]0,\xi], \\ \zeta\rho(-y;\lambda) & \text{if } x \in \,]\xi,\infty[, \end{cases} \tag{5.1.2}$$

where $m(x) = (1+\cos(\pi x/\xi))/2$;

ii) the density of the Lévy kernel of the process with infinite activity is a stable density with state-dependent intensities; in particular,

$$f(x,y) := \Big(\zeta_+(x)\mathbb{1}_{\mathbb{R}_+^*}(y) + \zeta_-(x)\mathbb{1}_{\mathbb{R}_-^*}(y)\Big)|y|^{-1-\alpha}, \tag{5.1.3}$$

where

$$\zeta_+(x) := \begin{cases} 2, & \text{if } x \in \,]-\infty,-\xi], \\ 2-(1+\cos(\pi x/\xi))/2 & \text{if } x \in \,]-\xi,0], \\ (1+\cos(\pi x/\xi))/2 & \text{if } x \in \,]0,\xi], \\ 0 & \text{if } x \in \,]\xi,\infty[, \end{cases}$$

$$\zeta_-(x) := 2 - \zeta_+(x).$$

Moreover, we have implemented the kernel density estimators $\hat{f}_t^\eta(x,y)$ based on the sample-path $\{X_s(\omega) : s \in [0,t]\}$ and $\hat{f}_n^{\Delta,\eta}(x,y)$ based on the sample $X_0(\omega)$, $X_\Delta(\omega), \dots, X_{n\Delta}(\omega)$ using the so-called *bi-weight kernel*

$$g(z) := \frac{15}{16}\left(1-z^2\right)^2 \mathbb{1}_{[-1,1]}(z). \tag{5.1.4}$$

Its roughness is given by $\xi_g = \int g(z)^2 \mathrm{d}z = 5/7$; its second moment by $\int z^2 g(z)\mathrm{d}z = 1/7$. To calculate asymptotic confidence intervals derived from Corollaries 3.2.11 and 3.3.8 which are non-negative, we invert a test-statistic following, for instance, Hansen (2009, p. 24). Let q_α denote the α-quantile of the normal distribution, then the estimated asymptotic confidence interval of level α for $f(x,y)$ is given by

$$\left\{ z \geq 0 : \left| \sqrt{\frac{\eta_1\eta_2\Delta \sum_{k=1}^{n} g^{\eta,x}(X_{(k-1)\Delta})}{\xi_g^2 z}} \left(\hat{f}_n^{\Delta,\eta}(x,y) - z \right) \right| \leq q_\alpha \right\} \tag{5.1.5}$$

and, analogously, for the estimator $\hat{f}_t^{\eta}(x)$.

To calculate the bias corrections $\hat{\gamma}_t^{\eta}(x,y)$ and $\hat{\gamma}_n^{\Delta,\eta}(x,y)$ for our estimators, we also estimate the derivatives of f and of the density μ' of the invariant measure: For the estimation of the first-order derivatives $\partial_x f(x,y)$ and $\mathrm{d}_x\mu'(x)$, we also use the bi-weight kernel eq. (5.1.4). For the estimation of the second-order derivatives $\partial_x^2 f(x,y)$ and $\partial_y^2 f(x,y)$, however, we use the so-called *tri-weight kernel*

$$h(z) := \frac{35}{32}\left(1 - z^2\right)^3 \mathbb{1}_{[-1,1]}(z). \tag{5.1.6}$$

5.1.1 The finite activity case

Firstly, we investigated the performance of the estimator $\hat{f}_t^{\eta}(x,y)$ based on the observation of a sample path $\{X_s(\omega) : s \in [0,t]\}$. We chose the parameters of the process with finite activity as reported in Table 5.1. The restriction of the Lévy density to the set $[-4,4] \times [-5,5]$ for these parameters is presented in Figure 5.1. We emphasise the discontinuity on the set $\mathbb{R}^* \times \{0\}$ and that f is not twice continuously differentiable on the set $\{-\xi,\xi\} \times \mathbb{R}$, which we indicated by the red dotted lines. We investigated the scenarios

c1) $t_1 = 100$, that is, 5000 jumps on average;

c2) $t_2 = 500$, that is, 25 000 jumps on average; and

c3) $t_3 = 2500$, that is, 125 000 jumps on average.

By construction, the jump-times $T_1, T_2, \ldots$ of the process form a Poisson random measure on $\mathbb{R}_+$ with intensity $\zeta\mathrm{d}t$. Given the value X_{T_k} for some $k \in \mathbb{N}$ (with the convention $T_0 = 0$), we simulated the left-limit $X_{T_{k+1}-}$ by an Euler step over the interval $[\![T_k, T_{k+1}[\![$. Next, we drew the jump $\Delta X_{T_{k+1}}$ from the distribution with density $y \mapsto \zeta^{-1} f(X_{T_{k+1}-}, y)$. Iteratively, we obtained approximate trajectories of our process sampled at the jump-times.

For one simulated trajectory of each scenario c1–c3, we present the jumps $(X_{T_k-}, \Delta X_{T_k})$ in Figure 5.2. The shape of the Lévy density as shown in Figure 5.1 is clearly visible in all three scenarios. We note that the number of jumps with left-limit such that $|X_{T_k-}| > 4$ is small compared to the total number of jumps. As the density f is an odd function which does not change in x for $|x| > \xi = 3$, we subsequently restrict our analysis to the set $[0,3] \times \mathbb{R}^*$.

At first, we compare our estimates $\hat{f}_t^\eta(x,y)$ pointwise. For each scenario c1–c3, each bandwidth $\eta \in \{0.4, 0.8\} \times \{0.2, 0.6\}$, each $y \in \{\pm 0.2, \pm 0.6, \ldots, \pm 3, \pm 4\}$, and each $x \in \{0, 1.5, 2.5\}$, we summarised our estimation results in Tables 5.2 to 5.4. Based on 100 trajectories per scenario, we give the empirical mean and the empirical root mean squared error of our estimator, and the empirical confidence level of the estimated 95%-confidence intervals defined by eq. (5.1.5). Likewise, we also state the same empirical quantities for the bias corrected estimators $\hat{f}_t^\eta(x,y) - \hat{\gamma}_t^\eta(x,y)$.

We observe the significant influence of the bandwidth choice on the bias of the estimates. In scenario c2 (Table 5.3), for $\eta = (0.4, 0.6)$ on the one hand, we observe an empirical bias of 0.124 (resp., of 0.052; resp., of 0.011) at $(0, 2.6)$ (resp., at $(0,3)$; resp., at $(0,4)$); for $\eta = (0.8, 0.6)$ on the other hand, we observe an empirical bias of 0.169 (resp., of 0.101; resp., of 0.041) at these points. In view of eq. (3.3.5), this phenomenon was certainly expected. Moreover, we observe that there are points where the bias correction does its job: In the former case, the bias reduces to 0.008 (resp., to -0.021; resp., to 0.005); in the latter case, the bias reduces to 0.01 (resp., to -0.014; resp., to 0.005). Nevertheless, we also observe a downside of the bias correction: The empirical standard deviation of the bias corrected estimator is increased compared to the estimator itself. At points and for bandwidths where the bias is small, this increased variability actually worsens the root mean squared error of the estimator. In addition, we observe that the empirical confidence level of the estimated (pointwise) confidence intervals is satisfactory. In scenarios c1 and c2, we observe levels of 90 – 99 % for points and bandwidths where the bias is rather small, and levels of 60 – 89 % for points and bandwidths where the bias is of significant order. In scenario c3 (Table 5.4), this distinction becomes more apparent: For the bandwidth choice $\eta = (0.8, 0.6)$, the empirical levels drop to 0 % for points such as $(0, 2.6)$ or $(2.5, 1.0)$ where the bias is dominant in front of the variance. Obviously, this large bandwidth choice is not appropriate in scenario c3. The bias correction improves the results to some extent.

Next, we compare our estimates $\hat{f}_t^{\eta}$ in terms of their functional properties. Due to the drawbacks of the bias correction observed in our pointwise analysis before, we only consider the uncorrected estimator in the following. For scenario c2 (resp., c3), the bandwidths $\eta = (0.4, 0.4), (0.4, 0.6), (0.6, 0.2), (0.6, 0.4)$ (resp., $\eta = (0.1, 0.4), (0.2, 0.4), (0.4, 0.2), (0.4, 0.4)$), and the points $x = 0, 0.75, 1.5, 2.25$, we summarised our estimation results in Table 5.5 (resp., Table 5.6). Based on 100 trajectories per scenario, we present the empirical mean (integrated) squared error (MSE) of our estimator on intervals of the form $[y_1, y_2] \subseteq [-3, 0[\,\cup\,]0, 3]$; that is,

$$\int_{y_1}^{y_2} \left| \hat{f}_t^{\eta}(x,y) - f(x,y) \right|^2 \mathrm{d}y.$$

Again, we observe the significant influence of the bandwidth choice on the MSE of the estimate. The "optimal" choice within the set of presented bandwidths varies with x as well as with $[y_1, y_2]$. In scenario c2 (Table 5.5), for $x = 0.75$ on the one hand, we have that $\eta = (0.6, 0.4)$ is better than the others in terms of the MSE on $[1.5, 3]$ but $\eta = (0.4, 0.6)$ is better than the others on $[-1.5, -0.6]$. For $x = 2.25$ on the other hand, $\eta = (0.4, 0.6)$ is better than the other three bandwidths on $[-3, -0.6] \cup [0.6, 3]$. In terms of the degree of smoothing and in terms of the relative error compared to the true value of the Lévy density, we notice the following: In scenario c2, for appropriate bandwidth choices, we obtain reasonable estimates (1) at $x = 0$ on the sets $\{0.3 < |y| \leq 2\}$ and $\{0.4 < |y| \leq 3\}$, and (2) at $x = 2.25$ on the sets $\{-4 < y < -0.6\}$ and $\{0.6 < y < 2\}$. In scenario c3, again for appropriate bandwidth choices, we obtain reasonable estimates (1) at $x = 0$ on the sets $\{0.2 < |y| < 2.5\}$ and $\{0.4 < |y| < 4\}$, and (2) at $x = 2.25$ on the sets $\{-5 < y < -0.4\}$ and $\{0.4 < y < 3\}$. We present the estimates corresponding to these observations in Figures 5.3 and 5.4.

Secondly, we investigated the performance of the estimator $\hat{f}_n^{\Delta,\eta}(x,y)$ based on the observation of the discrete sample $X_0(\omega), X_\Delta(\omega), \ldots, X_{n\Delta}(\omega)$. We kept the parameters as reported in Table 5.1. We studied the scenarios

d1) $t_1 = 500$ and $\Delta_1 = 0.01$, that is, 50 000 observations;

d2) $t_2 = 2500$ and $\Delta_2 = 0.01$, that is, 250 000 observations; and

d3) $t_3 = 2500$ and $\Delta_3 = 0.001$, that is, 2 500 000 observations.

We simulated the process in analogy to before. Given the value X_{T_k} for some $k \in \mathbb{N}$, we simulated Euler steps over the intervals $[\![T_k, (\lfloor T_k/\Delta \rfloor + 1)\Delta[\![$ up to $[\![\lfloor T_{k+1}/\Delta \rfloor \Delta, T_{k+1}[\![$. Next, we drew the jump ΔX_{T_k+1} from the distribution with density $y \mapsto \zeta^{-1} f(X_{T_{k+1}-}, y)$. Finally, we only kept the sample $X_0, X_\Delta, \ldots, X_{n\Delta}$.

For one simulated sample of each scenario d1–d3, we present the increments $(X_{(k-1)\Delta}, \Delta_k^n X)$ in Figure 5.5. In comparison to Figure 5.2, the influence of discretisation is clearly visible. Subsequently, we restrict our analysis to the same sets and bandwidths as in the continuous observation case.

At first, we compare our estimates $\hat{f}_n^{\Delta,\eta}(x, y)$ pointwise. For each scenario d1–d3, each bandwidth $\eta \in \{0.4, 0.8\} \times \{0.2, 0.6\}$, each $x \in \{0, 1.5, 2.5\}$, and each $y \in \{\pm 0.2, \pm 0.6, \ldots, \pm 3, \pm 4\}$, we summarised our estimation results in Tables 5.7 to 5.9. Based on 100 samples per scenario, we give the empirical mean and the empirical root mean squared error of our estimator, and the empirical confidence level of the estimated 95%-confidence intervals defined by eq. (5.1.5). Likewise, we also state the same empirical quantities for the bias corrected estimators $\hat{f}_n^{\Delta,\eta}(x, y) - \hat{\gamma}_n^{\Delta,\eta}(x, y)$.

In scenarios d1 and d2 where $\Delta = 0.01$, the bias due to discretisation is dominant. In scenario d3 where $\Delta = 0.001$, our estimates improve significantly; the drift component is much less influential for $|y|$ large. Certainly, the influence of the continuous martingale component for $|y|$ small is still present. Since the bias correction $\hat{\gamma}_n^{\Delta,\eta}$ only captures the bias due to the kernel smoothing, there is no significant improvement observable comparing the bias corrected estimates to the uncorrected ones. In the following, we focus on the uncorrected estimates only.

Next, we compare our estimates $\hat{f}_n^{\Delta,\eta}(x, y)$ in terms of their functional properties. For scenarios d2 and d3, the bandwidths $\eta = (0.1, 0.4), (0.2, 0.4), (0.4, 0.2), (0.4, 0.4)$ and the points $x = 0, 0.75, 1.5, 2.25$, we summarised our estimation results in Tables 5.10 and 5.11. Based on 100 samples per scenario, we present the empirical mean (integrated) squared error of our estimator on intervals of the form $[y_1, y_2] \subseteq [-3, 0[\cup]0, 3]$. Again, we observe the significant influence of the bandwidth choice on the MSE of the estimate. In terms of the degree of smoothing and in terms of the relative error compared to the true value of the Lévy density, for the same bandwidth choices as in scenario c3, we observe the following: In scenario d2, we only obtain reasonable estimates (1) at $x = 0$ on the set $\{1.75 \leq |y| \leq 3\}$, and (2) at $x = 2.25$ on the sets $\{-5 < y < -3\}$ and $\{1 < y < 3\}$. In scenario d3, nevertheless, we obtain reasonable estimates (1) at $x = 0$ on the sets $\{0.75 \leq |y| \leq 2\}$ and $\{0.5 \leq |y| \leq 4\}$, and (2) at $x = 2.25$ on the sets $\{-5 < y < -0.5\}$ and $\{0.5 < y < 3\}$. We present the estimates corresponding to these observations in Figures 5.6 and 5.7.

Table 5.1: Parameters for the characteristics $(B, C, \mathfrak{n})$ given by eqs. (5.1.1) and (5.1.2)

b	c	ζ	ξ	σ^2	λ
1	1	50	3	1	2

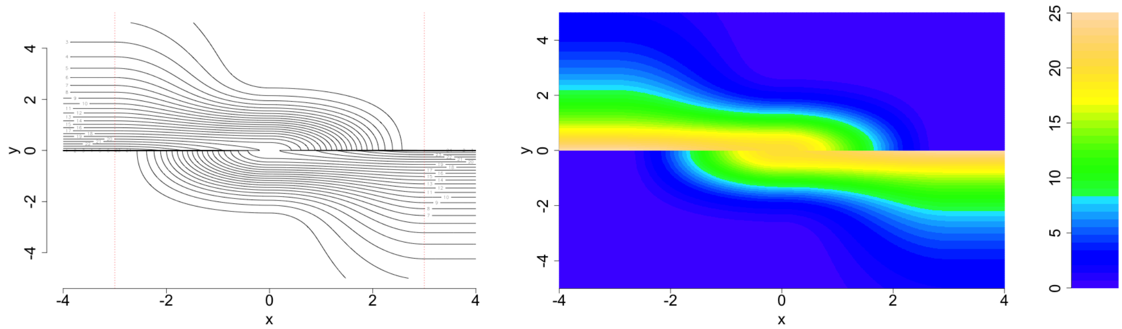

Figure 5.1: Contour plot (left) and topographical image plot (right) with legend (far right) of the restriction of the Lévy density $(x, y) \mapsto f(x, y)$ given by eq. (5.1.2) with parameters as in Table 5.1 to the set $[-4, 4] \times [-5, 5]$. The dotted red lines indicate the set $\{-\xi, \xi\} \times \mathbb{R}$ on which f is not twice continuously differentiable.

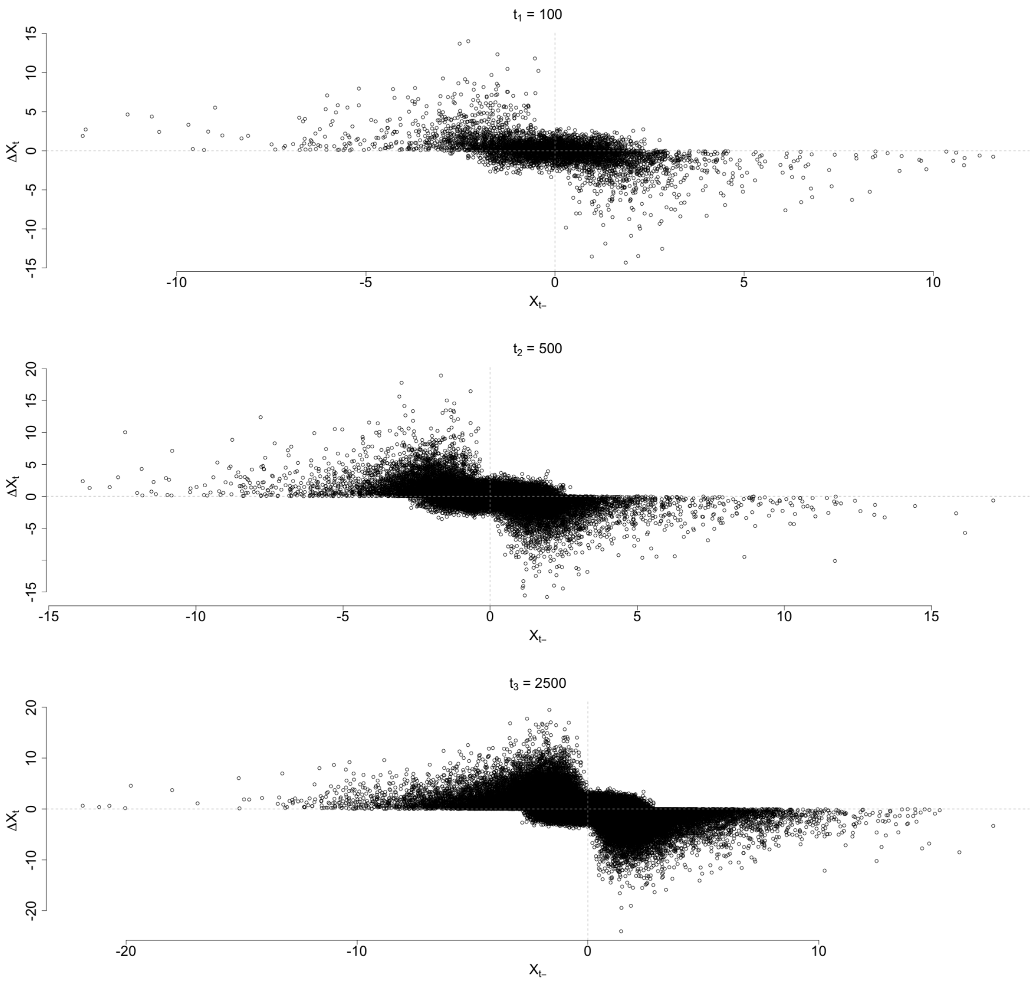

Figure 5.2: Jumps $(X_{T_k-}, \Delta X_{T_k})$ of one simulated trajectory of scenarios c1 (top), c2 (middle), and c3 (bottom).

Table 5.2: Scenario c1 — The empirical mean (columns 3, 6, 9, 12) of the estimator $\hat{f}_t^{\eta}(x,y)$ (resp., bias-corrected estimator $\hat{f}_t^{\eta}(x,y) - \hat{\gamma}_t^{\eta}(x,y)$) based on 100 trajectories (up to time $t = 100$) is compared to the true value (col. 2) of $f(x,y)$ given by eq. (5.1.2). In addition, the root mean squared error (rmse; cols. 4, 7, 10, 13) and the empirical confidence level (cl; cols. 5, 8, 11, 14) in percent of the estimated 95%-confidence interval given by eq. (5.1.5) are presented.

Estimation at $x = 0$

		$\eta = (0.4, 0.2)$			$\eta = (0.4, 0.6)$			$\eta = (0.8, 0.2)$			$\eta = (0.8, 0.6)$		
y	$f(x,y)$	mean	rmse	cl	mean	rmse	cl	mean	rmse	cl	mean	rmse	cl
−4.0	0.007	0.014	0.056	94	0.018	0.039	89	0.037	0.059	79	0.042	0.049	52
−3.0	0.222	0.236	0.268	92	0.280	0.156	91	0.266	0.199	93	0.324	0.148	77
−2.6	0.679	0.730	0.358	99	0.820	0.283	93	0.786	0.297	95	0.857	0.250	83
−2.2	1.774	1.973	0.745	93	2.024	0.486	88	1.902	0.510	94	2.011	0.375	88
−1.8	3.948	3.888	1.010	94	4.170	0.603	95	3.939	0.729	93	4.178	0.452	92
−1.4	7.486	7.420	1.273	97	7.597	0.778	96	7.414	0.945	95	7.585	0.521	95
−1.0	12.099	11.922	1.821	93	11.969	1.060	94	11.826	1.297	92	11.878	0.702	96
−0.6	16.661	16.632	2.277	92	16.362	1.221	96	16.434	1.614	94	16.208	0.955	91
−0.2	19.552	19.249	2.011	98	—	—	—	19.315	1.538	97	—	—	—
0.2	19.552	19.447	2.275	95	—	—	—	19.317	1.614	94	—	—	—
0.6	16.661	16.536	2.133	94	16.420	1.214	95	16.387	1.633	91	16.270	0.974	89
1.0	12.099	11.876	1.742	93	12.012	0.951	97	12.009	1.319	94	11.967	0.743	93
1.4	7.486	7.299	1.603	92	7.542	0.851	95	7.310	1.141	93	7.557	0.614	95
1.8	3.948	4.053	0.999	94	4.170	0.571	97	4.071	0.705	97	4.188	0.451	92
2.2	1.774	1.911	0.679	97	2.014	0.410	95	1.862	0.452	95	2.010	0.341	90
2.6	0.679	0.698	0.422	93	0.773	0.249	93	0.774	0.360	91	0.831	0.240	82
3.0	0.222	0.248	0.272	92	0.273	0.153	93	0.286	0.204	90	0.328	0.155	79
4.0	0.007	0.013	0.054	94	0.018	0.034	87	0.037	0.060	74	0.045	0.052	48

Estimation at $x = 0$ with bias correction

		$\eta = (0.4, 0.2)$			$\eta = (0.4, 0.6)$			$\eta = (0.8, 0.2)$			$\eta = (0.8, 0.6)$		
y	$f(x,y)$	mean	rmse	cl	mean	rmse	cl	mean	rmse	cl	mean	rmse	cl
−4.0	0.007	0.017	0.088	95	0.017	0.056	92	0.021	0.066	90	0.018	0.035	85
−3.0	0.222	0.249	0.397	88	0.220	0.227	88	0.232	0.293	89	0.215	0.171	74
−2.6	0.679	0.717	0.546	90	0.688	0.365	82	0.724	0.424	84	0.702	0.276	80
−2.2	1.774	2.100	1.168	75	1.941	0.676	73	1.886	0.794	75	1.820	0.476	72
−1.8	3.948	3.842	1.565	82	3.951	0.942	80	3.858	1.183	82	3.946	0.643	82
−1.4	7.486	7.422	2.090	80	7.468	1.289	74	7.373	1.440	82	7.484	0.856	79
−1.0	12.099	11.907	2.782	80	11.959	1.796	68	11.934	1.971	84	11.944	1.177	79
−0.6	16.661	16.638	3.676	70	16.726	1.913	80	16.619	2.690	68	16.649	1.409	75
−0.2	19.552	19.013	3.205	84	—	—	—	19.455	2.310	83	—	—	—
0.2	19.552	19.380	3.667	78	—	—	—	19.564	2.462	82	—	—	—
0.6	16.661	16.488	3.206	80	16.747	1.892	82	16.566	2.529	75	16.758	1.385	82
1.0	12.099	11.673	2.781	79	12.045	1.443	84	12.086	2.054	79	12.106	1.100	81
1.4	7.486	7.282	2.420	74	7.347	1.365	75	7.244	1.793	75	7.350	1.026	70
1.8	3.948	3.995	1.626	81	3.969	0.897	79	4.027	1.128	81	3.989	0.594	85
2.2	1.774	1.987	1.143	78	1.879	0.580	82	1.817	0.731	80	1.792	0.409	82
2.6	0.679	0.695	0.606	89	0.621	0.370	80	0.707	0.539	76	0.655	0.294	77
3.0	0.222	0.288	0.457	88	0.217	0.226	90	0.266	0.300	87	0.219	0.172	72
4.0	0.007	0.015	0.079	96	0.014	0.042	89	0.024	0.068	88	0.018	0.035	83

Table 5.2a: Scenario c1 (continued)

Estimation at $x = 1.5$

		$\eta = (0.4, 0.2)$			$\eta = (0.4, 0.6)$			$\eta = (0.8, 0.2)$			$\eta = (0.8, 0.6)$		
y	$f(x,y)$	mean	rmse	cl	mean	rmse	cl	mean	rmse	cl	mean	rmse	cl
−4.0	1.695	1.664	0.818	96	1.642	0.496	95	1.625	0.582	98	1.605	0.369	94
−3.0	2.900	2.918	1.004	99	2.948	0.597	98	2.811	0.707	97	2.817	0.407	97
−2.6	3.746	3.887	1.144	95	3.878	0.695	96	3.650	0.851	97	3.690	0.530	97
−2.2	5.048	5.216	1.652	92	5.164	0.847	93	4.954	1.145	90	4.973	0.605	95
−1.8	7.056	6.954	1.790	96	7.256	1.014	95	6.791	1.165	95	7.018	0.705	93
−1.4	9.951	10.236	2.119	93	10.205	1.279	95	9.942	1.471	93	10.006	0.898	95
−1.0	13.631	13.200	2.256	99	13.472	1.283	96	13.368	1.715	96	13.495	0.958	95
−0.6	17.591	17.750	2.444	98	17.565	1.395	96	17.711	1.733	95	17.529	1.072	96
−0.2	21.087	20.907	3.022	95	—	—	—	20.818	2.241	94	—	—	—
0.2	9.776	9.778	2.050	95	—	—	—	10.282	1.503	95	—	—	—
0.6	8.331	8.316	1.709	96	8.235	1.020	97	8.711	1.243	95	8.627	0.809	94
1.0	6.049	6.445	1.425	97	6.168	0.857	97	6.524	1.112	96	6.399	0.678	91
1.4	3.743	3.788	1.348	93	3.931	0.788	93	4.019	0.930	92	4.086	0.608	90
1.8	1.974	2.155	0.897	96	2.225	0.583	92	2.206	0.711	95	2.278	0.482	89
2.2	0.887	1.033	0.637	92	1.077	0.386	94	1.010	0.471	90	1.078	0.313	90
2.6	0.340	0.367	0.411	90	0.420	0.250	92	0.384	0.280	93	0.427	0.186	94
3.0	0.111	0.130	0.234	93	0.149	0.142	92	0.132	0.170	93	0.148	0.102	91
4.0	0.003	0.006	0.062	99	0.004	0.026	98	0.004	0.034	98	0.003	0.014	97

Estimation at $x = 1.5$ with bias correction

		$\eta = (0.4, 0.2)$			$\eta = (0.4, 0.6)$			$\eta = (0.8, 0.2)$			$\eta = (0.8, 0.6)$		
y	$f(x,y)$	mean	rmse	cl	mean	rmse	cl	mean	rmse	cl	mean	rmse	cl
−4.0	1.695	1.748	1.226	85	1.607	0.753	84	1.701	0.920	78	1.657	0.566	75
−3.0	2.900	2.963	1.675	82	2.946	0.971	79	2.995	1.155	82	2.922	0.629	85
−2.6	3.746	4.033	1.911	82	3.849	1.121	80	3.892	1.266	81	3.815	0.796	77
−2.2	5.048	5.353	2.637	74	5.093	1.359	80	5.225	1.881	76	5.081	0.951	81
−1.8	7.056	6.908	2.814	75	7.161	1.595	82	6.998	1.832	82	7.067	1.162	77
−1.4	9.951	10.388	3.190	82	10.353	2.083	78	10.124	2.436	77	10.157	1.495	72
−1.0	13.631	13.088	3.546	82	13.177	2.079	79	13.372	2.591	82	13.380	1.507	83
−0.6	17.591	17.581	4.052	82	17.722	2.332	79	17.842	2.760	82	17.810	1.637	80
−0.2	21.087	20.898	4.839	80	—	—	—	20.753	3.631	82	—	—	—
0.2	9.776	9.477	3.415	78	—	—	—	9.715	2.377	77	—	—	—
0.6	8.331	8.092	2.907	81	8.145	1.637	81	8.212	1.986	77	8.262	1.231	77
1.0	6.049	6.617	2.372	83	6.172	1.427	75	6.396	1.624	82	6.096	0.988	82
1.4	3.743	3.742	2.072	81	3.757	1.231	75	3.788	1.487	82	3.772	0.847	80
1.8	1.974	2.088	1.415	84	2.127	0.840	83	2.115	1.113	77	2.103	0.614	76
2.2	0.887	1.124	1.025	76	1.016	0.570	80	1.006	0.731	75	0.953	0.401	79
2.6	0.340	0.422	0.638	84	0.377	0.357	85	0.417	0.449	83	0.348	0.252	79
3.0	0.111	0.143	0.351	89	0.147	0.198	86	0.150	0.265	82	0.128	0.152	86
4.0	0.003	0.011	0.107	99	0.006	0.045	98	0.006	0.060	99	0.003	0.025	98

Table 5.2b: Scenario c1 (continued)

Estimation at $x = 2.5$

		$\eta = (0.4, 0.2)$			$\eta = (0.4, 0.6)$			$\eta = (0.8, 0.2)$			$\eta = (0.8, 0.6)$		
y	$f(x,y)$	mean	rmse	cl	mean	rmse	cl	mean	rmse	cl	mean	rmse	cl
−4.0	3.157	3.016	1.600	98	3.124	0.981	96	2.988	1.143	98	2.976	0.683	98
−3.0	5.219	4.952	1.955	98	5.115	1.279	95	5.041	1.482	99	5.075	0.911	95
−2.6	6.402	6.015	2.563	94	6.086	1.420	98	5.956	1.715	96	6.017	0.967	98
−2.2	7.883	7.505	2.750	98	7.658	1.733	96	7.341	1.804	97	7.471	1.150	96
−1.8	9.748	9.993	3.396	97	9.741	1.988	95	9.456	2.345	97	9.460	1.347	96
−1.4	12.085	11.983	3.217	98	11.932	1.957	94	11.810	2.368	96	11.854	1.342	95
−1.0	14.958	14.588	3.886	99	14.770	2.074	99	14.615	2.756	96	14.726	1.549	96
−0.6	18.396	17.922	4.051	98	18.300	2.534	96	18.251	2.887	95	18.285	1.843	97
−0.2	22.415	22.234	4.971	95	—	—	—	22.050	3.290	97	—	—	—
0.2	1.310	1.535	1.285	93	—	—	—	2.188	1.326	79	—	—	—
0.6	1.116	1.242	1.117	92	1.312	0.677	94	1.859	1.077	83	1.858	0.881	60
1.0	0.810	0.983	0.961	92	1.020	0.644	90	1.330	0.904	84	1.400	0.754	66
1.4	0.501	0.673	0.846	92	0.636	0.489	92	0.922	0.767	85	0.893	0.520	69
1.8	0.264	0.309	0.549	96	0.335	0.318	98	0.477	0.450	91	0.487	0.332	80
2.2	0.119	0.159	0.355	90	0.163	0.212	89	0.238	0.327	87	0.244	0.212	81
2.6	0.045	0.102	0.337	93	0.081	0.151	88	0.105	0.229	86	0.106	0.130	81
3.0	0.015	0.017	0.111	98	0.024	0.071	93	0.034	0.122	93	0.034	0.065	88
4.0	0.000	0.000	0.000	100	0.001	0.005	99	0.000	0.000	100	0.000	0.003	99

Estimation at $x = 2.5$ with bias correction

		$\eta = (0.4, 0.2)$			$\eta = (0.4, 0.6)$			$\eta = (0.8, 0.2)$			$\eta = (0.8, 0.6)$		
y	$f(x,y)$	mean	rmse	cl	mean	rmse	cl	mean	rmse	cl	mean	rmse	cl
−4.0	3.157	2.922	2.561	94	3.229	1.686	79	2.992	1.839	84	3.119	1.150	82
−3.0	5.219	4.729	3.176	83	4.946	1.982	81	5.176	2.300	87	5.289	1.389	82
−2.6	6.402	6.037	4.157	79	5.972	2.349	75	6.205	2.717	82	6.182	1.440	89
−2.2	7.883	7.591	4.308	82	7.541	2.927	77	7.674	2.949	84	7.672	1.772	83
−1.8	9.748	10.358	5.154	78	9.849	3.272	75	9.899	3.881	76	9.725	2.294	72
−1.4	12.085	11.777	5.257	80	11.714	3.167	78	11.791	3.761	83	11.972	2.094	81
−1.0	14.958	14.455	6.391	75	14.524	3.225	81	14.826	4.727	75	14.698	2.511	77
−0.6	18.396	17.613	6.520	82	18.176	4.021	75	18.040	4.924	77	18.210	2.950	73
−0.2	22.415	21.992	7.958	76	—	—	—	21.927	5.578	75	—	—	—
0.2	1.310	1.481	1.896	88	—	—	—	1.491	1.575	79	—	—	—
0.6	1.116	1.219	1.864	89	1.185	1.021	89	1.199	1.223	87	1.166	0.770	72
1.0	0.810	0.926	1.370	86	0.934	0.879	84	0.956	1.022	81	0.925	0.764	65
1.4	0.501	0.694	1.285	83	0.592	0.748	86	0.794	1.008	82	0.612	0.554	82
1.8	0.264	0.314	0.823	90	0.303	0.438	91	0.381	0.614	83	0.318	0.371	85
2.2	0.119	0.171	0.498	91	0.159	0.305	89	0.243	0.477	82	0.177	0.237	83
2.6	0.045	0.144	0.577	94	0.088	0.238	88	0.142	0.414	87	0.098	0.173	80
3.0	0.015	0.016	0.120	98	0.018	0.077	95	0.043	0.197	94	0.031	0.091	91
4.0	0.000	0.000	0.000	100	0.000	0.000	100	0.000	0.000	100	0.000	0.000	100

Table 5.3: Scenario c2 — The empirical mean (columns 3, 6, 9, 12) of the estimator $\hat{f}_t^{\eta}(x,y)$ (resp., bias-corrected estimator $\hat{f}_t^{\eta}(x,y) - \hat{\gamma}_t^{\eta}(x,y)$) based on 100 trajectories (up to time $t = 500$) is compared to the true value (col. 2) of $f(x,y)$ given by eq. (5.1.2). In addition, the root mean squared error (rmse; cols. 4, 7, 10, 13) and the empirical confidence level (cl; cols. 5, 8, 11, 14) in percent of the estimated 95%-confidence interval given by eq. (5.1.5) are presented.

Estimation at $x = 0$

		$\eta = (0.4, 0.2)$			$\eta = (0.4, 0.6)$			$\eta = (0.8, 0.2)$			$\eta = (0.8, 0.6)$		
y	$f(x,y)$	mean	rmse	cl	mean	rmse	cl	mean	rmse	cl	mean	rmse	cl
−4.0	0.007	0.023	0.034	76	0.023	0.022	65	0.048	0.051	34	0.049	0.046	7
−3.0	0.222	0.233	0.111	96	0.284	0.093	80	0.275	0.090	92	0.324	0.116	39
−2.6	0.679	0.726	0.222	88	0.804	0.170	82	0.764	0.177	85	0.849	0.190	38
−2.2	1.774	1.786	0.296	97	1.946	0.256	81	1.852	0.219	97	1.996	0.257	59
−1.8	3.948	3.965	0.456	96	4.160	0.335	84	3.982	0.325	92	4.175	0.293	83
−1.4	7.486	7.515	0.620	96	7.668	0.419	90	7.483	0.435	95	7.636	0.300	94
−1.0	12.099	12.122	0.713	99	12.114	0.414	98	12.071	0.527	97	12.041	0.305	95
−0.6	16.661	16.539	0.958	95	16.306	0.607	91	16.437	0.722	92	16.211	0.567	80
−0.2	19.552	19.433	0.915	98	—	—	—	19.334	0.683	98	—	—	—
0.2	19.552	19.487	0.902	98	—	—	—	19.325	0.695	95	—	—	—
0.6	16.661	16.722	0.936	95	16.421	0.600	92	16.584	0.698	95	16.295	0.533	88
1.0	12.099	12.021	0.806	96	12.060	0.459	92	12.023	0.549	98	12.015	0.323	96
1.4	7.486	7.546	0.549	98	7.653	0.394	94	7.520	0.418	97	7.644	0.308	90
1.8	3.948	4.008	0.424	96	4.199	0.358	85	4.002	0.307	96	4.204	0.318	68
2.2	1.774	1.849	0.294	96	1.981	0.267	79	1.881	0.230	94	2.008	0.262	55
2.6	0.679	0.723	0.185	95	0.803	0.170	81	0.766	0.162	88	0.848	0.190	44
3.0	0.222	0.213	0.107	97	0.274	0.085	84	0.273	0.094	88	0.323	0.112	43
4.0	0.007	0.019	0.030	81	0.018	0.019	75	0.048	0.052	36	0.048	0.044	3

Estimation at $x = 0$ with bias correction

		$\eta = (0.4, 0.2)$			$\eta = (0.4, 0.6)$			$\eta = (0.8, 0.2)$			$\eta = (0.8, 0.6)$		
y	$f(x,y)$	mean	rmse	cl	mean	rmse	cl	mean	rmse	cl	mean	rmse	cl
−4.0	0.007	0.022	0.045	79	0.017	0.024	70	0.026	0.040	66	0.016	0.022	73
−3.0	0.222	0.220	0.177	78	0.218	0.107	74	0.215	0.121	79	0.219	0.079	75
−2.6	0.679	0.714	0.346	73	0.697	0.177	80	0.702	0.246	74	0.696	0.139	73
−2.2	1.774	1.740	0.469	78	1.758	0.291	78	1.769	0.325	80	1.764	0.206	78
−1.8	3.948	3.946	0.736	77	3.951	0.419	82	3.909	0.494	80	3.924	0.288	81
−1.4	7.486	7.533	0.962	80	7.515	0.615	75	7.487	0.678	79	7.484	0.408	79
−1.0	12.099	12.153	1.134	85	12.196	0.644	85	12.163	0.897	77	12.190	0.512	78
−0.6	16.661	16.612	1.463	81	16.573	0.805	82	16.699	1.068	78	16.601	0.580	82
−0.2	19.552	19.468	1.526	79	—	—	—	19.514	1.079	84	—	—	—
0.2	19.552	19.606	1.445	84	—	—	—	19.550	1.014	83	—	—	—
0.6	16.661	16.698	1.563	75	16.709	0.909	73	16.849	1.161	81	16.758	0.655	82
1.0	12.099	11.973	1.259	78	12.023	0.745	80	12.095	0.861	78	12.101	0.512	77
1.4	7.486	7.607	0.918	79	7.504	0.567	82	7.545	0.623	90	7.491	0.412	73
1.8	3.948	4.001	0.704	84	3.992	0.395	83	3.973	0.488	82	3.968	0.309	77
2.2	1.774	1.843	0.486	76	1.804	0.272	81	1.836	0.336	82	1.803	0.192	79
2.6	0.679	0.698	0.280	82	0.687	0.178	76	0.698	0.212	80	0.689	0.137	79
3.0	0.222	0.188	0.167	75	0.201	0.113	70	0.195	0.127	78	0.208	0.080	69
4.0	0.007	0.018	0.040	83	0.012	0.021	84	0.022	0.035	72	0.012	0.018	79

Table 5.3a: Scenario c2 (continued)

Estimation at $x = 1.5$

y	$f(x,y)$	$\eta=(0.4,0.2)$ mean	rmse	cl	$\eta=(0.4,0.6)$ mean	rmse	cl	$\eta=(0.8,0.2)$ mean	rmse	cl	$\eta=(0.8,0.6)$ mean	rmse	cl
−4.0	1.695	1.668	0.383	94	1.688	0.225	93	1.590	0.285	92	1.610	0.180	91
−3.0	2.900	2.915	0.480	96	2.899	0.293	94	2.799	0.353	93	2.787	0.227	93
−2.6	3.746	3.696	0.534	95	3.769	0.345	94	3.570	0.401	94	3.630	0.266	89
−2.2	5.048	5.048	0.577	98	5.152	0.388	93	4.859	0.449	92	4.976	0.273	94
−1.8	7.056	7.109	0.847	95	7.197	0.506	92	6.903	0.576	96	7.028	0.313	97
−1.4	9.951	9.896	0.811	98	10.051	0.545	96	9.789	0.585	96	9.922	0.360	95
−1.0	13.631	13.811	1.092	93	13.694	0.587	97	13.679	0.713	97	13.588	0.391	95
−0.6	17.591	17.706	1.184	98	17.565	0.618	98	17.513	0.857	94	17.467	0.454	97
−0.2	21.087	21.081	1.304	96	—	—	—	21.002	0.934	98	—	—	—
0.2	9.776	9.835	0.915	96	—	—	—	10.338	0.843	87	—	—	—
0.6	8.331	8.380	0.798	93	8.290	0.411	96	8.778	0.725	90	8.659	0.433	89
1.0	6.049	6.136	0.712	96	6.122	0.369	97	6.392	0.600	92	6.380	0.418	82
1.4	3.743	3.877	0.570	93	3.913	0.353	92	4.004	0.463	88	4.066	0.390	76
1.8	1.974	2.014	0.412	94	2.127	0.299	86	2.104	0.309	94	2.224	0.306	70
2.2	0.887	0.930	0.289	91	0.995	0.210	86	0.976	0.228	87	1.043	0.202	75
2.6	0.340	0.335	0.155	98	0.392	0.105	91	0.366	0.111	95	0.414	0.103	78
3.0	0.111	0.123	0.099	93	0.132	0.053	95	0.123	0.072	94	0.140	0.049	86
4.0	0.003	0.002	0.008	99	0.005	0.012	91	0.003	0.010	95	0.005	0.008	91

Estimation at $x = 1.5$ with bias correction

y	$f(x,y)$	$\eta=(0.4,0.2)$ mean	rmse	cl	$\eta=(0.4,0.6)$ mean	rmse	cl	$\eta=(0.8,0.2)$ mean	rmse	cl	$\eta=(0.8,0.6)$ mean	rmse	cl
−4.0	1.695	1.722	0.626	78	1.711	0.367	70	1.666	0.421	82	1.701	0.243	84
−3.0	2.900	3.000	0.765	75	2.896	0.468	78	3.012	0.557	77	2.906	0.322	82
−2.6	3.746	3.744	0.895	77	3.734	0.537	76	3.744	0.585	81	3.723	0.369	81
−2.2	5.048	5.108	0.958	85	5.119	0.552	84	5.020	0.699	82	5.049	0.398	83
−1.8	7.056	7.194	1.414	68	7.144	0.835	68	7.110	0.943	73	7.070	0.525	76
−1.4	9.951	9.971	1.325	81	9.936	0.872	74	9.907	0.912	85	9.896	0.564	77
−1.0	13.631	13.846	1.764	77	13.689	0.985	82	13.893	1.189	82	13.686	0.633	81
−0.6	17.591	17.923	1.935	82	17.642	1.032	82	17.696	1.426	77	17.634	0.727	81
−0.2	21.087	21.174	2.033	79	—	—	—	21.094	1.443	78	—	—	—
0.2	9.776	9.530	1.456	73	—	—	—	9.713	1.047	74	—	—	—
0.6	8.331	8.302	1.295	79	8.303	0.705	87	8.320	0.918	83	8.316	0.477	83
1.0	6.049	6.053	1.181	76	6.047	0.620	81	6.051	0.773	81	6.046	0.441	81
1.4	3.743	3.890	0.934	75	3.793	0.500	80	3.797	0.689	77	3.780	0.365	82
1.8	1.974	1.979	0.701	78	1.973	0.401	77	1.982	0.455	82	1.989	0.270	76
2.2	0.887	0.888	0.438	78	0.892	0.280	74	0.908	0.325	75	0.899	0.199	74
2.6	0.340	0.305	0.261	80	0.319	0.137	84	0.331	0.177	83	0.329	0.104	80
3.0	0.111	0.135	0.150	84	0.108	0.074	86	0.127	0.114	84	0.103	0.061	75
4.0	0.003	0.000	0.005	99	0.005	0.016	93	0.003	0.014	96	0.005	0.011	85

Table 5.3b: Scenario c2 (continued)

Estimation at $x = 2.5$

		$\eta = (0.4, 0.2)$			$\eta = (0.4, 0.6)$			$\eta = (0.8, 0.2)$			$\eta = (0.8, 0.6)$		
y	$f(x,y)$	mean	rmse	cl	mean	rmse	cl	mean	rmse	cl	mean	rmse	cl
−4.0	3.157	3.021	0.731	99	3.066	0.442	96	2.969	0.519	98	2.986	0.345	95
−3.0	5.219	5.276	1.011	94	5.281	0.621	94	5.030	0.706	96	5.028	0.470	93
−2.6	6.402	6.247	1.199	94	6.395	0.739	96	6.007	0.893	91	6.125	0.580	88
−2.2	7.883	7.809	1.274	95	7.780	0.691	98	7.562	0.947	91	7.562	0.601	89
−1.8	9.748	9.498	1.524	92	9.600	0.877	89	9.286	1.068	91	9.432	0.656	91
−1.4	12.085	11.956	1.518	96	12.044	0.964	92	11.764	0.994	98	11.874	0.690	95
−1.0	14.958	14.828	1.946	94	14.895	1.054	97	14.778	1.360	91	14.815	0.758	94
−0.6	18.396	18.437	2.015	95	18.433	1.150	96	18.345	1.389	94	18.363	0.792	95
−0.2	22.415	22.711	2.091	96	—	—	—	22.475	1.398	97	—	—	—
0.2	1.310	1.549	0.609	93	—	—	—	2.213	0.993	36	—	—	—
0.6	1.116	1.315	0.562	93	1.291	0.339	88	1.892	0.871	44	1.881	0.801	6
1.0	0.810	1.033	0.499	90	0.974	0.296	89	1.451	0.735	42	1.399	0.623	7
1.4	0.501	0.612	0.342	93	0.632	0.224	89	0.857	0.437	64	0.885	0.413	14
1.8	0.264	0.413	0.292	89	0.372	0.186	84	0.511	0.307	65	0.504	0.266	32
2.2	0.119	0.157	0.171	94	0.172	0.120	90	0.208	0.151	85	0.232	0.139	62
2.6	0.045	0.052	0.097	92	0.060	0.062	93	0.083	0.095	85	0.088	0.067	76
3.0	0.015	0.013	0.050	96	0.018	0.035	92	0.022	0.042	92	0.031	0.034	84
4.0	0.000	0.000	0.000	100	0.001	0.006	99	0.000	0.003	98	0.001	0.005	96

Estimation at $x = 2.5$ with bias correction

		$\eta = (0.4, 0.2)$			$\eta = (0.4, 0.6)$			$\eta = (0.8, 0.2)$			$\eta = (0.8, 0.6)$		
y	$f(x,y)$	mean	rmse	cl	mean	rmse	cl	mean	rmse	cl	mean	rmse	cl
−4.0	3.157	3.008	1.188	85	3.031	0.641	84	3.060	0.844	77	3.089	0.493	78
−3.0	5.219	5.419	1.605	84	5.375	1.047	78	5.285	1.150	80	5.292	0.718	82
−2.6	6.402	6.210	1.870	80	6.439	1.188	71	6.241	1.367	75	6.362	0.831	72
−2.2	7.883	7.983	2.038	81	7.788	1.055	83	7.921	1.519	78	7.812	0.806	76
−1.8	9.748	9.650	2.327	75	9.533	1.378	75	9.562	1.649	78	9.598	0.964	77
−1.4	12.085	11.941	2.538	78	12.028	1.451	80	11.954	1.753	80	12.029	1.110	74
−1.0	14.958	14.831	3.214	73	14.744	1.701	77	14.904	2.382	72	14.850	1.253	73
−0.6	18.396	18.424	3.435	74	18.275	1.904	74	18.510	2.253	74	18.360	1.304	73
−0.2	22.415	22.939	3.558	79	—	—	—	22.770	2.482	79	—	—	—
0.2	1.310	1.344	0.903	76	—	—	—	1.304	0.639	76	—	—	—
0.6	1.116	1.141	0.813	74	1.099	0.444	82	1.111	0.700	62	1.135	0.397	73
1.0	0.810	0.958	0.667	76	0.851	0.408	78	0.903	0.566	69	0.841	0.320	72
1.4	0.501	0.542	0.499	86	0.525	0.263	86	0.547	0.406	71	0.509	0.229	73
1.8	0.264	0.450	0.445	75	0.342	0.256	73	0.379	0.300	77	0.321	0.182	71
2.2	0.119	0.178	0.264	82	0.147	0.157	87	0.153	0.171	85	0.129	0.124	85
2.6	0.045	0.053	0.146	91	0.051	0.080	89	0.073	0.131	81	0.055	0.071	81
3.0	0.015	0.015	0.079	96	0.015	0.048	93	0.020	0.057	91	0.021	0.036	88
4.0	0.000	0.000	0.000	100	0.001	0.009	99	0.000	0.000	100	0.001	0.006	98

Table 5.4: Scenario c3 — The empirical mean (columns 3, 6, 9, 12) of the estimator $\hat{f}_t^{\eta}(x,y)$ (resp., bias-corrected estimator $\hat{f}_t^{\eta}(x,y) - \hat{\gamma}_t^{\eta}(x,y)$) based on 100 trajectories (up to time $t = 2500$) is compared to the true value (col. 2) of $f(x,y)$ given by eq. (5.1.2). In addition, the root mean squared error (rmse; cols. 4, 7, 10, 13) and the empirical confidence level (cl; cols. 5, 8, 11, 14) in percent of the estimated 95%-confidence interval given by eq. (5.1.5) are presented.

Estimation at $x = 0$

		$\eta = (0.4, 0.2)$			$\eta = (0.4, 0.6)$			$\eta = (0.8, 0.2)$			$\eta = (0.8, 0.6)$		
y	$f(x,y)$	mean	rmse	cl	mean	rmse	cl	mean	rmse	cl	mean	rmse	cl
−4.0	0.007	0.019	0.018	67	0.021	0.016	36	0.049	0.045	0	0.051	0.045	0
−3.0	0.222	0.246	0.051	92	0.286	0.070	38	0.288	0.075	56	0.328	0.109	0
−2.6	0.679	0.695	0.083	94	0.792	0.123	34	0.747	0.090	81	0.841	0.165	0
−2.2	1.774	1.797	0.137	94	1.957	0.200	37	1.840	0.114	88	1.996	0.230	1
−1.8	3.948	3.974	0.208	93	4.192	0.273	42	3.977	0.153	96	4.195	0.261	15
−1.4	7.486	7.543	0.258	95	7.694	0.256	71	7.511	0.191	95	7.651	0.196	72
−1.0	12.099	12.073	0.335	94	12.071	0.212	94	11.976	0.270	93	11.990	0.189	84
−0.6	16.661	16.566	0.341	99	16.335	0.406	72	16.453	0.327	91	16.218	0.478	28
−0.2	19.552	19.557	0.409	98	—	—	—	19.354	0.354	90	—	—	—
0.2	19.552	19.481	0.433	96	—	—	—	19.315	0.394	93	—	—	—
0.6	16.661	16.555	0.396	97	16.322	0.411	67	16.453	0.332	89	16.212	0.475	22
1.0	12.099	12.008	0.348	96	12.050	0.199	96	11.970	0.270	93	11.991	0.170	93
1.4	7.486	7.534	0.307	91	7.667	0.240	81	7.509	0.205	92	7.648	0.196	77
1.8	3.948	4.012	0.213	94	4.190	0.277	46	4.020	0.159	93	4.204	0.272	16
2.2	1.774	1.829	0.144	94	1.981	0.223	25	1.864	0.128	87	2.014	0.247	2
2.6	0.679	0.727	0.093	92	0.811	0.141	25	0.765	0.106	67	0.853	0.178	0
3.0	0.222	0.242	0.051	92	0.285	0.070	37	0.289	0.076	51	0.330	0.111	0
4.0	0.007	0.018	0.017	71	0.020	0.015	28	0.049	0.044	2	0.050	0.044	0

Estimation at $x = 0$ with bias correction

		$\eta = (0.4, 0.2)$			$\eta = (0.4, 0.6)$			$\eta = (0.8, 0.2)$			$\eta = (0.8, 0.6)$		
y	$f(x,y)$	mean	rmse	cl	mean	rmse	cl	mean	rmse	cl	mean	rmse	cl
−4.0	0.007	0.012	0.017	78	0.009	0.011	75	0.015	0.019	67	0.010	0.011	67
−3.0	0.222	0.231	0.067	86	0.226	0.044	77	0.224	0.056	74	0.224	0.032	78
−2.6	0.679	0.657	0.125	78	0.672	0.082	73	0.664	0.093	76	0.678	0.058	71
−2.2	1.774	1.763	0.225	78	1.769	0.130	74	1.765	0.149	81	1.778	0.088	77
−1.8	3.948	3.936	0.325	80	3.978	0.193	74	3.926	0.251	73	3.967	0.137	76
−1.4	7.486	7.529	0.411	84	7.560	0.258	78	7.529	0.313	80	7.545	0.182	77
−1.0	12.099	12.084	0.571	78	12.098	0.312	84	12.092	0.387	85	12.106	0.242	76
−0.6	16.661	16.614	0.569	86	16.608	0.353	77	16.660	0.401	88	16.627	0.271	79
−0.2	19.552	19.763	0.722	76	—	—	—	19.644	0.489	80	—	—	—
0.2	19.552	19.597	0.697	80	—	—	—	19.561	0.494	82	—	—	—
0.6	16.661	16.581	0.620	77	16.604	0.411	78	16.648	0.436	81	16.623	0.266	79
1.0	12.099	11.982	0.558	79	12.065	0.329	77	12.028	0.390	82	12.090	0.210	85
1.4	7.486	7.532	0.506	75	7.517	0.265	76	7.532	0.336	70	7.528	0.184	82
1.8	3.948	4.000	0.324	72	3.968	0.203	75	3.988	0.233	74	3.970	0.151	67
2.2	1.774	1.802	0.229	70	1.798	0.134	72	1.795	0.153	80	1.795	0.094	78
2.6	0.679	0.704	0.133	77	0.696	0.082	71	0.695	0.091	78	0.695	0.061	68
3.0	0.222	0.226	0.078	78	0.218	0.042	80	0.226	0.051	79	0.223	0.034	75
4.0	0.007	0.012	0.017	82	0.009	0.009	82	0.014	0.017	69	0.010	0.009	66

Table 5.4a: Scenario c3 (continued)

Estimation at $x = 1.5$

		$\eta = (0.4, 0.2)$			$\eta = (0.4, 0.6)$			$\eta = (0.8, 0.2)$			$\eta = (0.8, 0.6)$		
y	$f(x,y)$	mean	rmse	cl	mean	rmse	cl	mean	rmse	cl	mean	rmse	cl
−4.0	1.695	1.672	0.185	91	1.676	0.100	93	1.595	0.152	87	1.602	0.112	72
−3.0	2.900	2.833	0.221	93	2.890	0.125	95	2.735	0.222	81	2.770	0.158	71
−2.6	3.746	3.706	0.243	95	3.769	0.148	96	3.550	0.261	83	3.628	0.158	77
−2.2	5.048	5.026	0.290	91	5.128	0.184	96	4.866	0.268	88	4.974	0.138	87
−1.8	7.056	7.024	0.323	95	7.155	0.224	92	6.873	0.293	89	7.010	0.145	93
−1.4	9.951	9.905	0.406	97	10.032	0.270	90	9.790	0.318	92	9.914	0.179	94
−1.0	13.631	13.616	0.518	91	13.633	0.291	93	13.528	0.388	93	13.559	0.218	91
−0.6	17.591	17.541	0.524	96	17.496	0.331	92	17.536	0.375	97	17.452	0.259	90
−0.2	21.087	21.140	0.608	93	—	—	—	21.045	0.434	95	—	—	—
0.2	9.776	9.793	0.433	93	—	—	—	10.288	0.596	57	—	—	—
0.6	8.331	8.463	0.382	96	8.339	0.199	97	8.845	0.573	56	8.709	0.404	31
1.0	6.049	6.136	0.343	92	6.130	0.186	93	6.414	0.427	63	6.406	0.373	20
1.4	3.743	3.771	0.243	98	3.870	0.184	89	3.955	0.278	78	4.052	0.325	12
1.8	1.974	2.002	0.193	94	2.107	0.175	74	2.098	0.189	77	2.209	0.249	12
2.2	0.887	0.903	0.120	96	0.986	0.122	69	0.954	0.111	89	1.037	0.158	15
2.6	0.340	0.360	0.082	94	0.403	0.077	69	0.374	0.067	87	0.423	0.088	28
3.0	0.111	0.117	0.041	97	0.139	0.038	79	0.123	0.032	92	0.146	0.039	47
4.0	0.003	0.003	0.008	93	0.005	0.005	89	0.003	0.005	94	0.005	0.004	88

Estimation at $x = 1.5$ with bias correction

		$\eta = (0.4, 0.2)$			$\eta = (0.4, 0.6)$			$\eta = (0.8, 0.2)$			$\eta = (0.8, 0.6)$		
y	$f(x,y)$	mean	rmse	cl	mean	rmse	cl	mean	rmse	cl	mean	rmse	cl
−4.0	1.695	1.712	0.304	73	1.683	0.167	76	1.694	0.200	74	1.677	0.105	79
−3.0	2.900	2.871	0.330	84	2.889	0.199	82	2.874	0.243	78	2.866	0.145	75
−2.6	3.746	3.768	0.394	80	3.753	0.234	78	3.704	0.282	76	3.715	0.171	74
−2.2	5.048	5.071	0.490	72	5.071	0.297	69	5.042	0.316	82	5.047	0.196	75
−1.8	7.056	7.063	0.532	79	7.068	0.323	81	7.016	0.363	83	7.036	0.232	79
−1.4	9.951	9.934	0.663	77	9.949	0.402	77	9.910	0.454	79	9.914	0.282	75
−1.0	13.631	13.648	0.762	77	13.621	0.452	78	13.598	0.604	71	13.598	0.356	72
−0.6	17.591	17.530	0.835	78	17.559	0.505	71	17.613	0.602	81	17.597	0.358	76
−0.2	21.087	21.227	0.959	80	—	—	—	21.186	0.695	75	—	—	—
0.2	9.776	9.568	0.723	74	—	—	—	9.726	0.501	74	—	—	—
0.6	8.331	8.367	0.562	82	8.375	0.338	81	8.401	0.413	78	8.412	0.243	77
1.0	6.049	6.026	0.511	82	6.050	0.291	78	6.058	0.376	76	6.070	0.204	80
1.4	3.743	3.680	0.376	73	3.716	0.221	79	3.731	0.271	80	3.743	0.162	80
1.8	1.974	1.951	0.273	84	1.964	0.187	72	1.961	0.229	70	1.978	0.131	75
2.2	0.887	0.876	0.189	83	0.877	0.110	80	0.890	0.139	80	0.890	0.077	84
2.6	0.340	0.347	0.132	71	0.345	0.076	73	0.349	0.092	73	0.349	0.051	76
3.0	0.111	0.115	0.066	81	0.111	0.042	77	0.115	0.046	80	0.113	0.028	82
4.0	0.003	0.004	0.014	93	0.004	0.007	87	0.004	0.008	88	0.004	0.005	91

Table 5.4b: Scenario c3 (continued)

Estimation at $x = 2.5$

		$\eta = (0.4, 0.2)$			$\eta = (0.4, 0.6)$			$\eta = (0.8, 0.2)$			$\eta = (0.8, 0.6)$		
y	$f(x,y)$	mean	rmse	cl	mean	rmse	cl	mean	rmse	cl	mean	rmse	cl
−4.0	3.157	3.200	0.354	97	3.170	0.213	94	3.052	0.263	94	3.036	0.199	87
−3.0	5.219	5.173	0.442	95	5.192	0.243	96	4.988	0.382	90	4.998	0.279	83
−2.6	6.402	6.328	0.501	97	6.353	0.297	94	6.087	0.451	84	6.140	0.327	77
−2.2	7.883	7.682	0.606	91	7.847	0.329	96	7.490	0.561	82	7.625	0.347	81
−1.8	9.748	9.769	0.690	93	9.801	0.417	94	9.511	0.521	89	9.557	0.336	88
−1.4	12.085	12.180	0.688	97	12.170	0.456	92	11.914	0.527	97	11.951	0.341	93
−1.0	14.958	14.910	0.883	95	14.994	0.466	96	14.781	0.633	91	14.854	0.337	93
−0.6	18.396	18.430	0.866	97	18.369	0.481	96	18.280	0.656	93	18.301	0.349	95
−0.2	22.415	22.088	1.007	94	—	—	—	22.137	0.752	90	—	—	—
0.2	1.310	1.568	0.345	82	—	—	—	2.262	0.970	0	—	—	—
0.6	1.116	1.308	0.288	90	1.298	0.225	70	1.897	0.804	0	1.881	0.772	0
1.0	0.810	0.963	0.242	86	0.955	0.185	70	1.385	0.594	0	1.386	0.583	0
1.4	0.501	0.618	0.220	81	0.622	0.156	70	0.870	0.391	6	0.889	0.395	0
1.8	0.264	0.315	0.114	91	0.334	0.095	79	0.455	0.206	27	0.481	0.222	0
2.2	0.119	0.149	0.084	91	0.156	0.059	83	0.212	0.108	54	0.225	0.111	5
2.6	0.045	0.055	0.043	95	0.063	0.032	87	0.077	0.047	79	0.090	0.049	27
3.0	0.015	0.016	0.026	93	0.022	0.018	88	0.026	0.024	88	0.031	0.020	63
4.0	0.000	0.000	0.000	100	0.001	0.002	96	0.000	0.002	99	0.001	0.002	90

Estimation at $x = 2.5$ with bias correction

		$\eta = (0.4, 0.2)$			$\eta = (0.4, 0.6)$			$\eta = (0.8, 0.2)$			$\eta = (0.8, 0.6)$		
y	$f(x,y)$	mean	rmse	cl	mean	rmse	cl	mean	rmse	cl	mean	rmse	cl
−4.0	3.157	3.259	0.575	80	3.223	0.324	84	3.241	0.411	78	3.201	0.254	80
−3.0	5.219	5.220	0.709	82	5.227	0.418	79	5.234	0.521	76	5.226	0.297	80
−2.6	6.402	6.418	0.854	74	6.380	0.497	79	6.359	0.519	79	6.373	0.337	76
−2.2	7.883	7.646	0.975	78	7.833	0.515	76	7.713	0.672	76	7.833	0.370	76
−1.8	9.748	9.889	1.107	79	9.842	0.640	83	9.809	0.800	72	9.782	0.456	74
−1.4	12.085	12.396	1.105	78	12.182	0.698	77	12.216	0.828	80	12.125	0.524	75
−1.0	14.958	14.922	1.484	68	14.919	0.762	74	14.958	1.013	75	14.928	0.564	73
−0.6	18.396	18.537	1.405	82	18.329	0.826	75	18.381	1.037	75	18.316	0.550	79
−0.2	22.415	21.976	1.549	73	—	—	—	22.138	1.231	73	—	—	—
0.2	1.310	1.354	0.373	83	—	—	—	1.347	0.319	67	—	—	—
0.6	1.116	1.101	0.338	79	1.114	0.216	74	1.102	0.276	75	1.128	0.169	71
1.0	0.810	0.841	0.304	78	0.805	0.186	74	0.794	0.245	70	0.804	0.142	72
1.4	0.501	0.539	0.294	68	0.527	0.169	66	0.520	0.228	56	0.512	0.124	66
1.8	0.264	0.273	0.178	74	0.264	0.099	78	0.260	0.130	70	0.253	0.075	72
2.2	0.119	0.134	0.125	84	0.120	0.071	73	0.131	0.097	65	0.110	0.055	68
2.6	0.045	0.056	0.061	88	0.050	0.039	88	0.049	0.050	84	0.040	0.032	66
3.0	0.015	0.018	0.040	88	0.019	0.025	84	0.018	0.030	87	0.015	0.017	83
4.0	0.000	0.000	0.000	100	0.000	0.001	97	0.000	0.002	99	0.001	0.002	95

Table 5.5: Scenario c2 — The empirical mean (integrated) squared error (columns 4, 6, 8, 10) on the interval $[y_1, y_2]$ of the estimator $\hat{f}_t^{\eta}(x, \cdot)$ for $f(x, y)$ given by eq. (5.1.2) based on 100 trajectories (up to time $t = 500$) is presented. In addition, the standard deviation (columns 5, 7, 9, 11) of the squared errors are shown.

			$x = 0$		$x = 0.75$		$x = 1.5$		$x = 2.25$	
η	y_1	y_2	mse	sd	mse	sd	mse	sd	mse	sd
$(0.4, 0.4)$	-3.0	-1.5	0.094	0.074	0.137	0.109	0.285	0.197	0.864	0.527
	-1.5	-0.6	0.253	0.218	0.314	0.274	0.487	0.483	1.086	0.960
	-0.6	-0.4	0.108	0.147	0.092	0.110	0.184	0.250	0.345	0.499
	0.4	0.6	0.111	0.153	0.081	0.095	0.075	0.088	0.052	0.064
	0.6	1.5	0.287	0.277	0.217	0.200	0.216	0.214	0.171	0.180
	1.5	3.0	0.108	0.098	0.086	0.079	0.068	0.078	0.041	0.036
$(0.4, 0.6)$	-3.0	-1.5	0.097	0.096	0.162	0.155	0.227	0.195	0.508	0.423
	-1.5	-0.6	0.185	0.186	0.188	0.170	0.393	0.437	0.585	0.550
	0.6	1.5	0.182	0.192	0.171	0.176	0.174	0.172	0.104	0.115
	1.5	3.0	0.106	0.102	0.078	0.075	0.072	0.062	0.033	0.040
$(0.6, 0.2)$	-3.0	-1.5	0.114	0.071	0.178	0.094	0.418	0.215	1.127	0.562
	-1.5	-0.6	0.307	0.218	0.430	0.288	0.705	0.484	1.411	0.823
	-0.6	-0.4	0.119	0.158	0.141	0.179	0.238	0.249	0.357	0.375
	-0.4	-0.2	0.143	0.186	0.124	0.122	0.277	0.299	0.442	0.569
	0.2	0.4	0.138	0.162	0.111	0.140	0.122	0.162	0.141	0.168
	0.4	0.6	0.126	0.182	0.101	0.107	0.103	0.113	0.117	0.153
	0.6	1.5	0.364	0.277	0.292	0.174	0.327	0.244	0.284	0.186
	1.5	3.0	0.109	0.064	0.099	0.049	0.094	0.063	0.079	0.064
$(0.6, 0.4)$	-3.0	-1.5	0.062	0.049	0.110	0.097	0.218	0.139	0.602	0.402
	-1.5	-0.6	0.186	0.159	0.234	0.218	0.311	0.257	0.835	0.694
	-0.6	-0.4	0.066	0.089	0.068	0.072	0.090	0.113	0.345	0.454
	0.4	0.6	0.070	0.083	0.060	0.077	0.055	0.076	0.078	0.097
	0.6	1.5	0.176	0.154	0.165	0.135	0.166	0.141	0.208	0.147
	1.5	3.0	0.071	0.053	0.063	0.051	0.058	0.056	0.050	0.046

Table 5.6: Scenario c3 — The empirical mean (integrated) squared error (columns 4, 6, 8, 10) on the interval $[y_1, y_2]$ of the estimator $\hat{f}_t^{\eta}(x, \cdot)$ for $f(x, y)$ given by eq. (5.1.2) based on 100 trajectories (up to time $t = 2500$) is presented. In addition, the standard deviation (columns 5, 7, 9, 11) of the squared errors are shown.

			$x = 0$		$x = 0.75$		$x = 1.5$		$x = 2.25$	
η	y_1	y_2	mse	sd	mse	sd	mse	sd	mse	sd
(0.1, 0.4)	−3.0	−1.5	0.086	0.076	0.111	0.081	0.214	0.151	0.813	0.579
	−1.5	−0.4	0.269	0.235	0.298	0.207	0.489	0.399	1.118	0.936
	0.4	1.5	0.27	0.226	0.275	0.203	0.199	0.168	0.123	0.088
	1.5	3.0	0.069	0.063	0.056	0.042	0.058	0.061	0.033	0.029
(0.2, 0.4)	−3.0	−1.5	0.043	0.042	0.051	0.04	0.124	0.099	0.294	0.183
	−1.5	−0.4	0.118	0.081	0.154	0.111	0.223	0.16	0.478	0.378
	0.4	1.5	0.16	0.123	0.139	0.122	0.118	0.095	0.079	0.06
	1.5	3.0	0.043	0.033	0.04	0.030	0.038	0.036	0.016	0.014
(0.4, 0.2)	−3.0	−1.5	0.035	0.021	0.056	0.033	0.138	0.067	0.361	0.167
	−1.5	−0.4	0.135	0.086	0.147	0.084	0.264	0.158	0.492	0.247
	−0.4	−0.2	0.045	0.053	0.038	0.043	0.077	0.092	0.121	0.235
	0.2	0.4	0.034	0.033	0.037	0.046	0.032	0.035	0.035	0.042
	0.4	1.5	0.144	0.098	0.128	0.074	0.125	0.069	0.097	0.067
	1.5	3.0	0.03	0.02	0.031	0.02	0.029	0.022	0.019	0.014
(0.4, 0.4)	−3.0	−1.5	0.027	0.023	0.033	0.028	0.061	0.049	0.190	0.143
	−1.5	−0.4	0.078	0.069	0.080	0.057	0.114	0.079	0.264	0.247
	0.4	1.5	0.075	0.057	0.077	0.063	0.064	0.058	0.055	0.041
	1.5	3.0	0.026	0.018	0.019	0.021	0.018	0.015	0.011	0.011

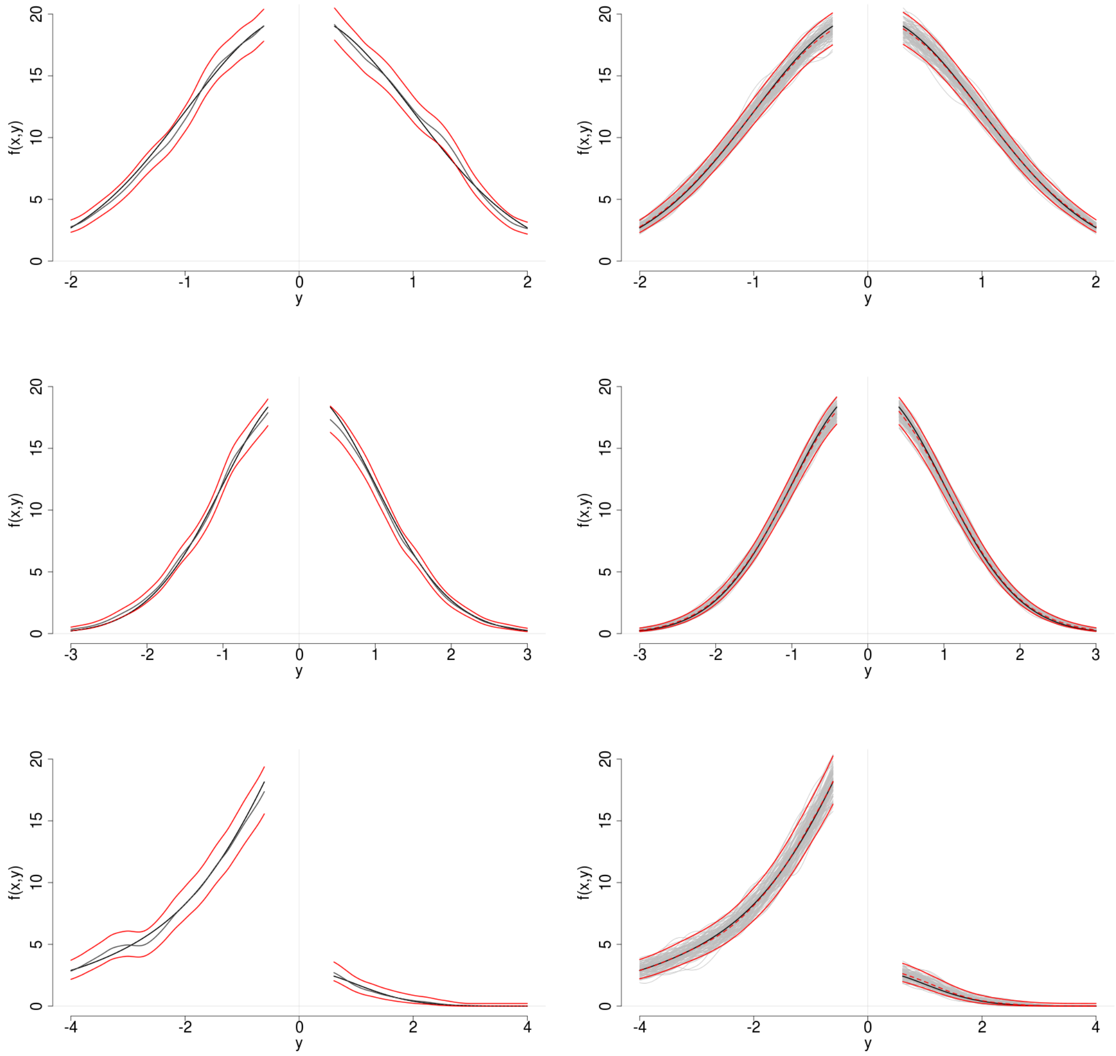

Figure 5.3: Scenario c2 — Estimation of the Lévy density $f(x, y)$ given by eq. (5.1.2) at $x = 0$ with $\eta = (0.6, 0.3)$ (top row), at $x = 0$ with $\eta = (0.4, 0.6))$ (middle row), and at $x = 2.25$ with $\eta = (0.6, 0.4)$ (bottom row) based on continuous observations up to time $t = 500$. Left: One typical estimate (grey) is compared to the true Lévy density (black). The upper and lower bounds of the estimated (pointwise) 95% -confidence intervals given by eq. (5.1.5) are shown in red. For $x = 2.25$, we note that the estimate is identically zero for $y > 2.7$ as there were no jumps. Right: Estimates based on 100 trajectories (grey) are compared to the true Lévy density (black). The (pointwise) mean of the estimates (red dashed line) and mean of the upper and lower bounds of the 95% -confidence intervals (red solid lines) are shown.

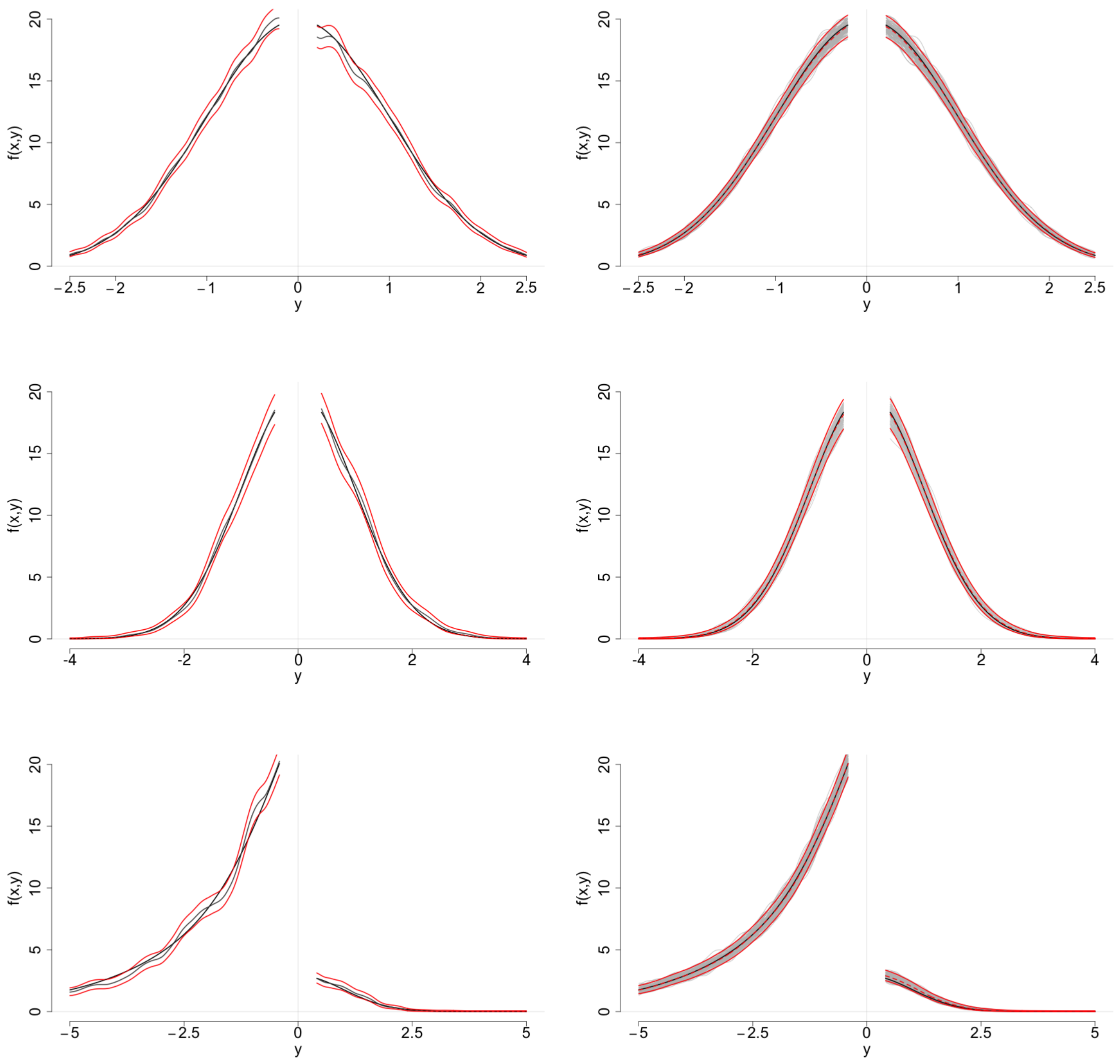

Figure 5.4: Scenario c3 — Estimation of the Lévy density $f(x, y)$ given by eq. (5.1.2) at $x = 0$ with $\eta = (0.4, 0.2)$ (top row), at $x = 0$ with $\eta = (0.1, 0.4)$ (middle row), and at $x = 2.25$ with $\eta = (0.4, 0.4)$ (bottom row) based on continuous observations up to time $t = 2500$. Left: One typical estimate (grey) is compared to the true Lévy density (black). The upper and lower bounds of the estimated (pointwise) 95%-confidence intervals given by eq. (5.1.5) are shown in red. For $x = 2.25$, we note that the estimate is identically zero for $y > 3.5$ as there were no jumps. Right: Estimates based on 100 trajectories (grey) are compared to the true Lévy density (black). The (pointwise) mean of the estimates (red dashed line) and mean of the upper and lower bounds of the 95%-confidence intervals (red solid lines) are shown.

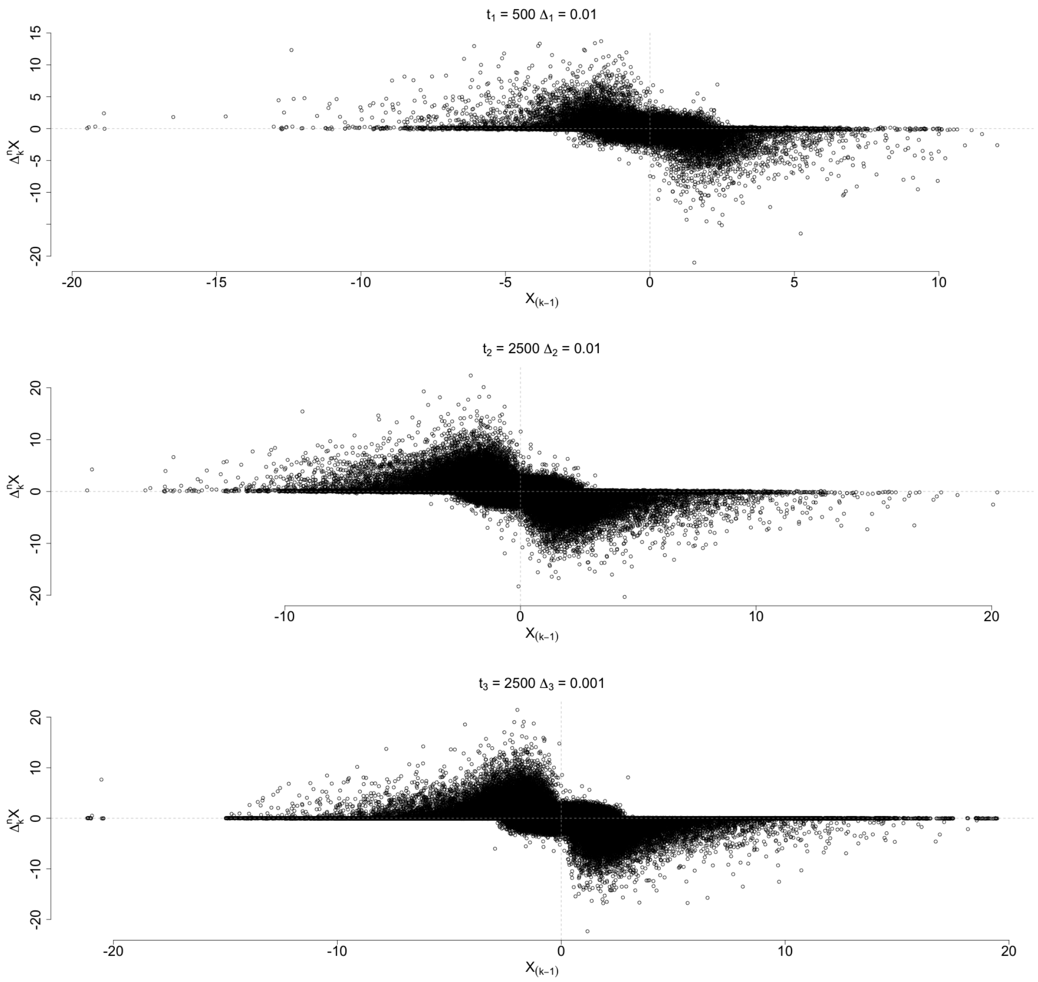

Figure 5.5: Increments $(X_{(k-1)\Delta}, \Delta_k^n X)$ of one simulated sample of scenarios d1 (top), d2 (middle), and d3 (bottom).

Table 5.7: Scenario d1 — The empirical mean (columns 3, 6, 9, 12) of the estimator $\hat{f}_n^{\Delta,\eta}(x,y)$ (resp., bias-corrected estimator $\hat{f}_n^{\Delta,\eta}(x,y) - \hat{\gamma}_n^{\Delta,\eta}(x,y)$) based on 100 samples (up to time $t = 500$ with $\Delta = 0.01$) is compared to the true value (col. 2) of $f(x,y)$ given by eq. (5.1.2). In addition, the root mean squared error (rmse; cols. 4, 7, 10, 13) and the empirical confidence level (cl; cols. 5, 8, 11, 14) in percent of the estimated 95%-confidence interval given by eq. (5.1.5) are presented.

Estimation at $x = 0$

		$\eta = (0.4, 0.2)$			$\eta = (0.4, 0.6)$			$\eta = (0.8, 0.2)$			$\eta = (0.8, 0.6)$		
y	$f(x,y)$	mean	rmse	cl	mean	rmse	cl	mean	rmse	cl	mean	rmse	cl
−4.0	0.007	0.061	0.077	44	0.067	0.069	10	0.088	0.092	5	0.092	0.089	0
−3.0	0.222	0.362	0.198	64	0.400	0.196	28	0.404	0.214	42	0.440	0.228	0
−2.6	0.679	0.817	0.229	91	0.867	0.221	57	0.856	0.225	68	0.917	0.253	18
−2.2	1.774	1.744	0.261	98	1.840	0.176	94	1.782	0.220	97	1.881	0.158	89
−1.8	3.948	3.463	0.624	82	3.597	0.417	76	3.485	0.556	73	3.602	0.386	53
−1.4	7.486	6.066	1.511	33	6.186	1.331	3	6.027	1.504	7	6.160	1.341	0
−1.0	12.099	9.292	2.881	2	9.321	2.805	0	9.264	2.874	0	9.262	2.851	0
−0.6	16.661	12.411	4.319	0	16.214	0.623	94	12.310	4.383	0	16.117	0.616	78
−0.2	19.552	69.575	50.048	0	—	—	—	69.574	50.034	0	—	—	—
0.2	19.552	69.664	50.132	0	—	—	—	69.696	50.157	0	—	—	—
0.6	16.661	12.462	4.253	0	16.269	0.574	95	12.367	4.323	0	16.188	0.560	83
1.0	12.099	9.445	2.739	7	9.351	2.777	0	9.337	2.798	0	9.271	2.841	0
1.4	7.486	6.124	1.498	38	6.218	1.318	7	6.065	1.493	13	6.165	1.345	0
1.8	3.948	3.471	0.625	84	3.598	0.420	75	3.468	0.570	62	3.597	0.390	54
2.2	1.774	1.751	0.315	94	1.861	0.192	91	1.782	0.219	93	1.893	0.175	83
2.6	0.679	0.830	0.247	87	0.881	0.233	52	0.869	0.239	71	0.920	0.257	15
3.0	0.222	0.358	0.184	70	0.397	0.192	23	0.391	0.189	43	0.434	0.218	0
4.0	0.007	0.071	0.085	34	0.072	0.073	5	0.091	0.097	11	0.094	0.091	0

Estimation at $x = 0$ with bias correction

		$\eta = (0.4, 0.2)$			$\eta = (0.4, 0.6)$			$\eta = (0.8, 0.2)$			$\eta = (0.8, 0.6)$		
y	$f(x,y)$	mean	rmse	cl	mean	rmse	cl	mean	rmse	cl	mean	rmse	cl
−4.0	0.007	0.060	0.099	59	0.055	0.069	34	0.059	0.079	40	0.054	0.061	26
−3.0	0.222	0.350	0.248	63	0.347	0.176	49	0.348	0.213	49	0.353	0.165	38
−2.6	0.679	0.806	0.315	77	0.761	0.198	71	0.809	0.260	69	0.785	0.171	61
−2.2	1.774	1.697	0.416	87	1.681	0.294	71	1.740	0.338	73	1.709	0.203	76
−1.8	3.948	3.438	0.794	72	3.450	0.617	53	3.459	0.683	65	3.456	0.568	37
−1.4	7.486	6.078	1.672	38	6.123	1.444	8	6.001	1.607	20	6.085	1.442	0
−1.0	12.099	9.281	3.012	11	9.391	2.777	1	9.310	2.892	1	9.368	2.770	0
−0.6	16.661	12.495	4.325	1	3.830	12.851	0	12.457	4.295	0	3.764	12.907	0
−0.2	19.552	52.097	32.656	0	—	—	—	52.242	32.749	0	—	—	—
0.2	19.552	52.228	32.757	0	—	—	—	52.479	32.975	0	—	—	—
0.6	16.661	12.472	4.328	1	3.824	12.858	0	12.486	4.243	0	3.842	12.829	0
1.0	12.099	9.556	2.772	21	9.424	2.747	0	9.491	2.713	3	9.378	2.759	0
1.4	7.486	6.162	1.652	40	6.156	1.452	14	6.089	1.573	25	6.104	1.442	0
1.8	3.948	3.462	0.798	69	3.457	0.621	50	3.433	0.691	55	3.441	0.579	30
2.2	1.774	1.734	0.508	74	1.719	0.289	75	1.726	0.345	77	1.734	0.208	75
2.6	0.679	0.815	0.358	70	0.790	0.209	71	0.830	0.277	68	0.799	0.175	59
3.0	0.222	0.345	0.239	62	0.342	0.177	46	0.341	0.177	66	0.341	0.145	36
4.0	0.007	0.067	0.101	52	0.064	0.077	26	0.070	0.090	36	0.062	0.068	13

Table 5.7a: Scenario d1 (continued)

Estimation at $x = 1.5$

y	$f(x,y)$	$\eta=(0.4,0.2)$ mean	rmse	cl	$\eta=(0.4,0.6)$ mean	rmse	cl	$\eta=(0.8,0.2)$ mean	rmse	cl	$\eta=(0.8,0.6)$ mean	rmse	cl
−4.0	1.695	1.474	0.393	96	1.484	0.270	89	1.401	0.370	87	1.418	0.303	54
−3.0	2.900	2.706	0.483	97	2.746	0.292	97	2.577	0.429	93	2.626	0.325	70
−2.6	3.746	3.489	0.600	93	3.601	0.311	95	3.371	0.548	82	3.460	0.352	78
−2.2	5.048	4.770	0.707	93	4.776	0.429	90	4.589	0.642	81	4.621	0.488	65
−1.8	7.056	6.266	1.103	85	6.289	0.872	65	6.146	1.059	65	6.183	0.918	16
−1.4	9.951	8.286	1.912	51	8.329	1.689	10	8.207	1.868	25	8.242	1.741	0
−1.0	13.631	10.778	2.967	13	10.782	2.894	1	10.718	2.968	1	10.689	2.964	0
−0.6	17.591	13.354	4.351	4	18.061	0.719	96	13.210	4.436	0	17.967	0.547	96
−0.2	21.087	81.052	60.008	0	—	—	—	80.689	59.623	0	—	—	—
0.2	9.776	53.306	43.558	0	—	—	—	54.028	44.267	0	—	—	—
0.6	8.331	6.252	2.178	24	9.370	1.104	38	6.572	1.826	10	9.699	1.396	0
1.0	6.049	4.453	1.742	41	4.480	1.603	1	4.662	1.474	19	4.699	1.372	0
1.4	3.743	2.672	1.177	46	2.785	0.995	15	2.824	0.974	25	2.930	0.834	4
1.8	1.974	1.418	0.648	74	1.502	0.508	52	1.522	0.516	65	1.586	0.415	35
2.2	0.887	0.650	0.333	88	0.706	0.231	80	0.689	0.259	87	0.751	0.169	75
2.6	0.340	0.273	0.142	99	0.302	0.092	98	0.287	0.110	99	0.321	0.067	97
3.0	0.111	0.100	0.084	98	0.122	0.054	97	0.103	0.061	99	0.125	0.042	93
4.0	0.003	0.012	0.031	89	0.015	0.022	67	0.013	0.022	79	0.015	0.018	60

Estimation at $x = 1.5$ with bias correction

y	$f(x,y)$	$\eta=(0.4,0.2)$ mean	rmse	cl	$\eta=(0.4,0.6)$ mean	rmse	cl	$\eta=(0.8,0.2)$ mean	rmse	cl	$\eta=(0.8,0.6)$ mean	rmse	cl
−4.0	1.695	1.523	0.579	77	1.500	0.340	76	1.492	0.420	78	1.487	0.289	64
−3.0	2.900	2.760	0.798	75	2.767	0.441	77	2.709	0.500	78	2.729	0.332	77
−2.6	3.746	3.483	0.897	76	3.598	0.473	81	3.482	0.707	72	3.556	0.382	73
−2.2	5.048	4.860	1.108	72	4.800	0.634	73	4.762	0.768	75	4.721	0.513	65
−1.8	7.056	6.337	1.445	69	6.213	1.079	54	6.322	1.157	64	6.225	0.955	37
−1.4	9.951	8.318	2.204	49	8.272	1.845	18	8.316	1.954	36	8.276	1.768	5
−1.0	13.631	10.795	3.123	28	10.825	2.924	6	10.824	2.949	9	10.785	2.913	0
−0.6	17.591	13.555	4.340	15	3.846	13.778	0	13.347	4.399	4	3.852	13.758	0
−0.2	21.087	65.378	44.470	0	—	—	—	65.439	44.434	0	—	—	—
0.2	9.776	35.181	25.519	0	—	—	—	35.362	25.656	0	—	—	—
0.6	8.331	6.104	2.442	26	0.000	8.331	0	6.167	2.295	7	0.000	8.331	0
1.0	6.049	4.384	2.016	37	4.389	1.752	6	4.400	1.846	21	4.422	1.675	0
1.4	3.743	2.600	1.404	40	2.665	1.172	16	2.585	1.266	20	2.666	1.120	4
1.8	1.974	1.355	0.820	69	1.401	0.648	42	1.383	0.709	54	1.394	0.625	11
2.2	0.887	0.620	0.455	73	0.620	0.352	54	0.631	0.373	62	0.628	0.304	38
2.6	0.340	0.269	0.209	89	0.258	0.155	76	0.265	0.175	79	0.262	0.124	69
3.0	0.111	0.097	0.124	94	0.110	0.082	83	0.095	0.095	91	0.102	0.062	78
4.0	0.003	0.015	0.048	90	0.018	0.033	71	0.013	0.031	82	0.015	0.022	59

Table 5.7b: Scenario d1 (continued)

Estimation at $x = 2.5$

y	$f(x,y)$	$\eta=(0.4,0.2)$ mean	rmse	cl	$\eta=(0.4,0.6)$ mean	rmse	cl	$\eta=(0.8,0.2)$ mean	rmse	cl	$\eta=(0.8,0.6)$ mean	rmse	cl
−4.0	3.157	2.786	0.894	95	2.846	0.571	91	2.689	0.694	92	2.736	0.519	77
−3.0	5.219	4.815	1.027	94	4.805	0.651	91	4.585	0.916	89	4.592	0.731	71
−2.6	6.402	5.801	1.312	94	5.804	0.903	80	5.552	1.135	82	5.595	0.913	56
−2.2	7.883	6.841	1.599	90	6.950	1.171	75	6.742	1.393	80	6.777	1.196	40
−1.8	9.748	8.222	1.951	88	8.206	1.737	51	8.063	1.887	57	8.057	1.771	17
−1.4	12.085	9.455	2.916	67	9.613	2.580	22	9.444	2.773	32	9.526	2.611	2
−1.0	14.958	11.418	3.839	49	11.318	3.744	3	11.260	3.820	11	11.251	3.753	0
−0.6	18.396	12.777	5.816	12	18.581	0.929	100	12.913	5.580	1	18.578	0.657	98
−0.2	22.415	88.398	66.075	0	—	—	—	87.616	65.231	0	—	—	—
0.2	1.310	41.612	40.369	0	—	—	—	42.541	41.260	0	—	—	—
0.6	1.116	1.192	0.409	98	3.909	2.807	0	1.619	0.592	73	4.361	3.251	0
1.0	0.810	0.783	0.359	99	0.814	0.180	100	1.093	0.407	81	1.113	0.338	63
1.4	0.501	0.532	0.288	96	0.532	0.164	96	0.703	0.297	87	0.700	0.236	68
1.8	0.264	0.202	0.202	99	0.260	0.123	97	0.319	0.162	94	0.360	0.134	85
2.2	0.119	0.140	0.167	93	0.123	0.098	96	0.166	0.126	88	0.167	0.082	89
2.6	0.045	0.045	0.098	94	0.056	0.059	94	0.062	0.070	94	0.074	0.048	90
3.0	0.015	0.024	0.066	93	0.025	0.039	85	0.030	0.056	88	0.032	0.033	84
4.0	0.000	0.016	0.060	91	0.013	0.034	83	0.012	0.035	85	0.010	0.021	75

Estimation at $x = 2.5$ with bias correction

y	$f(x,y)$	$\eta=(0.4,0.2)$ mean	rmse	cl	$\eta=(0.4,0.6)$ mean	rmse	cl	$\eta=(0.8,0.2)$ mean	rmse	cl	$\eta=(0.8,0.6)$ mean	rmse	cl
−4.0	3.157	2.773	1.447	71	2.874	0.848	74	2.815	0.962	75	2.863	0.562	79
−3.0	5.219	4.939	1.592	78	4.873	0.877	80	4.833	1.159	75	4.806	0.785	73
−2.6	6.402	5.975	1.947	78	5.881	1.253	66	5.756	1.412	75	5.807	0.937	57
−2.2	7.883	6.807	2.256	70	6.952	1.459	67	6.969	1.640	66	7.036	1.138	60
−1.8	9.748	8.239	2.609	73	8.244	1.977	52	8.332	2.057	63	8.272	1.702	35
−1.4	12.085	9.370	3.441	57	9.544	2.806	24	9.547	2.913	37	9.623	2.591	6
−1.0	14.958	11.590	4.181	53	11.383	3.877	18	11.369	3.943	23	11.350	3.741	2
−0.6	18.396	12.540	6.336	21	2.948	15.525	0	12.679	6.028	7	3.128	15.311	0
−0.2	22.415	74.752	52.710	0	—	—	—	73.994	51.703	0	—	—	—
0.2	1.310	23.691	22.773	0	—	—	—	23.471	22.329	0	—	—	—
0.6	1.116	1.042	0.703	82	0.000	1.116	0	1.069	0.539	75	0.000	1.116	0
1.0	0.810	0.678	0.572	86	0.682	0.312	87	0.707	0.500	69	0.697	0.283	75
1.4	0.501	0.464	0.464	91	0.479	0.284	80	0.497	0.344	77	0.466	0.220	72
1.8	0.264	0.188	0.289	93	0.196	0.186	90	0.201	0.232	93	0.200	0.149	74
2.2	0.119	0.150	0.233	85	0.122	0.150	89	0.142	0.175	85	0.102	0.110	91
2.6	0.045	0.053	0.156	90	0.052	0.081	89	0.057	0.108	88	0.049	0.062	86
3.0	0.015	0.029	0.113	93	0.028	0.061	82	0.033	0.086	88	0.028	0.043	78
4.0	0.000	0.022	0.096	93	0.019	0.056	84	0.016	0.060	90	0.015	0.035	76

Table 5.8: Scenario d2 — The empirical mean (columns 3, 6, 9, 12) of the estimator $\hat{f}_n^{\Delta,\eta}(x,y)$ (resp., bias-corrected estimator $\hat{f}_n^{\Delta,\eta}(x,y) - \hat{\gamma}_n^{\Delta,\eta}(x,y)$) based on 100 samples (up to time $t = 2500$ with $\Delta = 0.01$) is compared to the true value (col. 2) of $f(x,y)$ given by eq. (5.1.2). In addition, the root mean squared error (rmse; cols. 4, 7, 10, 13) and the empirical confidence level (cl; cols. 5, 8, 11, 14) in percent of the estimated 95%-confidence interval given by eq. (5.1.5) are presented.

Estimation at $x = 0$

		$\eta = (0.4, 0.2)$			$\eta = (0.4, 0.6)$			$\eta = (0.8, 0.2)$			$\eta = (0.8, 0.6)$		
y	$f(x,y)$	mean	rmse	cl	mean	rmse	cl	mean	rmse	cl	mean	rmse	cl
−4.0	0.007	0.063	0.061	3	0.068	0.063	0	0.087	0.083	0	0.092	0.086	0
−3.0	0.222	0.364	0.155	18	0.393	0.175	0	0.398	0.182	0	0.431	0.211	0
−2.6	0.679	0.802	0.157	67	0.875	0.204	5	0.841	0.177	25	0.915	0.239	0
−2.2	1.774	1.739	0.141	95	1.854	0.111	83	1.776	0.091	97	1.889	0.126	49
−1.8	3.948	3.461	0.512	28	3.584	0.377	10	3.474	0.487	6	3.596	0.357	1
−1.4	7.486	6.107	1.398	0	6.175	1.318	0	6.071	1.425	0	6.147	1.342	0
−1.0	12.099	9.309	2.803	0	9.301	2.801	0	9.261	2.845	0	9.252	2.849	0
−0.6	16.661	12.478	4.200	0	16.283	0.424	67	12.389	4.279	0	16.193	0.487	22
−0.2	19.552	69.854	50.307	0	—	—	—	69.744	50.195	0	—	—	—
0.2	19.552	69.782	50.235	0	—	—	—	69.689	50.139	0	—	—	—
0.6	16.661	12.494	4.179	0	16.289	0.422	72	12.423	4.244	0	16.207	0.474	14
1.0	12.099	9.349	2.769	0	9.341	2.762	0	9.288	2.821	0	9.282	2.819	0
1.4	7.486	6.075	1.433	0	6.180	1.314	0	6.032	1.463	0	6.146	1.344	0
1.8	3.948	3.462	0.519	31	3.592	0.373	13	3.472	0.493	7	3.598	0.357	0
2.2	1.774	1.728	0.118	99	1.852	0.102	90	1.768	0.082	99	1.888	0.122	49
2.6	0.679	0.809	0.159	70	0.874	0.201	0	0.845	0.178	21	0.913	0.237	0
3.0	0.222	0.352	0.141	26	0.389	0.171	0	0.388	0.171	2	0.425	0.205	0
4.0	0.007	0.063	0.061	2	0.067	0.062	0	0.089	0.084	0	0.092	0.085	0

Estimation at $x = 0$ with bias correction

		$\eta = (0.4, 0.2)$			$\eta = (0.4, 0.6)$			$\eta = (0.8, 0.2)$			$\eta = (0.8, 0.6)$		
y	$f(x,y)$	mean	rmse	cl	mean	rmse	cl	mean	rmse	cl	mean	rmse	cl
−4.0	0.007	0.054	0.060	25	0.055	0.054	5	0.058	0.059	8	0.057	0.053	0
−3.0	0.222	0.348	0.158	31	0.338	0.130	14	0.347	0.143	18	0.343	0.129	1
−2.6	0.679	0.782	0.180	61	0.782	0.132	42	0.780	0.152	55	0.789	0.127	23
−2.2	1.774	1.714	0.221	74	1.713	0.143	71	1.722	0.151	80	1.728	0.094	78
−1.8	3.948	3.438	0.571	30	3.427	0.545	5	3.437	0.546	11	3.436	0.522	0
−1.4	7.486	6.123	1.416	1	6.092	1.412	0	6.102	1.409	0	6.087	1.408	0
−1.0	12.099	9.326	2.811	0	9.307	2.803	0	9.365	2.754	0	9.344	2.760	0
−0.6	16.661	12.530	4.178	0	3.861	12.805	0	12.535	4.147	0	3.859	12.804	0
−0.2	19.552	52.608	33.081	0	—	—	—	52.605	33.065	0	—	—	—
0.2	19.552	52.589	33.058	0	—	—	—	52.573	33.030	0	—	—	—
0.6	16.661	12.536	4.158	0	3.893	12.772	0	12.585	4.092	0	3.887	12.776	0
1.0	12.099	9.370	2.776	0	9.392	2.719	0	9.362	2.765	0	9.383	2.723	0
1.4	7.486	6.080	1.470	3	6.094	1.411	0	6.031	1.480	0	6.073	1.421	0
1.8	3.948	3.428	0.595	32	3.450	0.529	9	3.446	0.547	19	3.452	0.511	1
2.2	1.774	1.693	0.202	78	1.709	0.122	75	1.700	0.153	73	1.726	0.084	77
2.6	0.679	0.791	0.185	64	0.786	0.134	48	0.787	0.156	51	0.788	0.121	23
3.0	0.222	0.331	0.142	45	0.333	0.122	14	0.333	0.128	24	0.334	0.117	1
4.0	0.007	0.056	0.063	21	0.054	0.052	4	0.059	0.059	7	0.056	0.052	1

Table 5.8a: Scenario d2 (continued)

Estimation at $x = 1.5$

		$\eta = (0.4, 0.2)$			$\eta = (0.4, 0.6)$			$\eta = (0.8, 0.2)$			$\eta = (0.8, 0.6)$		
y	$f(x,y)$	mean	rmse	cl	mean	rmse	cl	mean	rmse	cl	mean	rmse	cl
−4.0	1.695	1.454	0.287	69	1.472	0.241	35	1.403	0.314	31	1.416	0.286	2
−3.0	2.900	2.694	0.287	84	2.746	0.188	78	2.591	0.339	47	2.635	0.276	12
−2.6	3.746	3.578	0.307	87	3.597	0.204	85	3.421	0.368	52	3.461	0.300	22
−2.2	5.048	4.677	0.450	76	4.741	0.346	57	4.549	0.532	31	4.611	0.450	3
−1.8	7.056	6.250	0.881	41	6.309	0.771	2	6.137	0.954	3	6.195	0.871	0
−1.4	9.951	8.298	1.695	2	8.359	1.608	0	8.183	1.785	0	8.260	1.697	0
−1.0	13.631	10.717	2.944	0	10.746	2.895	0	10.689	2.958	0	10.681	2.955	0
−0.6	17.591	13.205	4.408	0	17.964	0.444	84	13.120	4.481	0	17.894	0.346	79
−0.2	21.087	80.692	59.612	0	—	—	—	80.392	59.309	0	—	—	—
0.2	9.776	53.619	43.851	0	—	—	—	54.223	44.451	0	—	—	—
0.6	8.331	6.326	2.031	0	9.445	1.126	0	6.596	1.750	0	9.743	1.418	0
1.0	6.049	4.460	1.612	0	4.483	1.573	0	4.684	1.381	0	4.701	1.353	0
1.4	3.743	2.719	1.046	0	2.794	0.956	0	2.856	0.901	0	2.935	0.812	0
1.8	1.974	1.445	0.545	10	1.518	0.464	0	1.531	0.453	0	1.603	0.376	0
2.2	0.887	0.666	0.245	55	0.718	0.180	28	0.698	0.204	46	0.760	0.134	25
2.6	0.340	0.267	0.098	85	0.300	0.060	87	0.289	0.069	88	0.319	0.038	92
3.0	0.111	0.099	0.041	96	0.117	0.021	98	0.107	0.029	97	0.125	0.021	88
4.0	0.003	0.015	0.020	68	0.016	0.015	36	0.016	0.017	52	0.016	0.014	15

Estimation at $x = 1.5$ with bias correction

		$\eta = (0.4, 0.2)$			$\eta = (0.4, 0.6)$			$\eta = (0.8, 0.2)$			$\eta = (0.8, 0.6)$		
y	$f(x,y)$	mean	rmse	cl	mean	rmse	cl	mean	rmse	cl	mean	rmse	cl
−4.0	1.695	1.463	0.338	64	1.467	0.274	40	1.475	0.285	52	1.476	0.244	25
−3.0	2.900	2.720	0.390	67	2.740	0.243	66	2.710	0.296	68	2.727	0.213	50
−2.6	3.746	3.639	0.441	73	3.607	0.271	63	3.580	0.333	67	3.567	0.238	53
−2.2	5.048	4.706	0.533	68	4.712	0.425	45	4.694	0.451	60	4.690	0.400	22
−1.8	7.056	6.269	0.950	41	6.267	0.843	9	6.272	0.891	30	6.258	0.826	1
−1.4	9.951	8.375	1.687	10	8.344	1.652	0	8.302	1.696	0	8.302	1.669	0
−1.0	13.631	10.687	3.023	0	10.789	2.868	0	10.732	2.942	0	10.769	2.876	0
−0.6	17.591	13.307	4.351	0	3.748	13.849	0	13.204	4.416	0	3.753	13.841	0
−0.2	21.087	64.880	43.818	0	—	—	—	64.892	43.819	0	—	—	—
0.2	9.776	35.630	25.888	0	—	—	—	35.718	25.961	0	—	—	—
0.6	8.331	6.232	2.170	2	0.000	8.331	0	6.247	2.120	0	0.000	8.331	0
1.0	6.049	4.399	1.703	2	4.384	1.684	0	4.410	1.670	0	4.415	1.645	0
1.4	3.743	2.657	1.141	8	2.675	1.084	0	2.658	1.115	1	2.677	1.075	0
1.8	1.974	1.408	0.606	17	1.411	0.581	1	1.414	0.579	1	1.417	0.565	0
2.2	0.887	0.656	0.286	51	0.642	0.263	17	0.650	0.267	33	0.640	0.257	0
2.6	0.340	0.261	0.135	72	0.260	0.106	49	0.259	0.108	59	0.255	0.098	28
3.0	0.111	0.093	0.063	77	0.099	0.035	85	0.096	0.049	73	0.097	0.028	77
4.0	0.003	0.015	0.026	69	0.015	0.019	49	0.017	0.022	53	0.014	0.016	37

Table 5.8b: Scenario d2 (continued)

Estimation at $x = 2.5$

		$\eta = (0.4, 0.2)$			$\eta = (0.4, 0.6)$			$\eta = (0.8, 0.2)$			$\eta = (0.8, 0.6)$		
y	$f(x,y)$	mean	rmse	cl	mean	rmse	cl	mean	rmse	cl	mean	rmse	cl
−4.0	3.157	2.896	0.383	98	2.918	0.297	84	2.760	0.442	69	2.774	0.402	18
−3.0	5.219	4.758	0.619	88	4.783	0.507	61	4.570	0.707	52	4.605	0.637	9
−2.6	6.402	5.767	0.802	78	5.786	0.683	49	5.566	0.895	38	5.587	0.839	2
−2.2	7.883	6.936	1.056	69	6.915	1.011	12	6.756	1.168	13	6.738	1.164	0
−1.8	9.748	8.145	1.703	28	8.173	1.609	0	8.003	1.790	3	8.025	1.738	0
−1.4	12.085	9.654	2.518	6	9.635	2.476	0	9.538	2.588	1	9.544	2.554	0
−1.0	14.958	11.249	3.768	0	11.305	3.674	0	11.267	3.721	0	11.294	3.676	0
−0.6	18.396	13.135	5.309	0	18.695	0.464	96	13.141	5.280	0	18.667	0.367	95
−0.2	22.415	87.664	65.268	0	—	—	—	87.293	64.886	0	—	—	—
0.2	1.310	41.572	40.277	0	—	—	—	42.510	41.209	0	—	—	—
0.6	1.116	1.193	0.237	95	3.880	2.769	0	1.621	0.533	12	4.344	3.230	0
1.0	0.810	0.780	0.176	97	0.791	0.104	95	1.084	0.301	46	1.099	0.299	3
1.4	0.501	0.482	0.140	96	0.481	0.083	95	0.658	0.190	67	0.669	0.179	25
1.8	0.264	0.250	0.096	97	0.258	0.059	97	0.345	0.109	85	0.360	0.105	39
2.2	0.119	0.110	0.067	97	0.121	0.039	96	0.151	0.059	91	0.168	0.057	63
2.6	0.045	0.051	0.046	95	0.058	0.031	88	0.073	0.047	75	0.077	0.038	59
3.0	0.015	0.023	0.031	87	0.027	0.023	81	0.028	0.026	81	0.034	0.024	53
4.0	0.000	0.008	0.020	81	0.009	0.013	53	0.009	0.016	65	0.009	0.012	31

Estimation at $x = 2.5$ with bias correction

		$\eta = (0.4, 0.2)$			$\eta = (0.4, 0.6)$			$\eta = (0.8, 0.2)$			$\eta = (0.8, 0.6)$		
y	$f(x,y)$	mean	rmse	cl	mean	rmse	cl	mean	rmse	cl	mean	rmse	cl
−4.0	3.157	2.953	0.542	80	2.969	0.348	75	2.916	0.400	77	2.924	0.312	59
−3.0	5.219	4.885	0.759	76	4.814	0.596	57	4.815	0.600	64	4.796	0.516	43
−2.6	6.402	5.816	1.028	68	5.841	0.740	49	5.812	0.806	62	5.805	0.686	31
−2.2	7.883	7.014	1.171	63	6.957	1.037	30	7.009	1.029	46	6.961	0.986	10
−1.8	9.748	8.209	1.804	42	8.179	1.659	7	8.212	1.665	14	8.173	1.622	0
−1.4	12.085	9.727	2.605	25	9.665	2.489	1	9.687	2.514	4	9.635	2.489	0
−1.0	14.958	11.268	3.844	3	11.280	3.729	0	11.284	3.752	0	11.317	3.672	0
−0.6	18.396	13.097	5.422	0	3.274	15.138	0	13.135	5.334	0	3.406	14.997	0
−0.2	22.415	73.311	50.960	0	—	—	—	73.582	51.201	0	—	—	—
0.2	1.310	23.602	22.369	0	—	—	—	23.505	22.238	0	—	—	—
0.6	1.116	1.054	0.359	74	0.000	1.116	0	1.064	0.270	74	0.000	1.116	0
1.0	0.810	0.679	0.326	73	0.667	0.222	61	0.680	0.244	72	0.678	0.179	53
1.4	0.501	0.427	0.234	80	0.400	0.160	67	0.424	0.184	69	0.398	0.146	48
1.8	0.264	0.222	0.158	79	0.204	0.112	71	0.218	0.130	76	0.207	0.090	61
2.2	0.119	0.106	0.102	91	0.089	0.070	69	0.094	0.086	65	0.088	0.057	68
2.6	0.045	0.045	0.066	91	0.049	0.042	85	0.054	0.058	77	0.047	0.036	63
3.0	0.015	0.026	0.047	81	0.025	0.029	78	0.024	0.034	80	0.022	0.021	71
4.0	0.000	0.009	0.029	86	0.010	0.019	66	0.009	0.022	74	0.009	0.015	46

Table 5.9: Scenario d3 — The empirical mean (columns 3, 6, 9, 12) of the estimator $\hat{f}_n^{\Delta,\eta}(x,y)$ (resp., bias-corrected estimator $\hat{f}_n^{\Delta,\eta}(x,y) - \hat{\gamma}_n^{\Delta,\eta}(x,y)$) based on 100 samples (up to time $t = 2500$ with $\Delta = 0.001$) is compared to the true value (col. 2) of $f(x,y)$ given by eq. (5.1.2). In addition, the root mean squared error (rmse; cols. 4, 7, 10, 13) and the empirical confidence level (cl; cols. 5, 8, 11, 14) in percent of the estimated 95%-confidence interval given by eq. (5.1.5) are presented.

Estimation at $x = 0$

		$\eta=(0.4,0.2)$			$\eta=(0.4,0.6)$			$\eta=(0.8,0.2)$			$\eta=(0.8,0.6)$		
y	$f(x,y)$	mean	rmse	cl	mean	rmse	cl	mean	rmse	cl	mean	rmse	cl
−4.0	0.007	0.024	0.023	55	0.026	0.022	14	0.052	0.048	0	0.056	0.050	0
−3.0	0.222	0.262	0.065	85	0.301	0.085	21	0.302	0.088	40	0.343	0.123	0
−2.6	0.679	0.726	0.087	94	0.815	0.145	20	0.766	0.102	71	0.857	0.181	0
−2.2	1.774	1.824	0.152	93	1.967	0.212	33	1.859	0.137	83	2.004	0.239	2
−1.8	3.948	3.961	0.213	96	4.125	0.212	67	3.971	0.141	94	4.141	0.210	31
−1.4	7.486	7.298	0.331	89	7.478	0.172	93	7.285	0.276	84	7.457	0.121	94
−1.0	12.099	11.804	0.451	87	11.747	0.397	60	11.728	0.448	70	11.673	0.449	16
−0.6	16.661	16.012	0.744	71	23.403	6.745	0	15.924	0.784	27	23.310	6.651	0
−0.2	19.552	189.703	170.153	0	—	—	—	189.604	170.053	0	—	—	—
0.2	19.552	189.826	170.275	0	—	—	—	189.672	170.121	0	—	—	—
0.6	16.661	16.124	0.647	80	23.479	6.822	0	15.972	0.743	32	23.332	6.673	0
1.0	12.099	11.785	0.506	85	11.761	0.393	64	11.673	0.501	55	11.670	0.451	16
1.4	7.486	7.356	0.300	91	7.493	0.161	90	7.342	0.253	86	7.472	0.121	94
1.8	3.948	3.938	0.213	93	4.115	0.215	71	3.953	0.153	94	4.135	0.210	45
2.2	1.774	1.808	0.139	96	1.954	0.196	33	1.846	0.124	89	1.992	0.226	4
2.6	0.679	0.723	0.092	93	0.809	0.139	24	0.770	0.107	63	0.855	0.179	0
3.0	0.222	0.252	0.057	90	0.296	0.080	27	0.302	0.090	36	0.342	0.123	0
4.0	0.007	0.025	0.023	47	0.027	0.022	8	0.054	0.049	0	0.056	0.050	0

Estimation at $x = 0$ with bias correction

		$\eta=(0.4,0.2)$			$\eta=(0.4,0.6)$			$\eta=(0.8,0.2)$			$\eta=(0.8,0.6)$		
y	$f(x,y)$	mean	rmse	cl	mean	rmse	cl	mean	rmse	cl	mean	rmse	cl
−4.0	0.007	0.017	0.021	68	0.013	0.014	64	0.018	0.021	54	0.015	0.015	53
−3.0	0.222	0.245	0.087	71	0.241	0.050	73	0.241	0.062	76	0.241	0.039	70
−2.6	0.679	0.715	0.125	83	0.702	0.077	77	0.698	0.082	82	0.697	0.055	76
−2.2	1.774	1.797	0.200	81	1.790	0.142	69	1.785	0.160	75	1.793	0.108	69
−1.8	3.948	3.939	0.343	75	3.918	0.189	81	3.936	0.228	78	3.931	0.130	78
−1.4	7.486	7.259	0.483	77	7.303	0.345	57	7.269	0.369	68	7.314	0.262	50
−1.0	12.099	11.833	0.620	78	11.817	0.418	62	11.853	0.471	71	11.806	0.370	53
−0.6	16.661	16.082	0.830	64	0.000	16.661	0	16.099	0.705	55	0.000	16.661	0
−0.2	19.552	0.000	19.552	0	—	—	—	0.000	19.552	0	—	—	—
0.2	19.552	0.000	19.552	0	—	—	—	0.000	19.552	0	—	—	—
0.6	16.661	16.222	0.714	78	0.000	16.661	0	16.163	0.655	56	0.000	16.661	0
1.0	12.099	11.866	0.698	65	11.807	0.448	62	11.783	0.546	62	11.782	0.389	45
1.4	7.486	7.326	0.457	79	7.328	0.293	74	7.367	0.333	76	7.361	0.233	69
1.8	3.948	3.904	0.337	70	3.893	0.215	72	3.914	0.258	71	3.908	0.155	66
2.2	1.774	1.783	0.209	79	1.771	0.125	76	1.774	0.159	77	1.776	0.087	79
2.6	0.679	0.700	0.128	76	0.697	0.083	69	0.699	0.089	81	0.697	0.060	70
3.0	0.222	0.231	0.071	79	0.233	0.050	69	0.240	0.061	76	0.237	0.038	74
4.0	0.007	0.017	0.021	69	0.016	0.014	56	0.019	0.021	61	0.016	0.014	44

Table 5.9a: Scenario d3 (continued)

Estimation at $x = 1.5$

		$\eta = (0.4, 0.2)$			$\eta = (0.4, 0.6)$			$\eta = (0.8, 0.2)$			$\eta = (0.8, 0.6)$		
y	$f(x,y)$	mean	rmse	cl	mean	rmse	cl	mean	rmse	cl	mean	rmse	cl
−4.0	1.695	1.636	0.184	95	1.649	0.108	94	1.568	0.170	82	1.579	0.132	61
−3.0	2.900	2.845	0.215	98	2.895	0.117	95	2.726	0.226	83	2.769	0.155	72
−2.6	3.746	3.724	0.202	99	3.772	0.127	98	3.578	0.221	91	3.630	0.146	80
−2.2	5.048	4.997	0.275	95	5.078	0.177	94	4.846	0.275	86	4.934	0.170	83
−1.8	7.056	6.884	0.404	90	7.037	0.227	92	6.752	0.402	73	6.901	0.222	80
−1.4	9.951	9.731	0.474	89	9.824	0.272	91	9.646	0.424	84	9.724	0.283	73
−1.0	13.631	13.314	0.534	95	13.321	0.404	83	13.216	0.511	80	13.246	0.422	50
−0.6	17.591	17.011	0.819	77	25.112	7.528	0	17.018	0.692	66	25.056	7.468	0
−0.2	21.087	203.550	182.466	0	—	—	—	203.120	182.035	0	—	—	—
0.2	9.776	168.806	159.032	0	—	—	—	169.593	159.818	0	—	—	—
0.6	8.331	8.175	0.409	92	15.142	6.816	0	8.581	0.375	81	15.531	7.202	0
1.0	6.049	5.952	0.357	92	5.953	0.225	88	6.226	0.296	84	6.231	0.228	72
1.4	3.743	3.620	0.264	95	3.731	0.163	93	3.816	0.183	95	3.917	0.207	57
1.8	1.974	1.958	0.186	95	2.034	0.129	88	2.039	0.149	92	2.130	0.176	44
2.2	0.887	0.867	0.112	99	0.958	0.103	78	0.912	0.089	94	1.001	0.127	42
2.6	0.340	0.333	0.071	94	0.384	0.063	83	0.353	0.051	98	0.403	0.071	45
3.0	0.111	0.112	0.042	96	0.133	0.037	83	0.117	0.027	98	0.140	0.035	65
4.0	0.003	0.004	0.007	95	0.005	0.005	88	0.005	0.006	94	0.006	0.005	83

Estimation at $x = 1.5$ with bias correction

		$\eta = (0.4, 0.2)$			$\eta = (0.4, 0.6)$			$\eta = (0.8, 0.2)$			$\eta = (0.8, 0.6)$		
y	$f(x,y)$	mean	rmse	cl	mean	rmse	cl	mean	rmse	cl	mean	rmse	cl
−4.0	1.695	1.655	0.286	74	1.657	0.169	75	1.647	0.192	78	1.655	0.117	76
−3.0	2.900	2.884	0.332	78	2.895	0.188	84	2.866	0.233	80	2.870	0.137	79
−2.6	3.746	3.758	0.346	79	3.759	0.201	84	3.736	0.237	83	3.728	0.146	84
−2.2	5.048	5.028	0.430	82	5.023	0.253	80	5.025	0.312	80	5.002	0.204	78
−1.8	7.056	6.897	0.578	75	6.937	0.360	73	6.891	0.434	76	6.913	0.293	71
−1.4	9.951	9.775	0.682	74	9.739	0.439	76	9.782	0.508	78	9.720	0.358	63
−1.0	13.631	13.356	0.741	78	13.333	0.519	70	13.285	0.581	73	13.286	0.446	56
−0.6	17.591	16.995	1.099	70	0.000	17.591	0	17.077	0.838	63	0.000	17.591	0
−0.2	21.087	0.000	21.087	0	—	—	—	0.000	21.087	0	—	—	—
0.2	9.776	0.000	9.776	0	—	—	—	0.000	9.776	0	—	—	—
0.6	8.331	7.980	0.697	65	0.000	8.331	0	8.128	0.489	67	0.000	8.331	0
1.0	6.049	5.844	0.592	74	5.864	0.373	70	5.900	0.412	75	5.916	0.262	70
1.4	3.743	3.506	0.443	64	3.551	0.315	62	3.583	0.309	65	3.604	0.225	60
1.8	1.974	1.914	0.292	78	1.898	0.190	75	1.920	0.229	75	1.907	0.152	71
2.2	0.887	0.828	0.184	75	0.855	0.121	74	0.849	0.135	79	0.862	0.090	72
2.6	0.340	0.314	0.120	77	0.327	0.073	75	0.323	0.080	81	0.329	0.050	77
3.0	0.111	0.109	0.066	80	0.108	0.048	70	0.106	0.046	78	0.105	0.032	75
4.0	0.003	0.004	0.011	92	0.004	0.006	85	0.005	0.009	88	0.004	0.005	86

Table 5.9b: Scenario d3 (continued)

Estimation at $x = 2.5$

y	$f(x,y)$	$\eta = (0.4, 0.2)$ mean	rmse	cl	$\eta = (0.4, 0.6)$ mean	rmse	cl	$\eta = (0.8, 0.2)$ mean	rmse	cl	$\eta = (0.8, 0.6)$ mean	rmse	cl
−4.0	3.157	3.101	0.397	93	3.127	0.234	94	2.991	0.309	89	3.000	0.219	80
−3.0	5.219	5.121	0.468	96	5.137	0.289	95	4.925	0.445	83	4.954	0.325	69
−2.6	6.402	6.268	0.547	95	6.318	0.329	95	6.047	0.523	82	6.108	0.370	63
−2.2	7.883	7.701	0.534	97	7.769	0.327	95	7.460	0.557	84	7.539	0.405	72
−1.8	9.748	9.482	0.721	94	9.590	0.391	94	9.270	0.677	77	9.363	0.462	63
−1.4	12.085	11.759	0.718	95	11.837	0.455	93	11.635	0.625	87	11.691	0.475	80
−1.0	14.958	14.623	0.821	97	14.586	0.558	89	14.513	0.698	88	14.516	0.548	74
−0.6	18.396	17.705	1.107	90	26.307	7.925	0	17.663	0.950	74	26.231	7.843	0
−0.2	22.415	213.415	191.005	0	—	—	—	212.659	190.246	0	—	—	—
0.2	1.310	153.312	152.006	0	—	—	—	154.508	153.200	0	—	—	—
0.6	1.116	1.305	0.299	84	7.993	6.879	0	1.879	0.783	0	8.582	7.467	0
1.0	0.810	0.936	0.231	90	0.939	0.173	80	1.358	0.570	2	1.355	0.553	0
1.4	0.501	0.576	0.175	89	0.592	0.134	79	0.831	0.355	14	0.853	0.359	0
1.8	0.264	0.314	0.120	93	0.326	0.093	80	0.432	0.188	39	0.461	0.203	0
2.2	0.119	0.122	0.062	97	0.147	0.047	92	0.190	0.088	73	0.215	0.101	9
2.6	0.045	0.054	0.046	89	0.060	0.029	89	0.082	0.050	72	0.088	0.047	35
3.0	0.015	0.022	0.031	87	0.024	0.019	88	0.028	0.025	84	0.033	0.022	53
4.0	0.000	0.002	0.009	94	0.001	0.004	93	0.002	0.007	90	0.002	0.004	84

Estimation at $x = 2.5$ with bias correction

y	$f(x,y)$	$\eta = (0.4, 0.2)$ mean	rmse	cl	$\eta = (0.4, 0.6)$ mean	rmse	cl	$\eta = (0.8, 0.2)$ mean	rmse	cl	$\eta = (0.8, 0.6)$ mean	rmse	cl
−4.0	3.157	3.114	0.674	69	3.146	0.377	74	3.150	0.435	73	3.146	0.261	74
−3.0	5.219	5.193	0.739	78	5.165	0.447	77	5.159	0.563	77	5.170	0.318	75
−2.6	6.402	6.315	0.863	74	6.340	0.531	74	6.309	0.610	76	6.356	0.363	76
−2.2	7.883	7.784	0.795	81	7.794	0.528	78	7.750	0.595	83	7.784	0.372	79
−1.8	9.748	9.527	1.101	69	9.601	0.580	79	9.545	0.812	69	9.571	0.464	66
−1.4	12.085	11.763	1.072	83	11.772	0.675	78	11.879	0.810	77	11.829	0.504	77
−1.0	14.958	14.673	1.174	77	14.543	0.774	74	14.681	0.923	74	14.603	0.642	67
−0.6	18.396	17.784	1.548	74	0.000	18.396	0	17.738	1.193	64	0.000	18.396	0
−0.2	22.415	0.000	22.415	0	—	—	—	0.000	22.415	0	—	—	—
0.2	1.310	0.000	1.310	0	—	—	—	0.000	1.310	0	—	—	—
0.6	1.116	1.121	0.398	70	0.000	1.116	0	1.117	0.290	69	0.000	1.116	0
1.0	0.810	0.805	0.297	80	0.799	0.177	82	0.811	0.264	63	0.799	0.143	63
1.4	0.501	0.474	0.246	78	0.487	0.162	65	0.482	0.214	66	0.484	0.125	66
1.8	0.264	0.281	0.156	79	0.269	0.111	71	0.244	0.147	66	0.247	0.093	64
2.2	0.119	0.096	0.096	91	0.107	0.059	81	0.101	0.081	69	0.104	0.053	68
2.6	0.045	0.052	0.066	88	0.046	0.042	85	0.055	0.054	81	0.043	0.033	68
3.0	0.015	0.025	0.047	82	0.020	0.027	80	0.022	0.031	82	0.018	0.019	83
4.0	0.000	0.002	0.013	96	0.001	0.006	94	0.003	0.012	94	0.002	0.006	86

Table 5.10: Scenario d2 — The empirical mean (integrated) squared error (columns 4, 6, 8, 10) on the interval $[y_1, y_2]$ of the estimator $\hat{f}_n^{\Delta,\eta}(x,\cdot)$ for $f(x,y)$ given by eq. (5.1.2) based on 100 samples (with lag $\Delta = 0.01$ and up to time $n\Delta = 2500$) is presented. In addition, the standard deviation (columns 5, 7, 9, 11) of the squared errors are shown.

			$x=0$		$x=0.75$		$x=1.5$		$x=2.25$	
η	y_1	y_2	mse	sd	mse	sd	mse	sd	mse	sd
(0.1, 0.4)	−3.0	−1.5	0.263	0.129	0.266	0.168	0.743	0.466	2.053	1.200
	−1.5	−0.5	8.164	1.405	7.497	1.520	9.463	2.405	12.818	3.771
	0.5	1.5	7.900	1.813	7.145	1.454	2.552	0.751	0.207	0.124
	1.5	3.0	0.284	0.172	0.469	0.222	0.246	0.123	0.035	0.026
(0.2, 0.4)	−3.0	−1.5	0.223	0.094	0.209	0.120	0.604	0.260	1.973	0.874
	−1.5	−0.5	8.086	1.123	7.473	1.048	8.951	1.700	12.613	2.850
	0.5	1.5	8.179	1.091	7.026	1.054	2.581	0.515	0.150	0.097
	1.5	3.0	0.246	0.097	0.421	0.139	0.231	0.083	0.024	0.016
(0.4, 0.2)	−3.0	−1.5	0.273	0.091	0.197	0.082	0.640	0.256	2.101	0.746
	−1.5	−0.5	8.663	0.895	7.924	0.934	9.948	1.175	14.536	2.077
	−0.5	−0.3	2.758	0.406	2.754	0.385	3.466	0.519	4.648	0.757
	0.3	0.5	2.624	0.422	1.958	0.294	0.764	0.134	0.748	0.171
	0.5	1.5	8.694	1.002	7.433	0.795	2.476	0.398	0.122	0.073
	1.5	3.0	0.263	0.100	0.459	0.112	0.234	0.062	0.024	0.013
(0.4, 0.4)	−3.0	−1.5	0.225	0.066	0.173	0.073	0.582	0.195	1.857	0.538
	−1.5	−0.5	8.169	0.774	7.368	0.784	9.119	1.028	13.020	1.768
	0.5	1.5	8.257	0.756	7.113	0.753	2.349	0.356	0.099	0.049
	1.5	3.0	0.219	0.082	0.405	0.101	0.199	0.060	0.015	0.010

Table 5.11: Scenario d3 — The empirical mean (integrated) squared error (columns 4, 6, 8, 10) on the interval $[y_1, y_2]$ of the estimator $\hat{f}_n^{\Delta,\eta}(x,\cdot)$ for $f(x,y)$ given by eq. (5.1.2) based on 100 samples (with lag $\Delta = 0.001$ and up to time $n\Delta = 2500$) is presented. In addition, the standard deviation (columns 5, 7, 9, 11) of the squared errors are shown.

			$x=0$		$x=0.75$		$x=1.5$		$x=2.25$	
η	y_1	y_2	mse	sd	mse	sd	mse	sd	mse	sd
(0.1, 0.4)	−3.0	−1.5	0.073	0.055	0.103	0.078	0.282	0.207	0.675	0.501
	−1.5	−0.5	0.368	0.279	0.415	0.316	0.534	0.473	1.070	0.850
	0.5	1.5	0.397	0.340	0.312	0.254	0.222	0.190	0.109	0.089
	1.5	3.0	0.078	0.068	0.071	0.077	0.053	0.040	0.028	0.025
(0.2, 0.4)	−3.0	−1.5	0.032	0.032	0.059	0.045	0.128	0.082	0.372	0.235
	−1.5	−0.5	0.271	0.242	0.262	0.198	0.330	0.254	0.654	0.493
	0.5	1.5	0.270	0.207	0.225	0.176	0.137	0.130	0.065	0.065
	1.5	3.0	0.035	0.028	0.033	0.029	0.025	0.022	0.015	0.011
(0.4, 0.2)	−3.0	−1.5	0.037	0.020	0.054	0.034	0.124	0.062	0.425	0.209
	−1.5	−0.5	0.291	0.156	0.239	0.130	0.341	0.210	0.689	0.404
	−0.5	−0.3	0.133	0.114	0.116	0.099	0.156	0.168	0.215	0.233
	0.3	0.5	0.124	0.111	0.098	0.119	0.031	0.037	0.026	0.032
	0.5	1.5	0.247	0.139	0.244	0.148	0.102	0.060	0.077	0.055
	1.5	3.0	0.034	0.022	0.034	0.021	0.030	0.020	0.016	0.012
(0.4, 0.4)	−3.0	−1.5	0.027	0.021	0.035	0.026	0.067	0.038	0.204	0.151
	−1.5	−0.5	0.215	0.136	0.205	0.156	0.273	0.198	0.397	0.305
	0.5	1.5	0.241	0.152	0.196	0.130	0.054	0.051	0.039	0.047
	1.5	3.0	0.023	0.017	0.015	0.012	0.013	0.010	0.009	0.007

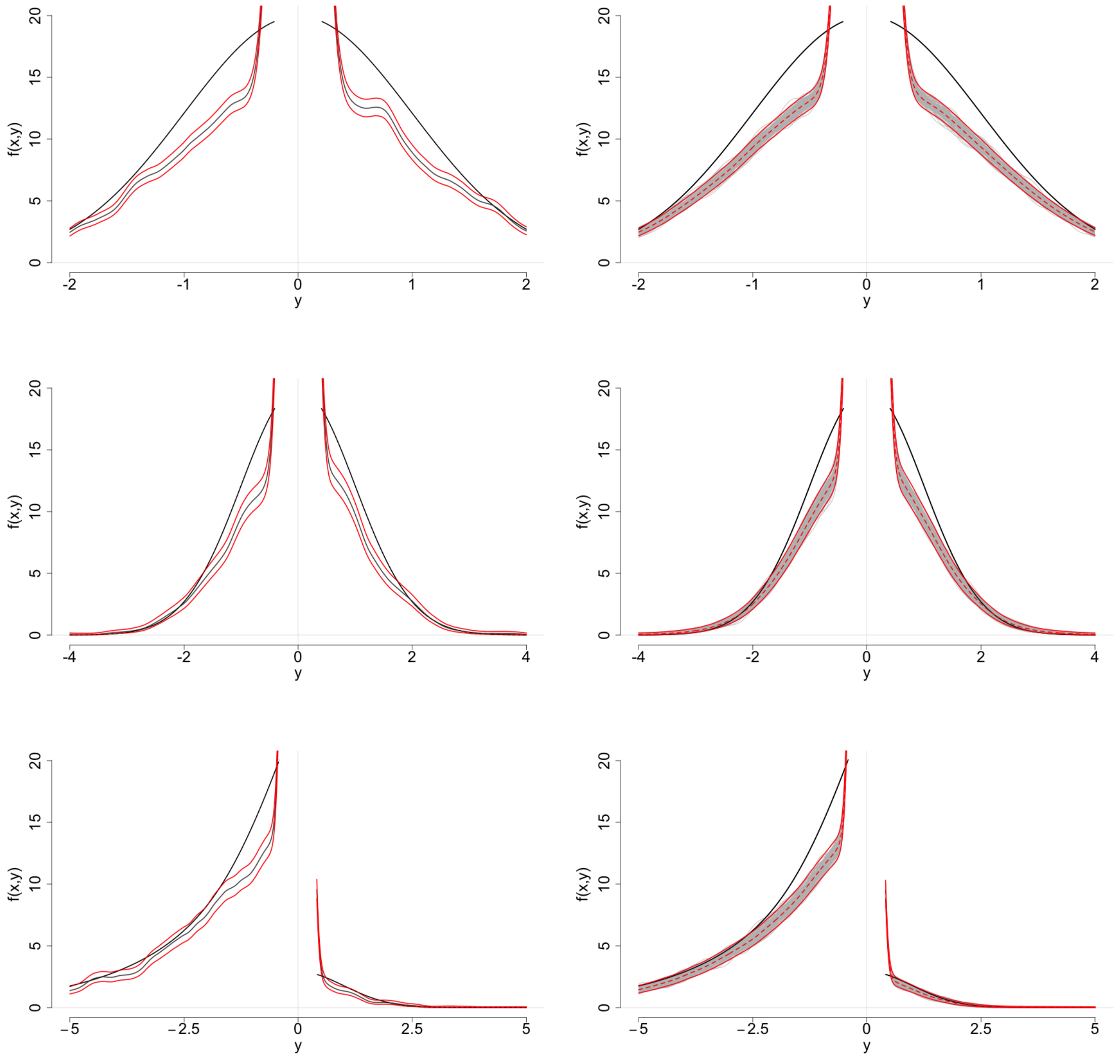

Figure 5.6: Scenario d2 — Estimation of the Lévy density $f(x,y)$ given by eq. (5.1.2) at $x = 0$ with $\eta = (0.4, 0.2)$ (top row), at $x = 0$ with $\eta = (0.1, 0.4)$ (middle row), and at $x = 2.25$ with $\eta = (0.4, 0.4)$ (bottom row) based on discrete observations with lag $\Delta = 0.01$ up to time $n\Delta = 2500$. Left: One typical estimate (grey) is compared to the true Lévy density (black). The upper and lower bounds of the estimated (pointwise) 95%-confidence intervals given by eq. (5.1.5) are shown in red. Right: Estimates based on 100 trajectories (grey) are compared to the true Lévy density (black). The (pointwise) mean of the estimates (red dashed line) and mean of the upper and lower bounds of the 95%-confidence intervals (red solid lines) are shown.

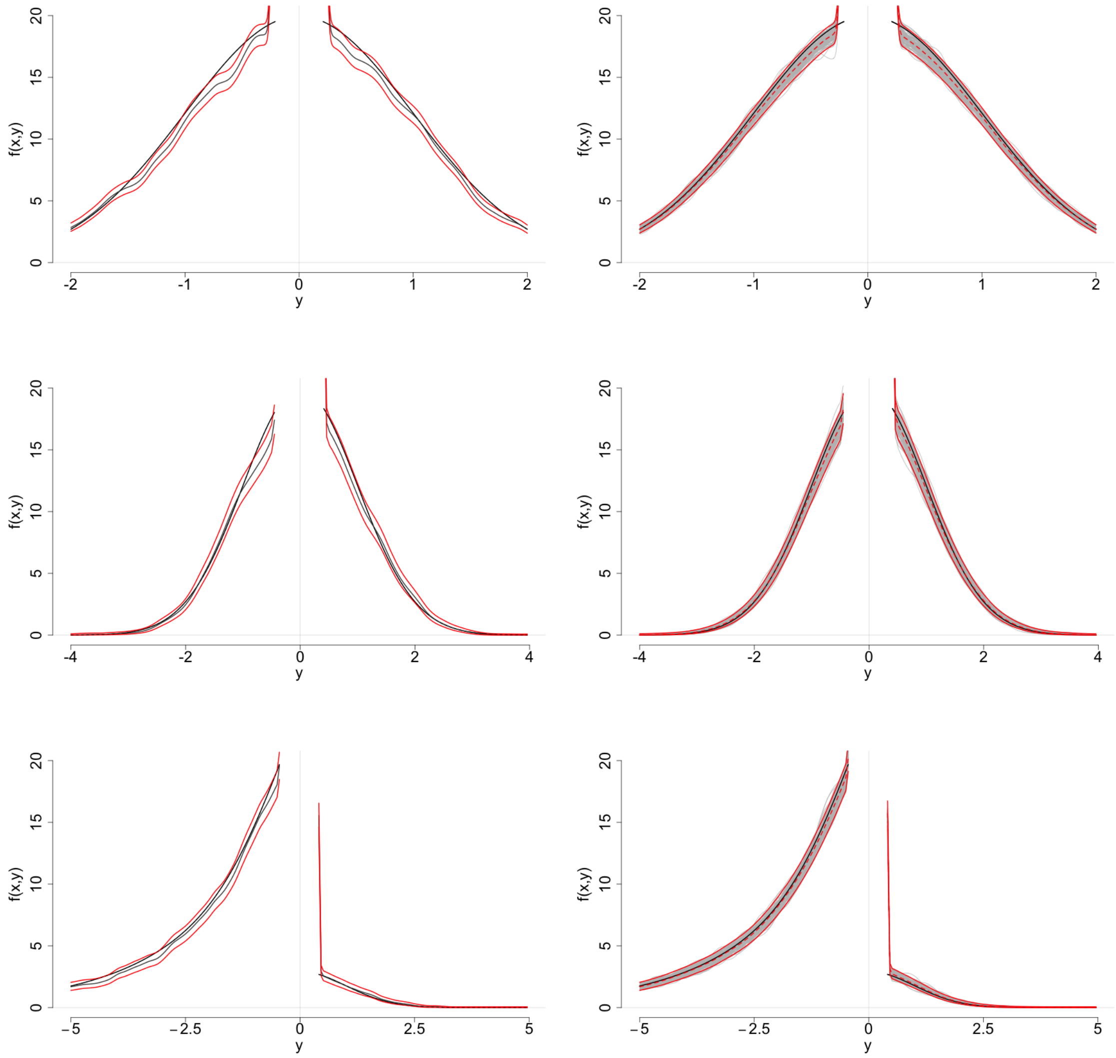

Figure 5.7: Scenario d3 — Estimation of the Lévy density $f(x, y)$ given by eq. (5.1.2) at $x = 0$ with $\eta = (0.4, 0.2)$ (top row), at $x = 0$ with $\eta = (0.1, 0.4)$ (middle row), and at $x = 2.25$ with $\eta = (0.4, 0.4)$ (bottom row) based on discrete observations with lag $\Delta = 0.001$ up to time $n\Delta = 2500$. Left: One typical estimate (grey) is compared to the true Lévy density (black). The upper and lower bounds of the estimated (pointwise) 95%-confidence intervals given by eq. (5.1.5) are shown in red. Right: Estimates based on 100 trajectories (grey) are compared to the true Lévy density (black). The (pointwise) mean of the estimates (red dashed line) and mean of the upper and lower bounds of the 95%-confidence intervals (red solid lines) are shown.

5.1.2 The infinite activity case

We investigated the performance of the estimator $\hat{f}_n^{\Delta,\eta}(x,y)$ based on the observation of the discrete sample $X_0(\omega), X_\Delta(\omega), \ldots, X_{n\Delta}(\omega)$. We chose the parameters of the process with infinite activity as reported in Table 5.12. The restriction of the Lévy density to the set $[-4,4]\times[-1,1]$ for these parameters is presented in Figure 5.8. We emphasise the singularity on the line $y=0$ and that f is not twice continuously differentiable on the set $\{-\xi,\xi\}\times\mathbb{R}$, which we indicated by the red dotted lines. We investigated the scenarios

d4) $t_4 = 1000$ and $\Delta_4 = 0.01$, that is 100 000 observations;

d5) $t_5 = 1000$ and $\Delta_5 = 0.0025$, that is 400 000 observations;

d6) $t_6 = 2500$ and $\Delta_6 = 0.0025$, that is 1 000 000 observations.

We simulated the process with the Euler scheme; as step length, we chose 1/10-th of the observation time-lag Δ. Given the value $X_{k\Delta/10}$, we simulated a stable increment with Lévy density $y \mapsto f(X_{k\Delta/10}, y)$ and a Brownian increment with drift $-bX_{k\Delta/10}$ and volatility c. Iteratively, we obtained an approximate sample $X_0, X_{\Delta/10}, \ldots, X_{n\Delta}$. Finally, we only kept every tenth observation.

For one simulated sample of each scenario d4–d6, we present the increments $(X_{(k-1)\Delta}, \Delta_k^n X)$ in Figure 5.9. Our first observation is with regard to the bias correction. We recall that we use a kernel which is of second-order. Since, for each x, the true Lévy density is convex as a function in y, on the one hand, we have that the bias from kernel smoothing is positive. On a neighbourhood of zero, depending on $\Delta > 0$, on the other hand, our estimator $\hat{f}_n^{\eta,\Delta}(x,\cdot)$ is always concave by construction. On this neighbourhood, consequently, the estimated bias correction always has the wrong sign. In the following, we focus on the uncorrected estimates only.

We compare our estimates $\hat{f}_n^{\Delta,\eta}(x,y)$ in terms of their functional properties. Just as in the finite activity case, we observe a significant influence of the bandwidth choice. In scenario d4, for instance, we observe that $\eta_1 > 0.2$ (resp., $\eta_1 > 0.3$) is necessary to obtain reasonable estimates at $x=0$ (resp., at $x=2$). On the set $\{|y| \le \eta_2 + 0.3\}$, the bias due to discretisation is dominant. At $x=0$, we obtain good estimates on the sets $\{0.5 \le |y| \le 1\}$ and $\{0.75 \le |y| \le 4\}$ for the bandwidth choices $\eta = (0.2, 0.2)$ and $\eta = (0.4, 0.4)$, respectively. At $x = 2$, we obtain good estimates on the sets $\{-3.5 < |y| < -0.75\}$ and $\{0.75 < y < 1.5\}$ for $\eta = (0.4, 0.4)$. In scenario d5, where the observation time-lag is one quarter of the time-lag of scenario d4, first, we observe that the bias due to discretisation is dominant on the set $\{|y| \leq \eta_2 + 0.2\}$. Apart from the improvement for $|y|$ small, the estimates in scenario d5 are similar to those of scenario d4. Finally, we observe that, for scenarios d5 and d6 where the observation time-lag is equal, the set on which the bias due to discretisation is dominant coincides. Nevertheless, the estimation for $|y|$ large improves significantly. At $x = 0$, we obtain very good estimates on the sets $\{0.4 < |y| < 3\}$ and $\{0.6 < |y| < 5\}$ for $\eta = (0.4, 0.2)$ and $\eta = (0.2, 0.4)$, respectively. At $x = 2$, we obtain very good estimates on the sets $\{-4 < y < -0.6\}$ and $\{0.6 < y < 2\}$ for $\eta = (0.2, 0.4)$. We present the estimates corresponding to these observations in Figures 5.10 to 5.12.

In summary, on the one hand, we have seen that larger bandwidths give better estimates in terms of variability and the degree of smoothing for $|y|$ large. On the other hand, smaller bandwidths allow for more reasonable estimates closer to zero than larger ones. Moreover, increasing the number of observations without reducing the observation time-lag does not give better estimates close to zero.

Table 5.12: Parameters for the characteristics $(B, C, \mathfrak{n})$ given by eqs. (5.1.1) and (5.1.3)

b	c	ξ	α
1	1	3	0.9

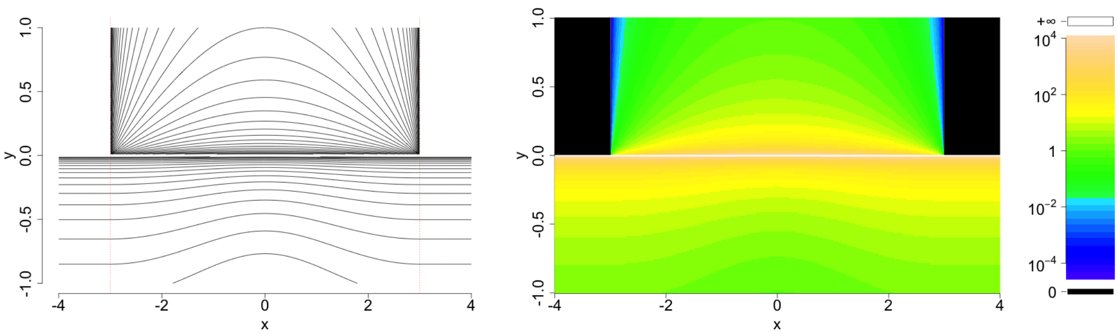

Figure 5.8: Contour plot (left) and topographical image plot (right) with legend (far right) of the restriction of the of the Lévy density $(x, y) \mapsto f(x, y)$ given by eq. (5.1.3) with parameters as in Table 5.12 to the set $[-4, 4] \times [-1, 1]$. The distance of the contour lines and the colour scheme are in logarithmic scale. The dotted red lines indicate the set $\{-\xi, \xi\} \times \mathbb{R}$ on which f is not twice continuously differentiable.

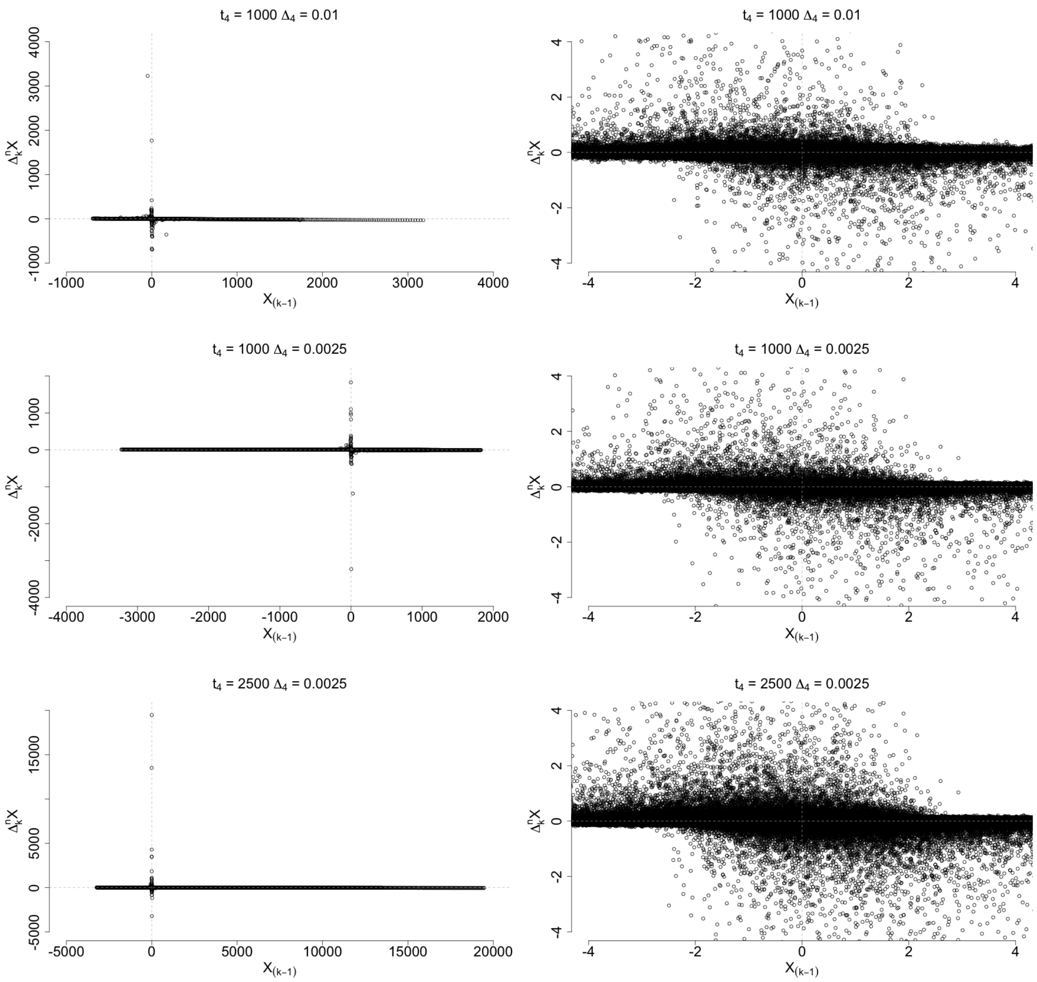

Figure 5.9: Increments $(X_{(k-1)\Delta}, \Delta_k^n X)$ of one simulated sample of scenarios d4 (top), d5 (middle), and d6 (bottom). Left: All increments. Right: Increments on the set $[-4,4] \times [-4,4]$.

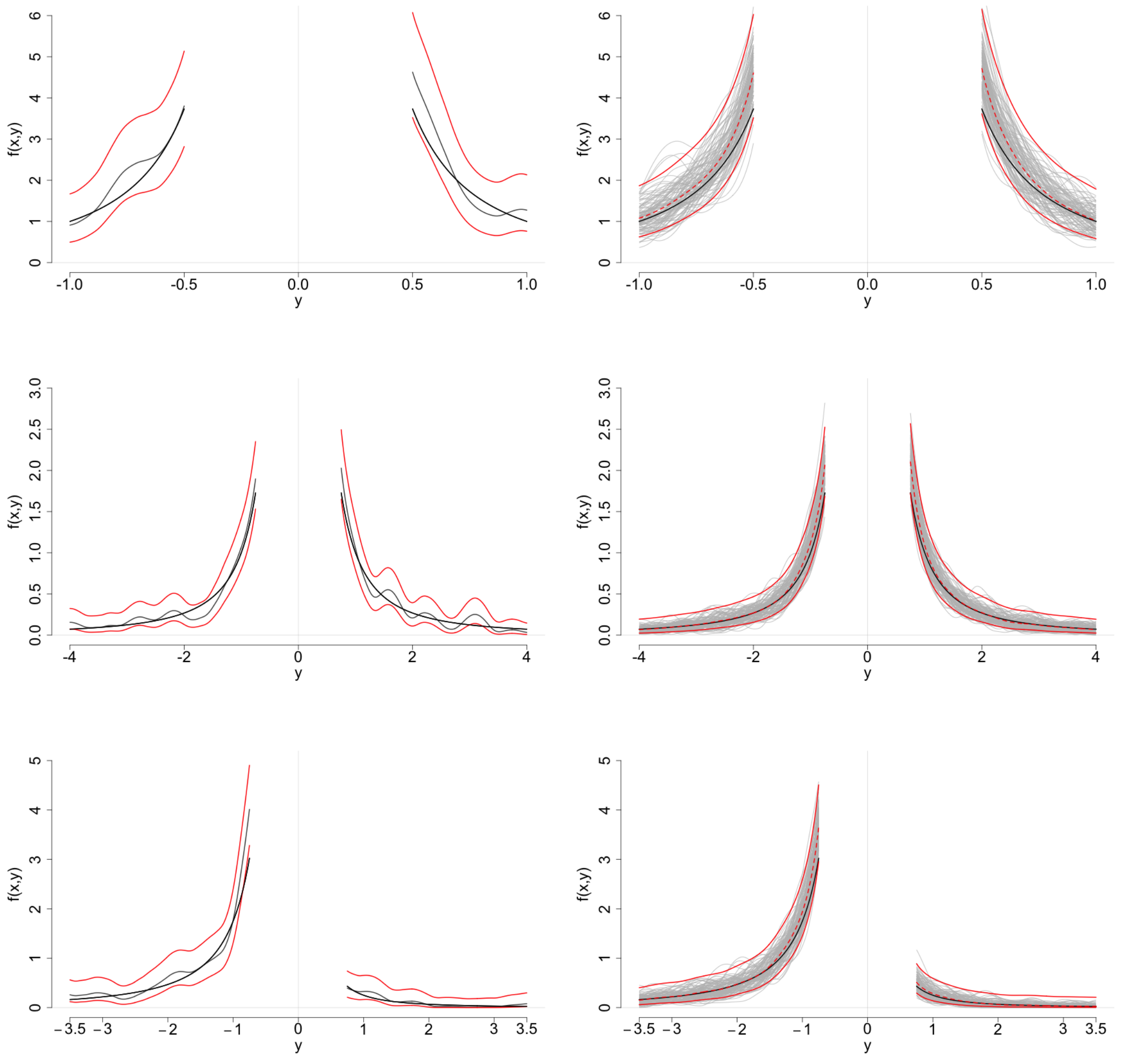

Figure 5.10: Scenario d4 — Estimation of the Lévy density $f(x,y)$ given by eq. (5.1.3) at $x = 0$ with $\eta = (0.2, 0.2)$ (top row), at $x = 0$ with $\eta = (0.4, 0.4)$ (middle row), and at $x = 2$ with $\eta = (0.4, 0.4)$ (bottom row) based on discrete observations with lag $\Delta = 0.01$ up to time $n\Delta = 1000$. Left: One typical estimate (grey) is compared to the true Lévy density (black). The upper and lower bounds of the estimated (pointwise) 95%-confidence intervals given by eq. (5.1.5) are shown in red. Right: Estimates based on 100 trajectories (grey) are compared to the true Lévy density (black). The (pointwise) mean of the estimates (red dashed line) and mean of the upper and lower bounds of the 95%-confidence intervals (red solid lines) are shown.

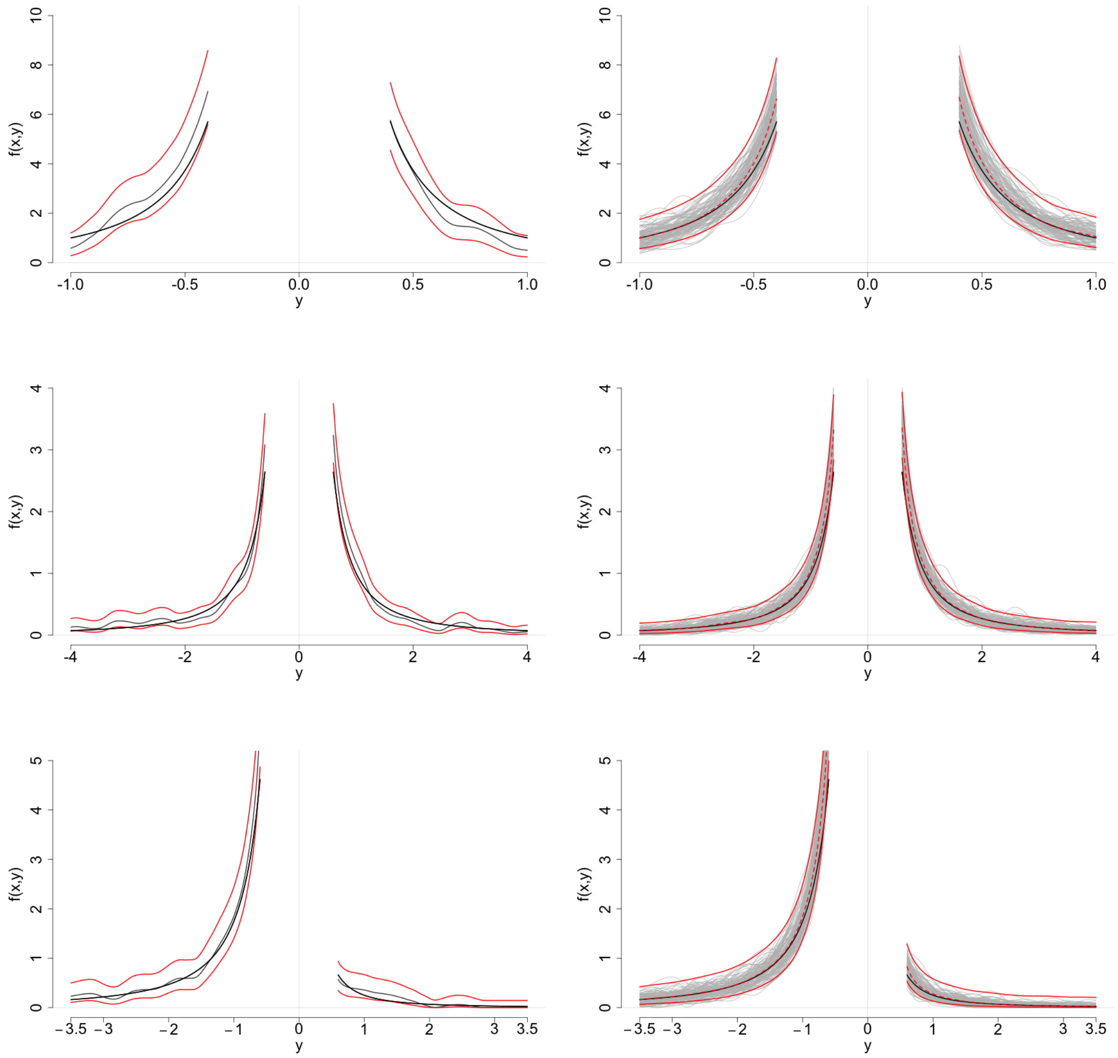

Figure 5.11: Scenario d5 — Estimation of the Lévy density $f(x,y)$ given by eq. (5.1.3) at $x=0$ with $\eta=(0.2,0.2)$ (top row), at $x=0$ with $\eta=(0.4,0.4)$ (middle row), and at $x=2$ with $\eta=(0.4,0.4)$ (bottom row) based on discrete observations with lag $\Delta=0.0025$ up to time $n\Delta=1000$. Left: One typical estimate (grey) is compared to the true Lévy density (black). The upper and lower bounds of the estimated (pointwise) 95%-confidence intervals given by eq. (5.1.5) are shown in red. Right: Estimates based on 100 trajectories (grey) are compared to the true Lévy density (black). The (pointwise) mean of the estimates (red dashed line) and mean of the upper and lower bounds of the 95%-confidence intervals (red solid lines) are shown.

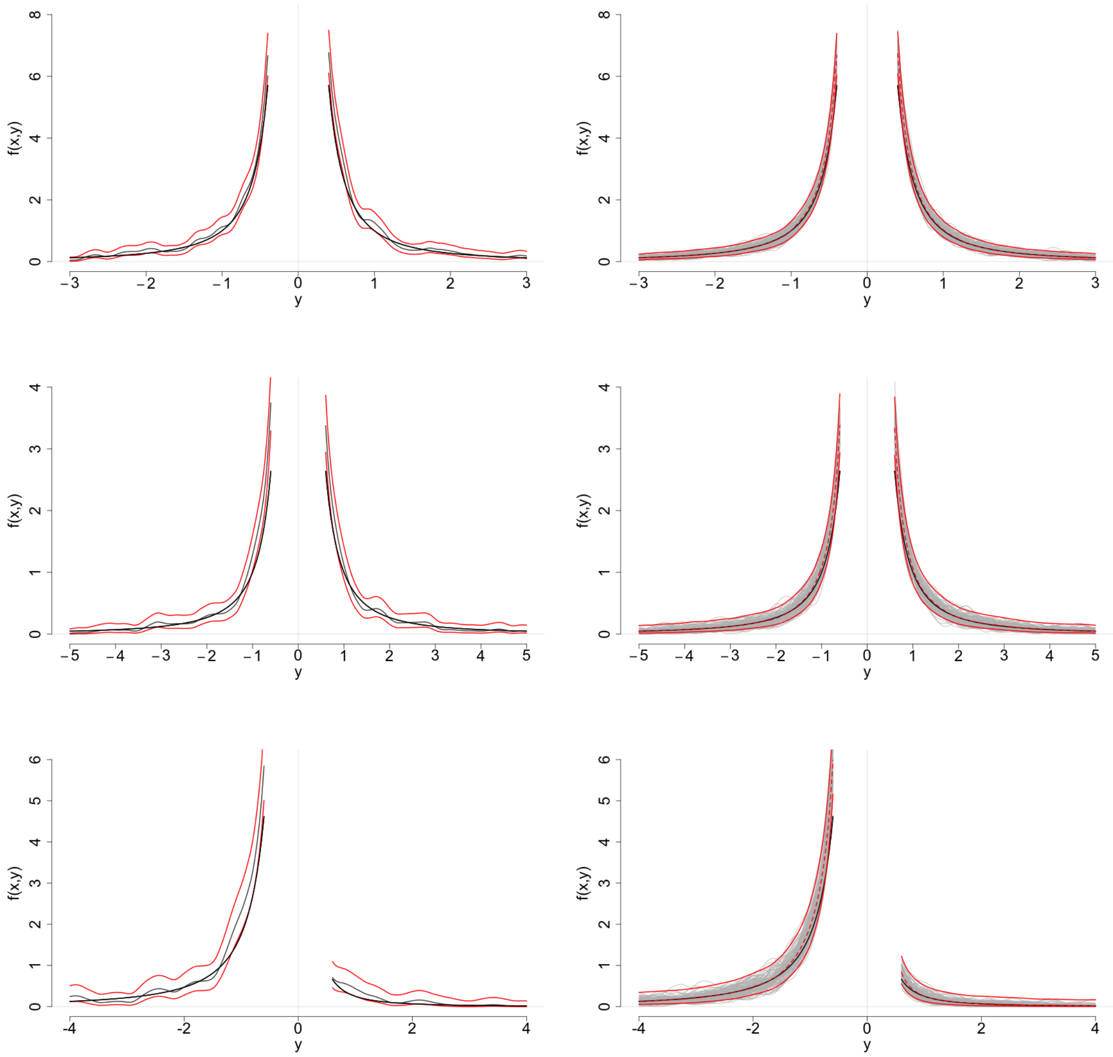

Figure 5.12: Scenario d6 — Estimation of the Lévy density $f(x, y)$ given by eq. (5.1.3) at $x = 0$ with $\eta = (0.4, 0.2)$ (top row), at $x = 0$ with $\eta = (0.2, 0.4)$ (middle row), and at $x = 2$ with $\eta = (0.2, 0.4)$ (bottom row) based on discrete observations with lag $\Delta = 0.0025$ up to time $n\Delta = 2500$. Left: One typical estimate (grey) is compared to the true Lévy density (black). The upper and lower bounds of the estimated (pointwise) 95%-confidence intervals given by eq. (5.1.5) are shown in red. Right: Estimates based on 100 trajectories (grey) are compared to the true Lévy density (black). The (pointwise) mean of the estimates (red dashed line) and mean of the upper and lower bounds of the 95%-confidence intervals (red solid lines) are shown.

5.2 Penalised projection estimation of the Lévy density of Lévy processes

This section is based on Section 4 of Ueltzhöfer and Klüppelberg (2011). The notation has been adjusted to fit the general notation of this thesis.

We have implemented the penalised projection estimation (PPE) method described in Section 2 of Ueltzhöfer and Klüppelberg (2011). Our estimator is based on piecewise quadratic polynomials: For every $m \in M = \mathbb{N}^*$, we denote by $\mathscr{D}_m$ the regular partition of a given domain of estimation $D \subseteq \mathbb{R}^*$ and define the sieve $\mathcal{S}_m$ by

$$\mathcal{S}_m := \Big\{ g \in L^2(D) : g_{|C} \text{ is a quadratic polynomial } \forall C \in \mathscr{D}_m \Big\}.$$

We note that the constants defined in Section 3.1 of Ueltzhöfer and Klüppelberg (2011) satisfy

$$\mathfrak{D}_m = 9m/\mathrm{vol}(D), \quad \mathfrak{D}'_m = 45m/\mathrm{vol}(D), \quad \text{and} \quad d_m = 3m.$$

Also, $M_n = \{1, \dots, \lfloor T_n \mathrm{vol}(D)/9 \rfloor\}$. In addition, the penalty constants in eq. (10) of Ueltzhöfer and Klüppelberg (2011) are set to $c_1 = 2$, $c_2 = 1$, $c_3 = 0.1$, and $c_4 = 0.5$. Although in practice, the penalty constants could be tuned to give better estimates in instances where Brownian motion is clearly present, here, we use the same constants whether Brownian motion is present or not. In doing so, we intend to emphasise the effect of Brownian motion on the PPE and the asymptotic behaviour of the PPE.

As a comparison, we also implemented the estimation procedure described in Sections 6 and 7 of Comte and Genon-Catalot (2009, 2011), respectively. We denote this estimator by SCE, which indicates the sinus cardinal (basis). Moreover, any notation referring to the latter procedure will be appended by the label *SC*. Let g^* denote the Fourier transform of a function g and let φ denote the sinus cardinal, that is, $\varphi(x) = \sin(\pi x)/(\pi x)$ with $\varphi(0) = 1$. For $m_{\mathrm{sc}} > 0$ the corresponding SC-projection space is given by $\mathcal{S}^{\mathrm{sc}}_{m_{\mathrm{sc}}} = \{g \in L^2(\mathbb{R} : \mathrm{supp}(g^*) \in [-\pi m_{\mathrm{sc}}, \pi m_{\mathrm{sc}}])\}$. The set $\{\varphi_{m_{\mathrm{sc}},k} : k \in \mathbb{Z}\}$, where $\varphi_{m_{\mathrm{sc}},k}(x) = \sqrt{m_{\mathrm{sc}}}\varphi(m_{\mathrm{sc}}x - k)$, forms an orthonormal basis of $\mathcal{S}^{\mathrm{sc}}_{m_{\mathrm{sc}}}$. Note that m_{sc} plays the role of a bandwidth and is unrelated to the m of our method.

Depending on whether Brownian motion is absent or present, the corresponding SCEs of $x \mapsto g^{\mathrm{sc}}(x) = xf(x)$ and $x \mapsto p^{\mathrm{sc}}(x) = x^3 f(x)$ are given by

$$\hat{g}^{\mathrm{sc}}_{m_{\mathrm{sc}}} = \sum_{k\in\mathbb{Z}} \hat{a}^{\mathrm{sc}}_{m_{\mathrm{sc}},k}\varphi_{m_{\mathrm{sc}},k} \quad \text{and} \quad \hat{p}^{\mathrm{sc}}_{m_{\mathrm{sc}}} = \sum_{k\in\mathbb{Z}} \hat{b}^{\mathrm{sc}}_{m_{\mathrm{sc}},k}\varphi_{m_{\mathrm{sc}},k},$$

respectively, where

$$\hat{a}^{\mathrm{sc}}_{m_{\mathrm{sc}},k} = \frac{1}{T_n}\sum_{j=1}^{T_n/\Delta_n} \Delta_j^n X \varphi_{m_{\mathrm{sc}},k}(\Delta_j^n X) \text{ and } \hat{b}^{\mathrm{sc}}_{m_{\mathrm{sc}},k} = \frac{1}{T_n}\sum_{j=1}^{T_n/\Delta_n} (\Delta_j^n X)^3 \varphi_{m_{\mathrm{sc}},k}(\Delta_j^n X).$$

The contrast values for the SCEs are equal to $-\sum_{k\in\mathbb{Z}}(\hat{a}^{\mathrm{sc}}_{m_{\mathrm{sc}},k})^2$ and $-\sum_{k\in\mathbb{Z}}(\hat{b}^{\mathrm{sc}}_{m_{\mathrm{sc}},k})^2$, and the respective penalty functions are defined by

$$\mathrm{pen}^{\mathrm{sc}}_n(m_{\mathrm{sc}}) = \frac{\kappa_{\mathrm{sc}} m_{\mathrm{sc}}}{T_n^2}\sum_{j=1}^{T_n/\Delta_n} (\Delta_j^n X)^2 \text{ and } \mathrm{pen}^{\mathrm{sc}}_n(m_{\mathrm{sc}}) = \frac{\kappa_{\mathrm{sc}} m_{\mathrm{sc}}}{T_n^2}\sum_{j=1}^{T_n/\Delta_n} (\Delta_j^n X)^6.$$

In analogy to Comte and Genon-Catalot (2009, 2011), we truncate the infinite sum in the definition of $\hat{g}^{\mathrm{sc}}_{m_{\mathrm{sc}}}$ and $\hat{p}^{\mathrm{sc}}_{m_{\mathrm{sc}}}$ to $\{k : |k| \leq 15\}$. In addition, m_{sc} is chosen from the set $\{0.1, 0.2, \ldots, 10\}$, and the constant in the penalties is set to $\kappa_{\mathrm{sc}} = 7.5$ if there is no Brownian motion and $\kappa_{\mathrm{sc}} = 3$ otherwise. As we are interested in the Lévy density itself, we transform the raw estimates $\hat{g}^{\mathrm{sc}}_{m_{\mathrm{sc}}}$ and $\hat{p}^{\mathrm{sc}}_{m_{\mathrm{sc}}}$ to $\hat{f}^{\mathrm{sc}}_{m_{\mathrm{sc}}}(x) = \hat{g}^{\mathrm{sc}}_{m_{\mathrm{sc}}}(x)/x$ and $\hat{f}^{\mathrm{sc}}_{m_{\mathrm{sc}}}(x) = \hat{p}^{\mathrm{sc}}_{m_{\mathrm{sc}}}/x^3$, respectively, and restrict them to the domain of estimation D from our method.

We simulated the following univariate models:

(i) a compound Poisson process with intensity 0.5 and exponentially distributed jumps with mean 1: $f(x) = 0.5e^{-x}\mathbb{1}_{\{x>0\}}$;

(ii) a superposition of (i) and Brownian motion with $\sigma = 0.5$;

(iii) a standard gamma process: $f(x) = x^{-1}e^{-x}\mathbb{1}_{\{x>0\}}$;

(iv) a superposition of (iii) and Brownian motion with $\sigma = 0.5$;

(v) a superposition of a bilateral gamma process with parameters $(\alpha^+, \beta) = (1, 1)$ and $(\alpha^-, \beta) = (0.7, 1)$ and Brownian motion with $\sigma = 0.5$:

$$f(x) = x^{-1}e^{-x}\mathbb{1}_{\{x>0\}} + x^{-1}e^{0.7x}\mathbb{1}_{\{x<0\}}.$$

Note that the parameters of the processes are taken as in Comte and Genon-Catalot (2009, 2011). In all cases, we investigated the scenarios

(1) $T_1 = 2500$, $\Delta_1 = 0.05$ (50 000 observations), and

(2) $T_2 = 5000$, $\Delta_2 = 0.02$ (250 000 observations).

Furthermore, we choose $D = [0.05, 10]$ in cases (i) and (iii), $D = [0.25, 10]$ in cases (ii) and (iv), and $D = [-10, -0.35] \cup [0.35, 10]$ in case (v).

As $f \in \mathcal{C}^\infty(D)$ in all cases (i–v), by Theorem 3.7 of Ueltzhöfer and Klüppelberg (2011), we expect the PPE based on piecewise quadratic polynomials to converge with rate $T^{-6/7}$. By Theorem 3.1 and Theorem 4.1 of Comte and Genon-Catalot (2009, 2011), respectively, we expect the SCE to converge with rates (i) $T^{-3/4}$, (ii) $T^{-7/8}$, (iii) $T^{-1/2}$, and (iv–v) $T^{-5/6}$ in the respective cases. We give a summary of the theoretical relative reductions corresponding to doubling T from scenario (1) to (2) in Table 5.13.

For the cases (ii) and (iv), moreover, we remark the probability for a purely Brownian increment to be bigger than the lower bound of D (0.25 in these cases) equals 1.27 % in scenario (1) and 0.02 % in scenario (2). Therefore, we expect significant distortions of the PPEs caused by Brownian motion in scenario (1), whereas these effects should remarkably diminish in scenario (2). In case (v), we have chosen D further away from the origin such that $\min_{x\in D}|x| = 0.35$. The probabilities that a purely Brownian increment falls into D, hence, are reduced to 0.174 % and $7.43 \cdot 10^{-5}$ %, respectively, in comparison to cases (ii) and (iv). Accordingly, we expect the impact of Brownian motion on the PPEs to be small in either scenario. We want to emphasise that the SCEs are based on all increments independent of their sizes. Hence, we do not expect a significant difference for the SCEs between cases (ii) and (iv) on the one hand, and case (v) on the other hand.

Results are given in Figure 5.13. Columns (a/b) correspond to the PPE, and columns (c/d) correspond to the SCE. Columns (a/c) show 50 estimated Lévy densities for scenario (1), and columns (b/d) show 50 estimated Lévy densities for scenario (2). On the y-axis, we restrict the plotted range to (i) $[0, 0.75]$, (ii) $[0, 1.5]$, (iii) $[0, 20]$, (iv) $[0, 6]$, and (v) $[0, 5]$. Near zero, some of the estimates fall out of this range and had to be truncated above. Nonetheless, all these cases are explicitly discussed below. Moreover, for the cases (iii–v) the Lévy densities and their estimates plotted over D are indistinguishable to the naked eye. However, there are notable differences over the range $D \cap [-2, 2]$ which we present here. In addition, for each scenario, we calculated the empirical MSE, that is, the mean of the empirical squared error of each estimate ($\|f - \hat{f}\|^2_{L^2(D)}$; cf. the definition at the beginning of Section 2.3 of Ueltzhöfer and Klüppelberg (2011)), and the mean of the estimated m and m_{sc} selected by penalisation. These are summarised in Table 5.14. In brackets, we give the standard deviation over 50 samples.

For (i), the pure compound Poisson process, we observe that all four plots exhibit high quality estimates with small variability. Near zero, the PPE follows the slope of the true Lévy density closely. The estimated values $\hat{f}^n_{\mathrm{pen}}(0.05)$ range between 0.36 and 0.50, and between 0.31 and 0.39 in scenarios (1) and (2), respectively. The conclusion that the true Lévy density is bounded (on $\mathbb{R}^*$) becomes obvious. For the SCE, this is not necessarily the case. The estimated values $\hat{f}^{\mathrm{sc}}(0.05)$ range between 0.79 and 1.25 in scenario (1), and between 1.54 and 1.95 in scenario (2). Compare these values with the true value $f(0.05) \approx 0.48$. Note also, the raw estimates $\hat{g}^{\mathrm{sc}}$ are, in general, non-zero at the origin. Without restriction to D, therefore, the SCEs of f have a pole at zero, whereas $f(x) \to 0.5$ as $x \to 0$. In contrast, the SCEs are smoother than the PPEs further away from zero. Moreover, the empirical mean squared errors of the PPEs and SCEs reduce by 52.6 %, and 95.2 % on average, respectively. For comparison, we refer to the asymptotic values summarised in Table 5.13.

For (ii), the superposition of (i) and Brownian motion, we observe highly unstable estimates in columns (a), (c) and (d), and high quality estimates in column (b) only. The distortions in the former cases are due to Brownian motion. However, in the latter case, the PPE behaves similar to case (i), where Brownian motion was absent. In particular, the PPE benefits considerably from the smaller observation time lag Δ_2. For the SCE this is not the case, as all observed increments are taken into account independent of their sizes. The values $\hat{f}^n_{\mathrm{pen}}(0.25)$ estimated by the PPE range between 4.47 and 6.01 in scenario (1), and between 0.36 and 0.48 in scenario (2). In contrast, the values $\hat{f}^{\mathrm{sc}}(0.25)$ estimated by the SCE range between -1.46 and +17.0, and between -0.50 and +7.95 in scenarios (1) and (2), respectively. The true value is $f(0.25) \approx 0.39$. Note that the raw estimates $\hat{p}^{\mathrm{sc}}$ are, in general, non-zero at the origin. Unrestricted, thus, the SCEs of f have a pole at zero, whereas $f(x) \to 0.5$ as $x \to 0$. Moreover, the defining property of Lévy densities, that is,

$$\int (|x|^2 \wedge 1)\hat{f}^{\mathrm{sc}}(x)\mathrm{d}x < \infty,$$

is violated.

For (iii), the standard gamma process, we observe that all four plots exhibit high quality estimates with small variability. The empirical mean squared error of the PPE is slightly smaller than the corresponding mean squared error of the SCE as the PPEs follow the slope near zero slightly closer. Further away from zero, though, the SCEs are smoother than the PPEs. We observe the empirical mean squared errors of the PPEs and SCEs reduce by 66.1 %, and 57.9 % on average, respectively. Again, we refer to the asymptotical values summarised in Table 5.13 for comparison.

For (iv), the superposition of (iii) and Brownian motion, similar to (ii) we observe unstable estimates in columns (a), (c) and (d), and estimates of higher quality in column (b) only. Once more, we observe distortions in the former cases due to Brownian motion. However, in the latter case, the PPE behaves very similar to case (iii), where Brownian motion was absent. The PPE benefits considerably from the smaller observation time lag Δ_2, whereas the SCE does not. The values $\hat{f}^n_{\mathrm{pen}}(0.25)$ estimated by the PPE range between 7.55 and 9.11 in scenario (1), and between 3.14 and 3.89 in scenario (2). In contrast, the values $\hat{f}^{\mathrm{sc}}(0.25)$ estimated by the SCE range between 3.50 and 11.6 with mean 8.02 in scenario (1), and between 3.49 and 10.4 with mean 6.78 in scenario (2). Compare these values to the true value $f(0.25) \approx 3.12$. Note also, the raw estimates $\hat{p}^{\mathrm{sc}}$ exhibit, in general, non-zero values at the origin for both scenarios (1) and (2). Analogously to case (ii), therefore, the unrestricted SCEs of f violate $\int (|x|^2 \wedge 1)\hat{f}^{\mathrm{sc}}(x)\mathrm{d}x < \infty$. Furthermore, the empirical mean squared errors of the PPEs and SCEs reduce by 98.3 %, and 57.9 % on average, respectively. For comparison, once more, we refer to the asymptotical values in Table 5.13.

For (v), the superposition of a bilateral gamma process and Brownian motion, we chose D further away from the origin in comparison to cases (ii) and (iv). The PPE exhibits a reasonable empirical mean squared error in both scenarios (1) and (2) as compared to case (iii), where Brownian motion was absent. Moreover, the PPEs are not too large to be plotted and, hence, not truncated. Although one may expect estimates like those in case (iv), changing D yields estimates like those in case (iii). The influence of purely Brownian increments is lowered considerably in comparison to case (iv). As for the SCE, in scenario (1) the estimated values $\hat{f}^{\mathrm{sc}}(-0.35)$ and $\hat{f}^{\mathrm{sc}}(0.35)$ range between 3.21 and 6.36, and between -1.87 and +3.25, respectively. In scenario (2), the SCEs' corresponding values range between 2.04 and 5.49, and between 1.43 and 4.35, respectively. Compare these values to the true values $f(-0.35) \approx 2.24$ and $f(0.35) \approx 2.01$. We note that the SCE does not benefit significantly from the change of D.

From a statisticians point of view, if Brownian motion is present, the choice of D appears to be crucial for a given scenario. In cases (ii) and (iv) above, if we choose a domain of estimation further away from the origin, e. g., $D = [0.35, 10]$, the distortions observed in scenarios (ii-1) and (iv-1) vanish and the plots look similar to cases (i-1) and (iii-1), respectively, where Brownian motion was absent. A practicable method, therefore, is to estimate σ first, e. g., as presented in Mancini (2005). Then, assuming $\hat{\sigma} = \sigma$, we determine D such that the probability for purely Brownian increments to fall into D is very small.

Having said that, there exists another provision despite changing D. Again for cases (ii) and (iv), we observe that the penalisation criterion chooses on average $m = 43.34$ and $m = 43.96$, respectively, in scenario (1), and $m = 3.50$ and $m = 25.58$ on average, respectively, in scenario (2). Although, the optimal m, that is, $m_n^\star$, increases with rate $T^{1/7}$ in these cases (cf. Proposition 3.5 of Figueroa-López (2009a)), the estimated m chosen by penalisation, in fact, decreases from scenario (1) to (2). Obviously, the relatively large amount of purely Brownian increments just above the threshold of 0.25 causes the penalised contrast to favour large m in scenario (1). Since we partition the domain equidistantly, only a few increments remain for each partition cell where a jump of corresponding size occurred. This increases the variance of our estimator significantly. If we increase the constants $c_1, \ldots, c_4$ in our penalty, the influence of Brownian motion is decreased such that smaller m, that is, coarser partitions, resulting in a smaller empirical mean squared error are chosen. In summary, not only the right choice of the domain of estimation D but the right balance between D and the penalty constants $c_1, \ldots, c_4$ is crucial.

Table 5.13: Summary of asymptotic rates of convergence (rows 1 and 3) and relative reduction of the MSE (rows 2 and 4) as T doubles from scenario (1) to (2) for the PPE (rows 1 and 2) and the SCE (rows 3 and 4) corresponding to the estimation of f for a CPP–Exp(1) with rate 0.5 (column i), a superposition of (i) and Brownian motion with $\sigma = 0.5$ (column ii), a standard gamma process (column iii), a superposition of (iii) and Brownian motion with $\sigma = 0.5$ (column iv), and a superposition of a bilateral gamma$(1, 1; 0.7, 1)$ process and Brownian motion with $\sigma = 0.5$ (column v).

	(i)	(ii)	(iii)	(iv)	(v)
PPE					
Asymptotic rate	$T^{-6/7}$	$T^{-6/7}$	$T^{-6/7}$	$T^{-6/7}$	$T^{-6/7}$
Rel. reduction ($T_2 = 2T_1$)	44.8 %	44.8 %	44.8 %	44.8 %	44.8 %
SCE					
Asymptotic rate	$T^{-3/4}$	$T^{-7/8}$	$T^{-1/2}$	$T^{-5/6}$	$T^{-5/6}$
Rel. reduction ($T_2 = 2T_1$)	40.5 %	45.5 %	29.3 %	43.9 %	43.9 %

Table 5.14: Summary of the estimation of f for a CPP–Exp(1) with rate 0.5 (row i), a superposition of (i) and Brownian motion with $\sigma = 0.5$ (row ii), a standard gamma process (row iii), a superposition of (iii) and Brownian motion with $\sigma = 0.5$ (row iv), and a superposition of a bilateral gamma(1,1;0.7,1) process and Brownian motion with $\sigma = 0.5$ (row v) by the PPE based on piecewise quadratic polynomials and the SCE.

X	(T_n, Δ_n)	$\overline{m_n^{\text{pen}}}$	$\overline{m_{\text{sc}}}$	$\overline{\text{se}}(\hat{f}^n_{\text{pen}})$	$\overline{\text{se}}(\hat{f}^{\text{sc}}_{m_{\text{sc}}})$	
(i)	$(2500, 0.05)$	2.92 (0.34)	0.96 (0.13)	0.876 (0.642)	8.065 (2.599)	$\times 10^{-3}$
	$(5000, 0.02)$	2.98 (0.14)	1.87 (0.35)	0.415 (0.209)	0.385 (0.271)	$\times 10^{-3}$
(ii)	$(2500, 0.05)$	43.34 (2.73)	0.47 (0.21)	0.752 (0.068)	1.527 (2.178)	
	$(5000, 0.02)$	3.50 (0.71)	0.55 (0.25)	0.007 (0.003)	5.397 (5.410)	$\times 10^{-1}$
(iii)	$(2500, 0.05)$	59.10 (4.40)	4.82 (0.41)	0.174 (0.052)	0.765 (0.133)	
	$(5000, 0.02)$	73.12 (8.68)	5.93 (0.31)	0.059 (0.018)	0.329 (0.053)	
(iv)	$(2500, 0.05)$	43.96 (4.54)	0.63 (0.27)	0.885 (0.091)	1.185 (0.747)	
	$(5000, 0.02)$	25.58 (5.40)	0.72 (0.24)	0.015 (0.005)	0.674 (0.488)	
(v)	$(2500, 0.05)$	26.36 (3.11)	0.29 (0.03)	0.137 (0.057)	5.679 (4.369)	$\times 10^{-1}$
	$(5000, 0.02)$	25.56 (3.00)	0.46 (0.23)	0.051 (0.012)	3.733 (2.003)	$\times 10^{-1}$

Notes: The empirical mean of the values for m chosen by penalisation, and the empirical MSE for each pair (T_n, Δ_n) are presented. Standard deviations over 50 samples are given within the brackets. The squared errors and their standard deviations are to be scaled by the factor in the last column.

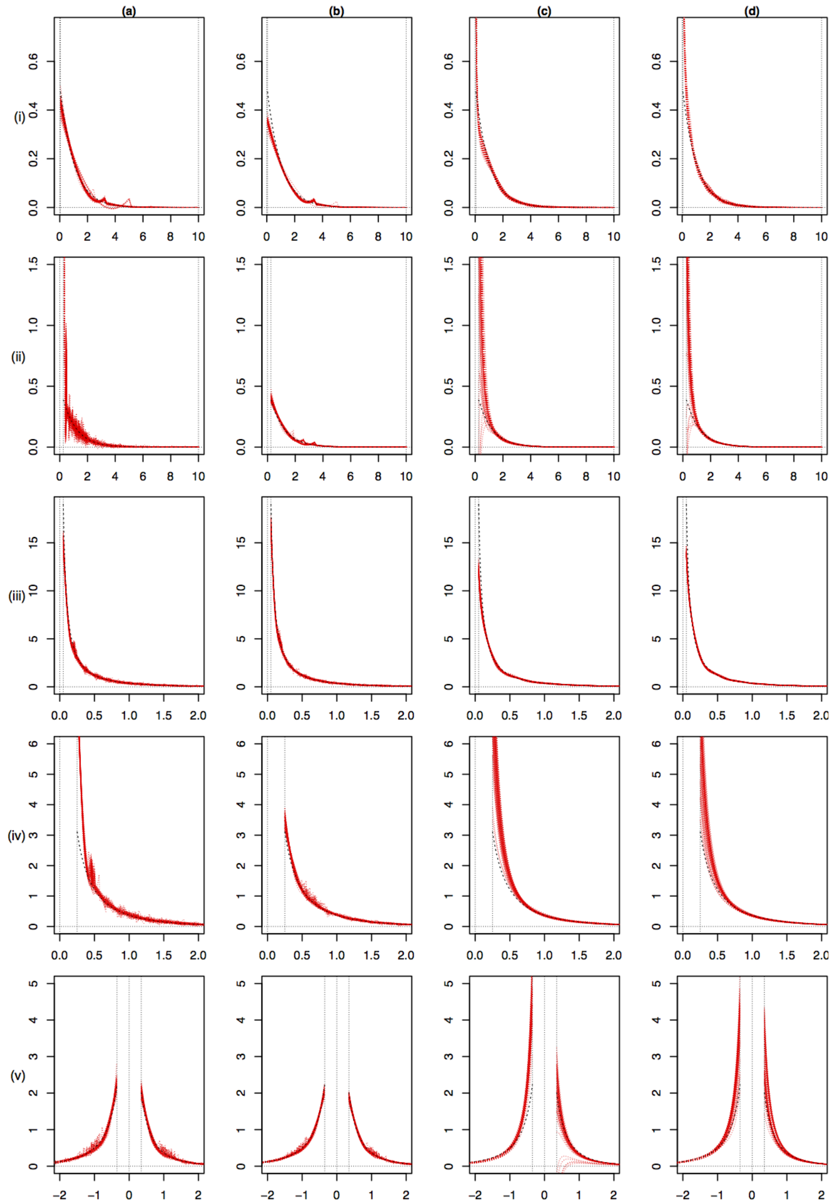

Figure 5.13: Estimation of f for a CPP–Exp(1) with intensity 0.5 (row i), a superposition of (i) and Brownian motion with $\sigma = 0.5$ (BM) (row ii), a standard gamma process (row iii), a superposition of (iii) and BM (row iv), and a superposition of a bilateral gamma$(1, 1; 0.7, 1)$ process and BM (row v). We present the true (dashed black) and 50 Lévy densities estimated (dotted red) by the PPE (columns a/b) and the SCE (column c/d), where $(T_n, \Delta_n) = (2500, 0.05)$ (columns a/c) and $(T_n, \Delta_n) = (5000, 0.02)$ (columns b/d).

6 Empirical modelling of the intermittency in atmospheric turbulence

This chapter is based on Ferrazzano and Ueltzhöfer (2012) and is joint work with Vincenzo Ferrazzano. The individual introduction in Section 6.1 has been edited for presentational purposes in view of the general introduction of this thesis (Chapter 1). Cross-references to the material presented in Chapter 2 have been added.

6.1 Introduction

Turbulence is the complex behaviour of a particle in a fluid, under certain conditions, described by its velocity. Its modelling is a long-standing problem in both physics and mathematics. The Navier–Stokes equations, the basic equations describing turbulence, are well-know since the 19th century. Actual comprehension of this phenomenon, however, is scarce. For an exhaustive account of the turbulence theory, we refer to the monographs of Frisch (1996) and Pope (2000). Since the seminal work of Kolmogorov (1941a,b, 1942, 1962), it is commonly accepted that turbulence can be regarded and analysed as a random phenomenon. In particular, the velocity of a turbulent flow can be modelled as a spatio-temporal stochastic process which preserves some statistical structure. The theory developed in a spatio-temporal setting is reduced to a time-series framework, utilising Taylor's frozen-field hypothesis (Pope, 2000, p. 223).

In this chapter, we focus on the modelling of the velocity $V = (V_t)_{t\in\mathbb{R}}$ of a weakly stationary turbulent flow along the main (longitudinal) flow direction at some fixed point in space. We note that virtually every observed turbulent flow displays several stylised facts. Experimental investigations highlighted that their magnitude depends only on a control parameter called the *Reynolds number*; it is proportional to the mean flow velocity over the kinematic viscosity. Our paramount aim is to advocate a statistical model, which is able to reproduce the following essential "intermittent" features of flows with a Reynolds number above a critical threshold, called *fully developed turbulent flows*: Firstly, the velocity increments display a distinctive clustering; the phenomenon originally called *intermittency*. In particular, the squared increments of turbulent flow velocities are significantly correlated; their auto-correlation function is positive and slowly decaying. Secondly, the velocity increments are semi-heavy tailed and display a distinctive scaling: On large time-scales, on the one hand, the distribution of the increments is approximately Gaussian. On small time-scales, on the other hand, the distribution develops exponential tails and is positively skewed. The skewness is given by *Kolmogorov's 4/5-law*. This law is a direct consequence of Navier-Stokes equations; it is one of the few exact and non-trivial results in turbulence theory.

Barndorff-Nielsen and Schmiegel (2008) proposed a causal continuous-time moving-average process (cCMA)

$$V_t = \bar{v} + \int_{-\infty}^{t} g_V(t-s)\mathrm{d}X_s, \tag{6.1.1}$$

driven by some – for a moment unspecified – normalised random orthogonal martingale measure $\mathrm{d}X$, as a suitable statistical model for a fully developed turbulent flow with mean $\bar{v} > 0$. In this model, the second-order properties depend only on the square-integrable moving-average kernel g_V; in particular, the auto-covariance γ_V and the (power) spectral density P_V have the simple forms

$$\gamma_V(t) = \int_0^\infty g_V(s+|t|)g_V(s)\mathrm{d}s \quad \text{and} \quad P_V(\omega) = \frac{1}{2\pi}\Big|\mathcal{F}g_V(\omega)\Big|^2,$$

where $\mathcal{F}h(\omega) := \int h(s)e^{i\omega s}\mathrm{d}s$ denotes the Fourier transform of $h \in L^2$. The driving noise X, henceforth called the *intermittency process*, accounts for all higher-order properties of V. In addition to the work of Barndorff-Nielsen and Schmiegel, Ferrazzano and Klüppelberg (2012) give a comprehensive study on the dependence of the moving-average kernel g_V on the Reynolds number of the turbulent flow. Thereupon, we build our intermittency model.

We advocate that the intermittency process $X = (X_t)_{t\in\mathbb{R}}$ is appropriately modelled by a two-sided, time-changed Lévy process

$$X_t := L_{\int_0^t Y_s \mathrm{d}s} \tag{6.1.2}$$

(cf. Section 2.2.3), where L is a purely discontinuous martingale with tempered stable Lévy measure (see Rosiński, 2007) and Y is itself a positive, ergodic, causal continuous-time moving average process – independent of L. In detail: We suppose there exists an $0 < \alpha < 2$ and there exist two completely monotone functions $q_+, q_- : \mathbb{R}_+ \to \mathbb{R}_+$, called the *tempering functions*, such that the Lévy measure of L is given by

$$F(\mathrm{d}x) = \frac{q_+(x)}{x^{1+\alpha}} \mathbb{1}_{\{x>0\}}(x)\mathrm{d}x + \frac{q_-(|x|)}{|x|^{1+\alpha}} \mathbb{1}_{\{x<0\}}(x)\mathrm{d}x; \tag{6.1.3}$$

and we suppose that Y is – again – a cCMA process

$$Y_t = \int_{-\infty}^{t} g_Y(t-s)\mathrm{d}Z_s, \tag{6.1.4a}$$

where the kernel belongs to the family

$$g_Y(t) = g_Y(t;\theta) = \begin{cases} C\sigma\zeta^{\nu-1}\exp(-\lambda\zeta) & \text{if } 0 < t < \zeta, \\ C\sigma t^{\nu-1}\exp(-\lambda t), & \text{if } t \geq \zeta, \end{cases} \tag{6.1.4b}$$

with strictly positive parameters $\theta = (\sigma, \nu, \lambda, \zeta)$, and $C = C_{\nu,\lambda,\zeta} > 0$ is a normalising constant such that $\int g_Y(t;\theta)^2 \mathrm{d}t = \sigma^2$; note that Z remains some unspecified Lévy subordinator with $\operatorname{Var} Z_1 = 1$ such that Y is independent of L, non-negative, and with finite fourth moment.

Recalling the “intermittent” features which we want to reproduce, we have a strong motivation for our model: The tempered stable distributions form a class of infinitely divisible distributions exhibiting the scaling behaviour observed in the intermittency. This class has been introduced to turbulence modelling by Mantegna and Stanley (1994) and Novikov (1994). In physics, the corresponding processes are known as *truncated Lévy flights*. Furthermore, since

$$\operatorname{Cov}\left[(X_{k\Delta} - X_{(k-1)\Delta})^2, (X_{(k+j)\Delta} - X_{(k+j-1)\Delta})^2\right] = \operatorname{Var}[L_1]^2 \operatorname{Cov}\left[\int_0^\Delta Y_{(k-1)\Delta+s}\mathrm{d}s, \int_0^\Delta Y_{(k+j-1)\Delta+r}\mathrm{d}r\right] \tag{6.1.5}$$

for every $k \in \mathbb{Z}$, $j \in \mathbb{N}$ and $\Delta > 0$, our model is, in principle, able to reproduce the clustering behaviour observed in the intermittency. Ferrazzano and Klüppelberg (2012), for instance, argues that the squared increments of the intermittency approximate the instantaneous rate of energy dissipation; that is, a quadratic functional of the spatial gradient of the three-dimensional velocity vector (Pope, 2000, eq. 5.128). In turbulence literature, such a proxy is often called *surrogate energy dissipation*. On a certain interval of time-lags, called the *scaling range*, the auto-correlation function of the true energy dissipation follows a power-law. Its exponent, called the *intermittency exponent*, measures the tendency of volatility increments to cluster (see Cleve, Greiner, Pearson, and Sreenivasan, 2004). The impact of surrogacy is studied, for example, by Cleve, Greiner, and Sreenivasan (2003). Their model (4) for the auto-covariance function of the surrogate energy dissipation inspired our parametric family for g_Y. The parameter ν in eq. (6.1.4b) is closely related to the intermittency exponent.

To model volatility clustering, time-changed Lévy processes have been introduced to mathematical finance by Geman, Madan, and Yor (2001) and Carr and Wu (2004). Likewise, these processes have been introduced to turbulence modelling by Barndorff-Nielsen, Blæsild, and Schmiegel (2004) and Barndorff-Nielsen and Schmiegel (2004, 2008). Since the processes $W \circ \int_0^\cdot Y_s \mathrm{d}s$ and $\int_0^\cdot Y_s^{1/2} \mathrm{d}W_s$ are indistinguishable in the case of a Brownian motion W, the relation to other stochastic volatility models like the BNS Ornstein–Uhlenbeck (Barndorff-Nielsen and Shephard, 2001, 2002) and the COGARCH model (Klüppelberg, Lindner, and Maller, 2004) is apparent.

We estimated our model from the so-called Brookhaven data set (Drhuva, 2000) which consists of measurements taken at the atmospheric boundary layer, about 35m above the ground. Brockwell et al. (2012) proposed a method to estimate the kernel g from an observed sample $V_0(\omega), V_\Delta(\omega) \ldots, V_{n\Delta}(\omega)$ of the velocity. Ferrazzano and Fuchs (2012) extended this method to estimate the increments $X_{k\Delta}(\omega) - X_{(k-1)\Delta}(\omega)$

of the intermittency process in addition. Treating these estimated increments as true observations, we estimated the time-change using a method of moment approach (see Kallsen and Muhle-Karbe, 2011). Next, we estimated the Lévy density of the Lévy process L combining the projection estimator of Figueroa-López (2009b, 2011) and the penalisation method which Ueltzhöfer and Klüppelberg (2011) studied in the case of a pure Lévy process. Under a constraint on the moments of the time-changed Lévy process, we also calculated least-squares fits of certain parametric families of tempered stable Lévy densities to our non-parametric estimate. We minimised an information criterion to find an optimal choice of parameters. In a simulation study, we compare a sample of increments from our intermittency model and the data. The fit of the empirical stationary distribution and the fit of the auto-correlation of the squared intermittency increments (that is, the clustering of large increments) is convincing.

We briefly outline this chapter: In Section 6.2 we present our model framework and its features; also, we describe the statistical methods which we apply for the estimation of the relevant quantities. In Section 6.3, we perform an empirical study of the Brookhaven data set. Finally, in Section 6.4, we compare our fitted model and the data set in a short simulation study.

6.2 The intermittency model and its estimation

In this section, we present our intermittency model from eqs. (6.1.2) to (6.1.4) in a rigorous manner. We outline its specific features in detail. In addition, we discuss the statistical methods for its estimation from discrete observations.

6.2.1 Modelling framework

On the filtered probability space $(\Omega, \mathscr{F}, (\mathscr{G}_t)_{t\in\mathbb{R}}, \mathbb{P})$, let $L = (L_t)_{t\in\mathbb{R}}$ be a two-sided, real-valued Lévy process without Gaussian part and Lévy measure F given by eq. (6.1.3). We suppose that $\mathbb{E}\, L_t = 0$; its characteristic exponent takes the form

$$\log \mathbb{E}\, \mathrm{e}^{\mathrm{i}tL_1} = \int \left(\mathrm{e}^{\mathrm{i}tx} - 1 - \mathrm{i}tx\right) F(\mathrm{d}x)$$

and, for $n \geq 2$, its cumulants are given by $c_n := c_n[L_1] := \int x^n F(\mathrm{d}x)$. In addition, let $Y = (Y_t)_{t \in \mathbb{R}}$ be the process given by eq. (6.1.4) – independent of L – such that

$$T_t := \int_0^t Y_s \mathrm{d}s \quad \text{is a } \mathscr{G}_t\text{-stopping time for all } t \in \mathbb{R}.$$

By Corollaire 10.12 of Jacod (1979) (recall Theorem 2.2.14), the time-changed Lévy process $X = (X_t)_{t \in \mathbb{R}}$ given by eq. (6.1.2) is a purely discontinuous martingale w. r. t. the filtration given by $\mathscr{F}_t := \mathscr{G}_{T_t}$; the process has càdlàg sample paths and $X_0 = 0$. We recall that the integer-valued random measure $\mathfrak{m}$ on $\mathbb{R} \times \mathbb{R}$ given by

$$\mathfrak{m}(\omega; \mathrm{d}t, \mathrm{d}x) := \sum_{\{s : \Delta X_s(\omega) \neq 0\}} \epsilon_{(s, \Delta X_s(\omega))}(\mathrm{d}t, \mathrm{d}x),$$

is called its *jump measure*. By Theorems 2.2.9 and 2.2.14, the increments $X_{t+s} - X_t$ can be represented as the stochastic integral

$$X_{t+s} - X_t = \iint_{]t, t+s] \times \mathbb{R}} x(\mathfrak{m} - \mathfrak{n})(\mathrm{d}r, \mathrm{d}x), \tag{6.2.1}$$

where $\mathfrak{n}(\omega; \mathrm{d}t, \mathrm{d}x) = Y_t(\omega)\mathrm{d}t F(\mathrm{d}x)$ is the predictable compensator of $\mathfrak{m}$. We call X the *intermittency process*.

The moments and auto-covariation function of the intermittency increments and their squares are determined by the cumulants of L and the mean and auto-covariation function of Y. Since the driving subordinator Z of Y satisfies $\operatorname{Var} Z_1 = 1$ by assumption, the auto-covariation function $\gamma_Y : \mathbb{R} \to \mathbb{R}_+$ of Y is given by the so-called auto-cross correlation of its moving-average kernel g_Y; that is

$$\gamma_Y(t; \theta) := \operatorname{Cov}[Y_0, Y_t] = \int_0^\infty g_Y(s; \theta) g_Y(|t| + s; \theta) \mathrm{d}s. \tag{6.2.2}$$

For $\Delta > 0$, we abbreviate $\gamma_Y^\Delta(k; \theta) := \int_{-\Delta}^{\Delta} |\Delta - s| \gamma_Y(k\Delta + s; \theta) \mathrm{d}s$. By construction, we have

$$\mathbb{E}\, X_t = 0 \quad \text{and} \quad \operatorname{Var} X_t = t c_2\, \mathbb{E}\, Y_0 \quad \text{for all } t \geq 0.$$

For identifiability, we suppose that $\mathbb{E}\, Y_0 = 1$; recalling that X is supposed to be the driving process of the cCMA model of eq. (6.1.1), furthermore, we assume that the intermittency process is normalised such that $c_2 = \operatorname{Var} X_1 = 1$. Under this assumptions, we note that

$$\mathbb{E}\, X_t^3 = tc_3 \quad \text{and} \quad \mathbb{E}\, X_t^4 = tc_4 + 3(\gamma_Y^t(0;\theta) + t^2) \tag{6.2.3}$$

(see, e. g., Barndorff-Nielsen and Shephard, 2006, Proposition 2). Dividing both sides of eq. (6.1.5) by $\operatorname{Var} X_\Delta^2$, moreover, we obtain that the auto-correlation $\rho_{X^2}^\Delta(k) := \operatorname{Corr}[(X_\Delta - X_0)^2, (X_{(k+1)\Delta} - X_{kt})^2]$ of the squared intermittency increments at lag $k \in \mathbb{N}^*$ is given by

$$\rho_{X^2}^\Delta(k;\theta) = \frac{\gamma_Y^\Delta(k;\theta)}{\Delta c_4 + 2\Delta^2 + 3\gamma_Y^\Delta(0;\theta)} \tag{6.2.4}$$

(see, e. g., Barndorff-Nielsen and Shephard, 2006, Proposition 5).

6.2.2 Estimation from discrete observations

We suppose to observe the intermittency process X on a discrete-time grid with sampling interval $\Delta > 0$. In particular for some $n \in \mathbb{N}^*$, we observe a realisation of the increments

$$\Delta_k^n X := X_{k\Delta} - X_{(k-1)\Delta}, \quad \text{for } k = 1, \ldots, n.$$

The jumps of the process and the time-change are latent.

Firstly, we turn to the estimation of the parameters $\theta = (\sigma, \nu, \lambda, \zeta)$ of the kernel g_Y given by eq. (6.1.4b). For typographical convenience, set $\mu_4 := \mathbb{E}\, X_\Delta^4$. Solving eq. (6.2.3) for c_4 and plugging it into eq. (6.2.4), we obtain

$$\rho_{X^2}^\Delta(k;\theta,\mu_4) = \frac{\gamma_Y^\Delta(k;\theta)}{\mu_4 - \Delta^2}, \tag{6.2.5}$$

where we emphasise the dependence on μ_4. We estimate the parameters θ, performing a least-squares fit of $\rho_{X^2}^\Delta(k;\theta,\mu_4)$ to its empirical version: In particular, let $\hat{\mu}_4$ (resp., $\hat{\rho}_{X^2}^\Delta$) denote the empirical fourth moment of the observed increments $\Delta_k^n X$ (resp., the empirical auto-correlation function of the squared increments $(\Delta_k^n X)^2$).

Then, our estimator of θ is given by

$$\hat{\theta} := \underset{\theta \in \mathbb{R}_+^4}{\arg\min} \sum_{k \in \mathbb{N}^*} \left| \hat{\rho}^{\Delta}_{X^2}(k) - \rho^{\Delta}_{X^2}(k; \theta, \hat{\mu}_4) \right|^2. \tag{6.2.6}$$

Secondly, we turn to the estimation of the Lévy density of the Lévy process L. The class of tempered stable Lévy measures, recall eq. (6.1.3), is truly of semi-parametric nature. By Bernstein (1929), every bounded, completely monotone function is the Laplace transform of some finite measure Q on $\mathbb{R}_+^*$; that is, $x \mapsto \int_0^\infty e^{-\lambda x} Q(d\lambda)$. In literature, parametric estimation of tempered stable Lévy densities is often based on the assumption that – for known orders $p_+, p_- \in \mathbb{N}^*$ – it belongs to the $2(p_+ + p_-) + 1$-parametric sub-family

$$f(x; \theta_{p_+,p_-}) = \begin{cases} x^{-1-\alpha} \sum_{k=1}^{p_+} c_k^+ \exp(-\lambda_k^+ x) & \text{for } x > 0, \\ |x|^{-1-\alpha} \sum_{k=1}^{p_-} c_k^- \exp(-\lambda_k^- |x|) & \text{for } x < 0, \end{cases} \tag{6.2.7}$$

where all parameters $\theta_{p_+,p_-} := (\alpha, (c_k^+, \lambda_k^+)_{k=1,\dots,p_+}, (c_k^-, \lambda_k^-)_{k=1,\dots,p_-})$ are strictly positive and, in addition, $\alpha < 2$. In view of the number of parameters, eq. (6.2.7) is frequently used for low orders. The issue of order selection is rarely addressed. We use a two-step approach to circumvent the latter issue: At first, we estimate the Lévy density employing an adaptive non-parametric method. Then, we calculate the least-squares fits of the parametric model eq. (6.2.7) to our non-parametric estimate for orders $p_+ + p_-$ up to some constant; we normalise our estimates so that the variance $\operatorname{Var} X_1$ of our fitted model is equal to one; and we penalise for deviations from the third and fourth empirical moments. Last, we minimise an information criterion to find our optimal choice for p_+ and p_-.

Various non-parametric estimators for the Lévy density of a Lévy process have been suggested in literature. Here, we focus on the projection estimator of Figueroa-López (2009b, 2011) which employs Grenander's method of sieves. In particular, let μ be some absolutely continuous Borel measure on $\mathbb{R}^*$, called the *reference measure*. We denote the μ-density of the Lévy measure F by f_μ; that is, $F(dx) = f_\mu(x)\mu(dx)$. Moreover, let $D \subset \mathbb{R}^*$ be a compact interval not containing zero, called the *domain of estimation*. Throughout, we suppose that f_μ is μ-square integrable over D. For each

$m \in \mathbb{N}^*$, let $K_m := \{y_{m,0}, \ldots, y_{m,m}\} \subset D$ be such that $\{[y_{m,k-1}, y_{m,k}] : k = 1, \ldots, m\}$ forms a μ-uniform partition of D. Then the space $\mathcal{S}_m := \mathcal{S}(3, K_m)$ of cubical $\mathcal{C}^2$-splines with control points K_m is an $m+3$-dimensional subspace of $\mathcal{L}^2(D, \mathscr{D}, \mu)$. The minimum contrast estimator $\hat{f}_\mu^m$ of f_μ w. r. t. the sieve $\mathcal{S}_m$, given by

$$\hat{f}_\mu^m := \operatorname*{arg\,min}_{h \in \mathcal{S}_m} \Big\{ -\frac{2}{n\Delta} \sum_{k=1}^n h(\Delta_k^n X) + \int \mathbb{1}_D(x) h(x)^2 \mu(\mathrm{d}x) \Big\}, \tag{6.2.8}$$

coincides with the respective projection estimator (cf. Lemma 2.1 of Ueltzhöfer and Klüppelberg, 2011).

By Figueroa-López (2009b), under some hypothesis on Y, the estimator $\hat{f}_\mu^m$ is consistent for the μ-density f_μ of the Lévy measure F if $n\Delta \to \infty$, $\Delta \to 0$ fast enough, and $m \to \infty$. For some related, pointwise central limit theorem, we refer to Figueroa-López (2011). For a numerically stable computation of $\hat{f}_\mu^m$, we construct the B-spline basis $\mathcal{B}_m := \{h_{m,j} : j = 1, \ldots, m+3\}$ of the space $\mathcal{S}_m$, and denote the Gramian matrix w. r. t. μ by $A = (a_{ij})_{i,j=1,\ldots,m+3}$; that is,

$$a_{ij} := \int \mathbb{1}_D(x) h_{m,i}(x) h_{m,j}(x) \mu(\mathrm{d}x).$$

Let $h_m : \mathbb{R} \to \mathbb{R}^{m+3}$ be the mapping with components $h_{m,j}$. Then the unique minimiser in eq. (6.2.8) is given by

$$\hat{f}_\mu^m(x) = \sum_{j=1}^{m+3} \hat{c}_{mj} h_{m,j}(x), \quad \text{where} \quad \hat{c}_m := A^{-1} \left(\frac{1}{n\Delta} \sum_{k=1}^n h_m(\Delta_k^n X) \right).$$

For each $m \in \mathbb{N}^*$, we are given an estimator $\hat{f}_\mu^m$ of f_μ on D; its associated contrast value is equal to $-(\hat{c}_m)^\intercal A \hat{c}_m$. As a data driven sieve selection method, we employ the penalisation method which Ueltzhöfer and Klüppelberg (2011) studied in the pure Lévy case. For $\zeta_1 \geq 1$ and $\zeta_2, \zeta_3, \zeta_4 > 0$, in particular, let $\mathrm{pen} : \mathbb{N}^* \to \mathbb{R}$ be the penalty function given by

$$\begin{aligned} \mathrm{pen}(m) := {} & \zeta_1 (n\Delta)^{-2} \operatorname{tr}\Big((h_m(\Delta_k^n X))_{k\leq n}^\intercal A^{-1} (h_m(\Delta_k^n X))_{k\leq n} \Big) \\ & + \zeta_2 \left(\frac{\mathfrak{D}_m}{n\Delta} \vee \frac{\mathfrak{D}_m^3}{(n\Delta)^4} \right) + \zeta_3 \left(\frac{\mathfrak{D}'_m}{n\Delta} \vee \frac{\mathfrak{D}'^3_m}{(n\Delta)^4} \right) + \zeta_4 \left(\frac{m+3}{n\Delta} \vee \frac{(m+3)^3}{(n\Delta)^4} \right), \end{aligned}$$

where

$$\mathfrak{D}_m := \sup_{h\in\mathcal{S}_m} \frac{\sup_{x\in D}|h(x)|^2}{\int_D h(x)^2\mu(\mathrm{d}x)}, \quad \text{and} \quad \mathfrak{D}'_m := \sup_{h\in\mathcal{S}_m} \frac{(\int_D |h'(x)|\mu(\mathrm{d}x))^2}{\int_D h(x)^2\mu(\mathrm{d}x)}. \tag{6.2.9}$$

Then the estimator $\hat{f}_\mu^{\hat{m}}$ where

$$\hat{m} := \operatorname*{arg\,min}_{m\in\mathbb{N}^*} \Big\{ -(\hat{c}_m)^{\intercal} A\hat{c}_m + \mathrm{pen}(m)\Big\}, \tag{6.2.10}$$

is called the *minimum penalised contrast estimator* of f_μ (w. r. t. the penalty pen).

In practice, we calculate an estimator $\hat{f}_\mu^{\hat{m}_+}$ on some domain $D_+ \subset \mathbb{R}_+^*$ and, separately, an estimator $\hat{f}_\mu^{\hat{m}_-}$ on some domain $D_- \subset \mathbb{R}_-^*$. For the Lebesgue density f of the Lévy measure F, we are thereby given the non-parametric estimate

$$\hat{f}(x) = \hat{f}_\mu^{\hat{m}_+}(x)\mu'(x)\mathbb{1}_{D_+}(x) + \hat{f}_\mu^{\hat{m}_-}(x)\mu'(x)\mathbb{1}_{D_-}(x). \tag{6.2.11}$$

In general, this estimate is not the restriction of a tempered stable Lévy density to the domain $D_+ \cup D_-$. For orders $p_+, p_- \in \mathbb{N}^*$ up to a specified order, we calculate the least-squares fit of the parametric family given by eq. (6.2.7) to our estimate given by eq. (6.2.11) under the constraint that the variance $\operatorname{Var} X_1$ of our fitted model equals one; and we penalise for deviations of the fitted third and fourth cumulant from the empirical ones (recall eq. (6.2.3)). In particular, our estimator of θ_{p_+,p_-} is given by

$$\begin{aligned}\hat{\theta}_{p_+,p_-} := \operatorname*{arg\,min}_{\{\theta_{p_+,p_-}: c_2(\theta_{p_+,p_-})=1\}} \Bigg\{ & \int_{D_+\cup D_-} |\hat{f}(x) - f(x;\theta_{p_+,p_-})|^2 \mathrm{d}x \\ & + \zeta\left(\left|\frac{c_3(\theta_{p_+,p_-})\Delta}{\hat{\mu}_3} - 1\right| + \left|\frac{c_4(\theta_{p_+,p_-})\Delta}{\hat{\mu}_4 - 3\Delta^2(\hat{\sigma}^2+1)} - 1\right|\right)\Bigg\},\end{aligned} \tag{6.2.12}$$

where

$$c_n(\theta_{p_+,p_-}) := \Gamma(n-\alpha)\sum_{k=1}^{p_+} c_k^+ (\lambda_k^+)^{\alpha-n} + (-1)^n\Gamma(n-\alpha)\sum_{k=1}^{p_-} c_k^-(\lambda_k^-)^{\alpha-n}$$

denotes the n-th cumulant of L_1 corresponding to the Lévy density $f(\,\cdot\,;\theta_{p_+,p_-})$, $\zeta > 0$ denotes some penalisation constant, $\hat{\mu}_n$ denotes the n-th empirical moment of the observed increments, and $\hat{\sigma}^2$ belongs to the fitted parameters of the kernel g_Y.

6.3 An empirical study of the Brookhaven wind speed data set

The Brookhaven turbulent wind speed data set consists of $n = 20 \cdot 10^6$ measurements taken at a frequency of 5000 Hz, covering a total time interval of 4000 s (66 min 40 s). A precise description of the data set is given in Drhuva (2000). We remark that the data set displays a Taylor's microscale Reynolds number of approximately 17 000 and is regarded a good representative of fully developed turbulence.

Ferrazzano and Fuchs (2012) proposed a method to estimate the increments $\Delta_k^n X(\omega)$ of the intermittency process from an observed sample $V_0(\omega), \ldots, V_{n\Delta}(\omega)$ of the velocity. This method – which, in principle, can be seen as applying an auto-regressive filter – has been employed to the Brookhaven data set. The filter has been chosen to involve measurements up to a time-lag of 78.8424 s; consequently, the estimated increments of the intermittency process cover a total time interval of 65 min 21.1276 s. For the remainder, we treat these estimates as if they were observed true increments of the intermittency process; henceforth, we refer to them as the "(intermittency) data".

We summarise the data in Figure 6.1 at the end of this chapter: At the top, we plotted the intermittency increments; the clustering of the increments is clearly observable. At the bottom, we present histograms of the increments $X_{k\Delta} - X_{(k-j)\Delta}$ of the intermittency process at time-lags $j\Delta$ for $j = 1, 1000, 10\,000$; for comparison, we also present the densities of a Gaussian random variable scaled to the empirical variance of the intermittency increments. At small-scale, we observe a heavy-tailed distribution; at large-scale, we observe an approximately Gaussian distribution.

For the estimation of the parameters θ of the moving-average kernel g_Y of the process Y given by eq. (6.1.4), we calculated the empirical auto-correlation function $\hat{\rho}^{\Delta}_{X^2}$ of the squared, observed intermittency increments $(\Delta_k^n X)^2$. We obtain from Cleve et al. (2004, Table I, data set "a2") that the surrogacy cutoff time is given by 0.5 ms, that is, 2.5Δ; for reasons stemming from physics, thus, we regard $\hat{\rho}^{\Delta}_{X^2}(k)$ reliable for $k \geq 3$ only. In addition, we observe a significant influence on the empirical auto-correlation function by non-stationary, large scale effects. For the estimation, thus, we consider $\hat{\rho}^{\Delta}_{X^2}(k)$ reliable up to one tenth of the de-correlation time – the lag $\hat{p} := 26\,698$ – only as well. We note that the empirical fourth moment of the increments is given by $\hat{\mu}_4 = 2.166 \cdot 10^{-6}$. With these considerations in mind, in practice, our estimator for the parameters θ (recall eq. (6.2.6)) is given by

$$\hat{\theta} := \underset{\theta \in \mathbb{R}_+^4}{\arg\min} \sum_{k=3}^{\hat{p}} \left| \hat{\rho}_{X^2}^{\Delta}(k) - \rho_{X^2}^{\Delta}(k;\theta,\hat{\mu}_4) \right|^2,$$

where $\rho_{X^2}^{\Delta}(k;\theta,\hat{\mu}_4)$ is given by eq. (6.2.5). We remark that no closed-form solution is known for the auto-correlation $\gamma_Y(\cdot;\theta)$ of Y given by eq. (6.2.2). In practice, hence, we utilised the convolution theorem, and employed the numerical approximation

$$\Big(\gamma_Y(k\Delta/100;\theta)\Big)_{|k|\leq 250\hat{p}} \approx \mathcal{D}^{-1} \left| \mathcal{D}\left[\Big(g_Y(k\Delta/100;\theta) \Big)_{k=0,\dots,500\hat{p}} \right] \right|^2,$$

where we sampled g_Y with a 100-times higher frequency and on a 5-times longer interval than used afterwards; $\mathcal{D}$ denotes the discrete Fourier transform, and $|\cdot|^2$ is understood component-wise. Our estimate is summarised in Figure 6.2 and Table 6.1. We present the empirical auto-correlation function $\hat{\rho}_{X^2}^{\Delta}(k)$ (black points) for the lags $k = 1,\dots,\hat{p}$ and compare it to the estimated auto-correlation function $\rho_{X^2}^{\Delta}(k;\hat{\theta},\hat{\mu}_4)$ (red solid line). We observe an excellent fit.

For the non-parametric estimation of the Lévy measure F of the Lévy process L, we chose $\mu(\mathrm{d}x) = x^{-4}\mathrm{d}x$ as reference measure. The main advantage of this particular choice over Lebesgue measure is that the μ-density f_μ of a tempered stable Lévy measure F does not have a pole at zero; in particular,

$$f_\mu(x) = q_+(x)x^{3-\alpha}\mathbb{1}_{\{x>0\}}(x) + q_-(|x|)|x|^{3-\alpha}\mathbb{1}_{\{x<0\}}(x).$$

We employed the minimum penalised contrast method as presented in eqs. (6.2.8) and (6.2.10) to estimate f_μ separately on $D_+ = [0.015, 0.8]$ and $D_- = [-0.8, -0.015]$; we chose the end points ± 0.8 as there are no observations with absolute value larger than 0.8 and we chose the end points $\pm 0.015 \approx \pm\sqrt{\Delta}$ to exclude an interval with a radius of about one standard deviation centred at the origin. As penalty coefficients we chose $\zeta_1 = 2$, $\zeta_2 = 1$, $\zeta_3 = 0.5$ and $\zeta_4 = 0.1$. As no closed-form solution is known for the constants $\mathfrak{D}_m$ and $\mathfrak{D}'_m$ in eq. (6.2.9), in practice, we replaced their true value by numerical approximations. In Table 6.2, we summarised the penalised contrast values (PCV) for the estimators $(\hat{f}_\mu^m)_{m=1,\dots,5}$ on D_+ and D_-.

We note that a local minimum is attained at $\hat{m}_+ = 4$ and $\hat{m}_- = 1$, respectively. For the Lebesgue density f of the Lévy measure F of the Lévy process L, we are given the non-parametric estimate

$$\hat{f}(x) := \hat{f}_\mu^{\hat{m}_+}(x)x^{-4}\mathbb{1}_{D_+}(x) + \hat{f}_\mu^{\hat{m}_-}(x)x^{-4}\mathbb{1}_{D_-}(x);$$

(recall eq. (6.2.11)). We observe that the non-parametric estimate oscillates around zero for $|x| > 0.3$; since no more than 591 observations – that is, 0.003% of the data – are larger in absolute value than 0.3, for the remainder, we consider our estimate reliable on the set $D = [-0.3, -0.015] \cup [0.015, 0.3]$ only.

For all orders $p_+ + p_- \leq 4$, we calculated the penalised least-squares estimator $\hat{\theta}_{p_+,p_-}$ defined in eq. (6.2.12); we replaced the integral over the set D by the discrete residual sum of squares given by

$$\mathrm{RSS}(\theta_{p_+,p_-}) := \sum_{k=15}^{300} \left|\hat{f}\,(x_k) - f(x_k;\theta_{p_+,p_-})\right|^2 + \left|\hat{f}\,(-x_k) - f(-x_k;\theta_{p_+,p_-})\right|^2,$$

where $x_k = k/1000$; and chose the penalty constant $\zeta = 5 \times 10^5$. To find an optimal choice for (p_+, p_-), we also calculated the corrected Akaike's information criterion

$$\mathrm{AIC_c}(p_+, p_-) := N\log(\mathrm{RSS}(\hat{\theta}_{p_+,p_-})/N) + 2K_{p_+,p_-} + \frac{2K_{p_+,p_-}(K_{p_+,p_-}+1)}{N - K_{p_+,p_-} - 1},$$

where $K_{p_+,p_-} := 2(p_+ + p_-) + 1$ is the number of parameters and $N := 572$ is the number of squared residuals evaluated for RSS.

Our results are summarised in Table 6.3. We observe that $\mathrm{AIC_c}$ is minimised for $p_+ = 1$ and $p_- = 2$. The fitted parameter $\hat{\alpha} = 1.39$ indicates that the paths of our process are of infinite variation. We present the corresponding estimated density in Figure 6.3. The parametric fit $f(x;\hat{\theta}_{1,2})$ (red solid line) is compared to the non-parametric estimate $\hat{f}(x)$ (black points). The estimates are indistinguishable to the eye.

6.4 Simulation study

This section is dedicated to a short simulation study. We simulate a sample of increments $\Delta_k^n X$ of our fitted intermittency model up to a terminal time of 1000 s and with a frequency of 5000 Hz. We specify the Lévy subordinator Z and simulate the moving-average process Y of the time change. Then we simulate the increments of the time-changed Lévy process $X_t = L(\int_0^t Y_s \mathrm{d}s)$ based on the realisation of Y. We compare our simulated path and the intermittency data.

In our model, the process Y given in eq. (6.1.4) is a causal continuous-time moving-average. We simulate from it approximating the stochastic integral defining the cCMA process by a stochastic Riemann sum: For $\Delta_1 > 0$, let $(\tilde{Y}_t^{\Delta_1})_{t\in\mathbb{R}}$ be given by

$$\tilde{Y}_t^{\Delta_1} := \sum_{k=-\infty}^{\lfloor t/\Delta_1 \rfloor} g_Y\Big((\lfloor t/\Delta_1 \rfloor - k)\, \Delta_1; \hat{\theta}\Big) \Big(Z_{k\Delta_1} - Z_{(k-1)\Delta_1} \Big) ; \tag{6.4.1}$$

then $\mathbb{E}\, |\tilde{Y}_t^{\Delta_1} - Y_t|^2 \to 0$ as $\Delta_1 \to 0$ for every $t \in \mathbb{R}$. To achieve a good approximation of Y on some time interval $[t_1, t_2]$, we simulate from the driving process Z on a much longer interval $[t_0, t_2]$ with $t_0 \ll t_1$ and with a smaller time-lag $\Delta_2 \ll \Delta_1$. Then we discard the samples on $[t_0, t_1]$ which are corrupted by numerical errors, and reduce the sampling frequency of the remainder.

We remark that the Lévy subordinator Z is left unspecified so far apart from its mean and variance. For this simulation study, we aim for a simple, yet likely choice for Z. For two reasons, we work with a Gamma process: Subordinators with infinite activity seem appropriate to us, since turbulent motion requires permanent injection of energy. And, the Gamma process is a well-understood subordinator which, moreover, is uniquely specified by its mean and variance.

We chose $\Delta_2 = 10^{-5}$. On the interval $] - 200, 1000]$, we simulated $1.2 \cdot 10^8$ independent and identically Gamma distributed increments $Z_{k\Delta_2} - Z_{(k-1)\Delta_2}$ with mean $\Delta_2/(\|g_Y(t;\hat{\theta})\|_1)$ and variance Δ_2, where $\|g_Y(t;\hat{\theta})\|_1 = 0.1385$. To calculate the convolution in eq. (6.4.1) we truncated the MA-kernel g_Y at $t^* = 200$, where $g_Y(t^*;\hat{\theta})/g_Y(0^+;\hat{\theta})) < 3.4 \cdot 10^{-7}$. We discarded the observations on the interval $]-120, 0]$ which are corrupted by numerical errors and down-sampled to a time-lag of $\Delta_1 = 1/5000$. Consequently, we obtained $5 \cdot 10^6$ (approximate) observations $\tilde{Y}_{k\Delta_1}$ on the interval $]0, 1000]$.

As a time-changed Lévy process, the intermittency process X has independent increments conditionally on Y. By eq. (6.2.1), moreover, we have

$$\log \mathbb{E}\left[\mathrm{e}^{\mathrm{i}u(X_{t+\Delta_1}-X_t)} \,\middle|\, Y\right] = \int_t^{t+\Delta_1} Y_s \mathrm{d}s \int \left(\mathrm{e}^{\mathrm{i}ux} - 1 - \mathrm{i}ux\right) f(x; \hat{\theta}_{p_+,p_-}) \mathrm{d}x.$$

For each k, approximating the increment of the time-change by $\Delta_1 \tilde{Y}_{k\Delta_1}$, we simulated the increment $X_{(k+1)\Delta_1} - X_{k\Delta_1}$ using the shot-noise representation (5.19) of Rosiński (2007). All jumps with absolute value larger than 10^{-6} where simulated exactly; the small jumps where approximated by a Gaussian random variable of appropriate variance. Consequently, we obtained a sample of $5 \cdot 10^6$ (approximate) increments $\Delta_k^n \tilde{X}$ on the interval $]0\,\mathrm{s}, 1000\,\mathrm{s}]$.

We present our simulation result in Figure 6.4. At the top, we plotted the increments of the intermittency at the sampling frequency of 5000 Hz. In comparison to the data as presented in Figure 6.1, we observe a convincing similarity. At the bottom, we compare the simulation and the data in more detail: On the left, we present a quantile-quantile plot comparing the empirical quantiles of the data (x-axis) to those of the simulation (y-axis). On the interval $[-0.3, 0.3]$, which carries more than 99.996% of the data, the fit is excellent. Since the least-square fitting of the Lévy density has been performed on the domain $[-0.3, -0.015] \cup [0.015, 0.3]$ only, we are very satisfied with the fit of the stationary distribution of the intermittency increments. On the right, we compare the empirical auto-correlation function of the squared intermittency data (black points) to the empirical auto-correlation of the square simulated increments (red solid line). Both axes are in logarithmic scale. Again, their agreement is excellent.

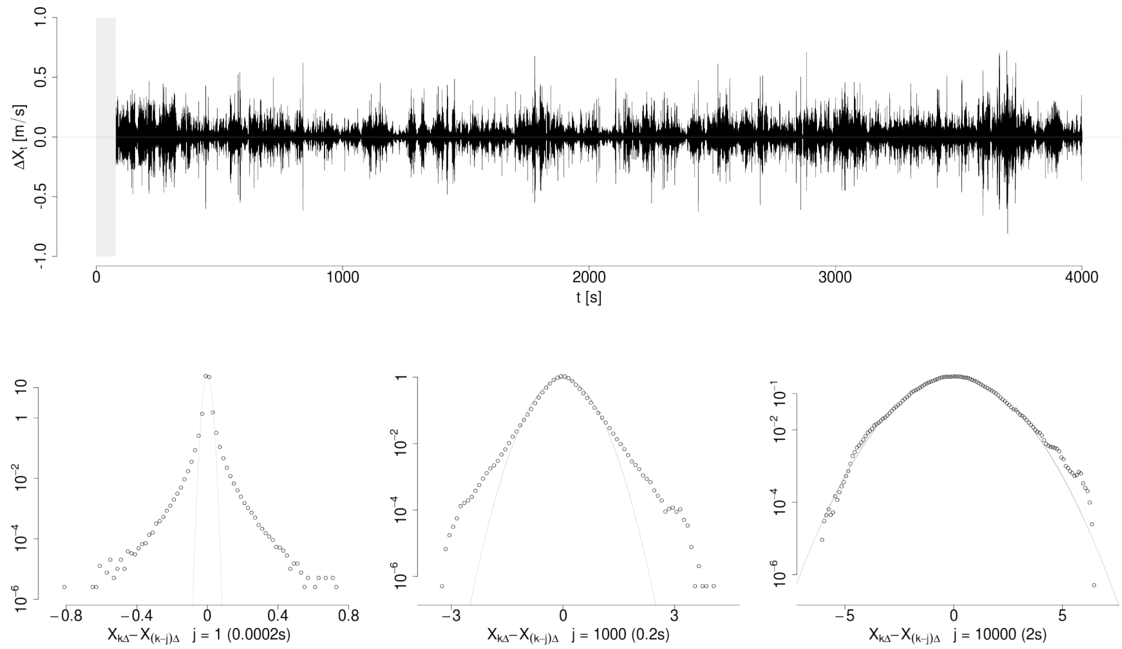

Figure 6.1: Increments of the intermittency process X. Top: (Estimated) increments $X_{k\Delta} - X_{(k-1)\Delta}$ of the intermittency process covering a total time interval of 65 min 21.1276 s. Bottom: Histograms of the intermittency increments at time-lags $j\Delta$ for $j = 1$ (left), $j = 1000$ (middle) and $j = 10\,000$ (right). The y-axes are in logarithmic scale. The solid grey line represents the Gaussian density scaled to the empirical variance of the intermittency increments.

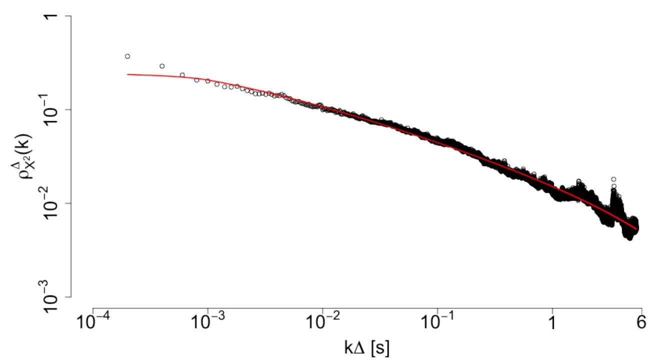

Figure 6.2: Comparison of the empirical auto-correlation $\hat{\rho}^{\Delta}_{X^2}$ of the squared intermittency increments $(\Delta^n_k X)^2$ (black points) for lags $k = 1, \dots, 26\,698$ corresponding to a time-lag of 5.3396s and of the parametric estimate $\rho^{\Delta}_{X^2}(\cdot;\hat{\theta})$ (red solid line). Both axes are in logarithmic scale. Right: Least-square estimates of the parameters $\hat{\theta}$ of $g_Y(\cdot;\theta_{\text{int}})$.

Table 6.1: Least-squares estimates of the parameters $\hat{\theta}$ of $g_Y(\cdot;\theta_{\text{int}})$

σ	ν	λ	ζ
3.6017	0.2881	0.0325	$1.152 \cdot 10^{-3}$

Table 6.2: Penalised contrast values (PCV) for the estimators $\hat{f}_\mu^m$ on D_+ and D_-

m	PCV on D_+	PCV on D_-
1	−1.283414	−1.016977
2	−1.283749	−1.016962
3	−1.283912	−1.016947
4	−1.283924	−1.016933
5	−1.283870	−1.016879

Table 6.3: (Penalised) least squares fitting of the parametric families $f(x;\theta_{p_+,p_-})$ in eq. (6.2.7) to the non-parametric estimate $\hat{f}(x)$ given by eq. (6.2.11).

p_+	$\hat{c}_k^+$	$\hat{\lambda}_k^+$	p_-	$\hat{c}_k^-$	$\hat{\lambda}_k^-$	$\hat{\alpha}$	AIC_c
1	2.542	14.35	1	3.101	24.17	1.314	4361.4
1	0.618	10.33	2	0.740	19.86	1.390	3911.4
				16.879	438.58		
2	0.177	6.67	1	0.219	17.17	1.487	4059.6
	16.156	1031.79					
2	63.279	37.81	2	76.977	47.44	0.701	5544.7
	782.867	2346.48		4.229	243.94		
1	0.012	9.57	3	0.014	19.68	1.411	3937.4
				0.001	162.75		
				0.539	518.30		
3	0.180	6.74	1	0.222	17.14	1.487	4062.6
	0.000	198.51					
	15.238	928.82					

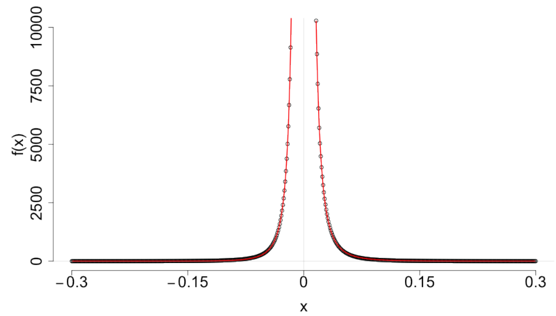

Figure 6.3: Comparison of the parametric estimate $f(x, \hat{\theta}_{1,2})$ (red solid line) and the non-parametric estimate $\hat{f}(x)$ (black points) on the domain $D = \{0.15 \leq |x| \leq 0.3\}$.

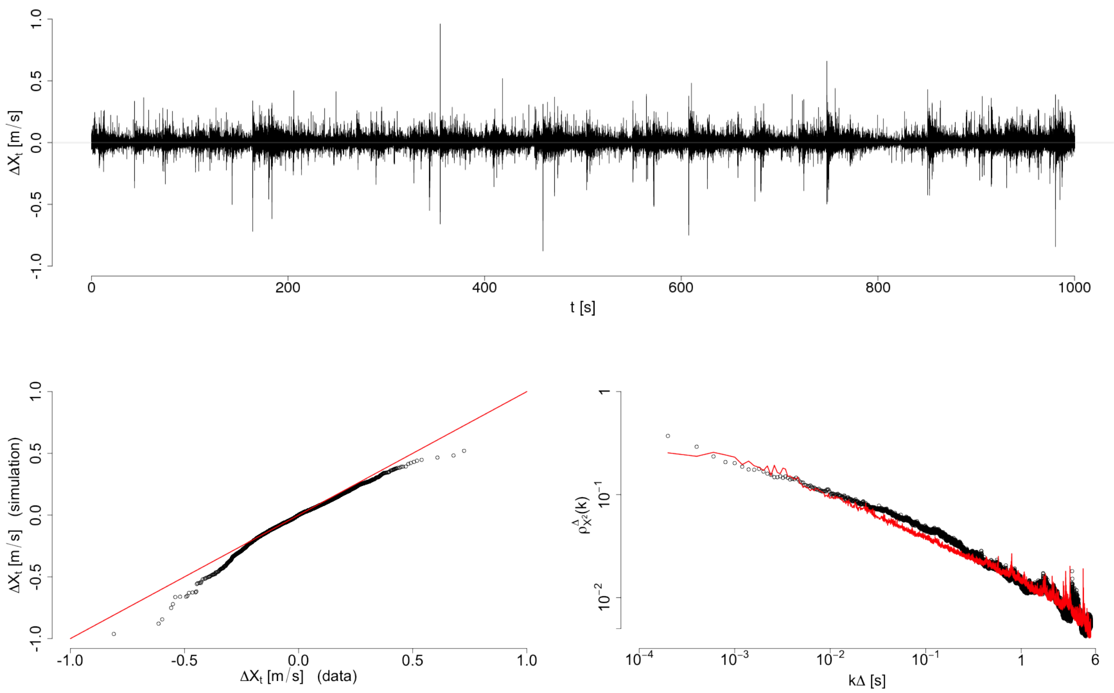

Figure 6.4: Simulation from the fitted intermittency model. Top: Simulated increments $X_{k\Delta} - X_{(k-1)\Delta}$ of the intermittency process on an interval of length 1000 s. Bottom-Left: Quantile-quantile plot (black points) of the observed increments $\Delta_k^n X$ of the data (x-axis) against the simulated increments (y-axis). The red line indicates the identity diagonal. The fit is excellent on $[-0.3, 0.3]$ which carries more than 99.996% of the data. Bottom-Right: Comparison of the empirical auto-correlation $\hat{\rho}_{X^2}^{\Delta}$ of the squared intermittency increments $(\Delta_k^n X)^2$ of the data (black points) for lags $k = 1, \ldots, 26\,698$ corresponding to a time-lag of 5.3396 s and of the empirical auto-correlation of the squared simulated intermittency increments (red solid line).

Bibliography

Akritas, M. G. (1982) Asymptotic theory for estimating the parameters of a Lévy process. *Ann. Inst. Statist. Math.*, 34:259–280.

Azéma, J., Kaplan-Duflo, M., and Revuz, D. (1967) Measure invariante sur les classes recurrents des processus de Markov. *Z. Wahrscheinlichkeitstheorie und Verw. Gebiete*, 8:157–181.

Barndorff-Nielsen, O. E., Blæsild, P., and Schmiegel, J. (2004) A parsimonious and universal description of turbulent velocity increments. *Eur. Phys. J. B*, 41:345–363.

Barndorff-Nielsen, O. E. and Schmiegel, J. (2004) Lévy based tempo-spatial modeling; with applications to turbulence. *Uspekhi Mat. Nauk*, 159:63–90. (in Russian).

— (2008) A stochastic differential equation framework for the timewise dynamics of turbulent velocities. *Theory Probab. Appl.*, 52:372–388.

Barndorff-Nielsen, O. E. and Shephard, N. (2001) Non-Gaussian Ornstein-Uhlenbeck based models and some of their uses in financial economics. *J. Roy. Statist. Soc. Ser. B*, 63:167–241.

— (2002) Econometric analysis of realised volatility and its use in estimating stochastic volatility models. *J. Roy. Statist. Soc. Ser. B*, 64:253–280.

— (2006) Impact of jumps on returns and realised variances: Econometric analisys of time-deformed Lévy processes. *J. Econometrics*, 131:217–252.

Basawa, I. V. and Brockwell, P. J. (1982) Non-parametric estimation for non-decreasing Lévy processes. *J. Roy. Statist. Soc. Ser. B*, 44:262–269.

Bass, R. F. (1979) Adding and subtracting jumps from Markov processes. *Trans. Amer. Math. Soc.*, 255:363–376.

Belomestny, D. (2011) Statistical inference for time-changed LÉvy processes via composite characteristic function estimation. *Ann. Statist.*, 39:2205–2242.

Benveniste, A. and Jacod, J. (1973) Systèmes de Lévy des processus de Markov. *Invent. Math.*, 21:183–198.

Bernstein, S. (1929) Sur les fonctions absolument monotones. *Acta Math.*, 52:1–66.

Billingsley, P. (1995) *Probability and Measure*. Wiley, New York, 3rd edition edn.

Blumenthal, R. M. and Getoor, R. K. (1964) Local times for Markov processes. *Z. Wahrscheinlichkeitstheorie und Verw. Gebiete*, 3:50–74.

— (1968) *Markov Processes and Potential Theory*, *Pure and Applied Mathematics*, vol. 29. Academic Press, New York.

Bochner, S. (1949) Diffusion equation and stochastic processes. *Proc. Nat. Acad. Sci. U.S.A.*, 35:368–370.

Brockwell, P. J., Ferrazzano, V., and Klüppelberg, C. (2012) High-frequency sampling and kernel estimation for continuous-time moving average processes. Submitted.

Carr, P., Geman, H., Madan, D., and Yor, M. (2003) Stochastic volatility for Lévy processes. *Math. Finance*, 13:345–382.

Carr, P. and Wu, L. (2004) Time-changed Lévy processes and option pricing. *Journal of Financial Economics*, 71:113–141.

Clark, P. K. (1973) A subordinated stochastic process model with finite variance for speculative prices. *Econometrica*, 41:135–155.

Cleve, J., Greiner, M., Pearson, B. R., and Sreenivasan, K. R. (2004) Intermittency exponent of the turbulent energy cascade. *Phys. Rev. E*, 69:066 316.

Cleve, J., Greiner, M., and Sreenivasan, K. R. (2003) On the effects of surrogacy of energy dissipation in determining the intermittency exponent in fully developed turbulence. *Europhys. Lett.*, 61:756–761.

Comte, F. and Genon-Catalot, V. (2009) Nonparametric estimation for pure jump Lévy processes based on high frequency data. *Stochastic Process. Appl.*, 119:4088–4123.

— (2011) Estimation for Lévy processes from high frequency data within a long time interval. *Ann. Statist.*, 39:803–837.

Darling, D. A. and Kac, M. (1957) On occupation times for Markoff processes. *Trans. Amer. Math. Soc.*, 84:444–458.

Down, D., Meyn, S. P., and Tweedie, R. L. (1995) Exponential and uniform ergodicity of Markov processes. *Ann. Appl. Probab.*, 23:1671–1691.

Drhuva, B. R. (2000) *An experimental study of high Reynolds number turbulence in the atmosphere*. Ph.D. thesis, Yale University.

Fan, J. and Yim, T. H. (2004) A crossvalidation method for estimating conditional densities. *Biometrika*, 91:819–834.

Ferrazzano, V. and Fuchs, F. (2012) On noise recovery for Lévy-driven CARMA processes and high-frequency behavior of approximating Riemann sums. Submitted for publication.

Ferrazzano, V. and Klüppelberg, C. (2012) Turbulence modeling by time-series methods. Working paper, Technische Universität München.

Ferrazzano, V. and Ueltzhöfer, F. A. J. (2012) Empirical modelling of the intermittency in atmospheric turbulence. Working paper, Technische Universität München.

Figueroa-López, J. E. (2009a) Non-parametric estimation for Lévy models based on discrete sampling. In Rojo, J. (ed.), *Optimality: The Third Erich L. Lehmann Symposium*, IMS Lecture Notes – Monograph Series, vol. 57, 117–146. Institute of Mathematical Statistics, Beachwood, OH.

— (2009b) Nonparametric estimation of time-changed Lévy models under high-frequency data. *Adv. Appl. Probability*, 41:1161–1188.

— (2011) Central limit theorems for the nonparametric estimation of time-changed Lévy models. *Scand. J. Statist.*, 38:748–765.

Frisch, U. (1996) *Turbulence: the Legacy of A. N. Kolmogorov*. Cambridge University Press, Cambridge.

Geman, H., Madan, D., and Yor, M. (2001) Time changes for Lévy processes. *Math. Finance*, 11:79–96.

Getoor, R. K. (1975) *Markov Processes: Ray Processes and Right Processes*, *Lecture Notes in Mathematics*, vol. 440. Springer, Berlin.

Greenwood, P. E. and Wefelmeyer, W. (1994) Nonparametric estimators for Markov step processes. *Stochastic Process. Appl.*, 52:1–16.

Grigelionis, B. (1971) On the representation of integer-valued random measures by means of stochastic integrals with respect to the poisson measure. *Lithuanian Math. J.*, 11:93–108.

Gugushvili, S., Klaassen, C. A. J., and Spreij, P. (eds.) (2010) *Special Issue: Statistical Inference for Lévy Processes with Applications to Finance*, *Stat. Neerl.*, vol. 64. Pp. 255–366.

Hall, P., Racine, J., and Li, Q. (2004) Cross-validation and the estimation of conditional probability densities. *J. Amer. Statist. Assoc.*, 99:1015–1026.

Hansen, B. E. (2009) Lecture notes on nonparametrics (Spring 2009, Chapters 1–2). Available at `http://www.ssc.wisc.edu/~bhansen/718/NonParametrics1.pdf`.

Höpfner, R. (1993) Asymptotic inference for Markov step processes: Observation up to a random time. *Stochastic Process. Appl.*, 48:295–310.

Höpfner, R., Jacod, J., and Ladelli, L. (1990) Local asymptotic normality and mixed normality for Markov statistical models. *Probab. Theory Related Fields*, 86:105–129.

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

Hunt, G. A. (1957) Markoff processes and potentials I. *Illinois J. Math.*, 1:44–93.

Jacod, J. (1979) *Calcul Stochastique et Problèmes de Martingales*, *Lecture Notes in Mathematics*, vol. 714. Springer, Berlin.

Jacod, J. and Protter, P. (2012) *Discretization of Processes*. Springer, 2012 edn.

Jacod, J. and Shiryaev, A. N. (2003) *Limit Theorems for Stochastic Processes*. Springer, Berlin. 2nd edition.

Kallsen, J. and Muhle-Karbe, J. (2011) Method of moments estimation in time-changed Lévy models. *Statistics and Decisions*, 28:169–194.

Karlsen, H. A. and Tjøstheim, D. (2001) Nonparametric estimation in null recurrent time series. *Ann. Statist.*, 29:372–416.

Klüppelberg, C., Lindner, A., and Maller, R. (2004) A continuous-time GARCH process driven by a Lévy process: stationarity and second-order behaviour. *J. Appl. Probab.*, 41:601–622.

Kolmogorov, A. N. (1941a) Dissipation of energy in locally isotropic turbulence. *Dokl. Akad. Nauk. SSSR*, 32:19–21. (in Russian).

— (1941b) The local structure of turbulence in incompressible viscous fluid for very large reynolds numbers. *Dokl. Akad. Nauk. SSSR*, 30:299–303. (in Russian).

— (1942) The equations of turbulent motion in an incompressible viscous fluid. *Izvestia Acad. Sci. USSR*, Phys. 6:56–58. (in Russian).

— (1962) A refinement of previous hypotheses concerning the local structure of turbulence in a viscous incompressible fluid at high Reynolds number. *J. Fluid Mech.*, 13:82–85.

Löcherbach, E. and Loukianova, D. (2008) On Nummelin splitting for continuous time Harris recurrent Markov processes and applications to kernel estimation for multi-dimensional diffusions. *Stochastic Process. Appl.*, 118:1301–1321.

— (2011) Polynomial deviation bounds for recurrent Harris processes having general state spaces. To appear in ESAIM Probab. Statist.

Mancini, C. (2005) Estimating the integrated volatility in stochastic volatility models with Lévy type jumps. Available at http://www.cmap.polytechnique.fr/SEM/finance/2004-2005/mancini1.pdf.

Mantegna, R. N. and Stanley, H. E. (1994) Stochastic processes with ultraslow convergence to a Gaussian: The truncated Lévy flight. *Phys. Rev. Lett.*, 73:2946–2949.

Meyn, S. P. and Tweedie, R. L. (1992) Stability of Markovian processes I: Criteria for discrete-time chains. *Adv. Appl. Probability*, 24:524–574.

— (1993) *Markov Chains and Stochastic Stability*. Springer, London. Online edition, 2005. Available at `http://probability.ca/MT/`.

Neveu, J. (1972) Potentiel markovien récurrent des chaînes de Harris. *Ann. Inst. Fourier (Grenoble)*, 22:85–130.

Novikov, E. A. (1994) Infinitely divisible distribution in turbulence. *Phys. Rev. E*, 50:R3303–R3305.

Pope, S. (2000) *Turbulent Flows*. Cambridge University Press, Cambridge.

Protter, P. (2005) *Stochastic Integration and Differential Equations*. Springer, Heidelberg, second edition edn.

Renyi, A. (1963) On stable sequences of events. *Sankhyā Ser. A*, 25:293–302.

Revuz, D. (1984) *Markov Chains*, *North-Holland Mathematical Library*, vol. 11. North-Holland, Amsterdam. Revised edition.

Rosiński, J. (2007) Tempering stable processes. *Stochastic Processes and their Applications*, 117:677 – 707.

Rubin, H. and Tucker, H. G. (1959) Estimating the parameters of a differential process. *Ann. Math. Statist.*, 30:641–658.

Sato, K. (1999) *Lévy Processes and Infinitely Divisible Distributions*. Cambridge University Press, Cambridge.

Sawyer, S. A. (1970) A formula for semigroups, with an application to branching diffusion processes. *Trans. Amer. Math. Soc.*, 152:1–38.

Sharpe, M. (1988) *General Theory of Markov Processes*, *Pure and Applied Mathematics*, vol. 133. Academic Press, San Diego.

Skorokhod, A. (1956) Limit theorems for stochastic processes. *Theory Probab. Appl.*, 1:261–290.

Touati, A. (1987) Théorèmes limites pour des processus de Markov récurrents. *C. R. Acad. Sci. Paris Sér. I Math.*, 305:841–844.

Ueltzhöfer, F. A. J. (2012) On non-parametric estimation of the Lévy kernel of Markov processes. Submitted.

Ueltzhöfer, F. A. J. and Klüppelberg, C. (2011) An oracle inequality for penalised projection estimation of Lévy densities from high-frequency observations. *J. Nonparametr. Stat.*, 23:967–989.

Watanabe, S. (1964) On discontinuous additive functionals and Lévy measures of a Markov process. *Japan. J. Math.*, 34:53–70.

Weil, M. (1971) Conditionnement par rapport au passé strict. *Séminaire de Probabiltés (Strasbourg)*, 5:362–372.