Statistical Analysis of Lévy Processes with Application in Finance

Dissertation

zur Erlangung des Doktorgrades Dr. rer. nat.
der Fakultät für Mathematik und Wirtschaftswissenschaften
der Universität Ulm

vorgelegt von
Achim Gegler
aus Stuttgart

2011
Amtierender Dekan: Prof. Dr. Paul Wentges  
1. Gutachter: Prof. Dr. Ulrich Stadtmüller  
2. Gutachter: Prof. Dr. Rüdiger Kiesel  
Tag der Promotion: 17. Oktober 2011
Notations

\(\mathbb{N}, \mathbb{R}\) are respectively the set of all natural numbers and the real numbers. \(\overline{\mathbb{R}}\) is the conjunction of \(\mathbb{R}\) and \([-\infty, \infty]\). \(\overline{\mathbb{N}}\) is the conjunction of \(\mathbb{N}\) and \(\{\infty\}\). \(X^d\) is the \(d\)-fold cartesian product of a set \(X\), it is always assumed that \(d \in \mathbb{N}\). \(\mathbb{R}^+\) is the set \((0, \infty)\) and \(\mathbb{R}_0^+\) the set \([0, \infty)\). \(\mathbb{N}_0\) is the set \(\mathbb{N} \cup \{0\}\).

\(\mathcal{B}(\Theta)\) is the Borel \(\sigma\)-algebra of a set \(\Theta\).

\(\mathbb{1}(.)\) is the indicator \(\sigma\)-function, with a boolean argument. If the argument is true, then the indicator function is equal to 1, if it is false then the indicator function is equal to 0.

\(a \land b := \min\{a, b\}\) and \(a \lor b := \max\{a, b\}\).

The convergence in probability of random variables, random vectors or stochastic processes is denoted by \(\overset{P}{\rightarrow}\). The convergence in distribution of random variables or random vectors is denoted by \(\overset{d}{\rightarrow}\). Weak convergence of a stochastic process is denoted by \(\overset{D}{\rightarrow}\).

Weak convergence of a probability measure is denoted by \(\Rightarrow\).

The identity matrix is denoted by \(I\).

The supremum norm is denoted by \(\|\cdot\|\) and the Euclidian norm by \(|\cdot|\).

The number of elements that are contained in a set \(A\) is denoted by \(\#A\).

The Landau symbols are denoted by \(O\) and \(o\).

w.l.o.g. = without loss of generality

\(b_n := \sqrt{2 \log n}\) and \(\bar{b}_n := (1, \ldots, d)^T b_n\)

\(B_n(D, \beta) := \{x \in \mathbb{R}^d : x^T D^{-1} x \leq \beta b_n^2\}\)

A nonstandard Wiener process is denoted by \(\mathcal{C}\mathcal{L}_t\).

A pure jump Lévy process is denoted by \(J\mathcal{L}_t\), which can be decomposed into its large jump part \(\mathcal{C}\mathcal{P}\mathcal{L}_t\) and into its small jump part \(\mathcal{S}J\mathcal{L}_t\).
## Contents

**Notations** iv

**1. Introduction** 1
   1.1. Motivation and summary ........................................... 1
   1.2. Overview of the thesis ............................................ 5

**2. Basics** 7
   2.1. Lévy processes ................................................... 7
      2.1.1. Definition and basic examples ............................... 7
      2.1.2. Infinite divisibility and Lévy Khinchin representation ... 11
      2.1.3. Poisson random measures and Lévy-Itô decomposition ... 15
      2.1.4. Further properties .......................................... 19
      2.1.5. Parametric models ........................................... 22
   2.2. Weak convergence in a generalized Skorohod space ............... 23
      2.2.1. Weak convergence of probability measures ................. 23
      2.2.2. Convergence in distribution ................................ 26
      2.2.3. The Skorohod space ......................................... 28
      2.2.4. A generalization of the Skorohod space in higher dimensions 32

**3. Estimation method** 35
   3.1. Estimation setting ................................................ 35
      3.1.1. Setting in the finite activity case ......................... 39
      3.1.2. Setting in the infinite activity case ..................... 40
   3.2. High frequency data of a Wiener process ......................... 42
   3.3. A continuously observed Lévy process ........................... 44
   3.4. Estimation of a finite activity Lévy process .................... 48
   3.5. Estimation of an infinite activity Lévy process ................. 55
   3.6. Estimation of the Blumenthal-Getoor index ...................... 61
   3.7. Estimation of the critical region and parameter choice ........ 66
      3.7.1. An iteration method for the critical region .............. 66
      3.7.2. Parameter choice ........................................... 72
   3.8. Correction method in the univariate and finite activity case ... 74
   3.9. Simulations and real data application ........................... 79
      3.9.1. In the univariate case ..................................... 80
      3.9.2. In the multivariate case ................................... 86
      3.9.3. Discussion .................................................. 89

**4. Estimation in the Kou model** 91
   4.1. Estimation setting ................................................. 91
1. Introduction

1.1. Motivation and summary

In financial mathematics stochastic processes are used to describe stock price developments. The increments of the logarithmic stock price are called log returns and are used for statistical investigations. The classical approach in the famous Black-Scholes model, introduced in Black and Scholes (1973), is to use just a Brownian motion to describe the logarithmic price development, so that the log returns are normally distributed. But various empirical studies revealed that this approach does not always match to the real data (cf. Mandelbrot, 1963; Cont, 2001). In particular the log returns of most financial data are not normally distributed, since they are skewed and have more probability mass in the tails (cf. Schoutens, 2003, Chapter 4). Furthermore the imperfectness of the Black-Scholes model also turns out by the difference between theoretical option prices that are computed by the Black-Scholes model and the observed market prices. This leads to the so called volatility smile, i.e. the implicit volatility, which is calculated by the Black-Scholes formula for given option market prices, is a convex function of the strike price. If the Black-Scholes model was correct, the implicit volatility should be constant and equal to the estimated historical volatility.

One possibility to improve the modeling is using a Lévy process instead of the Brownian motion. Lévy processes are a natural generalization of the Brownian motion and form a class of stochastic processes that have stationary and independent increments and are stochastically continuous. They are a combination of a Brownian motion and a pure jump process. According to Schoutens (2003, p.43), Lévy models are much more flexible, since they can take the skewness and the excess kurtosis into account. The jump process can be used to model seldom and sudden large movements of the stock price. These asymmetric jumps can cause asymmetric and heavy tails in the distribution of the log returns. However, not all drawbacks of the Black-Scholes model can be well explained by Lévy processes, e.g. for assets with volatility clusters in the price development, stochastic volatility models are a better choice (cf. Schoutens, 2003, Chapter 4.2).

In most cases the distribution of Lévy processes is not easy to handle, because it is not given in a closed form. However, it is possible to work with the characteristic function, since it can be explicitly represented and only depends on the so called Lévy triplet. The Lévy triplet consists of the variance of the Brownian motion, a drift term and the so called Lévy measure that describes the jump...
1. Introduction

behavior.

This thesis deals with the problem of estimating the Lévy triplet out of historical data. It is a semiparametric problem, because in general the Lévy measure cannot be modelled parametrically. One of the first papers tackling this problem is Rubin and Tucker (1959). There consistent estimators for the variance of the diffusion and a specific transformation of the Lévy measure are presented under the assumption that the process can be observed continuously. However, realistically only a finite number of observations at discrete time points is available, and usually an estimation setting is based on the increments of the process. Hence, e.g. the discontinuities cannot be observed exactly. This fact is the main issue of the estimation. It is hard to distinguish between small jumps and the continuous Brownian motion. That means, we have a so called ill-posed inverse problem, which is a situation, where completely different input parameters lead to output data that look almost the same.

In the recent years two different frameworks for the estimation of stochastic processes have been established, low frequency and high frequency data. In the context of asymptotic statistics low frequency means that the distance between two estimation points is asymptotically fixed. So an increasing sample size only leads to an increasing time horizon. For Lévy processes this means that a classical statistical situation with i.i.d. samples, which are independent of the sample size, is existent (because of the stationary and independent increments).

If the Lévy process is fully parameterized, a possibility is to use maximum likelihood procedures. However, in practice this approach has a lot of drawbacks. In most cases the transition density is not given in a closed from and the MLE can only be found numerically. It is a high dimensional problem and – since it is an ill-posed problem – different parameter constellations fit the data well. So the likelihood function may have several local maxima or is even unbounded (see Honoré 1998). This means that it is very difficult to use numerical methods. An overview of the maximum likelihood estimation of Lévy processes is given in Cont and Tankov (2004, Chapter 7.2.1). In contrast to the transition density, the characteristic function of a Lévy process is given in a closed from. So another method is to compare the theoretical characteristic function that depends on the unknown parameters with the empirical characteristic function that is calculated by the sample. Although this method is easier to implement than MLE, the resulting equation systems are still very complex and often only numerically solvable. But, since the problem is ill-posed, numerical methods are not easy to apply, either. In Riesner (2006) and Gegler (2007) the method is used for the Kou model, which has asymmetric double exponential distributed jump sizes. The method is embedded in the more general framework of Generalized Methods of Moments (GMM), which is introduced in Hansen (1982) and is a generalization of the methods of moments. Instead of considering the moments directly, a function of the moments is used. It turns out that the estimator only can be computed if some parameters are previously given, since the resulting equation systems are too complex. In Cont and Tankov (2004, Chapter 7.2.2) an overview of GMM for Lévy processes is given. In Aït-Sahalia (2004)
a maximum likelihood and a GMM estimator are considered for the Merton model, which has normally distributed jumps. In Neumann and Reiß (2009) a nonparametric estimator based on the characteristic function is used and it is shown that this estimator is rate-optimal. In Gugushvili (2009) an estimator for the Lévy triplet is constructed by using Fourier inversion.

The other framework that has been established is the high frequency setting. In the context of asymptotical statistics high frequency means that the distance between two observation points asymptotically tends to zero. The advantage of this approach is that it is easier to distinguish asymptotically between diffusion and jumps. The drawback is that the framework is outside the classical statistical situation. The increments are i.i.d., however they depend on the sample size \( n \), so a triangular array is of random variables is given.

The estimation method introduced in this thesis is based on the framework of high frequency. In Blanco and Soronow (2007) it is suggested to separate jumps from diffusion by just using three times the standard deviation of the continuous process, which is pre-estimated, as a threshold. Of course, this method is too easy to obtain consistent estimators, but in practice such simple methods are widely used. We pick up the idea of using a threshold to distinguish between diffusion and jumps and endow it with a whole statistical fundament that allows us to prove consistency and in some cases also asymptotic normality. While the proofs of the statistical properties are complex, the method itself remains easy to implement and so it is interesting for application.

The definition of our threshold is motivated by the modulus of continuity of the Brownian motion, because we have to ensure that the diffusion part of the increments is within the threshold bounds with high probability. The threshold also depends on the unknown standard deviation. So the threshold itself has to be estimated. This leads to an iteration procedure. Figure 1.1 shows an example of log returns and the estimated threshold that we have generated by the Merton model, which has normally distributed jump sizes. We will discuss this example in Section 3.9.1.

There are two different types of Lévy processes, finite and infinite activity. Finite activity means that on a compact time interval only a finite number of jumps occurs. So the Lévy process is simply a combination of a Brownian motion and a compound Poisson process. On the other hand infinite activity means that on a compact time interval infinitely many jumps occur, but the sum of the squared jump sizes is still finite. We obtain different results for the two cases.

The estimator for the variance of the diffusion part is based on the increments with an absolute value below the threshold. In case of finite activity we get an asymptotically normally distributed estimator for the variance of the continuous part. In the case of infinite activity only consistency is obtained. The increments with an absolute value above the threshold are used to estimate the Lévy measure. Similar to an empirical distribution function, we estimate the Lévy measure of an interval \((-\infty, z], z \in \mathbb{R}\), consider this quantity as a stochastic process itself and give a limit process. In the case of infinite activity
1. Introduction

![Figure 1.1: Example of log returns and estimated threshold, generated by the Merton model](image)

the Lévy measure of a neighborhood of the origin is infinite. Thus, we estimate a specific finite transformation of the Lévy measure, instead. In the case of finite activity we also give an asymptotically normally distributed estimator for the drift, whereas in the infinite activity case this is not possible with the threshold concept.

In Finance the development of different stock prices and their interactions are often considered simultaneously. Portfolios usually depend on different assets and a lot of options have more than one underlying. So we also extend our method to multivariate Lévy processes. In this case it is more suitable to talk about a critical region to decide whether an increment should be considered as diffusion or as a jump. Since the level curve of the density of the multivariate normal distribution forms an ellipsoid, we also choose an ellipsoid as the critical region. Obviously the critical region sensitively depends on the covariance of the diffusion process, which is also pre-estimated by an iteration method. We formulate the whole method in a multivariate framework, the univariate case is included as a special case. With the univariate case we deal in [Gegler and Stadtmüller (2010)](https://www.example.com). The multivariate case is considered in [Gegler (2011)](https://www.example.com).

A drawback of our method is the bias that occurs when the method is applied in practice. Although it is asymptotically unbiased, this does not hold true for a finite sample size. The reason is, that the small jumps that are lying between the threshold bounds, are ignored. We suggest a solution to that problem in the one dimensional case by introducing an extrapolation method for the estimator of the Lévy measure.

If special parametric models are considered, the Lévy measure is modelled by
1.2. Overview of the thesis

a parametric family. Then the estimation is easier, since the behavior of the small jumps can be derived from the behavior of the large jumps, because they are driven by the same parameters. Exemplarily we illustrate this method for the Kou model, where the jumps are assumed to be asymmetrically double exponentially distributed.

While we have developed our method, we realized that a similar threshold approach has been developed in [Mancini and Reno (2011), Mancini (2009, 2008, 2004, 2003, 2001) (univariate case) and Gobbi and Mancini (2010, 2007) (bivariate case)]. Although the main idea of these papers is related, the framework is different. Other processes are considered in different asymptotic settings and thus the complete Lévy triplet or a transformation of it cannot be estimated by their methods.

There are other methods proposed in the literature using the framework of high frequency. E.g. [Aït-Sahalia and Jacod (2007)] provide a very general method for estimating the volatility in a symmetric stable process that is perturbed by another Lévy process. This paper considers a suitable function \( k \) of the increments and compares it to a theoretical equivalent. If a truncation function for \( k \) is used this method is similar to our approach. The papers [Barndorff-Nielsen et al. (2006), Jacod (2008) and Woerner (2006)] use power and multipower variations to estimate the integrated variance. Barndorff-Nielsen and Shephard (2004a,b) deal with the multivariate case. Thereby the trick is to consider the product of succeeding increments. So the seldom large jumps are multiplied with small diffusion increments and are asymptotically devaluated that way. In [Figueroa-López (2009)] an estimator for the Lévy density is considered as well, but only if the estimation window lies outside of a neighborhood of the origin. In [Shimizu (2006)] a kernel-estimator of the Lévy density is considered in the case of finite activity. This estimator is rate optimal under certain assumptions.

1.2. Overview of the thesis

Chapter 2 contains the basic concepts that are needed in this thesis. In Section 2.1 we give a brief overview of the theory of Lévy processes. We start with the definition and consider some basic examples: the Wiener process, the Poisson process and the compound Poisson process. We give the Lévy Khinchin representation, the Lévy-Itô decomposition and some further properties. In addition we introduce the Merton model and the Kou model. Section 2.2 deals with the concept of weak convergence of probability measures. We start in a general metric space, afterwards we turn to the Skorohod space \( D[0,1] \) and finally we consider the space \( D[0,1]^d \) that is a generalization of the Skorohod space in a higher dimensions.

In Chapter 3 our estimation method is introduced. Section 3.1 contains the estimation setting. There we define the high-frequency framework as well as the critical region. In Section 3.2 estimators for the covariance matrix and the drift are developed under the assumption that the increments of the real diffusion
1. Introduction

part of the Lévy process are known. In Section 3.3 estimators are given for the Lévy measure and its specific transformation under the assumption that the process can be observed continuously. In Section 3.4 we prove the asymptotic properties of the estimators that are based on the critical region in the case of finite activity. Section 3.5 deals with the infinite activity case. In Section 3.6 the Blumenthal-Getoor index is estimated. The Blumenthal-Getoor index measures the behavior of the infinitely many small jumps in the case of infinite activity. An iteration method to find a critical region in practice is developed in Section 3.7. In Section 3.8 we introduce the correction method that deals with the bias that occurs in practice. Finally, in Section 3.9 we perform simulations and apply our method to some real data time series.

In Chapter 4 we use our estimation method to derive estimators for the parameters of the Kou model. In Section 4.1 we give the estimation setting and in Section 4.2 estimators are presented under the assumption that the exact jump times and sizes are known. The estimation method that is based on the critical region is introduced in Section 4.3. Section 4.4 contains a correction method for these estimators, where we can use the information that large and small jumps are controlled by the same parameters. In Section 4.5 we perform simulations and analyze our method.
2. Basics

2.1. Lévy processes

Lévy processes form a general class of stochastic processes that start at zero, have independent and identically distributed increments, and are stochastically continuous. The Wiener process as well as the compound Poisson process are famous examples. Lévy processes have a lot of nice properties, so they are interesting for application in financial mathematics. Lévy processes were named after the French mathematician Paul Lévy (1886–1971). He was one of the greatest mathematicians of the last century and his work had a significant influence on the development of the modern probability theory. He did pioneer work in the study of what we today call a Lévy process. The main work of Paul Lévy can be found in Lévy (1922, 1925, 1937, 1948). A good overview of the personal life and work of Paul Lévy can be found in Taylor (1975) and Cont and Tankov (2004). In this section we give an introduction to the theory of Lévy processes. We restrict ourselves to the main properties that are needed in the following chapters using the results and definitions presented in Sato (2007).

2.1.1. Definition and basic examples

In this subsection we give the definition of a Lévy process and introduce the most important examples and their properties. The Wiener process as well as the Poisson process and the compound Poisson process are considered as basic examples.

Throughout we use a suitable probability space $(\Omega, \mathcal{F}, \mathbb{P})$. We start with the definition of a Lévy process.

**Definition 2.1 (Lévy process)**

A stochastic process $\{L_t, t \geq 0\}$ with values in $\mathbb{R}^d$ ($d \in \mathbb{N}$) is called Lévy process, if it has the following properties

1. $L_0 = 0$ almost surely,

2. **Independent increments:** For all $n \in \mathbb{N}$ and for all $t_0 \leq t_1 \leq \cdots \leq t_n$ the random vectors $L_{t_0}, L_{t_1} - L_{t_0}, \ldots, L_{t_n} - L_{t_{n-1}}$ are independent,

3. **Stationary increments:** The distribution of $L_{t+h} - L_t$ with $t, h \geq 0$ does not depend on $t$,

4. **Stochastic continuity:** $\forall \varepsilon > 0$ and $t_0 \geq 0$: $\lim_{t \to t_0} \mathbb{P}(|L_t - L_{t_0}| > \varepsilon) = 0$. 

7
5. There exists $\Omega_0 \subset \Omega$ with $\mathbb{P}(\Omega_0) = 1$, such that for all $\omega \in \Omega_0$, $L_t(\omega)$ is right-continuous and has left limits.

Processes that satisfy condition 5) are also called cadlag. The expression derives from the French formulation continue à droite, limites à gauche, which means that with probability 1 the process is right continuous and left limits exists. It can be shown that a process satisfying conditions 1)–4) that is not cadlag has always a cadlag unique modification (cf. Protter 2004, Theorem 30). So condition 5) is no real restriction.

Next, we consider some very important examples of Lévy processes. We start with the $d$-dimensional Wiener process.

**Definition 2.2 (Wiener process)**

(i) A stochastic process $\{W_t, t \geq 0\}$ with values in $\mathbb{R}^d$ ($d \in \mathbb{N}$) is called standard Wiener process, if it has the following properties

1. It is a Lévy Process,
2. For all $t > 0$, $W_t$ is normally distributed with mean 0 and covariance matrix $tI$, where $I$ is the identity matrix.
3. $W_t$ is almost surely continuous.

(ii) Let $C$ be a symmetric and nonnegative $d \times d$ matrix and $\gamma \in \mathbb{R}^d$, then the process $\{CL_t, t \geq 0\}$ defined by

$$CL_t := C^{1/2}W_t + t\gamma$$

is called nonstandard Wiener process with drift $\gamma$ and covariance matrix $C$.

(iii) A nonstandard Wiener process with drift $\gamma = 0$ is denoted by $\{CL_t, t \geq 0\}$.

Obviously a nonstandard Wiener process with drift is also a Lévy process. The character "C" in the notation $CL_t$ indicates "Continuity". We will later see that every continuous Lévy process is a (nonstandard) Wiener process with drift. We will also see, that it is not necessary to require condition 3) in Definition 2.2, since it is automatically fulfilled, i.e. a Lévy process with normally distributed increments is automatically a.s. continuous (cf. Proposition 2.19, Proposition 2.23 and Sato (2007, p.22)).

**Proposition 2.3 (Characteristic function of a Wiener process)**

The characteristic function of a Wiener process $\{CL_t, t \geq 0\}$ with covariance matrix $C$ and drift $\gamma$ is given by

$$\Phi_{CL_t}(u) := \exp \left( t \left( -\frac{1}{2}u^T Cu + iu^T \gamma \right) \right).$$
2.1. Lévy processes

Proof. Since $C_t \sim \mathcal{N}(0, C_t)$, the characteristic function of the process coincides with the characteristic function of a normally distributed random variable.

An overview of Wiener processes can e.g. be found in [Hida (1980) and Applebaum (2004)]. In this section we only want to focus on two properties of Wiener processes. First, the so called selfsimilarity and second, the so called Lévy’s modulus of continuity theorem. These two properties play an important role within our estimation procedure.

According to [Cont and Tankov (2004) Definition 7.1] a stochastic process $X_t$ is called selfsimilar if there exists a $H \in \mathbb{R}^+$ such that for any scaling factor $c > 0$ the processes $(X_{ct})_{t \geq 0}$ and $(c^H X_t)_{t \geq 0}$ have the same law, i.e.

$$ (X_{ct})_{t \geq 0} \overset{d}{=} (c^H X_t)_{t \geq 0}. $$

The parameter $H$ is called selfsimilarity exponent. The next proposition shows that a Wiener process is selfsimilar with exponent $1/2$.

**Proposition 2.4** (Selfsimilarity of a Wiener process)

Let $\{C_t \mathcal{L}_t, t \geq 0\}$ be a nonstandard $\mathbb{R}$-valued Wiener process with variance $C$ and drift $\gamma = 0$ and let $c > 0$. Then,

$$ Y_t = C_t \mathcal{L}_{ct} / \sqrt{c} $$

is another nonstandard Wiener process with variance $C$ and drift $\gamma = 0$.

Proof. See e.g. [Sato (2007) Theorem 5.4].

Intuitively that means that a change of time scale has the same effect as some change of spatial scale (cf. Sato, 2007, p.69).

The selfsimilarity property is important if we estimate the covariance matrix of a nonstandard Wiener process. Reducing the distance between the observation points has the same effect as increasing the time horizon. The performance of the estimation only depends on the sample size (cf. Chapter 3).

Next, we turn to Lévy’s modulus of continuity Theorem. This theorem tells us how the modulus of continuity of a Wiener process behaves.

**Proposition 2.5** (Lévy’s modulus of continuity Theorem)

Let $\{W_t, t \in [0,1]\}$ be a standard Wiener process with values in $\mathbb{R}$. Then we have

$$ \sup_{0 \leq t \leq 1-h} \frac{W_{t+h} - W_t}{\sqrt{2h \log(h^{-1})}} \overset{a.s.}{\rightarrow} 1. $$

Proof. See e.g. [Csörgő and Révész (1981) Theorem 1.1.1.).
2. Basics

Our estimation method is based on the idea of separating the diffusion part from the jumps. Proposition 2.5 gives us a hint how to do this, since it tells us where the increments of the Wiener process are located.

Another very fundamental example of a Lévy process is the so called Poisson process.

**Definition 2.6** (Poisson process)

A stochastic process \( \{N_t, t \geq 0\} \) with values in \( \mathbb{R} \) is called Poisson process with intensity \( \lambda > 0 \) if it has the following properties

(i) It is a Lévy process,

(ii) For all \( t > 0 \), \( N_t \) is Poisson distributed with mean \( \lambda t \).

We can use a Poisson process to construct a compound Poisson process, which again is a Lévy process.

**Definition 2.7** (Compound Poisson process)

A stochastic process \( \{^{CP}L_t, t \geq 0\} \) with values in \( \mathbb{R}^d \) (\( d \in \mathbb{N} \)) is called compound Poisson process with jump size distribution function \( F \) and intensity \( \lambda > 0 \) if it has the following representation

\[
^{CP}L_t = \sum_{k=1}^{N_t} Y_k,
\]

where \( N_t \) is a Poisson process and \( Y_1, Y_2, \ldots : \Omega \rightarrow \mathbb{R}^d \) are i.i.d. random vectors that are independent from \( N_t \) and have a distribution function \( F \) that satisfies \( dF(0) = 0 \).

The assumption \( dF(0) = 0 \) is needed to obtain a unique representation. So no jumps of size zero can occur.

The next proposition gives the characteristic function of a compound Poisson process.

**Proposition 2.8** (Characteristic function of a compound Poisson process)

Let \( \{^{CP}L_t, t \geq 0\} \) be a compound Poisson process with intensity \( \lambda \) and jump size distribution \( F \). Then the characteristic function is given by

\[
\Phi^{CP}L_t(u) := \exp \left( t\lambda \int_{\mathbb{R}^d} \left( e^{iu^T x} - 1 \right) dF(x) \right).
\]

*Proof.* See e.g. [Cont and Tankov (2004), Proposition 3.4].
2.1. Lévy processes

The characteristic function of the compound Poisson process is important, since the characteristic function of a general Lévy process is based on it.

Next, we consider a non trivial combination of a Wiener process and a compound Poisson process.

**Definition 2.9** (Jump diffusion process)

Let \( \{C_L^t, t \geq 0\} \) be a nonstandard Wiener process with drift \( \gamma \) and covariance matrix \( C \neq 0 \) and \( \{CP^t, t \geq 0\} \) a compound Poisson process with jump size distribution \( F \) and intensity \( \lambda > 0 \), which is independent from \( \{C_L^t, t \geq 0\} \). The process \( \{JD^t, t \geq 0\} \) defined by

\[
JD^t = C^t + CP^t
\]

is called Jump diffusion process.

The next Proposition gives the characteristic function of a Jump diffusion process.

**Proposition 2.10** (Characteristic function of a Jump diffusion process)

Let \( \{JD^t, t \geq 0\} \) be a Jump diffusion process as in (2.1). Then the characteristic function is given by

\[
\Phi_{JD^t}(u) := \exp \left( t \left( -\frac{1}{2} u^T C u + i u^T \gamma + \lambda \int_{\mathbb{R}^d} (e^{iu^T x} - 1) dF(x) \right) \right)
\]

**Proof.** Obvious since Proposition 2.3 and 2.8.

2.1.2. Infinite divisibility and Lévy Khinchin representation

In this subsection we illustrate the connection between Lévy processes and infinite divisible distributions. Based on this, we give the characteristic function of a Lévy process. We start with the definition of the infinite divisible distribution.

**Definition 2.11** (Infinite divisible distribution)

A (multivariate) distribution \( P \) is infinite divisible if there exists, for all \( n \in \mathbb{N} \), a distribution \( P_n \) such that \( P \) is the \( n \)-fold convolution of \( P_n \).

The convolution of two distributions \( P_1 \) and \( P_2 \) on \( \mathbb{R}^d \) is defined by

\[
P(B) = \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} \mathbb{1}(x + y \in B) P_1(dx) P_2(dy), \quad B \in \mathcal{B}(\mathbb{R}^d).
\]

The convolution \( P \) is a distribution on \( \mathbb{R}^d \) again. For independent random vectors \( X_1, X_2 \) with distributions \( P_{X_1}, P_{X_2} \) the convolution of \( P_{X_1}, P_{X_2} \) is the distribution of the random vector \( X_1 + X_2 \) (cf. Sato [2007] p.8). So in other words, Definition 2.11 says that the distribution \( P \) of a random vector \( X \) is infinite...
divider if there exist, for all $n \in \mathbb{N}$, i.i.d. random vectors $X_{1,n}, \ldots, X_{n,n}$ such that $X \overset{d}{=} \sum_{k=1}^{n} X_{k,n}$.

Next, we illustrate the connection between infinite divisible distributions and Lévy processes.

**Proposition 2.12** (Connection infinite divisible distribution - Lévy process)

(i) Let $\{L_t, t \geq 0\}$ be a Lévy process with values in $\mathbb{R}^d$, then, for all $t \geq 0$, $L_t$ has an infinite divisible distribution and, for the characteristic function, we have

$$\Phi_{L_t} = (\Phi_{L_1})^t.$$

(ii) If $P$ is an infinite divisible distribution on $\mathbb{R}^d$, then there exists a $\mathbb{R}^d$-valued Lévy process $\{L_t, t \geq 0\}$ such that $P$ is the distribution of $L_1$.

*Proof.* See e.g. Sato (2007, Theorem 7.10). \qed

The clue is, that the characteristic function of a Lévy process is given by the characteristic function of the Lévy process at one single point that is normalized to $t = 1$.

In the following the supremum norm is denoted by $|| \cdot ||$ and by the Euclidian norm by $| \cdot |$.

The next proposition gives an explicit formula for the the characteristic function of a Lévy process, the so called Lévy Khinchin representation. In that formula the so called Lévy measure is used.

**Definition 2.13** (Lévy measure)

A measure $\nu$ on $\mathbb{R}^d$ satisfying

$$\nu(\{0\}) = 0 \quad \text{and} \quad \int_{\mathbb{R}^d} (||x||^2 \wedge 1) \nu(dx) < \infty \quad (2.2)$$

is called Lévy measure.

**Proposition 2.14** (Lévy Khinchin representation)

(i) Let $\{L_t, t \geq 0\}$ be a Lévy process with values in $\mathbb{R}^d$. Then the characteristic function of $L_t$ is given by

$$\phi_{L_t}(u) = \exp(t \psi_{L_1}(u)),$$

with the so called cumulant generating function

$$\psi_{L_1}(u) = -\frac{1}{2} u^T C u + i u^T \gamma$$

$$+ \int_{\mathbb{R}^d} (\exp(i u^T x) - 1 - i u^T x 1(||x|| \leq 1)) \nu(dx), \quad (2.3)$$
2.1. Lévy processes

where $C$ is a symmetric nonnegative definite $d \times d$ matrix, $\gamma \in \mathbb{R}^d$ and $\nu$ a Lévy measure.

(ii) The representation by (i) in $C$, $\gamma$ and $\nu$ is unique.

(iii) If $C$ is a symmetric nonnegative definite $d \times d$ matrix, $\gamma \in \mathbb{R}^d$ and $\nu$ a measure on $\mathbb{R}^d$ satisfying (2.2), then there exists a Lévy process $\{L_t, t \geq 0\}$ such that the characteristic function is given by (2.3).

Proof. In [Sato (2007), Theorem 8.1] statements (i)–(iii) are shown for infinite divisible distributions. Proposition 2.12 completes the proof. \qed

Remark 2.15

The term $1(\|x\| \leq 1)$ in (2.3) ensures that the integrand is integrable. Other terms are possible, in [Sato (2007), Theorem 8.1] the Euclidian norm is used. We sometimes work with $1(\|x\| \leq \xi)$, for $\xi > 0$. Then, the drift has to be adapted

$$\gamma[\xi] := \gamma - \int_{\mathbb{R}^d} (x1(\|x\| \leq 1) - x1(\|x\| \leq \xi)) \nu(dx).$$

See also [Sato (2007), Remark 8.4].

Remark 2.16 (Lévy triplet)

The triplet $(C, \gamma, \nu)$ is called "Lévy triplet" or "characteristics of the Lévy process". The Lévy triplet completely characterizes the Lévy process. It depends on the form of the integrand (Remark 2.15). If the triplet is given without further notation it refers to the representation in (2.3).

Remark 2.17

(i) If $\int_{\|x\| \leq 1} \|x\| \nu(dx) < \infty$, the representation (2.3) simplifies to

$$\psi_{L_1}(u) = -\frac{1}{2} u^T C u + iu^T \gamma[0] + \int_{\mathbb{R}^d} (\exp(iu^T x) - 1) \nu(dx),$$

where

$$\gamma[0] := \gamma - \int_{\mathbb{R}^d} x1(\|x\| \leq 1) \nu(dx),$$

since the integral exists (cf. [Sato (2007), p.39]).

(ii) If $\int_{\|x\| > 1} \|x\| \nu(dx) < \infty$, the representation (2.3) simplifies to

$$\psi_{L_1}(u) = -\frac{1}{2} u^T C u + iu^T \gamma[\infty] + \int_{\mathbb{R}^d} (\exp(iu^T x) - 1 - iu^T x) \nu(dx),$$

where

$$\gamma[\infty] := \gamma - \int_{\mathbb{R}^d} (x1(\|x\| \leq 1) - x) \nu(dx),$$

since the integral exists.
2. Basics

We distinguish between two types of Lévy processes, finite activity and infinite activity.

**Definition 2.18** (Finite and infinite activity)

(i) A Lévy process is said to have finite activity, if in $[0, 1]$ only a finite number of jumps occur, i.e.

$$\nu(\mathbb{R}^d) < \infty.$$  

(ii) A Lévy process is said to have infinite activity, if in $[0, 1]$ an infinite number of jumps occur, i.e.

$$\nu(\mathbb{R}^d) = \infty.$$  

Obviously, the characteristic function of every finite activity Lévy process has the form given in Remark 2.17.

According to their definitions, a Wiener process, a compound Poisson process and a Jump diffusion process are finite activity Lévy processes; Proposition 2.19 gives the corresponding Lévy triplets. Conversely, every finite Lévy process is a (possibly trivial) combination of a Wiener process and a compound Poisson process. This result is shown in the next subsection.

**Proposition 2.19**

(i) The Lévy triplet of a nonstandard Wiener process \(\{C_L, t \geq 0\}\) with covariance \(C\) and drift \(\gamma\) is given by \((C, \gamma, 0)\).

(ii) The Lévy triplet of a compound Poisson process \(\{CP_L, t \geq 0\}\) with jump size distribution function \(F\) and intensity \(\lambda\) is given by \((\lambda \int_{||x|| \leq 1} x F(x), 0, \lambda dF)\).

(iii) The Lévy triplet of a Jump diffusion process \(\{JD_L, t \geq 0\}\) with jump size distribution function \(F\), intensity \(\lambda\), covariance \(C\) and drift \(\gamma[0]\) is given by \((C, \gamma[0] + \lambda \int_{||x|| \leq 1} x F(x), \lambda dF)\).

(iv) The processes in (i)-(iii) have finite activity.

**Proof.** (i)–(iii) follows from Proposition 2.3, 2.8 and 2.10 (iv) is obvious.  

The next proposition can be proved by the Lévy Khinchin representation and states that a linear transformation of a Lévy process is again a Lévy process.

**Proposition 2.20** (Linear Transformation)

Let \(\{L_t, t \geq 0\}\) be a Lévy process on \(\mathbb{R}^d\) with Lévy triplet \((C, \gamma, \nu)\) and let \(U\) be
2.1. Lévy processes

an $n \times d$ matrix. Then $\{UL_t, t \geq 0\}$ is a Lévy process on $\mathbb{R}^n$ with Lévy triplet $(C_U, \gamma_U, \nu_U)$ given by

$$C_U = UCU^T,$$

$$\gamma_U = [\nu U^{-1}]_{\mathbb{R}^n \setminus \{0\}},$$

$$\nu_U = U\gamma + \int Ux (1(|x| \leq 1) - 1(|x| \leq 1))\nu(dx),$$

where $(\nu U^{-1})(B) = \nu(\{x : Ux \in B\})$ and $[\nu U^{-1}]_{\mathbb{R}^n \setminus \{0\}}$ is the restriction of the measure $\nu U^{-1}$ to $\mathbb{R}^n \setminus \{0\}$.

Proof. See e.g. [Sato (2007, Proposition 11.10)].

Proposition 2.20 implies that an one dimensional projection of a Lévy process is again an univariate Lévy process. This result is interesting for us, because we sometimes consider a single component of a Lévy process. We formulate this result as a corollary.

Corollary 2.21 (Projection of a Lévy process)

Let $\{Lt, t \geq 0\}$ be a Lévy process on $\mathbb{R}^d$ with Lévy triplet $(C, \gamma, \nu)$. Then the $i$-th ($i = 1, \ldots, d$) component $\{Lt,i, t \geq 0\}$ is a Lévy process on $\mathbb{R}$ with Lévy triplet $(c_{i,i}, \gamma_i, \tilde{\nu}_i)$, where $c_{i,i}$ is the $i$-th entry of the matrix $C$, $\gamma_i$ is the $i$-th component of $\gamma$ and $\tilde{\nu}_i$ is the restriction of the $i$-th margin $\nu_i = \nu(\mathbb{R}^{i-1} \times \cdot \times \mathbb{R}^{d-i-1})$ to $\mathbb{R} \setminus \{0\}$.

Proof. The proof directly follows from Proposition 2.20.

An interesting fact that turns out by Corollary 2.21 is, that the $i$-th margin $\nu_i$ of the Lévy measure $\nu$ and the Levy measure $\tilde{\nu}_i$ of the $i$-th component of the Lévy process are not the same. By construction (2.2) the Lévy measure can not have mass on $\{0\}$. With respect to $\nu$ that means that $\nu(\{0\}) \equiv 0$ ($\{0\} \subset \mathbb{R}^d$), but $\nu(B) > 0$ is possible for $B \subset \mathbb{R}^d$ with $B_i = \{0\}$ ($i$-th projection), so also $\nu_i(\{0\}) > 0$ is possible. In contrast, if we consider the one dimensional component of the Lévy process and its measure $\tilde{\nu}_i$, we have $\tilde{\nu}_i(\{0\}) \equiv 0$ for $\{0\} \subset \mathbb{R}$.

2.1.3. Poisson random measures and Lévy-Itô decomposition

In this subsection we turn to the famous Lévy-Itô decomposition, that tells us how the sample paths of a Lévy process look like. The result is that every finite activity Lévy process is a (possibly trivial) combination of a Wiener process and an independent compound Poisson process. A general Lévy process can be decomposed into a Wiener process and an independent pure jump process. The sum of the squared jump sizes on a compact time interval has to be finite, but the sum of the absolute jump sizes also can be infinite. So the pure jump process consists of a compound Poisson process that contains finitely many
"large" jumps (e.g. with absolute jump size $\geq 1$) and a square integrable martingale with infinitely many "small" jumps (e.g. with absolute jump size $< 1$). The martingale is defined as the limit of compound Poisson processes minus their compensating drift terms. The compensating drift term is necessary, since the limit of the compound Poisson processes alone does in general not exist. If $\int_{||x||<1} ||x|| \, dx < \infty$, then the sum of the absolute jump sizes on a compact time interval is finite, so the compensating drift term is not needed. The Lévy-Itô decomposition also gives us an easy intuitive interpretation of the Lévy measure, the Lévy measure of a Borel set is equal to the expected number of jumps in a time interval $[0,1]$ with jump sizes in the Borel set.

The decomposition is based on the so called Poisson random measure, which describes the jumps.

**Definition 2.22** (Poisson random measure)
Let $(\Theta, \mathcal{B}(\Theta), \rho)$ be a $\sigma$–finite measure space. A family of $\mathbb{N}_0$–valued random variables $\{N(B) : B \in \mathcal{B}(\Theta)\}$ is called a Poisson random measure on $\Theta$ with intensity measure $\rho$ if

(i) for all $B \in \mathcal{B}$, $N(B) \sim \text{Poi}(\nu(B))$,

(ii) for all $B_1, ..., B_n \in \mathcal{B}$ disjoint, $N(B_1), ..., N(B_n)$ are independent,

(iii) for all $\omega \in \Omega$, $N(\cdot, \omega)$ is a measure on $\Theta$.

We use the following notation: If a random variable $X \sim \text{Poi}(\infty)$, then $X = \infty$ a.s. and if $X \sim \text{Poi}(0)$, then $X = 0$ a.s..

In Proposition 2.23 and 2.26 we state the Lévy-Itô decomposition.

**Proposition 2.23**
Let $\{L_t, t \geq 0\}$ be a $\mathbb{R}^d$–valued Lévy process over $(\Omega, \mathcal{F}, P)$ with Lévy triplet $(C, \gamma, \nu)$. We use $\Omega_0 \subset \Omega$ of Definition 2.1 (on $\Omega_0$, $L_t$ is cadlag, $P(\Omega_0) = 1$), and define, for all $(S, B) \in \mathcal{B}(\mathbb{R} \times \mathbb{R}^d)$,

$$J((S, B), \omega) := \begin{cases} \#\{t : (t, \Delta L_t(\omega)) \in (S, B)\} & \omega \in \Omega_0 \\ 0 & \omega \notin \Omega_0. \end{cases}$$

Then we have

(i) $\{J((S, B)) : (S, B) \in \mathcal{B}(\mathbb{R} \times \mathbb{R}^d)\}$ is a Poisson random measure on $\mathbb{R} \times \mathbb{R}^d$ with intensity $l(S) \cdot \nu(B)$, where $l$ denotes the Lebesgue measure.

(ii) Define

$$JL_t(\omega) := \int_0^t \int_{||x||>1} x J((ds, dx), \omega) \, ds \, dx + \lim_{\varepsilon \to 0} \int_0^t \int_{||x||\in(\varepsilon,1]} (x J((ds, dx), \omega) - x \, ds \, \nu(dx)).$$
Then there exists $\Omega_1 \in \mathbb{F}$ with $\mathbb{P}(\Omega_1) = 1$ such that, for all $\omega \in \Omega_1$, $J_t$ is defined for all $t \in [0, \infty)$ and the convergence is uniform in $t$ on every bounded interval. $\{J_t, t \geq 0\}$ is a Lévy process on $\mathbb{R}^d$ with Lévy triplet $(0, 0, \nu)$.

(iii) Define $C_Lt(\omega) := L_t(\omega) - J_t(\omega)$.

Then $\{C_Lt, t \geq 0\}$ is a Wiener process with covariance matrix $C$ and drift $\gamma$.

(iv) The processes $\{J_t, t \geq 0\}$ and $\{C_Lt, t \geq 0\}$ are independent.

Proof. In [Sato 2007, Theorem 19.2] the proof is given for an additive process that is more general than a Lévy process (without the stationary increment assumption). But the proof can be easily adapted to Lévy processes. The idea of the proof is to start with the construction of a Poisson random measure and a Lévy process that has the Lévy-Itô decomposition for a given Lévy triplet. Afterwards it has to be shown that the original Lévy process has the same decomposition. \hfill $\square$

Remark 2.24

(i) The process

$$\int_0^t \int_{||x||>1} x J((ds,dx),\omega) =: C_P L_t[1]$$

in Proposition 2.23 is either a compound Poisson process with intensity

$$\nu (\mathbb{R}^d\{|x| \leq 1\})$$

and jump size distribution function

$$F(x) = \nu((-\infty,x]\{|x| \leq 1\})/\nu(\mathbb{R}^d\{|x| \leq 1\})$$

or the zero process that is a.s. equal to zero (cf. [?, Theorem 2.3.7 and Theorem 2.3.9])

(ii) The process

$$\lim_{\varepsilon \to 0} \int_0^t \int_{||x|| \in (\varepsilon,1]} (x J((ds,dx),\omega) - x \, ds \, \nu(dx)) =: sJ L_t[1]$$

in Proposition 2.23 is a square integrable martingale. The term

$$\int_0^t \int_{||x|| \in (\varepsilon,1]} x J((ds,dx),\omega)$$

is a compound Poisson process, but letting $\varepsilon \to 0$, the limit does in general not exist. Hence the compensated drift term

$$\int_0^t \int_{||x|| \in (\varepsilon,1]} x \, ds \, \nu(dx)$$
2. Basics

is used \cite{Cont and Tankov 2004, p.81}.

(iii) The processes in (i) and (ii) are independent (cf. \cite{Cont and Tankov 2004, Proposition 3.7}).

Remark 2.25

Analogously to Remark 2.15, we also can decompose the process $J L_t$ at another point $\xi > 0$, then we have

\[ L_t(\omega) = C L_t(\omega) - \gamma t + \gamma [\xi] t + J L_t[\xi](\omega) =: C L_t(\omega) - \gamma t + \gamma [\xi] t + \int_0^t \int_{||x|| > \xi} (x J((ds, dx), \omega)) \\
+ \lim_{\varepsilon \to 0} \int_0^t \int_{||x|| \in (\varepsilon, \xi]} (x J((ds, dx), \omega)) - x ds \nu(dx), \]

where $\gamma [\xi]$, that has been defined in Remark 2.15 and $\{C L_t, t \geq 0\}$ is a non-standard Wiener process with covariance matrix $C$ and drift $\gamma$.

If $\int_{||x|| < 1} ||x|| \nu(dx) < \infty$, then the Lévy-Itô decomposition can be simplified, since the compensating drift term is not needed anymore. The next Proposition gives the result.

Proposition 2.26

Assume that the process $\{L_t, t \geq 0\}$ of Proposition 2.23 satisfies $\int_{||x|| < 1} ||x|| \nu(dx) < \infty$. Then there exists a $\Omega_2 \in \mathbb{F}$ with $P(\Omega_2) = 1$ such that, for any $\omega \in \Omega_2$,

\[ C P L_t[0](\omega) := \int_0^t \int_{x \in \mathbb{R}^d} (x J((ds, dx), \omega)) \]

is defined for all $t \geq 0$. The process $\{C P L_t[0], t \geq 0\}$ is a Lévy process with triplet $(0, \int_{||x|| \leq 1} x \nu(dx), \nu)$. Define

\[ C L_t(\omega) := L_t(\omega) - C P L_t[0](\omega) \quad \text{for } \omega \in \Omega_2. \]

Then the process $\{C L_t, t \geq 0\}$ is a Wiener process with covariance matrix $C$ and drift $\gamma - \int_{||x|| < 1} x \nu(dx)$. The processes $\{C P L_t[0], t \geq 0\}$ and $\{C L_t, t \geq 0\}$ are independent.

Proof. See e.g. \cite[Theorem 19.3]{Sato 2007}.

Remark 2.27

If $\nu(\mathbb{R}^d) < \infty$, the process $\{C P L_t[0], t \geq 0\}$ is either the zero process or a compound Poisson process with intensity $\nu(\mathbb{R}^d)$ and jump size distribution function $F(x) = \nu((-\infty, x]) / \nu(\mathbb{R}^d)$. In that case we denote $C P L_t[0] = C P L_t$ (cf. \cite{Cont and Tankov 2004, Proposition 3.3 and Proposition 3.8}).
Remark 2.28
Note that in Proposition 2.23 \(\{C_{L_t}, t \geq 0\}\) is a nonstandard Wiener process with covariance matrix \(C\) and drift \(\gamma = \gamma[1]\), whereas in Proposition 2.26 it is a nonstandard Wiener process with covariance matrix \(C\) and drift \(\gamma - \int_{||x||<1} x \nu(dx) = \gamma[0]\).

From the Lévy-Itô decomposition follows an easy interpretation of the Lévy measure.

Corollary 2.29
For the Lévy measure of a Lévy process \(\{L_t, t \geq 0\}\) we have
\[
\nu(B) := \mathbb{E}\left[\#\{t \in [0,1], \Delta L_t \neq 0, \Delta L_t \in B\}\right], \quad B \in \mathcal{B}(\mathbb{R}^d).
\]

Proof. The proof follows directly from Proposition 2.23 since the expectation of the Poisson random measure is equal to the intensity measure.

2.1.4. Further properties

The cumulants and the moments, respectively can be easily computed by the Lévy Khinchin representation.

Proposition 2.30 (Moments and cumulants)
Let \(\{L_t, t \geq 0\}\) a \(\mathbb{R}\)-valued Lévy process with triplet \((C, \gamma, \nu)\) and \(l \in \mathbb{N}\). Then we have
\[
(i) \quad \mathbb{E}\left[|L_t|^l\right] < \infty \text{ for all } t > 0, \text{ if and only if }
\]
\[
\int_{|x| \geq 1} |x|^l \nu(dx) < \infty.
\]
\[
(ii) \quad \text{If the absolute moments exist, the cumulants are given by}
\]
\[
\mathbb{E}[L_t] = t \left(\gamma_1 + \int_{|x| \geq 1} x_1 \nu(dx)\right),
\]
\[
\text{Var}[L_t] = C + \int_{\mathbb{R}} x^2 \nu(dx),
\]
\[
c_l(h) = t \int_{\mathbb{R}} x^l \nu(dx) \quad \text{for } l \geq 3.
\]

Proof. See e.g. Cont and Tankov (2004, Proposition 3.13).
2. Basics

Remark 2.31
Let \( \{L_t, t \geq 0\} \) be an \( \mathbb{R}^d \)-valued Lévy process with triplet \((C, \gamma, \nu)\) and \( i, j \in \{1, \ldots, d\} \). If \( \mathbb{E}[|L_{1,i}|^2] < \infty \) and \( \mathbb{E}[|L_{1,j}|^2] < \infty \), the covariance can be computed by the characteristic function and is equal to
\[
\text{Cov}[L_{1,i}, L_{1,j}] = C_{i,j} + \int_{\mathbb{R}^d} x_i x_j \nu(dx).
\]

In the next chapter we also consider a transformation of the Lévy measure. Therefore we need the moments of a random variable, that changes the jump sizes of the Lévy process.

Proposition 2.32
Let \( \{L_t, t \geq 0\} \) a Lévy process on \( \mathbb{R}^d \) and \( f: \mathbb{R}^d \to \mathbb{R}^+ \) measurable, such that \( \int_{\mathbb{R}^d} f(x) \nu(dx) < \infty \), \( \int_{\mathbb{R}^d} (f(x))^2 \nu(dx) < \infty \) and \( f(0) = 0 \) and let \( B_1, B_2 \in \mathbb{B}(\mathbb{R}^d) \) disjoint. Define
\[
Y_i(\omega) := \sum_{0 \leq t \leq t_0} f(\Delta L_t(\omega)) \mathbf{1}(\Delta L_t(\omega) \in B_i)
\]
for \( i = \{1, 2\} \), \( t_0 \in \mathbb{R}^+ \). Then, \( Y_i \) is a random variable and
\[
\mathbb{E}[Y_i] = t_0 \int_{B_i} f(x) \nu(dx),
\]
\[
\text{Var}[Y_i] = t_0 \int_{B_i} (f(x))^2 \nu(dx).
\]
for \( i \in \{1, 2\} \). \( Y_1 \) and \( Y_2 \) are independent.

Proof. We define \( f_n: \mathbb{R}^d \to \mathbb{R}_0^+ \), \( x \mapsto f(x) \mathbf{1}(|x| \geq 1/n) \). Then, \( f_n(x) \) converges monotonously to \( f(x) \), for all \( x \in \mathbb{R}^d \). We define
\[
Y_{i,n}(\omega) := \sum_{0 \leq t \leq t_0} f_n(\Delta L_t(\omega)) \mathbf{1}(\Delta L_t(\omega) \in B_i).
\]
By Sato (2007, Proposition 19.5) \( Y_{i,n} \) is a random variable,
\[
\mathbb{E}[Y_{i,n}] = t_0 \int_{B_i} f_n(x) \nu(dx),
\]
\[
\text{Var}[Y_{i,n}] = t_0 \int_{B_i} (f_n(x))^2 \nu(dx)
\]
and the families \( \{Y_{1,n}, n \in \mathbb{N}\} \) and \( \{Y_{2,n}, n \in \mathbb{N}\} \) are independent. But we have by the Monotone Convergence Theorem
\[
Y_i(\omega) = \lim_{n \to \infty} Y_{i,n}(\omega).
\]
Thus, \( Y_i \) is also a random variable and the moments can be computed by the Monotone Convergence Theorem. The independence follows from Sato (2007, Proposition 1.13).

20
The next proposition deals with the quadratic variation of a Lévy process.

**Proposition 2.33 (Quadratic variation)**
For a \( \mathbb{R} \)-valued Lévy process the quadratic variation is given by

\[
[L, L]_t = C_t + \sum_{s \in [0,t]} (\Delta L_s)^2.
\]

The corresponding moments are equal to

\[
\mathbb{E} [L, L]_t = C_t + \int_{\mathbb{R}} x^2 \nu(dx),
\]

\[
\text{Var} [L, L]_t = \int_{\mathbb{R}} x^4 \nu(dx).
\]

**Proof.** See e.g. [Cont and Tankov (2004, Example 8.5)] and Proposition 2.32.

The quadratic covariation is obtained analogously.

**Proposition 2.34 (Quadratic covariation)**
For a \( \mathbb{R}^d \)-valued Lévy process the quadratic covariation between the \( i \)-th and \( j \)-th component (\( i, j \in \{1, \ldots , d\} \)) is given by

\[
[L_i, L_j]_t = C_{i,j} t + \sum_{s \in [0,t]} \Delta L_{i,s} \Delta L_{j,s}.
\]

The corresponding moments are equal to

\[
\mathbb{E} [L_i, L_j]_t = C_{i,j} t + \int_{\mathbb{R}^2} x_i x_j \nu(dx),
\]

\[
\text{Var} [L_i, L_j]_t = \int_{\mathbb{R}^2} x_i^2 x_j^2 \nu(dx).
\]

**Proof.** See e.g. [Cont and Tankov (2004, Example 8.10)] and Proposition 2.32.

In the case of infinite activity it is interesting to investigate the behavior of the infinitely many small jumps. Since \( \int_{|x| \leq 1} |x|^2 \nu(dx) < \infty \), the sum of the squared jump sizes has always to be finite on a compact time interval. However, it is possible that e.g. the sum of all absolute jump sizes is finite, too. We introduce the **Blumenthal-Getoor index** and use [Cont and Tankov (2004, p.228)].

**Definition 2.35 (Blumenthal-Getoor index)**
The Blumenthal-Getoor index of an \( \mathbb{R}^d \)-valued Lévy process with Lévy measure \( \nu \) is defined by

\[
\gamma := \inf \{ \gamma' > 0, \int_{|x| \leq 1} |x|^{\gamma'} \nu(dx) < \infty \}.
\]
So the Blumenthal-Getoor index tells us how the Lévy measure behaves near the origin. Obviously, we always have $\gamma \in [0, 2]$. We have used the Euclidian norm, instead of the supremum norm, because in Cont and Tankov (2004) the Euclidian norm is used, too.

### 2. Basics

2.1.5. Parametric models

In this subsection we consider the Merton model and the Kou model. In both models the Lévy process has finite activity and the jump size distribution is given in a parametric form.

The Merton model is introduced in Merton (1976) as one of the first extensions of the Black-Scholes model. It is assumed, that the logarithmic stock price follows a Jump diffusion with normally distributed jump sizes, i.e. the jump density is of the form

$$f(x) := \frac{1}{\sqrt{2\pi \eta}} \exp \left( -\frac{1}{2} \left( \frac{x - \beta}{\eta} \right)^2 \right),$$

where $\beta \in \mathbb{R}$ is the expectation and $\eta > 0$ the standard deviation of the jump size distribution. The Lévy measure is given by $\nu(dx) := \lambda f(x)dx$, where $\lambda > 0$ is the jump intensity. This univariate model can be simply extended to a multivariate model by using a multivariate normal distribution for the jump sizes and a multivariate Wiener process for the diffusion part.

In the Kou model, introduced in Kou (2002), it is assumed that the logarithmic stock price follows a Jump diffusion process with asymmetric double exponential distributed jump sizes, i.e. the jump density is of the form

$$f(x) := p \eta_1 \exp(-\eta_1 x) \mathbb{I}(x \geq 0) + (1 - p) \eta_2 \exp(\eta_2 x) \mathbb{I}(x < 0),$$

where $\eta_1 > 0$, $\eta_2 > 0$ and $p \in (0, 1)$. The Lévy measure is given by $\nu(dx) := \lambda f(x)dx$ with jump intensity $\lambda > 0$.

The Kou model is widely used in financial mathematics. Kou (2002) mentions explicitly two empirical phenomena that can be explained well by the Kou model. First, the asymmetric leptokurtic feature, which means that the return distribution is skewed to the left and has a higher peak and heavier tails than those of the normal distribution. Second, the volatility smile, which means that in reality it is widely recognized that the implied volatility is a convex curve of the strike price. The Kou model can produce such a “smile”. Furthermore an advantage of the Kou model is, that it leads to analytical solutions for many option price problems, e.g. for European put and call options (see Kou, 2002, Kou and Wang, 2004).

Besides the Merton and the Kou model, in literature fully parameterized pure jump Lévy processes are widely used. A large class are the Generalized hyperbolic models that are introduced in Barndorff-Nielsen and Halgreen (1977). Here the probability density follows a Generalized Hyperbolic distribution. The
2.2. Weak convergence in a generalized Skorohod space

Process has no diffusion part and in general it is of infinite variation. That class contains various popular parameterized processes as special cases, e.g., the Variance Gamma process, introduced in Madan and Seneta (1987), the Hyperbolic process (cf. Bingham and Kiesel 2001a,b) and the Normal Inverse Gaussian process (cf. Barndorff-Nielsen 1995, 1997, Rydberg 1996a,b, 1997). Another example of a pure jump Lévy process is the CGMY process, introduced in Carr et al. (2002). This process is parameterized by four parameters, that allow the process to have finite or infinite activity and to be of finite or infinite variation. The Lévy measure of the CGMY process is given explicitly. Another example is the Meixner process (cf. Schoutens and Teugels 1998, Schoutens 2000, Grigelionis 1999). It is a pure jump process of infinite variation and driven by three parameters. A good overview over parametric, pure jump Lévy processes can be found in Schoutens (2003, Chapter 5.3). In this thesis we focus on disentangling jumps from diffusion, so we do not study these pure jump processes in detail.

2.2. Weak convergence in a generalized Skorohod space

In this section we introduce the concept of weak convergence of probability measures. The classical setting of this theory is formulated in Billingsley (1968). Convergence in distribution of random vectors is equivalent to or even defined by weak convergence of their distributions that are probability measures on \( \mathbb{R}^d \). Convergence in distribution of stochastic processes can analogously be defined by weak convergence of the corresponding probability measures on a suitable function space. For our purpose those function spaces are interesting that contain càdlàg functions. The space of all càdlàg functions from \([0,1]\) to \(\mathbb{R}\) is called Skorohod space and is denoted by \(D[0,1]\). First topologies for \(D[0,1]\) were provided in Prohorov (1953, 1956), Skorokhod (1956b, 1955, 1956a, 1957) and Kolmogorov (1956). In Billingsley (1968) the topology for \(D[0,1]\) is presented that is called the Skorohod topology today. The space \(D[0,1]\) was extended in several directions. The space \(D[0,1]^d\) is considered in Bickel and Wichura (1971), Neuhaus (1971) and Straf (1972), the space \(D[0,\infty)\) in Lindvall (1973) and Billingsley (1999). Bass and Pyke (1983) consider weak convergence of processes indexed by a collection of subsets of \([0,1]^d\).

In Subsection 2.2.1 we start with the general theory of weak convergence of probability measures on an arbitrary metric space. In Subsection 2.2.2 the concept of convergence in distribution of random elements is introduced that is based on weak convergence of the corresponding distributions. In Subsection 2.2.3 we consider the classical Skorohod space \(D[0,1]\) in detail and Subsection 2.2.4 deals with the generalization \(D[0,1]^d\) that is needed in Chapter 3.

2.2.1. Weak convergence of probability measures

In this subsection we follow Billingsley (1999, Chapter 1). We consider a metric space \(S\) and denote its Borel \(\sigma\)-algebra by \(\mathcal{B}(S)\). All probability measures we...
consider in this subsection are defined on $\mathcal{B}(S)$. We start with the definition of weak convergence.

**Definition 2.36** (Weak convergence) 
A sequence of probability measures $P_n$ converges weakly to a probability measure $P$, if
\[ \int_S f \, dP_n \rightarrow \int_S f \, dP, \]
for every bounded, continuous and real function $f$ on $S$. The weak convergence is denoted by $P_n \Rightarrow P$.

The next proposition shows that the probability measure is determined by the values of $\int_S f \, dP$ for bounded, continuous and real functions on $S$.

**Proposition 2.37** 
Probability measures $P$ and $Q$ coincide if
\[ \int_S f \, dP = \int_S f \, dQ \]
for every bounded, continuous and real function $f$ on $S$.

*Proof.* See [Billingsley (1999, Theorem 1.2)].

Proposition 2.37 implies that a sequence of probability measures cannot converge weakly to each of two different limits.

The next proposition gives equivalent conditions for weak convergence that are sometimes used as definitions themselves.

**Proposition 2.38** (Portmanteau Theorem) 
The following conditions are equivalent

1. $P_n \Rightarrow P$.
2. $\lim_{n \to \infty} \int_S f \, dP_n = \int_S f \, dP$ for all bounded, uniformly continuous real $f$.
3. $\limsup_{n \to \infty} P_n(F) \leq P(F)$ for all closed $F \in \mathcal{B}(S)$.
4. $\liminf_{n \to \infty} P_n(G) \geq P(G)$ for all open $G \in \mathcal{B}(S)$.
5. $\lim_{n \to \infty} P_n(A) = P(A)$ for all $P$-continuity sets $A \in \mathcal{B}(S)$, i.e. $P(\delta A) = 0$, where $\delta A$ is the boundary of $A$.

*Proof.* See [Billingsley (1999, Theorem 2.1)].

The next definition defines special subclasses of $\mathcal{B}(S)$ that play an important role for proving weak convergence.
2.2. Weak convergence in a generalized Skorohod space

Definition 2.39 (Separating and convergence-determining class)

(i) A subclass $A$ of $\mathcal{B}(S)$ is called convergence-determining class, if for every $P$ and every sequence $\{P_n\}$, convergence $P_n(A) \to P(A)$ for all $P$–continuity sets $A \in A$ implies $P_n \Rightarrow P$.

(ii) A subclass $A$ of $\mathcal{B}(S)$ is called separating class if $P(A) = Q(A)$ for all $A \in A$ implies $P(A) = Q(A)$ for all $A \in \mathcal{B}(S)$.

In Billingsley (1999) conditions can be found guaranteeing that a subclass is a separating or a convergence-determining class.

In the Euclidian space $\mathbb{R}^d$ with $d \in \mathbb{N}$ it can be shown that the sets $\{y \in \mathbb{R}^d : y \leq x, x \in \mathbb{R}^d\}$ form a convergence-determining class and $\{y \in \mathbb{R}^d : y \leq x\}$ is a continuity set if and only if the distribution function $F(x) := P\{y \in \mathbb{R}^d : y \leq x\}$ is continuous at $x$ (cf. Billingsley, 1999, p.18 and 19). So it is sufficient to show that the distribution functions $F_n$ converge at every continuity point to $F$.

Showing weak convergence in a function space is not that easy. The strategy that is introduced in Subsection 2.2.3 is to show weak convergence of the finite-dimensional distributions – which alone is not enough – and to use additional tightness conditions. To introduce this theory in Subsection 2.2.3 we need some further properties and definitions.

Proposition 2.40

$\{P_n\}$ converges weakly to $P$ if and only if, each subsequence contains a further subsequence that converges weakly to $P$.


Definition 2.41 (Relative Compactness)

A family $\Pi$ of probability measures on a metric space $(S, \mathcal{B}(S))$ is relative compact, if every sequence of elements of $\Pi$ contains a weakly convergent subsequence.

Definition 2.42 (Tightness)

A family $\Pi$ of probability measures on a metric space $(S, \mathcal{B}(S))$ is tight, if for all $\varepsilon > 0$ there exists a compact set $K_\varepsilon \in B(S)$ such that $P(K_\varepsilon) > 1 - \varepsilon$ for all $P \in \Pi$.

The next Proposition states the connection between relative compactness and tightness. The first part is called direct half, the second is called converse half.
Proposition 2.43 (Prohorov’s Theorem)

(i) Let $\Pi$ be a family of probability measures. If $\Pi$ is tight, then it is relative compact.

(ii) Suppose that $S$ be separable and complete. If a family of probability measures $\Pi$ is relative compact, it is complete.

Proof. See Billingsley (1999, Theorem 5.1 and Theorem 5.2).

So summarizing we obtain the following obvious corollary.

Corollary 2.44

If $\{P_n, n \in \mathbb{N}\}$ is tight, and if all weakly convergent subsequences converge to the same limit $P$, then $P_n \Rightarrow P$.


Next, we introduce the so called mapping theorem. Let $(S', d')$ be another metric space and $h : S \rightarrow S'$ a $\mathcal{B}(S) / \mathcal{B}(S')$-measurable function. Then a probability measure on $(S', \mathcal{B}(S'))$ is defined by

\[ P' := P \circ h^{-1}. \]

The mapping theorem shows under which conditions weak convergence in $S$ implies weak convergence of the corresponding probability measures on $S'$.

Proposition 2.45 (Mapping Theorem)

Let $(S', d')$ be another metric space and $h : S \rightarrow S'$ a $\mathcal{B}(S) / \mathcal{B}(S')$-measurable function. Define

\[ D_h := \{x \in S : h(x) \text{ discontinuous}\}. \]

Then $D_h \in \mathcal{B}(S)$. Let $\{P_n, n \in \mathbb{N}\}$, $P$ be probability measures on $S$. If $P_n \Rightarrow P$ and $P(D_h) = 0$, then

\[ P_n \circ h^{-1} \Rightarrow P \circ h^{-1}. \]

Proof. See Billingsley (1999, Theorem 2.7)).

2.2.2. Convergence in distribution

The theory of convergence in distribution is closely connected to the theory of weak convergence. We give a brief overview.

Definition 2.46 (Random element)

Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space. A random element with domain $\Omega$ and range $S$ is a mapping $X : \Omega \rightarrow S$, such that $X$ is $\mathcal{F} / \mathcal{B}(S)$-measurable.
2.2. Weak convergence in a generalized Skorohod space

In case of $S = \mathbb{R}$, the random element is a random variable, in case of $S = \mathbb{R}^d$ a random vector and if $S$ is a function space, $X$ is a stochastic process.

**Definition 2.47** (Distribution)
The distribution of a random element $X$ is a probability measure on $(S, \mathcal{B}(S))$ defined by
\[ P(A) = \mathbb{P}(X \in A), \quad A \in \mathcal{B}(S). \]

Convergence of random elements is defined by weak convergence of their distributions.

**Definition 2.48** (Convergence in distribution)
A sequence of random elements $\{X_n, n \in \mathbb{N}\}$ converges in distribution to a random element $X$ if
\[ P_n \Rightarrow P, \]
where $P_n$ is the distribution of $X_n$ and $P$ the distribution of $X$, respectively. In case of $S = \mathbb{R}^d$ we denote the convergence in distribution by $X_n \overset{d}{\rightarrow} X$ and if $S$ is a function space by $X_n \overset{D}{\rightarrow} X$.

Using Definition 2.36, convergence in distribution means that
\[ \mathbb{E}[f(X_n)] \rightarrow \mathbb{E}[f(X)], \]
for every bounded, continuous and real function $f$ on $S$.

Convergence in distribution follows from convergence in probability, this result is given in Proposition 2.50. Therefore, we use a definition of convergence in probability, that can be found in Kallenberg (1997, p.40).

**Proposition 2.50**
Let $\{X_n, n \in \mathbb{N}\}$ and $X$ be random elements with domain $\Omega$ and range $S$. Then $X_n \overset{P}{\rightarrow} X$ implies $X_n \overset{d}{\rightarrow} X$.

The converse site holds true if $X$ is a constant.
2. Basics

Proof. See e.g. Kallenberg (1997, Lemma 3.7).

2.2.3. The Skorohod space

In this subsection we introduce weak convergence in the classical Skorohod space. We follow Billingsley (1999, Chapter 3). The Skorohod space is denoted by $D[0,1]$ and defined as follows.

**Definition 2.51** (Skorohod space)
The space $D[0,1]$ is the space of all real valued functions $x$ on $[0,1]$ that are right-continuous and have left-hand limits, i.e.

1. for $0 \leq t < 1$: $\lim_{s \downarrow t} x(s) = x(t)$,
2. for $0 \leq t \leq 1$: $\lim_{s \uparrow t} x(s)$ exists.

A function has a discontinuity of the first kind at $t$, if $\lim_{s \downarrow t} x(s)$ and $\lim_{s \uparrow t} x(s)$ exist, but are different and $x(t)$ lies between them. So $D[0,1]$ is the space of all functions whose discontinuities are of the first kind. The condition $\lim_{s \downarrow t} x(s) = x(t)$ is just a standardization. Of course $C[0,1]$, the space of all real continuous functions on $[0,1]$, is a subspace of $D[0,1]$.

In $C[0,1]$ the modulus of continuity is defined as follows

$$w_x(\delta) := \sup_{s,t \in [0,1], s-t<\delta} |x(s) - x(t)|.$$  

In the Skorohod space the following generalization is considered.

**Definition 2.52** (Modulus)
We define

$$w'_x(\delta) := \inf \max_{1 \leq i \leq r} w_x[t_{i-1}, t_i],$$

where

$$w_x[t_{i-1}, t_i] := \sup_{s,t \in [t_{i-1}, t_i]} |x(s) - x(t)|$$

and the infimum is taken over all finite sets of points satisfying

$$0 = t_0 < t_1 \cdots < t_r = 1, \quad \text{and} \quad t_i - t_{i-1} > \delta.$$  

A necessary and sufficient condition for $x \in D[0,1]$ is that $\lim_{\delta \to 0} w'_x(\delta) = 0$ (cf. Billingsley 1999, p.123). From this it follows that an element of $D[0,1]$ can only have a finite number of jumps with an absolute jump size that exceeds a given positive number and thus the overall number of discontinuities is at most countable.

Next, a metric $d$ for $D[0,1]$ is introduced.
2.2. Weak convergence in a generalized Skorohod space

**Definition 2.53** (Skorohod metric)

Let \( \Lambda \) be the class of all strictly increasing continuous functions from \([0, 1]\) to \([0, 1]\), such that for all \( \lambda \in \Lambda \),
\[
\lambda(0) = 0 \quad \text{and} \quad \lambda(1) = 1.
\]

The identity map is denoted by \( \lambda_I \). Then we define the metric of \( D[0, 1] \) by
\[
d(x, y) := \inf_{\lambda \in \Lambda} \{||\lambda - \lambda_I|| \vee ||x \circ \lambda||\},
\]
where \( ||\cdot|| \) denotes the supremum norm.

In [Billingsley, 1999, p.124] it is shown that \( d \) is a metric. Skorohod convergence implies that \( x_n(t) \to x(t) \) holds true for every continuity point \( t \) of \( x \).

The Skorohod metric defines the so called Skorohod topology. If the Skorohod topology is relativized to \( C[0, 1] \), it coincides with the uniform topology there (cf. [Billingsley, 1999, p.124]).

Intuitively a function \( x \) is near to a function \( y \) in the Skorohod metric if for every \( t \) the value \( x(t) \) is equal to \( y(t) \pm \varepsilon \) (small change of space) or equal to \( y(t \pm \varepsilon) \) (small change of time). In contrast, the uniform metric that is used in \( C[0, 1] \) only allows to change the space a little bit.

However, under \( d \) the space \( D[0, 1] \) is not complete. In [Billingsley, 1999] an equivalent metric \( d_0 \) is constructed by using a different norm on \( \Lambda \), such that \( D[0, 1] \) is complete in \( d_0 \). However, we do not introduce this concept in detail. Furthermore it can be shown that \( D[0, 1] \) is separable under \( d \) and \( d_0 \) (cf. [Billingsley, 1999, Theorem 12.2]).

In \( C[0, 1] \) the Arzelà-Ascoli theorem gives a characterization of relative compact subsets. (cf. [Billingsley, 1999, Theorem 7.2]). A subset \( A \) is relative compact, if its closure is compact. The following proposition is an analogue in \( D[0, 1] \). The proposition is fundamental for proving weak convergence.

**Proposition 2.54**

A set \( A \subset D[0, 1] \) is relative compact in the Skorohod topology, if and only if,
\[
\sup_{x \in A} ||x|| < \infty \quad \text{and} \quad \lim_{\delta \to 0} \sup_{x \in A} u'_x(\delta) = 0.
\]

**Proof.** See [Billingsley, 1999, Theorem 12.3]. \( \square \)

The strategy of proving weak convergence is to show weak convergence of the finite-dimensional distributions and to use tightness conditions. The finite-dimensional distributions are defined via the natural projections.

**Definition 2.55** (Natural projections)

For a finite subset \( \{t_1, \ldots, t_k\} \subset [0, 1] \) we define the natural projection \( \pi_{t_1, \ldots, t_k} \) from \( D[0, 1] \) to \( \mathbb{R}^k \) by
\[
\pi_{t_1, \ldots, t_k}(x) := (x(t_1), \ldots, x(t_k)).
\]
The next proposition tells us, when the natural projections are continuous and shows that the natural projections are measurable.

**Proposition 2.56**

(i) The projections $\pi_0$ and $\pi_1$ are continuous. For $0 < t < 1$, $\pi_t$ is continuous at $x$ if and only if $x$ is continuous at $t$.

(ii) Each $\pi_{t_1 \ldots t_k}$ is $\mathbb{B}(D[0,1])/\mathbb{B}(\mathbb{R}^k)$-measurable.

*Proof.* See [Billingsley (1999), Theorem 12.5].

The inverse images of the natural projections are called *finite-dimensional sets*. Next, we define the *finite-dimensional distribution*.

**Definition 2.57** (Finite-dimensional distribution)

Let $P$ be a probability measure on $D[0,1]$. For a finite subset $\{t_1, \ldots, t_k\} \subset [0,1]$ a probability measure on $(\mathbb{R}^k, \mathbb{B}(\mathbb{R}^k))$ is defined by

$$P \circ \pi_{t_1 \ldots t_k}^{-1}$$

and called finite-dimensional distribution of $P$.

If $P$ is the distribution on a stochastic process $X$, the finite-dimensional distribution $P \circ \pi_{t_1 \ldots t_k}^{-1}$ is the distribution of the random vector

$$(X(t_1), \ldots, X(t_k)) := \pi_{t_1 \ldots t_k} \circ X.$$ 

An interesting question is whether weak convergence of a sequence of probability measures on $D[0,1]$ implies weak convergence of the finite-dimensional distributions.

**Proposition 2.58**

Let $\{P_n, n \in \mathbb{N}\}$ be a sequence of probability measures on $D[0,1]$ that converge weakly to $P$. If

$$(t_1, \ldots, t_k) \subset \{t \in [0,1] : P(x : x \text{ is discontinuous at } t) = 0\},$$

then

$$P_n \circ \pi_{t_1 \ldots t_k}^{-1} \Rightarrow P \circ \pi_{t_1 \ldots t_k}^{-1}.$$ 

*Proof.* The proof is given in [Billingsley (1999), p.138]. The Mapping theorem (Proposition 2.45) is used as well as Proposition 2.56.

For our purpose the converse part is more interesting. Usually it is easy to prove weak convergence of the finite-dimensional distributions.
2.2. Weak convergence in a generalized Skorohod space

**Proposition 2.59**
If $\{P_n\}$ is tight, and if

$$P_n \circ \pi_{t_1,\ldots,t_k} \Rightarrow P \circ \pi_{t_1,\ldots,t_k}$$

for all $\{t_1,\ldots,t_k\} \subset [0,1]$, then $P_n \Rightarrow P$.

**Proof.** The proof is given in [Billingsley (1999, Theorem 13.1)](#). Since $\{P_n\}$ is tight, every subsequence has a further weakly convergent subsequence. The limit measures of those subsequences have to coincide with $P$ on a separating class of $D[0,1]$ that is constructed by the finite-dimensional sets of the natural projections and Proposition 2.58. So $P$ and the limit measures of the subsequences have to be identical. \(\square\)

Next, we have to find suitable conditions such that $\{P_n\}$ is tight. The following proposition is based on Proposition 2.54.

**Proposition 2.60**
The sequence $\{P_n\}$ is tight if and only if

$$\lim_{a \to \infty} \limsup_{n \in \mathbb{N}} P_n(x : \|x\| \geq a) = 0$$

and for all $\varepsilon > 0$

$$\lim_{\delta \to 0} \limsup_{n \in \mathbb{N}} P(x : w'_x(\delta) \geq \varepsilon) = 0.$$

**Proof.** See [Billingsley (1999, Theorem 13.2)](#). \(\square\)

The next proposition is formulated directly for stochastic processes with range $D[0,1]$ and gives a more sophisticated criterion for convergence in distribution. The result is based on Proposition 2.60.

**Proposition 2.61**
Suppose that

$$(X_n(t_1),\ldots,X_n(t_k)) \overset{d}{\to} (X(t_1),\ldots,X(t_k)), \ (n \to \infty)$$

for all $\{t_1,\ldots,t_k\} \subset [0,1]$ and that

$$X(1) - X(1-\delta) \overset{d}{\to} 0, \ (\delta \to 0)$$

and that, for $r \leq s \leq t$, $n \in \mathbb{N}$,

$$\mathbb{E}[\|X_n(s) - X_n(r)\|^{2\beta} | X_n(t) - X_n(s)\|^{2\beta}] \leq (F(t) - F(r))^{2\alpha},$$

where $\beta \geq 0$ and $\alpha > \frac{1}{2}$, and $F$ is a nondecreasing continuous function on $[0,1]$. Then,

$$X_n \overset{D}{\to} X.$$

**Proof.** See [Billingsley (1999, Theorem 13.5)](#). \(\square\)
2.2.4. A generalization of the Skorohod space in higher dimensions

This subsection deals with weak convergence in the space $D[0,1]^d$. We follow Bickel and Wichura (1971) and start with the definition of $D[0,1]^d$.

**Definition 2.62**

Let $t = (t^{(1)}, \ldots, t^{(d)}) \in [0,1]^d$ and $R_i$ be one of the relations $<$ or $\geq$ for all $i \in \{1, \ldots, d\}$ and the quadrant $Q_{R_1, \ldots, R_d}(t)$ is defined by

$$Q_{R_1, \ldots, R_d}(t) := \left\{ s = (s^{(1)}, \ldots, s^{(d)}) \in [0,1]^d : s^{(i)} R_i t^{(i)}, i = 1, \ldots, d \right\}.$$

The space $D[0,1]^d$ contains all $x$ that satisfy

$$\lim_{s \to t, s \in Q_{R_1, \ldots, R_d}(t)} x(s)$$

exists for all $t \in [0,1]^d$ and for all $R_1, \ldots, R_d$ and

$$x(t) = \lim_{s \to t, s \in Q_{\geq, \ldots, \geq}(t)} x(s)$$

for all $t \in [0,1]^d$.

Next, we define a metric on $D[0,1]^d$, that coincides to the Skorohod metric, if $d = 1$.

**Definition 2.63**

Let $\Lambda$ be the set of all functions $\lambda : [0,1]^d \to [0,1]^d$ of the form $\lambda(t^{(1)}, \ldots, t^{(d)}) = (\lambda_1(t^{(1)}), \ldots, \lambda_d(t^{(d)}))$ such that $\lambda_i : [0,1] \to [0,1]$ is continuous, strictly increasing and fixes zero and one. The identity map is denoted by $\lambda_I$. Then the metric of $D[0,1]^d$ is defined by

$$d(x,y) = \inf_{\lambda \in \Lambda} (||x - y \circ \lambda|| \wedge ||\lambda - \lambda_I||),$$

where $|| \cdot ||$ denotes the supremum norm.

The strategy of proving weak convergence is developed analogously to the space $D[0,1]$. We directly turn to random elements with range $D[0,1]^d$. An analogous result of Proposition 2.61 is presented in Bickel and Wichura (1971, Theorem 3). However, we use a version that is given in Heinrich and Schmidt (1985, Lemma 3). Therefore we need the definition of neighboring blocks.

**Definition 2.64** (Neighboring blocks)

A block $B$ in $[0,1]^d$ is a subset of $[0,1]^d$ of the form

$$\{ s \in [0,1]^d : s^{(i)} = t^{(i)} \}$$

where $s = (s^{(1)}, \ldots, s^{(d)})$ and $t = (t^{(1)}, \ldots, t^{(d)})$. The $p$-th face of $B = \{ s \in [0,1]^d : s^{(i)} = t^{(i)} \}$ is defined by

$$\{ s \in [0,1]^d : s^{(i)} = t^{(i)} \}.$$
2.2. Weak convergence in a generalized Skorohod space

Disjoint blocks $B$ and $C$ are neighbors, if they have the same $p$-th face for some $p \in \{1, \ldots, d\}$.

Now we consider a random element $X$ on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ with range $D[0,1]^d$ and for each block $B = (s, t]$ we define the increment of $X$ around $B$ by

$$X(B) := \sum_{i=1}^{d} \sum_{\varepsilon_i \in \{0,1\}} (-1)^{d - \sum_{i=1}^{d} \varepsilon_i} X(s^{(1)} + \varepsilon_1(t^{(1)} - s^{(1)}), \ldots, s^{(d)} + \varepsilon_d(t^{(d)} - s^{(d)})).$$

**Proposition 2.65**

Let $\{X_n, n \in \mathbb{N}\}$ and $X$ be random elements with range $D[0,1]^d$. Suppose that the following conditions are satisfied.

(i) For all finite subsets $\{t_1, \ldots, t_k\} \subset [0,1]^d$ we have

$$(X_n(t_1), \ldots, X_n(t_k)) \xrightarrow{d} (X(t_1), \ldots, X(t_k)).$$

(ii) For all $\varepsilon > 0$ and all $p \in \{1, \ldots, d\}$ we have

$$\lim_{t^{(p)} \uparrow 1} \mathbb{P}(\|X(t^{(1)}, \ldots, t^{(p)}, \ldots, t^{(d)}) - X(t^{(1)}, \ldots, 1, \ldots, t^{(d)})\| \geq \varepsilon) = 0.$$

(iii) There exists a $n_0 \in \mathbb{N}$ such that

$$\mathbb{E}[|X_n(B)|^\gamma |X_n(C)|^\gamma] = (\mu(B))^{\alpha} (\mu(C))^{\alpha}$$

for every pair $(B, C)$ of neighboring blocks in $[0,1]^d$ and $n \geq n_0$, where $\gamma > 0$, $\alpha > \frac{1}{2}$ and $\mu$ is a finite non-negative measure on $[0,1]^d$.

Then, as $n \to \infty$,

$$X_n \xrightarrow{d} X.$$ 

**Proof.** See Heinrich and Schmidt (1985, Lemma 3). □

33
3. Estimation method

In the following chapter we introduce the estimation method. In Section 3.1 we give the estimation setting. The high-frequency framework and the critical region are defined. In Section 3.2 we estimate the covariance matrix and the drift of a Wiener process based on high frequency data. In Section 3.3 a Lévy process is considered under the assumption that the process can be continuously observed. Section 3.4 deals with the estimators that are based on the critical region in the case of finite activity. The infinite activity case is considered in Section 3.5. In Section 3.6 the Blumenthal-Getoor index is estimated, that measures the behavior of the infinitely many small jumps. Section 3.7 deals with the practical choice of the critical region and in Section 3.8 a correction method for the univariate and finite activity case is introduced. Finally, in Section 3.9 simulations and real data applications are performed.

3.1. Estimation setting

As in Chapter 2 we use a suitable probability space \((\Omega, \mathcal{F}, \mathbb{P})\). On this probability space we define an \(\mathbb{R}^d\)-valued Lévy process \(\{L_t, t \geq 0\}\). Its Lévy triplet is given by \((C, \gamma, \nu)\). The Lévy process can be used to describe a logarithmic stock price development, i.e.

\[
L_t = \log S_t, \quad t \geq 0,
\]

where \(\{S_t, t \geq 0\}\) is the stock price.

We do not require that the Lévy measure \(\nu\) is absolute continuous in the multivariate framework. However, we require that the Lévy measures \(\tilde{\nu}_i\) of the corresponding univariate component Lévy processes \(\{L_{t,i}, t \geq 0\}\) are absolute continuous. Keep in mind that \(\tilde{\nu}_i\) is different to the margins \(\nu_i\) of the Lévy measure \(\nu\) (cf. Corollary 2.21).

**Assumption 3.1**

For all \(i \in \{1, \ldots, d\}\), the Lévy measure \(\tilde{\nu}_i\) of the univariate component Lévy process \(\{L_{t,i}, t \geq 0\}\) is absolute continuous with density \(f_i\) and, for all \(\varepsilon > 0\), there exists \(\kappa_\varepsilon > 0\) such that

\[
f_i(x) < \kappa_\varepsilon, \quad x \in (-\varepsilon, \varepsilon)^c.
\]

Furthermore we get technical problems if the covariance matrix \(C\) is singular. It would be easy to incorporate this trivial case. However, we do not consider this case to avoid technical overload.
3. Estimation method

Assumption 3.2
The covariance matrix $C$ is not singular.

We impose that we can observe the process at high frequent, equidistant time points. In the context of asymptotic statistics high frequency means that the distance between two time points tends to zero.

Definition 3.3 (High frequency time points)
The time points are defined by
\[ t_k := h k, \quad k = 1, \ldots, n, \quad h := \kappa n^{-\alpha}, \quad (3.1) \]
where
\[ \alpha \in (0, 1], \quad \kappa > 0. \]

The time horizon is given by
\[ T = \frac{1}{\kappa} n^{1-\alpha}. \]

For simplification in the following we set w.l.o.g. $\kappa = 1$. In most cases we assume
\[ \alpha \in (1/2, 1] \]
to ensure that $T \to \infty$ and $h/T \to 0$. If we estimate the covariance matrix $C$, we will also allow $\alpha = 1$, due to the self-similarity property of the Wiener process (Proposition 2.4).

Next we define the increments of the process at this time points.

Definition 3.4 (Data points)
For $k = 1, \ldots, n$, we define
\[ X_{h,k} := L_{kh} - L_{k-1}h \quad k = 1, \ldots, n. \]

Since $L_t$ is a Lévy process, $X_{h,1}, \ldots, X_{h,n}$ are i.i.d. for all $h > 0$.

Our strategy to develop estimators is as follows. In Section 3.2 we think about estimators for the diffusion parameters under the unrealistic assumption that we know the increments of the diffusion part of the process which are defined as follows
\[ C X_{h,k} := C L_{kh} - C L_{k-1}h \quad k = 1, \ldots, n, \]
where $\{C L_t, t \geq 0\}$ is a nonstandard Wiener process, which is obtained by the Lévy-Itô decomposition (Proposition 2.23 and 2.26). Next, in Section 3.3 we develop estimators for the Lévy measure under the unrealistic assumption that we can observe the process continuously. That means we know the exact jump times and sizes
\[ \{\Delta L_t, t \in [0, T]\}, \quad \text{where} \quad \Delta L_t := L_t - \lim_{s \to t^-} L_s. \]
3.1. Estimation setting

We call the estimators that are based on these unrealistic assumptions **pre-estimators**. It is relatively straightforward to show their asymptotic properties.

However, in reality we do not know the increments of the diffusion part of the process and cannot observe the process continuously. To disentangle jumps from diffusion, we choose the following approach. We introduce a critical set $B_n \subset \mathbb{R}^d$. If we have $X_{k,h} \notin B_n$, we consider $X_{k,h}$ as a jump. If we have $X_{k,h} \in B_n$, we assign $X_{k,h}$ to the continuous part. We choose the character $B$ for **Bound** and use the following definition.

**Definition 3.5** (Critical region)

For all $n$ we define

$$B_n(D, \beta_n) := \{x \in \mathbb{R}^d : x^T D_n^{-1} x \leq \beta_n b_n^2\},$$

where $D$ is a symmetric $d \times d$–matrix, $\beta > 0$ and

$$b_n := \sqrt{2h \log n}.$$

We call a sequence of random sets $\{B_n, n \in \mathbb{N}\}$ a sequence of critical regions, if there exists $1 < \beta' < \beta'' < \infty$ such that

$$B_n(C, \beta') \subseteq B_n[\omega] \subseteq B_n(I, \beta''),$$

for all $n \in \mathbb{N}$ and $\omega \in \Omega$ ($I$ is the identity matrix). An element $B_n$ of the sequence is called critical region. For $\{\beta_n, n \in \mathbb{N}\}$ satisfying $\inf_{n \in \mathbb{N}} \beta_n > 1$ and $\sup_{n \in \mathbb{N}} \beta_n < \infty$, we call $\{B_n = B_n(C, \beta_n), n \in \mathbb{N}\}$ a sequence of true critical regions.

The matrix $I$ in the upper boundary $B_n(I, \beta'')$ is chosen for technical reasons. Every matrix with full rank would be suitable. In the following we assume w.l.o.g. that $B_n$ is deterministic to avoid technical overload. Important for the proofs are the bounds $B_n(C, \beta')$ and $B_n(I, \beta'').$

The true critical regions form ellipsoids in $\mathbb{R}^d$. We choose that definition, since the level curve of the density of the multivariate normal distribution forms an ellipsoid, too.

The critical set $B_n$ is chosen in such a way that it contains asymptotically the diffusion increments. In the univariate case, we could simply use Lévy’s modulus of continuity Theorem (Proposition 2.5). It is formulated for a finite time horizon. However a formulation for an infinite time horizon is easily obtained by time scaling (Proposition 2.4). Then the result formulated in our notation is

$$\sup_{0 \leq t \leq T-h} \frac{C L_{t+h} - C L_t}{\sqrt{2h C \log n}} \overset{a.s.}{\to} 1,$$
3. Estimation method

where $C$ is the variance of the process. Because we are in a multivariate setting, we prove directly that $B_n$ contains the diffusion increments asymptotically.

**Proposition 3.6**

Let $\{C_{X_{h,k}} \mid k = 1, \ldots, n\}$ be the increments of a Wiener process with covariance matrix $C$ and drift $\gamma$ and let $\alpha \in (1/2, 1]$ in $[3.7]$. Define

$$M_1^{(n)} := \{\exists k \in \{1, \ldots, n\}, s.t. C_{X_{h,k}} \notin B_n\},$$

then, as $n \to \infty$,

$$P(M_1^{(n)}) \to 0.$$

**Proof.** We use the increments of the centered version of the Wiener process $C' L_t$ (Definition 2.2(iii)) and have

$$P(C_{X_{h,1}} \notin B_n) = P(C' X_{h,1} + h \gamma \notin B_n) \leq P(C' X_{h,1} + h \gamma \notin B_n(C', \beta'))$$

$$= P(C' X_{h,1} + h \gamma \notin \{x : x^T C^{-1} x \leq \beta' b_n^2\})$$

$$= P(C' X_{h,1} \notin \{x : (x - h \gamma)^T C^{-1}(x - h \gamma) \leq \beta' b_n^2\}).$$

Because $b_n^2 = 2h \log n$, there exists a $\tilde{\beta}' \in (1, \beta')$ such that an upper bound is given by

$$P(C' X_{h,1} \notin \{x : x^T C^{-1} x \leq \tilde{\beta}' b_n^2\}) = P(C' X_{h,1}^T C^{-1} C' X_{h,1} > \tilde{\beta}' b_n^2)$$

$$= P(C' X_{h,1}^T C^{-1} C' X_{h,1} > \tilde{\beta}' 2 \log n),$$

where we have use the selfsimilarity property given in Proposition 2.4. The term

$$C' X_{h,1}^T C^{-1} C' X_{h,1}$$

is $\chi_d^2$-distributed. Thus, the probability form above is equal to

$$\int_{\tilde{\beta}' 2 \log n}^{\infty} \frac{t^{d/2-1} e^{-t/2}}{2^{d/2} \Gamma(d/2)} dt$$

and there exists a $\xi \in (1/\tilde{\beta}', 1)$ such that an upper bound is given by

$$\int_{\tilde{\beta}' 2 \log n}^{\infty} \exp \left(-\frac{\xi t}{2}\right) dt = \frac{2}{\xi} \exp \left(-\xi \tilde{\beta}' \log n\right) = \frac{2}{\xi} n^{-\xi \tilde{\beta}'}.$$

By the Binomial distribution and Bernoulli’s inequality we have

$$P(M_1^{(n)}) = 1 - (1 - P(C X_{h,1} \notin B_n))^n \leq n P(C X_{h,1} \notin B_n)$$

$$\leq \frac{2}{\xi} n^{1 - \xi \tilde{\beta}'} \to 0.$$

The strategy is to define estimators that use the critical set to separate jumps from diffusion. We show that these estimators behave asymptotically similarly to the pre-estimators.
We obtain different results for finite and infinite Lévy processes. So we consider both cases separately. Tests to find out whether a Lévy process has finite or infinite activity are presented in Aït-Sahalia and Jacod (2010).

### 3.1.1. Setting in the finite activity case

We start with the case of finite activity and formulate an additional assumption.

**Assumption 3.7**

For all $i \in \{1, \ldots, d\}$, the density $f_i$ of the Lévy measure $\tilde{\nu}_i$ of the univariate component Lévy process $\{L_{t,i}, t \geq 0\}$ is bounded, i.e. there exists $\kappa > 0$ such that

$$f_i(x) < \kappa, \quad x \in \mathbb{R}.$$ 

To estimate the drift, we consider the simplified representation of the characteristic function given in Remark 2.17,

$$\phi_{L_i}(u) = \exp(t\psi_{L_i}(u))$$

where

$$\psi_{L_i}(u) = -\frac{1}{2} u^T C u + iu^T \gamma[0] + \int_{\mathbb{R}^d} (\exp(iu^T x) - 1) \nu(dx).$$

So we consider the drift term $\gamma[0]$ and not the term $\gamma = \gamma[1]$. The pre-estimators for $C$ and $\gamma[0]$ are based on the increments of the diffusion part $C L_t$ of the Lévy process that is obtained by Proposition 2.26. So $C L_t$ is a nonstandard Wiener process with covariance $C$ and drift $\gamma[0]$. Furthermore we estimate the following function of the Lévy measure $v$:

$$v : \mathbb{R}^d \to \mathbb{R}^+, \quad z \mapsto \nu(z), \quad \text{with } \nu(z) := \nu((-\infty, z]),$$

where we have used the following notation

$$(-\infty, z] := (-\infty, z^{(1)}) \times \ldots \times (-\infty, z^{(d)}].$$

The pre-estimator of this function is based on the real jumps.

**Definition 3.8** (Pre-estimators)

The pre-estimators are defined as follows

$$\hat{C}_{pre} := \frac{1}{T} \sum_{k=1}^n C X_{h,k} C X_{h,k}^T,$$

$$\hat{\gamma}[0]_{pre} := \frac{1}{T} \sum_{k=1}^n C X_{h,k},$$

$$\hat{\nu}_{pre}(z) := \frac{1}{T} \sum_{0 < t \leq T} 1(\Delta L_t \leq z, \Delta L_t \neq 0), \quad z \in \mathbb{R}^d.$$
3. Estimation method

Theorem 3.16, Theorem 3.17 and Theorem 3.19 show the asymptotic properties of these pre-estimators.

Next, we define estimators that do not need such unrealistic assumptions. The critical region is used, instead.

**Definition 3.9** (Estimators)

The estimators are defined as follows

\[
\hat{C} := \frac{1}{T} \sum_{k=1}^{n} X_{h,k}X_{h,k}^T \mathbb{1}(X_{h,k} \in B_n),
\]

\[
\hat{\gamma}[0] := \frac{1}{T} \sum_{k=1}^{n} X_{h,k} \mathbb{1}(X_{h,k} \in B_n),
\]

\[
\hat{\nu}(z) := \frac{1}{T} \sum_{k=1}^{n} \mathbb{1}(X_{h,k} \notin B_n, X_{h,k} \leq z), \quad z \in \mathbb{R}^d.
\]

In Theorem 3.22, 3.23 and 3.25 the asymptotic properties of the estimators are given. The properties coincide with those of the pre-estimators.

3.1.2. Setting in the infinite activity case

We turn to the infinite activity case. For the covariance matrix \(C\) we use the same estimator as in the case of finite activity,

**Definition 3.10** (Estimator for the covariance matrix)

The estimator is defined as follows

\[
\hat{C} := \frac{1}{T} \sum_{k=1}^{n} X_{h,k}X_{h,k}^T \mathbb{1}(X_{h,k} \in B_n)
\]

While \(\hat{C}\) is asymptotically normally distributed in the case of finite activity, in the case of infinite activity, it is only consistent. This is shown in Theorem 3.27.

Since the Lévy measure is infinite, we estimate a specific finite transformation, instead.

**Definition 3.11** (Transformation of the Lévy measure)

Let \(A \in \mathbb{R}^{d \times d}, A \neq 0, z \in \mathbb{R}^d\) and \(l \in \mathbb{N}\). We assume that \(\mathbb{E}_1 \left[ L_{1,l}^{2l} \right] < \infty, \text{ for all } i \in \{1, \ldots, d\}\). Then, we define

\[
\mu(z) = \mu_{A,l} := \int_{\{x \leq z\}} (x^T A x)^l \nu(dx).
\]
3.1. Estimation setting

From Definition 2.13 follows directly, that $\mu(z) < \infty$ for all $z \in \mathbb{R}^d$. We define a
pre-estimator for $\mu$ based on the assumption that the exact jumps are known.

**Definition 3.12** (Pre-estimator for the transformation of the Lévy measure)
The pre-estimator is defined as follows

$$\hat{\mu}_{\text{pre}}(z) = \hat{\mu}_{A,l,\text{pre}}(z) := \frac{1}{T} \sum_{0 \leq t \leq T} (\Delta L_t^T A \Delta L_t)^l \mathbf{1}(\Delta L_t \leq z) \quad z \in \mathbb{R}^d.$$ 

In Theorem 3.21 the asymptotic properties of $\hat{\mu}_{\text{pre}}$ are given. Afterwards we
define an estimator that is based on the critical region and behaves similar to
the pre-estimator.

**Definition 3.13** (Estimator for the transformation of the Lévy measure)
The estimator for the quantity $\mu_{A,l}(z)$ is defined as follows

$$\hat{\mu}(z) = \hat{\mu}_{A,l}(z) := \frac{1}{T} \sum_{k=1}^{n} (X_{h,k}^T A X_{h,k})^l \mathbf{1}(X_{h,k} \notin B_n, X_{h,k} \leq z), \quad z \in \mathbb{R}^d.$$ 

The asymptotic properties of $\hat{\mu}$ are shown in Theorem 3.29. In general we only
obtain consistency. To get better rates we need the following additional assumption.

**Assumption 3.14**
For all $i \in \{1, \ldots, d\}$, we have for the density $f_i$ of the Lévy measure $\tilde{\nu}_i$ of the
univariate component Lévy process $\{L_{t,i}, t \geq 0\}$

$$f_i(x) = \mathcal{O}(x^{-3}), \quad x \to 0.$$ 

Since $\int_{|x_i| \leq 1} x_i^2 f_i(x_i) dx_i < \infty$, the assumption is e.g. satisfied, if there exists an
$\varepsilon > 0$ such that $f_i$ is monotone in both $(0, \varepsilon)$ and $(-\varepsilon, 0)$.

Then the obtained rates of convergence depend on $\alpha$, the exponentiation $l$ and
additional moment conditions.

Of course, we also can estimate $\mu$ in the case of finite activity, the asymptotic
properties are shown in Theorem 3.26.

The estimation of the drift is completely different to the finite activity case,
where a separation of the jump and drift term is existent. Because of the compen-
sation of the infinitely many small jumps, we cannot use the critical region. Instead we assume

$$\mathbb{E}[|L_{t,i}|] < \infty, \quad \forall i \in \{1, \ldots, d\}$$

and consider the following version of the characteristic function (see Remark 2.15)

$$\phi_{L_i}(u) = \exp(t \psi_{L_i}(u)),$$
3. Estimation method

where
\[
\psi_{L_1}(u) = -\frac{1}{2} u^T C u + i u^T \gamma[\infty] + \int_{\mathbb{R}^d} (\exp(iu^T x) - 1 - iu^T x) \nu(dx).
\]

We estimate the drift \(\gamma[\infty]\) that can be interpreted as an overall trend term, since the jumps are fully compensated.

**Definition 3.15 (Estimator for the drift)**
The estimator is defined as follows
\[
\hat{\gamma}[\infty] := \frac{1}{T} \sum_{k=1}^{n} X_{h,k}.
\]

In Theorem 3.30 it is shown that this estimator is strong consistent and asymptotically normally distributed.

3.2. High frequency data of a Wiener process

Let \(\{C_L_t, t \geq 0\}\) be a multivariate nonstandard Wiener process with covariance matrix \(C \in \mathbb{R}^{d \times d}\) and drift \(\gamma \in \mathbb{R}^d\). The corresponding increments are denoted by \(C_{X_{h,1}}, \ldots, C_{X_{h,n}}\). Recall, the estimators are defined as follows
\[
\hat{C}_{pre} := \frac{1}{T} \sum_{k=1}^{n} C_{X_{h,k}} C_{X_{h,k}}^T \quad \text{and} \quad \hat{\gamma}_{pre} := \frac{1}{T} \sum_{k=1}^{n} C_{X_{h,k}}.
\]

In this section we show that these estimators are asymptotically normally distributed.

**Theorem 3.16**

Let \(1/2 < \alpha \leq 1\) in (3.1). Then, as \(n \to \infty\),
\[
\sqrt{n} \left( \hat{C}_{pre} - C \right) \xrightarrow{d} Z,
\]
where
\[
Z = \begin{pmatrix}
Z_{1,1} & \ldots & Z_{1,d} \\
\vdots & \ddots & \vdots \\
Z_{d,1} & \ldots & Z_{d,d}
\end{pmatrix}
\]
is a normally distributed random vector with values in \(\mathbb{R}^{d^2}\). Expectation and covariance are given by
\[
E[Z_{i,j}] = 0 \quad \text{and} \quad \text{Cov}[Z_{i,j}, Z_{k,m}] = C_{i,k} C_{j,m} + C_{i,m} C_{j,k}.
\]

**Proof.** First, we show that the drift vector \(\gamma\) has no effect. We use the centered version of the Wiener process
\[
C^\gamma_L_t = C_L_t - \gamma t
\]
3.2. High frequency data of a Wiener process

and the corresponding centered increments \( \{ \gamma^\prime X_{h,k}, \, k = 1, \ldots, n \} \), then we have

\[
\sqrt{n} \left( \frac{1}{T} \sum_{k=1}^{n} C X_{h,k} C' X_{h,k}^T - \frac{1}{T} \sum_{k=1}^{n} C' X_{h,k} C^T X_{h,k} \right) = \sqrt{n} \frac{1}{T} \sum_{k=1}^{n} \left( h \gamma^\prime X_{h,k} \gamma + h \gamma X_{h,k}^T \gamma^T + h^2 \gamma \gamma^T \right) \overset{L}{\rightarrow} 0,
\]

since

\[ E[\gamma^\prime X_{h,k}] = 0, \, \forall k = 1, \ldots, n \quad \text{and} \quad \sqrt{n} \cdot n \cdot h^2 / T = n^{1/2 - \alpha} \rightarrow 0, \, n \rightarrow \infty. \]

By the selfsimilarity property of the Wiener process (Proposition 2.4) we have

\[
C^\prime X_{h,k} \overset{d}{=} \sqrt{h} C^\prime X_{1,k}, \quad \forall k = 1, \ldots, n.
\]

That means

\[
\sqrt{n} \left( \frac{1}{T} \sum_{k=1}^{n} C^\prime X_{h,k} C^\prime X_{h,k}^T - C \right) \overset{d}{=} \sqrt{n} \left( \frac{1}{n} \sum_{k=1}^{n} C^\prime X_{1,k} C^\prime X_{1,k}^T - C \right) \overset{d}{\rightarrow} Z,
\]

where the convergence directly follows from the Central Limit Theorem. The covariance can be computed by Isserlis’ Theorem (see Isserlis, 1918),

\[
\text{Cov} [Z_{i,j}, Z_{k,m}] = E \left[ C^\prime X_{1,1,j} C^\prime X_{1,1,k} C^\prime X_{1,1,m} C^\prime X_{1,1,i} \right] - C_{i,j} C_{k,m} = C_{i,k} C_{j,m} + C_{i,m} C_{j,k}.
\]

The trick of the proof is to use the selfsimilarity property. Then, we obtain a statistical problem which has sample variables that are independent from the sample size \( n \). As a result we obtain a rate \( \sqrt{n} \).

**Theorem 3.17**

Let \( 1/2 < \alpha < 1 \) in (3.1). Then, as \( n \rightarrow \infty \),

\[
\sqrt{T} \left( \hat{\gamma}_{pre} - \gamma \right) \overset{d}{\rightarrow} Z \sim N(0, C).
\]

**Proof.** We set w.l.o.g. \( T \in \mathbb{N} \), otherwise we consider \( \lfloor T \rfloor \) and show that \( \lfloor T \rfloor - T \) has no effect. Then, we have

\[
\frac{1}{T} \sum_{k=1}^{n} C X_{h,k} - \gamma = C L_T - \gamma = \frac{1}{T} \sum_{k=1}^{T} C X_{1,k} - \gamma.
\]

The random vectors \( \{ C X_{1,k}, \, k = 1, \ldots, T \} \) are independent and \( N(\gamma, C) \)-distributed. So the Central Limit Theorem can be used.

The estimator for the drift is equal to the value of the process at the end of the observation period minus the value of the process at the beginning (= 0) divided by the time horizon \( T \). As a result we obtain the rate \( \sqrt{T} \) of convergence.
3. Estimation method

3.3. A continuously observed Lévy process

In this section we assume that we can observe the Lévy process continuously, so we know the exact jump times and sizes
\[ \{\Delta L_t, \ t \in [0, T]\}. \]
To deal with the real jumps we introduce the following definition.

Definition 3.18 (Number of real jumps)
Let \(0 < t_1 \leq t_2 < \infty\).

(i) Let \(B \in \mathcal{B}(\mathbb{R}^d)\) and \(z_1, z_2 \in \mathbb{R}^d\). The number of real jumps is defined by
\[
N_{(t_1, t_2]}(B) := \sum_{t_1 < t \leq t_2} 1(\Delta L_t \in B, \Delta L_t \neq 0).
\]
Furthermore we denote
\[
N_{(t_1, t_2]}(z_1, z_2) := N_{(t_1, t_2]}(\{x \in \mathbb{R}^d : z_1 \leq x \leq z_2\})
\]
\[
N_{(t_1, t_2]}(z_1) := N_{(t_1, t_2]}(-\infty, z_1),
\]
\[
N_{t_1}(\cdot) := N_{(-\infty, t_1]}(\cdot).
\]

(ii) Let \(i \in \{1, \ldots, d\}\) and \(B \in \mathcal{B}(\mathbb{R})\). The number of real jumps of the component process is defined by
\[
N_{(t_1, t_2]}(z_1) := \sum_{t_1 < t \leq t_2} 1(\Delta L_{t,i} \in B, \Delta L_{t,i} \neq 0).
\]
For \(z_1, z_2 \in \mathbb{R}\), the quantities \(N_{(t_1, t_2]}(z_1, z_2)\), \(N_{(t_1, t_2]}(z_1)\) and \(N_{t_1}(\cdot)\) are defined analogously to (i).

By Definition 2.22 and Proposition 2.23 we have
\[
N_{(t_1, t_2]}(B) \sim \text{Poi}(t_2 - t_1, \nu(B)),
\]
where we use the convention \(N_{(t_1, t_2]}(B) = \infty\) a.s. if \(N_{(t_1, t_2]}(B) \sim \text{Poi}(\infty)\). Furthermore we obtain that, for \(B_1, B_2\) disjoint, the random variables \(N_{(t_1, t_2]}(B_1)\) and \(N_{(t_1, t_2]}(B_2)\) are independent.

The quantity \(N_{T,i}\) refers to the i-th component process. Please note, \(N_{T,i}\) is different from the i-th component of \(\tilde{N}_T\), just as \(\tilde{\nu}_i\) is different from \(\nu_i\) (cf. Corollary 2.21).

We start with the case of finite activity. Recall, the estimator for the function \(z \mapsto \nu(z)\) is defined as follows
\[
\hat{\nu}_{\text{pre}}(z) := \frac{1}{T} N_T(z).
\]
So we count the jumps that are component-by-component less than or equal to \(z\).
In the next theorem we show that this estimator is asymptotically normally distributed for a fixed \( z \in \mathbb{R}^d \) (or even for several \( z_1, \ldots, z_m \in \mathbb{R}^d \)).

Furthermore we consider the quantity
\[
\sqrt{T} \left( \hat{\nu}_{\text{pre}}(z) - \nu(z) \right)
\]
as a random field in \( z \). Since function spaces from \( \mathbb{R}^d \) to \( \mathbb{R} \) are very uncommon in literature, we consider a transformation, instead. Therefore we use a function \( \vec{g} : [0, 1]^d \to \mathbb{R}^d \), where \( \vec{g}(y) = \left( g \left( y^{(1)} \right), \ldots, g \left( y^{(d)} \right) \right)^T \)

with a bijective and strictly monotone function \( g \) that maps from \([0, 1]\) to \( \mathbb{R} \).

Then, we consider the following transformed quantity
\[
\sqrt{T} \left( \hat{\nu}_{\text{pre}} \circ \vec{g} - \nu \circ \vec{g} \right),
\]
which is an element of the space \( D[0, 1]^d \), because the conditions of Definition 2.62 are obviously satisfied. We show that this random field converges in distribution (weak convergence of the distribution) to a Gaussian process.

**Theorem 3.19**

Let \( \nu(\mathbb{R}^d) < \infty \).

(i) Let \( m \in \mathbb{N} \) and \( z_1, \ldots, z_m \in \mathbb{R}^d \), Then, as \( T \to \infty \),
\[
\sqrt{T} \left( \hat{\nu}_{\text{pre}}(z_1) - \nu(z_1), \ldots, \hat{\nu}_{\text{pre}}(z_m) - \nu(z_m) \right)^T \xrightarrow{d} Z \sim N(0, \Sigma),
\]
where
\[
\Sigma_{i,j} := \nu \left( (z_i^{(1)} \wedge z_j^{(1)}), \ldots, z_i^{(d)} \wedge z_j^{(d)} \right)^T, \quad i, j \in \{1, \ldots, m\}.
\]

(ii) Define
\[
\vec{g} : [0, 1]^d \to \mathbb{R}^d, \quad \vec{g}(y) = \left( g \left( y^{(1)} \right), \ldots, g \left( y^{(d)} \right) \right)^T,
\]
where \( g : [0, 1] \to \mathbb{R} \) is a strictly increasing and bijective function. Then, as \( T \to \infty \),
\[
\sqrt{T} \left( \hat{\nu}_{\text{pre}} \circ \vec{g} - \nu \circ \vec{g} \right) \xrightarrow{D} G,
\]
where \( G \) is a Gaussian process over \([0, 1]^d\) with
\[
\mathbb{E} \left[ G(y) \right] = 0,
\]
\[
\text{Cov} \left[ G(y_1), G(y_2) \right] = \nu \left( \left( g \left( y_1^{(1)} \wedge y_2^{(1)} \right), \ldots, g \left( y_1^{(d)} \wedge y_2^{(d)} \right) \right)^T \right),
\]
for \( y, y_1, y_2 \in [0, 1]^d \).
3. Estimation method

Proof. Ad (i), we have for $z \in \mathbb{R}^d$
\[
\hat{v}_{pre}(z) = \frac{1}{T} N_T(z) = \frac{1}{T} \sum_{k=1}^{T} N_{k-1,k}(z).
\]
The random variables $N_{k-1,k}(z)$ are i.i.d. and $\text{Poi}(\nu(z))$-distributed. So the convergence follows form the Central Limit theorem. To calculate the covariance we define
\[
A_i := \{ x \in \mathbb{R}^d : x^{(1)} \leq z_i^{(1)}, \ldots, x^{(d)} \leq z_i^{(d)} \}
\]
\[
A_j := \{ x \in \mathbb{R}^d : x^{(1)} \leq z_j^{(1)}, \ldots, x^{(d)} \leq z_j^{(d)} \}
\]
\[
A_{i \wedge j} := A_i \cap A_j = \{ x \in \mathbb{R}^d : x^{(1)} \leq z_i^{(1)} \wedge z_j^{(1)}, \ldots, x^{(d)} \leq z_i^{(d)} \wedge z_j^{(d)} \}.
\]
Then we have
\[
\Sigma_{i,j} = \text{Cov} [N_1(z_i), N_1(z_j)] = \text{Cov} [N_1(A_i), N_1(A_j)]
\]
\[
= \mathbb{E} [N_1(A_i) N_1(A_j)] - \mathbb{E} [N_1(A_i)] \mathbb{E} [N_1(A_j)]
\]
\[
= \mathbb{E} [N_1^2(A_{i \wedge j})] + \mathbb{E} [N_1(A_{i \wedge j}) N_1((A_i \cup A_j) \setminus A_{i \wedge j})]
\]
\[
- \mathbb{E} [N_1(A_{i \wedge j})]^2 - \mathbb{E} [N_1(A_{i \wedge j})] \mathbb{E} [N_1((A_i \cup A_j) \setminus A_{i \wedge j})]
\]
\[
= \text{Var} [N_1(A_{i \wedge j})]
\]
\[
= \text{Var} \left[ N_1 \left( (z_i^{(1)} \wedge z_j^{(1)}), \ldots, z_i^{(d)} \wedge z_j^{(d)} \right)^T \right]
\]
\[
= \nu \left( (z_i^{(1)} \wedge z_j^{(1)}), \ldots, z_i^{(d)} \wedge z_j^{(d)} \right)^T,
\]
for $i, j \in \{1, \ldots, d\}$. We have used that $N_1(A)$ and $N_1(B)$ are independent for disjoint sets $A$ and $B$.

Ad (ii), we use Proposition 2.65. The first condition of Proposition 2.65 is verified by (i). We turn to the second condition. For all $\varepsilon > 0$ we have
\[
\lim_{y^{(p)} \to 1} \mathbb{P} \left( \left| G(y^{(1)}, \ldots, y^{(p)}, \ldots, y^{(d)}) - G(y^{(1)}, \ldots, 1, \ldots, y^{(d)}) \right| > \varepsilon \right)
\]
\[
\leq \lim_{y^{(p)} \to 1} \frac{1}{\varepsilon^2} \mathbb{E} \left[ \left( G(y^{(1)}, \ldots, y^{(p)}, \ldots, y^{(d)}) - G(y^{(1)}, \ldots, 1, \ldots, y^{(d)}) \right)^2 \right]
\]
\[
= \lim_{y^{(p)} \to 1} \frac{1}{\varepsilon^2} \left( \mathbb{E} \left[ G^2(y^{(1)}, \ldots, y^{(p)}, \ldots, y^{(d)}) \right] + \mathbb{E} \left[ G^2(y^{(1)}, \ldots, 1, \ldots, y^{(d)}) \right] - 2 \mathbb{E} \left[ G(y^{(1)}, \ldots, y^{(p)}, \ldots, y^{(d)}) G(y^{(1)}, \ldots, 1, \ldots, y^{(d)}) \right] \right)
\]
\[
= \lim_{y^{(p)} \to 1} \frac{1}{\varepsilon^2} \left( \nu \left( \left( g \left( y^{(1)} \right), \ldots, g \left( y^{(p)} \right), \ldots, g \left( y^{(d)} \right) \right)^T \right)
\]
\[
- \nu \left( \left( g \left( y^{(1)} \right), \ldots, g \left( y^{(p)} \right), \ldots, g \left( y^{(d)} \right) \right)^T \right) \right) = 0.
\]
We turn to condition (iii) of Proposition 2.65. For disjoint sets $A$ and $B$ we
3.3. A continuously observed Lévy process

have
\[ \mathbb{E} \left[ T \left( \tilde{\nu}_{\text{pre}}(\tilde{g}(A)) - \nu(\tilde{g}(A)) \right)^2 \cdot T \left( \tilde{\nu}_{\text{pre}}(\tilde{g}(B)) - \nu(\tilde{g}(B)) \right)^2 \right] \\
= \mathbb{E} \left[ T \left( \tilde{\nu}_{\text{pre}}(\tilde{g}(A)) - \nu(\tilde{g}(A)) \right)^2 \right] \cdot \mathbb{E} \left[ T \left( \tilde{\nu}_{\text{pre}}(\tilde{g}(B)) - \nu(\tilde{g}(B)) \right)^2 \right] \\
= \nu(\tilde{g}(A))\nu(\tilde{g}(B)), \]
where we have used that \( N_T(A) \) and \( N_T(B) \) are independent and that \( \tilde{g} \) is bijective. In case \( d = 1 \) we can use directly Proposition 2.61 instead of Proposition 2.65.

Remark 3.20

(i) An explicit example for the function \( g : [0, 1] \to \mathbb{R} \) is given by
\[
g(y) := \log(2y)1_{\{y \in [0, \frac{1}{2}]\}} - \log(2(1 - y))1_{\{y \in (\frac{1}{2}, 1]\}},
\]
where we use the convention \( \log(0) = -\infty \).

(ii) If \( d = 1 \) and \( \nu(\cdot) \) is strictly increasing and \( \nu(\mathbb{R}) \neq 0 \), we can choose
\[
g(y) = \nu^{-1}(y\nu(\mathbb{R})).
\]
Then the obtained limit process is equal to \( \nu(\mathbb{R})W_t \), where \( W \) is a Wiener process on \([0, 1]\).

We turn to the general case where infinite activity is possible. Recall, for \( A \in \mathbb{R}^{d \times d}, A \neq 0 \) and \( l \in \mathbb{N} \), we consider the following transformation of the Lévy measure
\[
\mu(z) = \mu_{A,l} := \int_{\{x \leq z\}} (x^T A x)^l \nu(dx),
\]
where we assume \( \mathbb{E} \left[ L_{1,i}^{2l} \right] < \infty \), for all \( i \). The estimator is defined as follows
\[
\hat{\mu}_{\text{pre}}(z) = \hat{\mu}_{A,l,\text{pre}}(z) := \frac{1}{T} \sum_{0 \leq t \leq T} (\Delta L_t^T A \Delta L_t)^l 1(\Delta L_t \leq z).
\]

The next theorem shows the asymptotic properties of this estimator.

Theorem 3.21

Let \( l \in \mathbb{N} \) and \( \mathbb{E} \left[ L_{1,i}^{4l} \right] < \infty \), for all \( i \).

(i) Let \( m \in \mathbb{N} \) and \( z_1, \ldots, z_m \in \mathbb{R}^d \), Then, as \( T \to \infty \),
\[
\sqrt{T} \left( \hat{\mu}_{\text{pre}}(z_1) - \mu(z_1), \ldots, \hat{\mu}_{\text{pre}}(z_m) - \mu(z_m) \right)^T \xrightarrow{d} Z \sim N(0, \Sigma),
\]
where
\[
\Sigma_{i,j} := \int_{\{x^{(1)}_1 \leq z_1^{(1)} \wedge \ldots \wedge z_m^{(1)} \wedge \ldots \wedge x^{(d)}_1 \leq z_1^{(d)} \wedge \ldots \wedge x^{(d)}_m \}} (x^T A x)^{2l} \nu(dx)
\]

(ii) Define \( \tilde{g} \) as in Proposition 3.19. Then, as \( T \to \infty \),
\[
\sqrt{T} \left( \hat{\mu}_{\text{pre}} \circ \tilde{g} - \mu \circ \tilde{g} \right) \xrightarrow{d} G,
\]

47
3. Estimation method

where $G$ is a Gaussian process over $[0, 1]^d$ with

$$E[G(y)] = 0,$$

$$Cov[G(y_1), G(y_2)] = \int_{\{x^{(1)} \leq y^{(1)}_1 \wedge y^{(1)}_2, \ldots, x^{(d)} \leq y^{(d)}_1 \wedge y^{(d)}_2\}} (x^T Ax)^2 \nu(dx),$$

for $y, y_1, y_2 \in [0, 1]^d$.

Proof. Ad (i), we have for $z \in \mathbb{R}^d$

$$\hat{\mu}_{pre}(z) = \frac{1}{T} \sum_{0 \leq t \leq T} (\Delta L_t^T A \Delta L_t)^T \mathbf{1}(\Delta L_t \leq z) \sum_{k=1}^{T} \sum_{k-1 \leq t \leq k} (\Delta L_t^T A \Delta L_t)^T \mathbf{1}(\Delta L_t \leq z).$$

The random variables

$$\sum_{k-1 \leq t \leq k} (\Delta L_t^T A \Delta L_t)^T \mathbf{1}(\Delta L_t \leq z)$$

are i.i.d. The expectation is given by Proposition 2.32 and is equal to

$$\int_{x \leq z} (x^T Ax)^T \nu(dx) = \mu(z).$$

So the convergence follows from the Central Limit theorem. The covariance can be computed analogously to Proposition 3.19 by using the sets $A_i, A_j$ and $A_{i\wedge j}$. Furthermore Proposition 2.32 is used.

Ad (ii), the statement can be shown analogously to Proposition 3.19 by using Proposition 2.32 for the moments and the independence.

3.4. Estimation of a finite activity Lévy process

In this section we show the asymptotic properties of the estimators in the case of finite activity. The estimators have been defined in Section 3.1.1.

The idea of the proofs is to break up the estimators in such a manner that we obtain the pre-estimators and some remainder terms. Therefore we use the version of the Lévy-Itô decomposition that is given in Proposition 2.26 and obtain

$$L_t = C L_t + CP L_t \quad t \geq 0,$$

where $\{C L_t, t \geq t\}$ is a nonstandard Wiener process with covariance matrix $C$ and drift $\gamma[0]$. The process $\{CP L_t, t \geq t\}$ is a compound Poisson process with intensity $\nu(\mathbb{R}^d)$ and jump size distribution $\nu(\cdot)/\nu(\infty)$. By

$$\{C X_{h,k} = (C X_{h,k,1}, \ldots, C X_{h,k,d})^T, \quad k = 1, \ldots, n\},$$

$$\{CP X_{h,k} = (CP X_{h,k,1}, \ldots, CP X_{h,k,d})^T, \quad k = 1, \ldots, n\}.$$
we denote the corresponding increments, so we have
\[ X_{h,k} = C X_{h,k} + CP X_{h,k}, \quad k = 1, \ldots, n. \]

We start with the estimation of \( C \). Recall, the estimator and its pre-estimator are given by
\[
\hat{C}_{\text{pre}} := \frac{1}{T} \sum_{k=1}^{n} C X_{h,k} C X_{h,k}^T \quad \text{and} \quad \hat{C} := \frac{1}{T} \sum_{k=1}^{n} X_{h,k} X_{h,k}^T \mathbb{1}(X_{h,k} \in B_n).
\]

In Theorem 3.16 we have seen that the pre-estimator is asymptotically normally distributed. Theorem 3.22 shows that the difference \( \hat{C}_{\text{pre}} - \hat{C} \) converges to zero with rate \( \sqrt{n} \). That means that \( \hat{C} \) is also asymptotically normally distributed with the same covariance.

**Theorem 3.22**

Let \( 1/2 < \alpha \leq 1 \) in (3.1). Then, as \( n \to \infty \),
\[
\sqrt{n} \left( \hat{C} - \hat{C}_{\text{pre}} \right) \xrightarrow{P} 0 \quad \text{and hence} \quad \sqrt{n} \left( \hat{C} - C \right) \xrightarrow{d} Z,
\]
where
\[
Z = \begin{pmatrix}
Z_{1,1} & \cdots & Z_{1,d} \\
\vdots & \ddots & \vdots \\
Z_{d,1} & \cdots & Z_{d,d}
\end{pmatrix}
\]
is a normally distributed random vector with values in \( \mathbb{R}^{d^2} \). Expectation and covariance are given by
\[
E[Z_{i,j}] = 0 \quad \text{and} \quad \text{Cov}[Z_{i,j}, Z_{k,m}] = C_{i,k}C_{j,m} + C_{i,m}C_{j,k}.
\]

**Proof.** By We have
\[
\hat{C} - \hat{C}_{\text{pre}} = \frac{1}{T} \sum_{k=1}^{n} C X_{h,k} C X_{h,k}^T \mathbb{1}(X_{h,k} \in B_n) - \frac{1}{T} \sum_{k=1}^{n} C X_{h,k} C X_{h,k}^T \\
= \frac{1}{T} \sum_{k=1}^{n} C X_{h,k} C X_{h,k}^T \mathbb{1}(X_{h,k} \in B_n) - \frac{1}{T} \sum_{k=1}^{n} C X_{h,k} C X_{h,k}^T \\
+ \frac{1}{T} \sum_{k=1}^{n} C X_{h,k} CP X_{h,k}^T \mathbb{1}(X_{h,k} \in B_n) + \frac{1}{T} \sum_{k=1}^{n} CP X_{h,k} C X_{h,k}^T \mathbb{1}(X_{h,k} \in B_n) \\
+ \frac{1}{T} \sum_{k=1}^{n} CP X_{h,k} CP X_{h,k}^T \mathbb{1}(X_{h,k} \in B_n) \\
= \frac{1}{T} \sum_{k=1}^{n} C X_{h,k} CP X_{h,k}^T \mathbb{1}(X_{h,k} \in B_n) + \frac{1}{T} \sum_{k=1}^{n} CP X_{h,k} C X_{h,k}^T \mathbb{1}(X_{h,k} \in B_n)
\]
(3.2)
\[
+ \frac{1}{T} \sum_{k=1}^{n} CP X_{h,k} CP X_{h,k}^T \mathbb{1}(X_{h,k} \in B_n) \quad (3.3)
\]
3. Estimation method

\[ + \frac{1}{T} \sum_{k=1}^{n} C P X_{h,k} C P X_{h,k}^T \mathbb{1}(X_{h,k} \in B_n) \quad (3.4) \]

\[ - \frac{1}{T} \sum_{k=1}^{n} C X_{h,k} C X_{h,k}^T \mathbb{1}(X_{h,k} \notin B_n). \quad (3.5) \]

We have to proof that the terms tend to zero in probability with rate \( \sqrt{n} \).

Ad (3.2) and (3.3), Lemma B.2 shows that

\[ \sqrt{n} \frac{1}{T} \sum_{k=1}^{n} \left| C X_{h,k,i_1} \right| \left| C P X_{h,k,i_2} \right| \mathbb{1}(X_{h,k} \in B_n) \overset{P}{\to} 0, \]

for all \( i_1, i_2 \in \{1, \ldots, d\} \). So the terms tend to zero in probability with rate \( \sqrt{n} \) in every component.

Ad (3.4), we only have to consider the diagonal elements of the square matrix, since for all \( i_1, i_2 \in \{1, \ldots, d\} \)

\[ \frac{1}{T} \sum_{k=1}^{n} \left| C P X_{h,k,i_1} \right| \left| C P X_{h,k,i_2} \right| \mathbb{1}(X_{h,k} \in B_n) \]

\[ \leq \frac{1}{T} \sum_{k=1}^{n} (C P X_{h,k,i_1}^2 + C P X_{h,k,i_2}^2) \mathbb{1}(X_{h,k} \in B_n) \]

The convergence is shown in Lemma B.3(i).

Ad (3.5), only the diagonal elements are considered, too. This is shown in Lemma B.3(i).

Next, we consider the estimation of \( \gamma[0] \). In Section 3.1.1 we have defined the estimator and its pre-estimator as follows

\( \hat{\gamma}[0]_{pre} := \frac{1}{T} \sum_{k=1}^{n} C X_{h,k} \) and \( \hat{\gamma}[0] := \frac{1}{T} \sum_{k=1}^{n} X_{h,k} \mathbb{1}(X_{h,k} \in B_n) \).

In Theorem 3.22 we have proved that \( \hat{\gamma}[0]_{pre} \) is asymptotically normally distributed. In Theorem 3.23 we show that the difference \( \hat{\gamma}[0]_{pre} - \hat{\gamma}[0] \) converges in probability to zero with rate \( \sqrt{T} \). Thus, \( \hat{\gamma}[0] \) has the same asymptotic properties as \( \hat{\gamma}[0]_{pre} \).

\[ \text{Theorem 3.23} \]

Let \( 1/2 < \alpha < 1 \) in (3.1). Then, as \( n \to \infty \),

\[ \sqrt{T} (\hat{\gamma}[0] - \hat{\gamma}[0]_{pre}) \overset{P}{\to} 0 \text{ and hence } \sqrt{T} (\hat{\gamma}[0] - \gamma[0]) \overset{d}{\to} Z \sim N(0, C). \]
3.4. Estimation of a finite activity Lévy process

Proof. We have

\[ \hat{\gamma}[0] - \hat{\gamma}[0]_{\text{pre}} = \frac{1}{T} \sum_{k=1}^{n} C X_{h,k} \mathbb{1}(X_{h,k} \in B_n) - \frac{1}{T} \sum_{k=1}^{n} C X_{h,k} \]

\[ = \frac{1}{T} \sum_{k=1}^{n} C P X_{h,k} \mathbb{1}(X_{h,k} \in B_n) \] (3.6)

\[ - \frac{1}{T} \sum_{k=1}^{n} C X_{h,k} \mathbb{1}(X_{h,k} \notin B_n). \] (3.7)

We have to prove that the terms tend to zero in probability with rate \( \sqrt{T} \). The term \((3.6)\) is considered component-by-component in Lemma B.4(ii) and \((3.7)\) in Lemma B.3(ii).

We turn to the estimation of the function \( z \mapsto \nu(z) \). In Section 3.1.1 we have defined the estimator and its pre-estimators as follows

\[ \hat{\nu}_{\text{pre}}(z) := \frac{1}{T} \sum_{0<t\leq T} \mathbb{1}(\Delta L_t \leq z, \Delta L_t \neq 0), \]

\[ \hat{\nu}(z) := \frac{1}{T} \sum_{k=1}^{n} \mathbb{1}(X_{h,k} \notin B_n, X_{h,k} \leq z), \quad z \in \mathbb{R}^d. \]

In Theorem 3.19 we have shown the asymptotic properties of \( \hat{\nu}_{\text{pre}} \). Next, we have to prove that the difference between both estimators converges in probability to zero with rate \( \sqrt{T} \). As in the theorems above we break up \( \hat{\nu} \) in such a manner that \( \hat{\nu}_{\text{pre}} \) and some remainder terms are obtained. The pre-estimator is based on the real jumps

\{\Delta L_t, t \in [0,T]\}.

These real jumps are contained in \( \{C P X_{h,k}, k = 1, \ldots, n\} \). However, these quantities do not coincide. The next proposition shows that each \( C P X_{h,k} \) contains asymptotically at most one jump. Thus, \( \{C P X_{h,k}, k = 1, \ldots, n\} \) and \( \{\Delta L_t, t \in [0,T]\} \) are very similar.

Proposition 3.24
Let \( \alpha \in (1/2, 1] \) in (3.1). Define

\[ M_2^{(n)} := \{ \exists k \in \{1, \ldots, n\} s.t. N_{(k-1)h,kh} > 1 \}, \]

then, as \( n \to \infty \)

\[ \mathbb{P} \left( M_2^{(n)} \right) \to 0. \]

Proof. By the Poisson distribution, the probability that at least two jumps
3. Estimation method

occur in a time interval with length $h$ is equal to

$$(1 - \exp(-\nu(R^d) h)(1 + \nu(R^d) h)).$$

Hence,

$$P \left( M_2^{(n)} \right) \leq n(1 - \exp(-\nu(R^d) h)(1 + \nu(R^d) h))$$

$$\leq n(1 - (1 - \nu(R^d)) h)(1 + \nu(R^d) h) = n h^2 \nu^2(R^d) \to 0.$$

\[ \square \]

**Theorem 3.25**

Let $1/2 < \alpha < 1$ in (3.1). Then, as $n \to \infty$,

$$\sup_{z \in \mathbb{R}^d} \sqrt{T} |\hat{\nu}(z) - \hat{\nu}_{pre}(z)| \to P 0.$$

Hence, for

$$\tilde{g} : [0,1]^d \to \mathbb{R}^d, \quad \tilde{g}(y) = \left(g \left(y^{(1)}\right), \ldots, g \left(y^{(d)}\right)\right)^T,$$

where $g : [0,1] \to \mathbb{R}$ is a strictly increasing and bijective function, we have, as $n \to \infty$,

$$\sqrt{T} \left( \hat{\nu} \circ \tilde{g} - \nu \circ \tilde{g} \right) \to G,$$

where $G$ is a Gaussian process over $[0,1]^d$ with

$$E[G(y)] = 0,$$

$$\text{Cov} [G(y_1), G(y_2)] = \nu \left( \left( g \left(y_1^{(1)} \land y_2^{(1)}\right), \ldots, g \left(y_1^{(d)} \land y_2^{(d)}\right)\right)^T \right),$$

for $y, y_1, y_2 \in [0,1]^d$.

**Proof.** By Theorem 3.19 and Proposition 2.50 we only have to prove

$$\sqrt{T} \sup_{z \in \mathbb{R}^d} |\hat{\nu}(z) - \hat{\nu}_{pre}(z)|$$

$$= \sqrt{T} \sup_{z \in \mathbb{R}^d} \left| \frac{1}{T} \sum_{k=1}^{n} \mathbb{I}(X_{h,k} \notin B_n, X_{h,k} \leq z) - \frac{1}{T} \sum_{0 \leq t \leq T} \mathbb{I}(\Delta L_t \leq z, \Delta L_t \neq 0) \right| \to P 0.$$

This term is considered in Lemma 3.3 (iii), where we use the decomposition of $X_{h,k}$ and Proposition 3.24. \[ \square \]

Next, we consider the specific transformation of the Lévy, measure which we have defined in Section 3.1.1 as follows

$$\mu(z) = \int_{\{x \leq z\}} (x^T A x)^l \nu(dx), \quad A \in \mathbb{R}^{d \times d}, A \neq 0, \ l \in \mathbb{N}, \ z \in \mathbb{R}^d.$$
3.4. Estimation of a finite activity Lévy process

We have introduced this quantity for the infinite activity case. However, we can also use it in the case of finite activity. The estimator and its pre-estimator are defined as follows

\[ \hat{\mu}_{\text{pre}}(z) = \frac{1}{T} \sum_{0 \leq t \leq T} (\Delta L_t^T A \Delta L_t)^l \mathbb{1}(\Delta L_t \leq z), \]

\[ \hat{\mu}(z) = \frac{1}{T} \sum_{k=1}^{n} (X^T_{h,k} A X_{h,k})^l \mathbb{1}(X_{h,k} \notin B_n, X_{h,k} \leq z). \]

The next theorem deals with the difference between both estimators. However, we only obtain the rate \( \sqrt{T} \) of convergence under additional conditions.

**Theorem 3.26**

(i) Let \( l \in \mathbb{N}, r \in (0, (1 - \alpha)/2) \) and let condition 1.) or 2.) from below be satisfied.

1.) Assume \( 1/2 < \alpha \leq 2/3 \) and \( \mathbb{E} \left[ L_{1,i}^{2(2\lor p)} \right] < \infty \) for all \( i \), where

\[ p := \left\lfloor \frac{l \alpha}{2 - r} + 1 \right\rfloor. \]

2.) Assume \( 2/3 < \alpha < 1 \) and \( \mathbb{E} \left[ L_{1,i}^4 \right] < \infty \) for all \( i \).

Then, as \( n \to \infty \),

\[ n^r \sup_{z \in \mathbb{R}^d} |\hat{\mu}(z) - \hat{\mu}_{\text{pre}}(z)| \to^P 0 \quad \text{and hence} \quad n^r \sup_{x \in \mathbb{R}^d} |\hat{\mu}(z) - \mu(z)| \to^P 0. \]

(ii) Let \( l \in \mathbb{N} \) and let condition 1.) or 2.) from below be satisfied.

1.) Assume \( 1/2 < \alpha \leq 2/3 \) and \( \mathbb{E} \left[ L_{1,i}^{2(2\lor p)} \right] < \infty \) for all \( i \), where

\[ p := \left\lfloor \frac{l \alpha}{2 - r} + 1 \right\rfloor. \]

2.) Assume \( 2/3 < \alpha < 1 \) and \( \mathbb{E} \left[ L_{1,i}^{4+2} \right] < \infty \) for all \( i \).

Then, as \( n \to \infty \)

\[ \sqrt{T} \sup_{z \in \mathbb{R}^d} |(\hat{\mu}(z) - \mu_{\text{pre}}(z))| \to^P 0. \]

Hence, for

\[ \bar{g} : [0, 1]^d \to \mathbb{R}^d, \quad \bar{g}(y) = \left( g\left(y^{(1)}\right), \ldots, g\left(y^{(d)}\right)\right)^T, \]

where \( g : [0, 1] \to \mathbb{R} \) is a strictly increasing and bijective function, we have, as \( n \to \infty \)

\[ \sqrt{T} \ (\hat{\mu} \circ \bar{g} - \mu \circ \bar{g}) \to^D G, \]
3. Estimation method

where $G$ is a Gaussian process over $[0, 1]^d$ with

$E[G(y)] = 0,$

$\text{Cov}[G(y_1), G(y_2)]$

$$= \int \{x^{(1)} \leq g(y_1^{(1)}), \ldots, x^{(d)} \leq g(y_d^{(1)} \wedge g(y_d^{(2)}))\} (x^T Ax)^2 \nu(dx),$$

for $y, y_1, y_2 \in [0, 1]^d$.

**Proof.** We use the given $r$ in (i), in (ii) we set $r = \frac{1-\alpha}{2}$. We use Theorem 3.21 and Proposition 2.50, so we only have to prove

$$n^r \sup_{z \in \mathbb{R}^d} \left| \hat{\mu}(z) - \hat{\mu}_{\text{pre}}(z) \right|$$

$$= n^r \sup_{z \in \mathbb{R}^d} \left| \frac{1}{T} \sum_{k=1}^{n} (X_{h,k}^T A X_{h,k})^l 1(X_{h,k} \notin B_n, X_{h,k} \leq z) \right|$$

$$- \frac{1}{T} \sum_{0 \leq t \leq T} (\Delta L_t^T A \Delta L_t)^l 1(\Delta L_t \leq z)$$

$$n^r \sup_{z \in \mathbb{R}^d} \left| \frac{1}{T} \sum_{k=1}^{n} \left( C P X_{h,k}^T A C P X_{h,k} + C P X_{h,k}^T A C X_{h,k} \right)^l 1(X_{h,k} \notin B_n, X_{h,k} \leq z) \right|$$

$$- \frac{1}{T} \sum_{0 \leq t \leq T} (\Delta L_t^T A \Delta L_t)^l 1(\Delta L_t \leq z) \xrightarrow{P} 0.$$

By Lemma A.1 (iii), we obtain an upper bound that is up to a positive constant equal to

$$n^r \sup_{z \in \mathbb{R}^d} \left| \frac{1}{T} \sum_{k=1}^{n} \left( C P X_{h,k}^T A C P X_{h,k} \right)^l 1(X_{h,k} \notin B_n, X_{h,k} \leq z) \right|$$

$$- \frac{1}{T} \sum_{0 \leq t \leq T} (\Delta L_t^T A \Delta L_t)^l 1(\Delta L_t \leq z)$$

$$+ n^r \frac{1}{T} \sum_{m=1}^{l} \sum_{k=1}^{n} \left( C P X_{h,k}^T A C X_{h,k} \right)^{l-m} \left| C X_{h,k}^T A C P X_{h,k} \right|^m 1(X_{h,k} \notin B_n) \quad (3.8)$$

$$+ n^r \frac{1}{T} \sum_{m=1}^{l} \sum_{k=1}^{n} \left( C P X_{h,k}^T A C X_{h,k} \right)^{l-m} \left| C X_{h,k}^T A C X_{h,k} \right|^m 1(X_{h,k} \notin B_n) \quad (3.9)$$

$$+ n^r \frac{1}{T} \sum_{m=1}^{l} \sum_{k=1}^{n} \left( C P X_{h,k}^T A C X_{h,k} \right)^{l-m} \left( C X_{h,k}^T A C X_{h,k} \right)^m 1(X_{h,k} \notin B_n). \quad (3.10)$$

$$+ n^r \frac{1}{T} \sum_{m=1}^{l} \sum_{k=1}^{n} \left( C P X_{h,k}^T A C X_{h,k} \right)^{l-m} \left( C X_{h,k}^T A C X_{h,k} \right)^m 1(X_{h,k} \notin B_n). \quad (3.11)$$
The term in (3.8) is considered in Lemma B.4(v) for condition 1.) and B.4(v) for condition 2.). There we use Proposition 3.24. For (3.9) and (3.10) we use the following upper bound

$$\max_{i_1, i_2 \in \{1, \ldots, d\}} |A_{i_1, i_2}| \cdot n^r \frac{1}{T} \sum_{m=1}^{l} \sum_{k=1}^{n} \left( \sum_{i_1, i_2 = 1}^{d} |C P X_{h, k, i_1}| |C P X_{h, k, i_2}| \right)^{l-m} \cdot 2 \left( \sum_{i_1, i_2 = 1}^{d} |C X_{h, k, i_1}| |C P X_{h, k, i_2}| \right)^{m} \mathbb{1}(X_{h, k} \notin B_n).$$

We apply Lemma A.1(ii) and A.1(i) and obtain an upper bound that is up to a positive constant equal to

$$\max_{i_1, i_2 \in \{1, \ldots, d\}} |A_{i_1, i_2}| \cdot n^r \frac{1}{T} \sum_{m=1}^{l} \sum_{k=1}^{n} \sum_{i_1, i_2, i_3 = 1}^{d} |C P X_{h, k, i_1}| |C X_{h, k, i_2}| \left( \sum_{i_1, i_2, i_3 = 1}^{d} |C X_{h, k, i_3}| \right)^{m} \mathbb{1}(X_{h, k} \notin B_n).$$

This term tends to zero by Lemma B.2. For (3.11) we analogously have

$$\max_{i_1, i_2 \in \{1, \ldots, d\}} |A_{i_1, i_2}| \cdot n^r \frac{1}{T} \sum_{m=1}^{l} \sum_{k=1}^{n} \sum_{i_1, i_2 = 1}^{d} |C X_{h, k, i_1}| \left( \sum_{i_1, i_2, i_3 = 1}^{d} |C X_{h, k, i_2}| \right)^{2l-2m} \mathbb{1}(X_{h, k} \notin B_n).$$

This term tends to zero by Lemma B.3(i) for $l = 1$, $m = 1$ and Lemma A.3(iv) for $l > 1$, $m = l$ and Lemma B.2 for all other summands.

### 3.5. Estimation of an infinite activity Lévy process

In this section we show the asymptotic properties of the estimators in the case of infinite activity. The estimators have been defined in Section 3.1.2.

The idea of the proofs is similar to the finite activity case. We break up the estimators in such a manner that we obtain the pre-estimators and some remainder terms. Therefore we use the general version of the Lévy-Itô decomposition that is given in Proposition 2.23 and obtain

$$L_t = C L_t + J L_t, \quad t \geq 0,$$

where $\{C L_t, t \geq 0\}$ is a nonstandard Wiener process with covariance matrix $C$ and drift $\gamma$. The process $\{J L_t, t \geq 0\}$ is a pure jump Lévy process with Lévy measure $\nu$. The corresponding increments are denoted by

$$\{C X_{h, k} = (C X_{h, k, 1}, \ldots, C X_{h, k, d})^T, k = 1, \ldots, n\},$$
$$\{ J X_{h, k} = (J X_{h, k, 1}, \ldots, J X_{h, k, d})^T, k = 1, \ldots, n\}.$$
3. Estimation method

and its pre-estimator as follows

$$\hat{C} := \frac{1}{T} \sum_{k=1}^{n} X_{h,k} X_{h,k}^T \mathbb{1}(X_{h,k} \in B_n),$$

$$\hat{C}_{\text{pre}} := \frac{1}{T} \sum_{k=1}^{n} C X_{h,k} C X_{h,k}^T.$$  

In Theorem 3.16 we have shown that $\hat{C}_{\text{pre}}$ is asymptotically normally distributed. Next, we prove that the difference between both estimators converges in probability to zero. However, we cannot show any rate of convergence, so we only obtain consistency for $\hat{C}$.

**Theorem 3.27**

Let $1/2 < \alpha \leq 1$ in (3.1). Then, as $n \to \infty$,

$$\hat{C} - \hat{C}_{\text{pre}} \overset{P}{\to} 0 \quad \text{and hence} \quad \hat{C} - C \overset{P}{\to} 0.$$  

**Proof.** We have

$$\frac{1}{T} \sum_{k=1}^{n} X_{h,k} X_{h,k}^T \mathbb{1}(X_{h,k} \in B_n) - \hat{C}_{\text{pre}}$$

$$= \frac{1}{T} \sum_{k=1}^{n} C X_{h,k} C X_{h,k}^T \mathbb{1}(X_{h,k} \notin B_n) \quad (3.12)$$

$$+ \frac{1}{T} \sum_{k=1}^{n} J X_{h,k} J X_{h,k}^T \mathbb{1}(X_{h,k} \in B_n) \quad (3.13)$$

$$+ \frac{1}{T} \sum_{k=1}^{n} C X_{h,k} J X_{h,k}^T \mathbb{1}(X_{h,k} \in B_n) \quad (3.14)$$

$$+ \frac{1}{T} \sum_{k=1}^{n} J X_{h,k} C X_{h,k}^T \mathbb{1}(X_{h,k} \in B_n). \quad (3.15)$$

Analogously to Theorem 3.22 we only consider the diagonal elements of the matrices in (3.12) and (3.13). In Lemma C.7 and Lemma C.8 it is shown that they converges to zero in probability. For the entries of the matrix in (3.14) and (3.15) by the Cauchy Schwartz inequality we have

$$\frac{1}{T} \sum_{k=1}^{n} C X_{h,k,i_1} J X_{h,k,i_2} \mathbb{1}(X_{h,k} \in B_n)$$

$$\leq \left(\frac{1}{T} \sum_{k=1}^{n} C X_{h,k,i_1}^2 \right)^{1/2} \left(\frac{1}{T} \sum_{k=1}^{n} J X_{h,k,i_2}^2 \right)^{1/2} \mathbb{1}(X_{h,k} \in B_n),$$

where $i_1, i_2 \in \{1, \ldots, d\}$. By Lemma A.3 the expectation of the first factor is uniform bounded and by Lemma C.8 the second factor converges to zero in
3.5. Estimation of an infinite activity Lévy process

Next, we turn to the specific transformation of the Lévy measure that we have defined in Section 3.1.2 as follows

\[ \mu(z) = \int \{ x^T A x \leq z \} \nu(dx), \quad A \in \mathbb{R}^{d \times d}, A \neq 0, \ l \in \mathbb{N}, \ z \in \mathbb{R}^d. \]

The estimator and its pre-estimator are given by

\[ \hat{\mu}_{\text{pre}}(z) = \frac{1}{T} \sum_{0 \leq t \leq T} (\Delta L_t^T A \Delta L_t)^I (\Delta L_t \leq z), \]

\[ \hat{\mu}(z) = \frac{1}{T} \sum_{k=1}^n (X_{h,k}^T A X_{h,k})^I (X_{h,k} \notin B_n, X_{h,k} \leq z). \]

The pre-estimator is based on the real jumps \( \{ \Delta L_t, t \in [0, T] \} \), which are contained in \( \{ J X_{h,k}, k = 1, \ldots, n \} \). However, because of the infinitely many jumps, it is not easy to handle the interaction between \( \{ \Delta L_t, t \in [0, T] \} \) and \( \{ J X_{h,k}, k = 1, \ldots, n \} \). We choose the following approach. We decompose \( J L_t \) into its large jump part and small jump part along a suitable positive zero-sequence \( \{ \xi_n, n \in \mathbb{N} \} \).

The quantities are defined in Remark 2.25. For the corresponding increments we have

\[ J X_{h,k} = C P X_{h,k} [\xi_n] + s J X_{h,k} [\xi_n] - \gamma h + \gamma [\xi_n] h, \quad k = 1, \ldots, n. \]

The next proposition gives conditions on \( \xi_n \) that ensure that asymptotically each \( C P X_{h,k} [\xi_n] \) contains at most one jump. Since \( \xi_n \) converges to zero, \( \{ C P X_{h,k} [\xi_n], k = 1, \ldots, n \} \) and the real jumps \( \{ \Delta L_t, t \in [0, T] \} \) are very similar.

**Proposition 3.28**

Let \( \alpha \in (1/2, 1] \) in \( (3.1) \). Define with some \( u_n > 0 \)

\[ M_2^{(n)}(u_n) := \left\{ \exists k \in \{1, \ldots, n\} \text{ s.t. } N_{(k-1)h, kh} (\{ x \in \mathbb{R}^d : ||x|| \geq u_n \}) > 1 \right\}. \]

(i) Assume that

\[ n^{1-2\alpha} \nu^2 (\{ x \in \mathbb{R}^d : ||x|| \geq \xi_n \}) \rightarrow 0, \quad n \rightarrow \infty. \]

Then, as \( n \rightarrow \infty \),

\[ \mathbb{P} \left( M_2^{(n)}(\xi_n) \right) \rightarrow 0. \]

(ii) Assume that

\[ n^{1-2\alpha}/\xi_n^4 \rightarrow 0, \quad n \rightarrow \infty. \]
3. Estimation method

and that Assumption \[3.14\] is satisfied. Then, as \( n \to \infty \),
\[
P \left( M_2^{(n)}(\xi_n) \right) \to 0.
\]

(iii) Let \( \xi > 0 \) be a constant, then, as \( n \to \infty \),
\[
P \left( M_2^{(n)}(\xi) \right) \to 0.
\]

**Proof.** Ad (i), analogously to Lemma \[3.24\] we obtain
\[
P \left( M_2^{(n)}(\xi_n) \right) \leq nh^2 \nu^2 \left( \{ x \in \mathbb{R}^d : ||x|| \geq \xi_n \} \right)^2 \to 0.
\]

Ad (ii), because of Assumption \[3.14\] there exists a \( \kappa > 0 \) such that
\[
nh^2 \nu^2 \left( \{ x \in \mathbb{R}^d : ||x|| \geq \xi_n \} \right) \leq nh^2 \left( \sum_{i=1}^d \tilde{\nu}_i ((-\xi_n,\xi_n)^c) \right)^2 \leq \kappa nh^2 \left( \int_{(-\xi_n,\xi_n)^c} \frac{1}{x^3} dx \right)^2 = \kappa nh^2 (1/\xi_n^2)^2 \to 0.
\]

Ad (iii), analogously to (i) and (ii). \( \square \)

**Theorem 3.29**

(i) Let \( 1/2 < \alpha < 1 \) in \[3.1\] \( l \in \mathbb{N} \) and and assume \( \mathbb{E} \left[ L_{1,i}^{(l)} \right] < \infty \) for all \( i \). Then, as \( n \to \infty \),
\[
\sup_{z \in \mathbb{R}^d} |\hat{\mu}_{pre}(z) - \mu(z)| \to^p 0 \quad \text{and hence} \quad \sup_{z \in \mathbb{R}^d} |\hat{\mu}(z) - \mu(z)| \to^p 0.
\]

(ii) Let \( 1/2 < \alpha < 1 \) in \[3.1\] We choose
\[
r \in \begin{cases} 
0 \lor 1/2 - 2\alpha/3, \alpha/3) & \text{if } \alpha \leq 3/5 \\
0 \lor 1/2 - 2\alpha/3, (1-\alpha)/2) & \text{if } \alpha > 3/5 
\end{cases}
\]

define
\[
l := \left[ \frac{3r}{\alpha - 3r} \right] + 2, \quad \text{and} \quad p := \frac{l_0}{\alpha - 3r} + 1 \frac{l}{l_0}
\]

and require
\[
\mathbb{E} \left[ L_{1,i}^{(2l(p))} \right], \quad i \in \{1, \ldots, d\}
\]

and that Assumption \[3.14\] is satisfied. Then, as \( n \to \infty \),
\[
n^r \sup_{z \in \mathbb{R}^d} |\hat{\mu}(z) - \mu(z)| \to^p 0 \quad \text{and hence} \quad n^r \sup_{z \in \mathbb{R}^d} |\hat{\mu}(z) - \mu(z)| \to^p 0.
\]
3.5. Estimation of an infinite activity Lévy process

(iii) Let $3/5 < \alpha < 1$ in (3.1). We define

$$l := \left\lfloor \frac{1 - \alpha}{5\alpha/3 - 1} \right\rfloor + 2,$$
and $p := \left\lfloor \frac{2l^{2(2\alpha)}}{5-3\alpha} + 1 \right\rfloor$

and require

$$\mathbb{E} \left[ L_1^{2l(2\alpha)} \right],$$
and that Assumption 3.14 is satisfied. Then, as $n \to \infty$

$$\sqrt{T} \sup_{z \in \mathbb{R}^d} |\hat{\mu}(z) - \mu(z)| \overset{P}{\longrightarrow} 0.$$

Hence, for

$$\tilde{g} : [0, 1]^d \to \mathbb{R}^d, \quad \tilde{g}(y) = \left( g \left( y^{(1)} \right), \ldots, g \left( y^{(d)} \right) \right)^T,$$
where $g : [0, 1] \to \mathbb{R}$ is a strictly increasing and bijective function, we have, as $n \to \infty$

$$\sqrt{T} \left( \hat{\mu} \circ \tilde{g} - \mu \circ \tilde{g} \right) \overset{D}{\longrightarrow} G,$$
where $G$ is a Gaussian process over $[0, 1]^d$ with

$$\mathbb{E} [G(y)] = 0,$$
$$\text{Cov} [G(y_1), G(y_2)] = \int_{\{x^{(1)} \leq g(y^{(1)}_1) \land \ldots \land x^{(d)} \leq g(y^{(d)}_1) \}} \cdots \{x^{(1)} \leq g(y^{(1)}_2) \land \ldots \land x^{(d)} \leq g(y^{(d)}_2) \}} (x^T Ax)^{2l} \nu(dx),$$
for $y, y_1, y_2 \in [0, 1]^d$

Proof. The proof is similar to Theorem 3.26. We use the given $r$ in (ii) and set $r = 0$ in (i) and $r = \frac{1 - \alpha}{2}$ in (iii). We use Theorem 3.21 and Proposition 2.50, so we only have to prove

$$n^r \sup_{z \in \mathbb{R}^d} \left| \frac{1}{T} \sum_{k=1}^{n} (X_{h,k}^T A X_{h,k})^l 1(X_{h,k} \notin B_n, X_{h,k} \leq z) - \mu_{\text{pre}}(z) \right|$$

$$= n^r \sup_{z \in \mathbb{R}^d} \left| \frac{1}{T} \sum_{k=1}^{n} (X_{h,k}^T A X_{h,k})^l 1(X_{h,k} \notin B_n, X_{h,k} \leq z) \right.$$

$$- \frac{1}{T} \sum_{0 \leq t \leq T} (\Delta L_t^T A \Delta L_t)^l 1(\Delta L_t \leq z)$$

$$n^r \sup_{z \in \mathbb{R}^d} \left| \frac{1}{T} \sum_{k=1}^{n} \left( lX_{h,k}^T A X_{h,k} + 2^l X_{h,k}^T A X_{h,k} + C X_{h,k}^T A^C X_{h,k} \right)^l 
1(X_{h,k} \notin B_n, X_{h,k} \leq z) \right.$$

$$- \frac{1}{T} \sum_{0 \leq t \leq T} (\Delta L_t^T A \Delta L_t)^l 1(\Delta L_t \leq z) \right| \overset{P}{\longrightarrow} 0.$$

By Lemma A.1 (iii), we obtain an upper bound that is up to a positive constant
3. Estimation method

equal to

\[
\begin{align*}
&n^r \sup_{z \in \mathbb{R}^d} \left| \frac{1}{T} \sum_{k=1}^{n} \left( J X_{h,k}^T A J X_{h,k} \right)^l \mathbb{1} (X_{h,k} \notin B_n, X_{h,k} \leq z) \\
&\quad - \frac{1}{T} \sum_{0 \leq t \leq T} (\Delta L^T t A \Delta L_t)^l \mathbb{1} (\Delta L_t \leq z) \right| \tag{3.16} \\
&+ n^r \frac{1}{T} \sum_{m=1}^{l} \sum_{m=1}^{n} \left( J X_{h,k}^T A J X_{h,k} \right)^{l-m} \left| C X_{h,k}^T A J X_{h,k} \right|^m \mathbb{1} (X_{h,k} \notin B_n) \tag{3.17} \\
&+ n^r \frac{1}{T} \sum_{m=1}^{l} \sum_{m=1}^{n} \left( J X_{h,k}^T A J X_{h,k} \right)^{l-m} \left| J X_{h,k}^T A C X_{h,k} \right|^m \mathbb{1} (X_{h,k} \notin B_n) \tag{3.18} \\
&+ n^r \frac{1}{T} \sum_{m=1}^{l} \sum_{m=1}^{n} \left( J X_{h,k}^T A J X_{h,k} \right)^{l-m} \left| C X_{h,k}^T A C X_{h,k} \right|^m \mathbb{1} (X_{h,k} \notin B_n). \tag{3.19}
\end{align*}
\]

The term in (3.16) is considered in Lemma C.10(i) for (i) and C.10(ii) for (ii) and (iii). There we use Proposition 3.28. For (3.17) and (3.18) we use the following upper bound

\[
\begin{align*}
\max_{i_1,i_2 \in \{1,\ldots,d\}} \left| A_{i_1,i_2} \right| \cdot n^r \frac{1}{T} \sum_{m=1}^{l} \sum_{m=1}^{n} \left( \sum_{i_1,i_2=1}^{d} \left| J X_{h,k,i_1} \right| \left| J X_{h,k,i_2} \right| \right)^{l-m} \\
+ 2 \left( \sum_{i_1,i_2=1}^{d} \left| C X_{h,k,i_1} \right| \left| J X_{h,k,i_2} \right| \right) \mathbb{1} (X_{h,k} \notin B_n).
\end{align*}
\]

We apply Lemma A.1(ii) and A.1(i) and obtain an upper bound that is up to a positive constant equal to

\[
\begin{align*}
n^r \max_{i_1,i_2 \in \{1,\ldots,d\}} \left| A_{i_1,i_2} \right| \cdot \frac{1}{T} \sum_{m=1}^{l} \sum_{m=1}^{n} \sum_{i_1,i_2,i_3=1}^{d} \left| J X_{h,k,i_1} \right|^{2l-2m} \left| J X_{h,k,i_2} \right|^m \\
\left| C X_{h,k,i_3} \right|^m \mathbb{1} (X_{h,k} \notin B_n).
\end{align*}
\]

That term tends to zero by Lemma C.6. For (3.19) we have analogously

\[
\begin{align*}
2 n^r \max_{i_1,i_2 \in \{1,\ldots,d\}} \left| A_{i_1,i_2} \right| \frac{1}{T} \sum_{m=1}^{l} \sum_{m=1}^{n} \sum_{i_1,i_2,i_3=1}^{d} \left| J X_{h,k,i_1} \right|^{2l-2m} \left| C X_{h,k,i_2} \right|^{2m} \mathbb{1} (X_{h,k} \notin B_n)
\end{align*}
\]

This term tends zero by Lemma C.7 for \( l = 1, m = 1 \), Lemma A.3(i) for \( l > 1, m = 1 \) and Lemma C.6 for all other summands. \( \square \)

As mentioned in Section 3.1.2 we estimate the drift term \( \gamma_{[\infty]} \). This estimator is not based on the critical region. Theorem 3.30 gives the asymptotic properties.
3.6. Estimation of the Blumenthal-Getoor index

Theorem 3.30
Let $0 < \alpha < 1$ in (3.1). Then, as $n \to \infty$,
\[
\hat{\gamma}[\infty] \xrightarrow{a.s.} \gamma[\infty] \quad \text{and} \quad \sqrt{T} (\hat{\gamma}[\infty] - \gamma[\infty]) \xrightarrow{d} Z \sim N \left( 0, C + \int_{\mathbb{R}^d} xx^T \nu(dx) \right).
\]

Proof. We assume w.l.o.g. $T \in \mathbb{N}$, then,
\[
\hat{\gamma}[\infty] := \frac{1}{T} \sum_{k=1}^{n} X_{h,k} = \frac{1}{T} L_T = \frac{1}{T} \sum_{k=1}^{T} X_{1,k}.
\]
The random variables $X_{1,k}$, $k = 1, \ldots, n$ are i.i.d. and the moments are given in Proposition 2.30. So we can apply the Strong Law of Large Numbers and the Central Limit theorem.

3.6. Estimation of the Blumenthal-Getoor index

Recall, the Blumenthal-Getoor index of a Lévy process with Lévy measure $\nu$ is defined by
\[
\inf \{ \gamma' > 0, \int_{|x| \leq 1} |x| \gamma' \nu(dx) < \infty \}.
\]
In this section we introduce an estimation method for this quantity. Thereby the strategy is to consider every single component of the Lévy process separately and to use the estimator $\hat{\mu}$ along suitable zero-sequences. Therefor we need the following assumption on the behavior of the Lévy densities $f_i$ of the univariate component processes $\{L_{t,i}, t \geq 0\}$ in a neighborhood of the origin.

Assumption 3.31
There exists $\theta \in (-\infty, 2]$ and $\varepsilon > 0$ such that each $f_i$ satisfies at least one of the conditions (i) – (iii) and there exists at least one $f_i$ that satisfies (ii) or (iii).

(i) $f_i(x_i) = o \left( 1 / x_i^{1+\theta+\varepsilon} \right)$, $|x_i| \to 0$,

(ii) $f_i(x_i) = \frac{\kappa}{x_i^{1+\theta}} \mathbf{1}(x > 0) + o \left( 1 / x_i^{1+\theta+\varepsilon} \right)$, $|x_i| \to 0$, for some $\kappa > 0$,

(iii) $f_i(x_i) = \frac{\kappa}{x_i^{1+\theta}} \mathbf{1}(x < 0) + o \left( 1 / x_i^{1+\theta+\varepsilon} \right)$, $|x_i| \to 0$, for some $\kappa > 0$.

We define
\[
\gamma := \theta \lor 0.
\]
The following proposition gives an equivalent definition of the Blumenthal-Getoor index that refers to the single components of the Lévy process.
3. Estimation method

**Proposition 3.32**
Let $\gamma' > 0$, we have
\[
\int_{|x| \leq 1} |x|^{\gamma'} \nu(dx) < \infty \quad \text{if and only if} \quad \int_{|x_i| \leq 1} |x_i|^{\gamma'} \tilde{\nu}_i(dx_i) < \infty,
\]
for all $i \in \{1, \ldots, d\}$.

**Proof.** We have
\[
\int_{|x| \leq 1} |x|^{\gamma'} \nu(dx) = \int_{|x| \leq 1} \left( \sum_{i=1}^{d} x_i^2 \right)^{\gamma'/2} \nu(dx) = \int_{|x| \leq 1} \|x\|_{2/\gamma'}^{\gamma'}/2 \nu(dx) = \int_{|x| \leq 1} \left( \sum_{i=1}^{d} |x_i|^{\gamma'}/2 \right)^{\gamma'/2} \nu(dx),
\]
where $\|\cdot\|_{2/\gamma'}$ is the $2/\gamma'$-norm. This norm is equivalent to the $1-$norm. Thus,
\[
\int_{|x| \leq 1} \|x\|_{2/\gamma'}^{\gamma'} \nu(dx) < \infty \iff \int_{|x| \leq 1} \sum_{k=1}^{d} |x_i|^{\gamma'} \nu(dx) < \infty.
\]
Next,
\[
\int_{|x| \leq 1} \sum_{k=1}^{d} |x_i|^{\gamma'} \nu(dx) = \sum_{k=1}^{d} \int_{|x| \leq 1} |x_i|^{\gamma'} \nu(dx) = \sum_{k=1}^{d} \int_{|x_i| \leq 1} \sum_{|x| > 1} |x_i|^{\gamma'} \nu(dx).
\]
The second term is always finite, because $\nu(\mathbb{R}^d) < \infty$. The first term is equal to
\[
\sum_{k=1}^{d} \int_{|x_i| \leq 1} |x_i|^{\gamma'} \tilde{\nu}_i(dx_i).
\]
And the sum is finite if and only if every summand is finite. \(\square\)

**Proposition 3.33**
The $\gamma$ of Assumption 3.31 is equal to the Blumenthal-Getoor index.

**Proof.** We define the set
\[
S := \left\{ \gamma' > 0 : \int_{|x| \leq 1} |x|^{\gamma'} \nu(dx) < \infty \right\}.
\]
The Blumenthal-Getoor index is defined as $\inf S$, so we have to show
\[
(i) \, \gamma \geq \inf S \quad \text{and} \quad (ii) \, \gamma \leq \inf S.
\]
Ad (i), because of Assumption 3.31 there exists a $\kappa > 0$ such that,
3.6. Estimation of the Blumenthal-Getoor index

for all \(\gamma' > \gamma\),

\[
\sum_{i=1}^{d} \int_{|x_i| \leq 1} |x_i|^\gamma f_i(x_i)dx_i \leq \int_{|x_i| \leq 1} x_i^\gamma \frac{K}{x_i^{1+\gamma}}dx_i < \infty.
\]

By Proposition 3.32, we obtain \(\gamma' \in S\). Thus \(\gamma' \geq \inf S\) for all \(\gamma' > \gamma\). That means that also \(\gamma \geq \inf S\).

Ad (ii), we assume that \(\gamma > \inf S > 0\). Then \(\theta = \gamma > 0\) in Assumption 3.31. Thus there exists a \(i \in \{1, \ldots, d\}\) such that

\[
f_i(x_i) = \frac{K}{|x_i|^{1+\gamma}} + o(1/|x_i|^{1+\gamma}).
\]

By Proposition 3.32, there exists a \(\gamma' \in (\inf S, \gamma)\) such that

\[
\int_{|x_i| \leq 1} |x_i|^\gamma f_i(x_i)dx_i < \infty.
\]

This is not possible, because \(1 + \gamma - \gamma' > 1\) and thus the term cannot be integrable.

Next, we define the estimator for the Blumenthal-Getoor index. The strategy is to estimate \(\gamma\) of Assumption 3.31. Recall, we have defined the transformation of the Lévy measure the following way

\[
\mu(z) = \int_{\{x \leq z\}} \left(x^TAx\right)^l \nu(dx), \quad A \in \mathbb{R}^{d \times d}, \ A \neq 0, \ l \in \mathbb{N}, \ z \in \mathbb{R}^d.
\]

Its estimator has been defined by

\[
\tilde{\mu}(z) = \frac{1}{T} \sum_{k=1}^{n} (X_{h,k}^TA_{X_{h,k}})^l \mathbb{1}(X_{h,k} \notin B_n, X_{h,k} \leq z).
\]

We choose suitable matrices \(A\) and consider \(\tilde{\mu}(z)\) along zero sequences.

**Definition 3.34** (Estimator for the Blumenthal-Getoor index)

Let \(\alpha \in (1/2, 1)\) and choose

\[
l := \left|\frac{1 - \alpha}{4\alpha - 1}\right| + 2, \quad \delta_n := n^{(\alpha-1)/9(l+1)}
\]

and define zero sequences

\[
x_{1,n} := -\sqrt{\delta_n}, \quad x'_{1,n} := -2\sqrt{\delta_n}, \quad x_{2,n} := \sqrt{\delta_n}, \quad x'_{2,n} := 2\sqrt{\delta_n}.
\]

Then, for \(i \in \{1, \ldots, d\}\), we define the sequence \(z_{i,1,n}\) in \(\mathbb{R}^d\) by

\[
z_{i,1,n}^{(i)} := x_{1,n}, \quad z_{i,1,n}^{(j)} := \infty, \quad j \neq i.
\]

For \(i \in \{1, \ldots, d\}\) the sequences \(z'_{i,1,n}, z_{i,2,n}\) and \(z'_{i,2,n}\) are defined analogously.

Similar, the sequences \(\overline{Z}_{i,n} \in \mathbb{R}^d, \ i \in \{1, \ldots, d\}\) are defined by

\[
\overline{Z}_{i,n}^{(i)} := \delta_n, \quad \overline{Z}_{i,n}^{(j)} := 0, \quad j \neq i.
\]
3. Estimation method

Next, we choose, for \( i \in \{1, \ldots, d\} \), the matrices \( A_i \in \mathbb{R}^{d \times d} \) as follows
\[
A_{i,i} := 1, \quad A_{j_1,j_2} := 0 \quad j_1 \neq i \quad \text{or} \quad j_2 \neq i.
\]

Then, we define
\[
\hat{\gamma}_{1,i} := \frac{\log \left( \frac{\tilde{\mu}_{A_i,l}(z_{i,1,n}) - \tilde{\mu}_{A_i,l}(z_{i,1,n} - \delta_{i,n})}{\delta_n x_{i,1,n}^l} \right) - \log \left( \frac{\tilde{\mu}_{A_i,l}(z'_{i,1,n}) - \tilde{\mu}_{A_i,l}(z'_{i,1,n} - \delta_{i,n})}{\delta_n x_{i,1,n}^l} \right)}{\log \left( \frac{x_{i,1,n}}{x_{i,1,n}'} \right)} - 1,
\]
\[
\hat{\gamma}_{2,i} := \frac{\log \left( \frac{\tilde{\mu}_{A_i,l}(z_{i,2,n}) - \tilde{\mu}_{A_i,l}(z_{i,2,n} - \delta_{i,n})}{\delta_n x_{i,2,n}^l} \right) - \log \left( \frac{\tilde{\mu}_{A_i,l}(z'_{i,2,n}) - \tilde{\mu}_{A_i,l}(z'_{i,2,n} - \delta_{i,n})}{\delta_n x_{i,2,n}^l} \right)}{\log \left( \frac{x_{i,2,n}}{x_{i,2,n}'} \right)} - 1.
\]

Finally, the estimator for the Blumenthal-Getoor index is given by
\[
\hat{\gamma} := \max_{i=1,\ldots,d} \left\{ 0 \lor \hat{\gamma}_{1,i} \lor \hat{\gamma}_{2,i} \right\}.
\]

Theorem 3.35 shows consistency of this estimator. Because we consider the logarithms of \( \tilde{\mu}(z) \), it would be only possible to prove logarithmic rates of convergence.

\textbf{Theorem 3.35}

Let \( 1/2 < \alpha < 1 \) in (3.1) and Assumption 3.31 be satisfied. Then,
\[
\hat{\gamma} \xrightarrow{P} \gamma.
\]

\textbf{Proof.} We start with \( \hat{\gamma}_{1,i} \), \( i \in \{1, \ldots, d\} \). At first, we use Theorem 3.29 to replace \( \tilde{\mu} \) by \( \mu \).

\[
\frac{\left| \tilde{\mu}_{A_i,l}(z_{i,1,n}) - \tilde{\mu}_{A_i,l}(z_{i,1,n} - \delta_{i,n}) - (\mu_{A_i,l}(z_{i,1,n}) - \mu_{A_i,l}(z_{i,1,n} - \delta_{i,n})) \right|}{\delta_n x_{i,1,n}^l} \leq 2\delta_n^{-(1+l)} \left| \tilde{\mu}_{A_i,l}(z_{i,1,n}) - \tilde{\mu}_{A_i,l}(z_{i,1,n} - \delta_{i,n}) - (\mu_{A_i,l}(z_{i,1,n}) - \mu_{A_i,l}(z_{i,1,n} - \delta_{i,n})) \right|.
\]

Because Assumption 3.31 implies Assumption 3.14, the definition of \( l \) above leads to a speed of convergence \( r = (1 - \alpha)/9 \) in Theorem 3.29. The convergence follows form
\[
\delta_n^{-(1+l)} = n^r.
\]
3.6. Estimation of the Blumenthal-Getoor index

Analogously we have

\[
\left| \frac{\widehat{\mu}_{A,l}(z_{i,1,n}) - \mu_{A,l}(z_{i,1,n} - \delta_{i,n}) - \left( \mu_{A,l}(z_{i,1,n}) - \mu_{A,l}(z_{i,1,n} - \delta_{i,n}) \right)}{\delta_n x_{1,n}^{2l}} \right| \xrightarrow{P} 0.
\]

Secondly, we want to reduce the term to the densities. By the First mean value theorem for integration there exists \( \zeta_{1,n} \in (x_{1,n} - \delta_n, z_{1,n}) \) such that

\[
\frac{\mu_{A,l}(z_{i,1,n}) - \mu_{A,l}(z_{i,1,n} - \delta_{i,n})}{\delta_n x_{1,n}^{2l}} = \frac{\int_{x_{1,n} - \delta_n \leq x \leq x_{1,n}} x_i^{2l} f_i(x_i)dx_i}{\delta_n x_{1,n}^{2l}} = \frac{\zeta_{1,n}^{2l}}{x_{1,n}^{2l}} f_i(\zeta_{1,n}).
\]

Analogously we have

\[
\frac{\mu_{A,l}(z_{i,1,n}) - \mu_{A,l}(z_{i,1,n} - \delta_{i,n})}{\delta_n x_{1,n}^{2l}} = \frac{\zeta_{1,n}^{2l}}{x_{1,n}^{2l}} f_i(\zeta_{1,n}).
\]

Next, we have to distinguish between two cases (i)

\[
f_i(x_i) = \frac{\kappa}{x_i^{1+\gamma}} + o\left(1/x_i^{1+\gamma}\right), \quad x_i \in (-\varepsilon, 0)
\]

and (ii)

\[
f_i(x_i) = o\left(1/x_i^{1+\gamma}\right), \quad |x_i| \to 0,
\]

Ad (i), we have

\[
\log \left( \frac{\zeta_{1,n}^{2l}}{x_{1,n}^{2l}} f_i(\zeta_{1,n}) \right) - \log \left( \frac{\kappa}{\zeta_{1,n}^{1+\gamma_1}} \right) \leq \log \left( \frac{\kappa}{\zeta_{1,n}^{1+\gamma_1}} + o\left( \frac{\kappa}{\zeta_{1,n}^{1+\gamma_1}} \right) \right) + \log \left( \frac{\zeta_{1,n}^{2l}}{x_{1,n}^{2l}} \right).
\]

Obviously both summands tend to zero. Analogously we have

\[
\log \left( \frac{\zeta_{1,n}^{2l}}{x_{1,n}^{2l}} f_i(\zeta_{1,n}) \right) - \log \left( \frac{\kappa}{\zeta_{1,n}^{1+\gamma_1}} \right) \to 0.
\]

Altogether we have

\[
\log \left( \frac{\kappa}{\zeta_{1,n}^{1+\gamma_1}} \right) - \log \left( \frac{\kappa}{\zeta_{1,n}^{1+\gamma_1}} \right) - 1 = (\gamma + 1) \log \left( \frac{\zeta_{1,n}^{1+\gamma_1}}{\zeta_{1,n}^{1+\gamma_1}} \right) - 1.
\]

This term tends obviously to \( \gamma \).

Ad (ii), there exists a positive function \( c(x) \to 0, x \to 0 \) and an \( \varepsilon > 0 \) such that

\[
f_i(x_i) = \frac{c(x_i)}{x_i^{1+\theta}} \quad x_i \in (-\varepsilon, 0).
\]
3. Estimation method

So analogously to (i) we obtain
\[
\limsup_{n \to \infty} \frac{\log \left( c(\zeta_n)/c(\zeta'_n) \right) (\gamma + 1) \log \left( \zeta'_{1,n}/\zeta_{1,n} \right)}{\log \left( x'_{1,n}/x_{1,n} \right)} - 1 \leq \gamma.
\]

We obtain analogous results for \( \widehat{\gamma}_{2,i} \), \( i \in \{1, \ldots, d\} \). So, by Assumption 3.31, \( \max_{i=1,\ldots,d}(0 \lor \widehat{\gamma}_{1,i} \lor \widehat{\gamma}_{2,i}) \xrightarrow{P} \gamma. \)

3.7. Estimation of the critical region and parameter choice

We recall the definition of the critical set (Definition 3.5 of Section 3.1). At first, for all \( n \), we have defined
\[
B_n(D, \beta) := \{ x \in \mathbb{R}^d : x^T D^{-1} x \leq \beta b_n^2 \},
\]
where \( D \) is a symmetric \( d \times d \)-matrix, \( \beta \in \mathbb{R}^+ \) and
\[
b_n := \sqrt{2h \log n}.
\]
We have called a sequence of random sets \( \{B_n, n \in \mathbb{N}\} \) a sequence of critical regions, if there exists \( 1 < \beta' < \beta'' < \infty \) such that
\[
B_n(C, \beta') \subseteq B_n[\omega] \subseteq B_n(I, \beta'')
\]
for all \( n \in \mathbb{N} \) and \( \omega \in \Omega \) (\( I \) is the identity matrix). An element \( B_n \) of the sequence is called critical region. For \( \{\beta_n, n \in \mathbb{N}\} \) satisfying \( \inf_{n \in \mathbb{N}} \beta_n > 1 \) and \( \sup_{n \in \mathbb{N}} \beta_n < \infty \), we have called \( \{B_n = B_n(C, \beta_n), n \in \mathbb{N}\} \) a sequence of true critical regions.

In this section we discuss how to find a suitable critical region. From a practical point of view, for a given sequence \( \{\beta_n, n \in \mathbb{N}\} \), we want to choose a sequence of critical regions that is as similar as possible to the corresponding sequence of true critical regions. However, the true critical regions depend on the unknown covariance matrix \( C \) of the diffusion part. Our strategy is to use the estimator \( C \), instead. This estimator in turn depends on the critical region. So we introduce in Subsection 3.7.1 an iteration method to estimate \( B_n \) and \( C \) simultaneously. Furthermore, in Subsection 3.7.2 we define a sequence \( \{\beta_n, n \in \mathbb{N}\} \) and give an intuitive interpretation of it.

3.7.1. An iteration method for the critical region

If we use an estimator to generate a critical region, it cannot always be guaranteed that this set satisfies the condition of Definition 3.5 for all \( \omega \in \Omega \). But in the next theorem we prove that it is sufficient for the results of the estimators, defined in the previous, that Definition 3.5 hold true with probability tending to one. This result is based on the fact that the pre-estimators are in-
3.7. Estimation of the critical region and parameter choice

dependent of the critical region and that the difference between the estimator and the pre-estimator converges to zero.

**Theorem 3.36**
The asymptotic results of Theorem 3.22, 3.23, 3.25, 3.26, 3.27 and 3.29 do not change if we use a sequence of random sets \( \{ B_n, n \in \mathbb{N} \} \) satisfying

\[
P(B_n(C, \beta') \subseteq \widetilde{B}_n \subseteq B_n(I, \beta'')) \xrightarrow{P} 1,
\]

for some \( 1 < \beta' < \beta'' < \infty \), instead of a sequence of critical regions. An element \( \widetilde{B}_n \) of such a sequence is called asymptotic critical region.

**Proof.** We give the proof w.l.o.g. for \( \hat{C} \) in the case of finite activity. We consider \( \hat{C} \) component-by-component and define a critical region

\[
B_n^{(2)} := \begin{cases} 
\widetilde{B}_n & B_n(C, \beta') \subseteq \widetilde{B}_n \subseteq B_n(I, \beta'') \\
B_n & \text{else}
\end{cases}
\]

Then, for all \( i, j \in \{1, \ldots, d\} \), we set

\[
\hat{\theta}_n(B_n) := \sqrt{n} \frac{1}{T} \sum_{k=1}^{n} X_{h,k,i} X_{h,k,j} 1(X_{h,k} \in \widetilde{B}_n),
\]

\[
\hat{\theta}_n(B_n^{(2)}) := \sqrt{n} \frac{1}{T} \sum_{k=1}^{n} X_{h,k,i} X_{h,k,j} 1(X_{h,k} \in B_n^{(2)}),
\]

\[
\bar{\theta}_n := \sqrt{n} \frac{1}{T} \sum_{k=1}^{n} C X_{h,k,i} C X_{h,k,j}.
\]

Then, as \( n \to \infty \),

\[
|\hat{\theta}_n(B_n) - \bar{\theta}_n| \leq |\hat{\theta}_n(B_n) - \bar{\theta}_n| 1(B_n(C, \beta') \subseteq \widetilde{B}_n \subseteq B_n(C, \beta'')) \]

\[
+ |\hat{\theta}_n(B_n) - \bar{\theta}_n| 1(B_n(C, \beta') \not\subseteq \widetilde{B}_n \text{ or } \widetilde{B}_n \not\subseteq B_n(C, \beta''))
\]

\[
\leq |\hat{\theta}_n(B_n^{(2)}) - \bar{\theta}_n| + |\hat{\theta}_n(B_n) - \bar{\theta}_n| 1(B_n(C, \beta') \not\subseteq \widetilde{B}_n \text{ or } \widetilde{B}_n \not\subseteq B_n(C, \beta'')) \xrightarrow{P} 0,
\]

because \( B_n^{(2)} \) fulfills Definition 3.5 and the probability that the second summand is unequal to zero converges to zero. That means Theorem 3.22 remains true. All other estimators are of the same structure.

The next proposition implies that if we have a consistent estimator for \( C \), we also have an asymptotic critical region that is similar to the true critical region.

**Proposition 3.37**
Let \( \hat{D} \) be a consistent estimator for a symmetric \( d \times d \) matrix \( D \) which satisfies

\[
D - C \quad \text{is nonnegative definite.}
\]
3. Estimation method

Furthermore, let \( \{\beta_n, n \in \mathbb{N}\} \) satisfy \( \inf_{n \in \mathbb{N}} \beta_n > 1 \) and \( \sup_{n \in \mathbb{N}} \beta_n < \infty \). Then, for all \( \varepsilon > 0 \) and as \( n \to \infty \),

\[
P \left( B_n(D, \beta_n(1 - \varepsilon)^2) \subseteq B_n(\hat{D}, \beta_n) \subseteq B_n(D, \beta_n(1 + \varepsilon)^2) \right) = P \left( (1 - \varepsilon) B_n(D, \beta_n) \subseteq B_n(\hat{D}, \beta_n) \subseteq (1 + \varepsilon) B_n(D, \beta_n) \right) \to 1.
\]

Proof. We have to show, that for all \( \varepsilon > 0 \)

\[
P \left( (1 - \varepsilon) \{ x \in \mathbb{R}^d : x^T D^{-1} x \leq \beta_n b_n^2 \} \subseteq \{ x \in \mathbb{R}^d : x^T \hat{D}^{-1} x \leq \beta_n b_n^2 \} \right) \subseteq (1 + \varepsilon) \{ x \in \mathbb{R}^d : x^T D^{-1} x \leq \beta_n b_n^2 \} \to 1,
\]

for all consistent estimators \( \hat{D} \) of \( D \). We only show w.l.o.g.

\[
P \left( (1 - \varepsilon) \{ x \in \mathbb{R}^d : x^T D^{-1} x \leq 1 \} \subseteq \{ x \in \mathbb{R}^d : x^T \hat{D}^{-1} x \leq 1 \} \right) \to 1.
\]

We define

\[
\xi := (1 - \varepsilon)^{-2} > 1,
\]

then, we have

\[
P \left( (1 - \varepsilon) \{ x \in \mathbb{R}^d : x^T D^{-1} x \leq 1 \} \right) \to 1,
\]

since every Leading Principal Minor of

\[
\xi D^{-1} - \hat{D}^{-1}
\]

converges in probability to the Leading Principal Minor of

\[
(\xi - 1) D^{-1} > 0,
\]

since the Leading Principal Minors are continuous functions of the entries of the matrix.

\[\square\]

**Corollary 3.38**

Let \( \hat{D} \) be a consistent estimator for a symmetric \( d \times d \)-matrix \( D \) which satisfies

\[
D - C \quad \text{is nonnegative definite.}
\]

Furthermore let \( \{\beta_n, n \in \mathbb{N}\} \) satisfy \( \inf_{n \in \mathbb{N}} \beta_n > 1 \) and \( \sup_{n \in \mathbb{N}} \beta_n < \infty \).
3.7. Estimation of the critical region and parameter choice

A sequence \( \{ \tilde{B}_n, n \in \mathbb{N} \} \) given by

\[
\tilde{B}_n := B_n(\hat{D}, \beta_n)
\]

fulfills the condition of Theorem \[3.36\].

For an arbitrary random set \( S_n, S_n[\omega] \subseteq \mathbb{R}^d \), we extend the notation of the estimator of the covariance matrix as follows

\[
\hat{C}(S_n) := \frac{1}{T} \sum_{k=1}^{n} X_{h,k}X_{h,k}^T \mathbb{1}(X_{h,k} \in S_n).
\]

That means \( \hat{C}(S_n) \) is a consistent estimator for \( C \) if \( S_n \) is an asymptotic critical region.

Next, we introduce the iteration method. The idea is to start with a set that is too big in any case. Subsequently we shrink it in such a way that an asymptotic critical region is obtained that is similar to the true critical region.

**Definition 3.39** (Iteration method)

Let \( \{ \beta_n, n \in \mathbb{N} \} \) satisfy \( \inf_{n \in \mathbb{N}} \beta_n > 1 \) and \( \sup_{n \in \mathbb{N}} \beta_n < \infty \). We define

\[
\hat{B}_{n,0}(\beta_n) := \mathbb{R}^d
\]

and, for all \( i = 1, \ldots, n \),

\[
\hat{B}_{n,i}(\beta_n) := B_n(\hat{C}(\hat{B}_{n,i-1}(\beta_n)), \beta_n) \quad \text{and} \quad \hat{B}_{n}^*(\beta_n) := \hat{B}_{n,n}(\beta_n).
\]

**Proposition 3.40** (Properties of the iteration method)

Let \( n \in \mathbb{N} \) and \( \omega \in \Omega \) be fixed .

(i) Let \( i \in \{ 0, \ldots, n-1 \} \), if

\[
\hat{B}_{n,i}(\beta_n)[\omega] = \hat{B}_{n,i+1}(\beta_n)[\omega] \quad \text{then} \quad \hat{B}_{n,j}(\beta_n)[\omega] = \hat{B}_{n,i}(\beta_n)[\omega],
\]

for all \( j \in \{ i, \ldots, n \} \).

(ii) Let \( i \in \{ 0, \ldots, n \} \). If

\[
X_{h,k}(\omega) \notin \{ \hat{B}_{n,i-1}(\beta_n)[\omega] \cup \hat{B}_{n,i+1}(\beta_n)[\omega] \} \setminus \{ \hat{B}_{n,i}(\beta_n)[\omega] \cap \hat{B}_{n,i+1}(\beta_n)[\omega] \}
\]

for all \( k \in \{ 1, \ldots, n \} \), then,

\[
\hat{B}_{n,i+2}(\beta_n)[\omega] = \hat{B}_{n,i+1}(\beta_n)[\omega].
\]

(iii) \( \hat{B}_{n,i}(\beta_n)[\omega] \) is monotone decreasing in \( i \), i.e.

\[
\hat{B}_{n,i}(\beta_n)[\omega] \supseteq \hat{B}_{n,i+1}(\beta_n)[\omega]
\]

for all \( i \in \{ 0, \ldots, n-1 \} \).
3. Estimation method

(iv) \( \hat{B}_{n,i}(\beta_n)[\omega] \) converges after at most \( n \) steps, i.e.
\[
\hat{B}_{n,n+1}(\beta_n)[\omega] = \hat{B}_{n,n}(\beta_n)[\omega].
\]

(v) Let \( B_n \) be a critical region. Then from
\[
B_n(\omega) \subseteq \hat{B}_{n,i}(\beta_n)[\omega] \quad \text{follows} \quad B_n(\hat{C}(B_n(\omega)), \beta_n) \subseteq \hat{B}_{n,i+1}(\beta_n)[\omega].
\]

Proof. Ad (i), we have to show that from
\[
\hat{B}_{n,i}(\beta_n)[\omega] = \hat{B}_{n,i+1}(\beta_n)[\omega] \quad \text{follows} \quad \hat{B}_{n,i+1}(\beta_n)[\omega] = \hat{B}_{n,i+2}(\beta_n)[\omega].
\]
This is obtained by
\[
\hat{B}_{n,i+2}(\beta_n)[\omega] = \{ x \in \mathbb{R}^d : x^T \hat{C}^{-1}(\hat{B}_{n,i+1}(\beta_n)[\omega]) x \leq \beta_n b_n^2 \}
\]
\[
= \{ x \in \mathbb{R}^d : x^T \hat{C}^{-1}(\hat{B}_{n,i}(\beta_n)[\omega]) x \leq \beta_n b_n^2 \} = \hat{B}_{n,i+1}(\beta_n)[\omega].
\]

Ad (ii), we have
\[
\hat{C}(\hat{B}_{n,i+1}(\beta_n)[\omega]) = \frac{1}{T} \sum_{k=1}^{n} X_{h,k}(\omega)X_{h,k}^T(\omega) \mathbb{1}(X_{h,k}(\omega) \in \hat{B}_{n,i+1}(\beta_n)[\omega])
\]
\[
= \frac{1}{T} \sum_{k=1}^{n} X_{h,k}(\omega)X_{h,k}^T(\omega) \mathbb{1}(X_{h,k}(\omega) \in \hat{B}_{n,i}(\beta_n)[\omega]) = \hat{C}(\hat{B}_{n,i}(\beta_n)[\omega])
\]
That means \( \hat{B}_{n,i+2}(\omega) = \hat{B}_{n,i+1}(\omega) \).

Ad (iii), since \( B_{n,0}(\beta_n)[\omega] = \mathbb{R}^d \) we have \( B_{n,0}(\beta_n)[\omega] \supseteq B_{n,1}(\beta_n)[\omega] \). If
\[
\hat{B}_{n,i}(\beta_n)[\omega] \subseteq \hat{B}_{n,i-1}(\beta_n)[\omega],
\]
holds true, we also have
\[
\hat{B}_{n,i+1}(\beta_n)[\omega] = \{ x \in \mathbb{R}^d : x^T \hat{C}^{-1}(\hat{B}_{n,i}(\beta_n)[\omega]) x \leq \beta_n b_n^2 \}
\]
\[
= \{ x \in \mathbb{R}^d : x^T \left( \frac{1}{T} \sum_{k=1}^{n} X_{h,k}(\omega)X_{h,k}^T(\omega) \mathbb{1}(X_{h,k}(\omega) \in \hat{B}_{n,i}(\beta_n)[\omega]) \right)^{-1} x \leq \beta_n b_n^2 \}
\]
\[
\subseteq \{ x \in \mathbb{R}^d : x^T \left( \frac{1}{T} \sum_{k=1}^{n} X_{h,k}(\omega)X_{h,k}^T(\omega) \mathbb{1}(X_{h,k}(\omega) \in \hat{B}_{n,i}(\beta_n)[\omega]) \right)^{-1} x \leq \beta_n b_n^2 \}
\]
\[
+ \frac{1}{T} \sum_{k=1}^{n} X_{h,k}(\omega)X_{h,k}^T(\omega) \mathbb{1}(X_{h,k}(\omega) \in \hat{B}_{n,i-1}(\beta_n)[\omega] \setminus \hat{B}_{n,i}(\beta_n)[\omega])^{-1} x \leq \beta_n b_n^2 \}
\]
\[
= \hat{B}_{n,i+1}(\beta_n)[\omega].
\]
We have used the fact, that for two positive definite and symmetrical matrices \( A_1 \) and \( A_2 \), the inequality
\[
x^T(A_1 + A_2)^{-1} x \leq x^T A_1^{-1} x
\]
holds true. This fact can be easily proven by showing that
\[
A_1^{-1} - (A_1 + A_2)^{-1}
\]
3.7. Estimation of the critical region and parameter choice

is positive definite.

Ad (iv), we assume that

\[ \hat{B}_{n,n+1}(\beta_n)[\omega] \neq \hat{B}_{n,n}(\beta_n)[\omega]. \]

Then, by (i), we have

\[ \hat{B}_{n,i}(\beta_n)[\omega] \neq \hat{B}_{n,i+1}(\beta_n)[\omega] \]

d for all \( i \in \{0, ..., n\} \). Then, by (ii) and (iii) we obtain that there exists at least \( n + 1 \) different indices \( k_0 \neq k_1 \neq ... \neq k_n \) such that for all \( i \in \{0, ..., n\} \)

\[ X_{h,k_i}(\omega) \in \{ \hat{B}_{n,i}(\beta_n)[\omega] \cup \hat{B}_{n,i+1}(\beta_n)[\omega] \} \setminus \{ \hat{B}_{n,i}(\beta_n)[\omega] \cap \hat{B}_{n,i+1}(\beta_n)[\omega] \}. \]

But that is not possible, since \( k_1, k_2, ..., k_{n+1} \in \{1, ..., n\} \) (we have only \( n \) observations).

Ad (v), we have

\[ \hat{B}_{n,i+1}(\beta_n)[\omega] = \left\{ x \in \mathbb{R}^d : x^T \tilde{C}^{-1}(\hat{B}_{n,i}(\beta_n)[\omega]) x \leq \beta_n b_n^2 \right\} \]

\[ = \left\{ x \in \mathbb{R}^d : x^T \left( \frac{1}{T} \sum_{k=1}^{n} X_{h,k}(\omega) X_{h,k}^T(\omega) [X_{h,k}(\omega) \in \tilde{B}_{n,i}(\beta_n)[\omega]] \right)^{-1} x \leq \beta_n b_n^2 \right\} \]

\[ = \left\{ x \in \mathbb{R}^d : x^T \left( \frac{1}{T} \sum_{k=1}^{n} X_{h,k}(\omega) X_{h,k}^T(\omega) [X_{h,k}(\omega) \in \hat{B}_{n,i}(\beta_n)[\omega] \setminus B_n(\omega)] \right)^{-1} x \leq \beta_n b_n^2 \right\} \]

\[ \sup \left\{ x \in \mathbb{R}^d : x^T \left( \frac{1}{T} \sum_{k=1}^{n} X_{h,k}(\omega) X_{h,k}^T(\omega) [X_{h,k}(\omega) \in B_n(\omega)] \right)^{-1} x \leq \beta_n b_n^2 \right\} \]

\[ = B_n(\hat{C}(B_n(\omega)), \beta_n). \]

The next theorem proves that the iteration method gives us an asymptotic critical region that is similar to the true critical region.

**Theorem 3.41**

Let \( \{\beta_n, n \in \mathbb{N}\} \) satisfy

\[ \inf_{n \in \mathbb{N}} \beta_n > 1 \quad \text{and} \quad \sup_{n \in \mathbb{N}} \beta_n < \infty. \]

Then, for all \( \epsilon > 0 \) and as \( n \to \infty \),

\[ \mathbb{P} \left( B_n(C, \beta_n(1 - \epsilon)^2) \subseteq \hat{B}_n^*(\beta_n) \subseteq B_n(C, \beta_n(1 + \epsilon)^2) \right) \]

\[ = \mathbb{P} \left( (1 - \epsilon) B_n(C, \beta_n) \subseteq \hat{B}_n^*(\beta_n) \subseteq (1 + \epsilon) B_n(C, \beta_n) \right) \to 1. \]
3. Estimation method

Proof. At first, by Theorem 3.27 and 3.29 we have
\[
\frac{1}{T} \sum_{k=1}^{n} X_{h,k} X_{h,k}^T \overset{P}{\to} C + \int_{\mathbb{R}^d} x x^T \nu(dx).
\]
Then, we have
\[
\hat{B}_{n,1}(\beta_n) = B_n \left( \frac{1}{T} \sum_{k=1}^{n} X_{h,k} X_{h,k}^T, \beta_n \right).
\]
By Corollary 3.38 there exists a \( \beta'' > 1 \) such that
\[
P(\hat{B}_{n,1}(\beta_n) \subseteq B_n(I, \beta'')) \to 1.
\]
Because of Proposition 3.40(iii) we also have
\[
P(\hat{B}_{n}^*(\beta_n) \subseteq B_n(I, \beta'')) \to 1.
\]
Next, we use Proposition 3.40(iii), (v) and \( B_{n,0}(\beta_n) = \mathbb{R}^d \). For all \( \beta' \) satisfying \( 1 < \beta' < \inf_n \beta_n \) we have
\[
P\left( B_n(C, \beta') \nsubseteq \hat{B}_{n}^*(\beta_n) \right) \\
\leq P\left( \exists i \in \{0, \ldots, n\} : B_n(C, \beta') \nsubseteq \hat{B}_{n,i}(\beta_n), B_n(C, \beta') \nsubseteq \hat{B}_{n,i+1}(\beta_n) \right) \\
\leq P\left( \exists i \in \{0, \ldots, n\} : B_n(C, \beta') \nsubseteq \hat{B}_{n,i}(\beta_n), B_n(C, \beta') \nsubseteq \hat{B}_{n,i+1}(\beta_n), B_n(C, \beta') \nsubseteq \hat{B}_{n,i}(\beta_n) \right) \\
\leq P\left( B_n(C, \beta') \nsubseteq \hat{B}_{n,i}(\beta_n) \right) \\
\leq P\left( \sqrt{\beta' / \beta_n} B_n(C, \beta_n) \nsubseteq B_n(\hat{C}(B_n(C, \beta'), \beta_n)) \right) \\
\leq P\left( \sqrt{\beta' / \inf_n \beta_n} B_n(C, \beta_n) \nsubseteq B_n(\hat{C}(B_n(C, \beta'), \beta_n)) \right) \to 0,
\]
because \( \beta' / \inf_n \beta_n < 1 \) and Proposition 3.37

That means the condition in Theorem 3.36 is satisfied and thus \( \hat{C}(\hat{B}_{n}^*(\beta_n)) \) is a consistent estimator of \( C \) and because of Proposition 3.40(iv) we have
\[
\hat{B}_{n}^*(\beta_n) = B_n(\hat{C}(\hat{B}_{n}^*(\beta_n))).
\]
So we can apply Proposition 3.37

3.7.2. Parameter choice

In this subsection we consider the sequence \( \{\beta_n, n \in \mathbb{N}\} \) that has a nice interpretation. To illustrate this interpretation we choose the following approach. We define an alternative critical set
\[
B'_n(E_n) := \{ x \in \mathbb{R}^d : x^T C^{-1} x \leq h z_{1-E_n/n} \}, \tag{3.20}
\]
where
\[
E_n := \tau n^{-\rho}, \quad \tau > 0, \quad \rho > 0.
\]
3.7. Estimation of the critical region and parameter choice

The $1 - E_n/n$ quantile of the $\chi^2_d$-distribution is denoted by $z_{1-E_n/n}$. Obviously, by the binomial distribution we have

$$E \left[ \#k \in \{1, \ldots, n\} : C' X_{h,k} \notin B'_n(E_n) \right] = E_n,$$

where $\{C' X_{h,k}, k = 1, \ldots, n\}$ are the centered diffusion increments. This means we interpret the quantity $E_n$ as the expected number of data points, where we erroneously anticipate a jump. In doing so we have ignored the drift term, that has asymptotically no effect.

The connection to Definition 3.5 is simple, we have

$$B'_n(E_n) = B_n \left( C, \frac{z_{1-E_n/n}}{2 \log n} \right),$$

So we define

$$\beta_n := \frac{z_{1-E_n/n}}{2 \log n}$$

and choose $E_n$ instead of $\beta_n$ directly. The next theorem shows how $\beta_n$ that is defined this way behaves asymptotically. The result implies that there exists

$$1 < \beta' < \beta'' < \infty,$$

such that

$$\beta_n \in (\beta', \beta'')$$

for almost every $n$. That means we can use the iteration method.

**Theorem 3.42**
We define

$$\beta_n := \frac{z_{1-E_n/n}}{2 \log n}, \quad \beta := 1 + \rho > 1, \quad E_n := \tau n^{-\rho}, \quad \rho > 0, \quad \tau > 0.$$

Then,

$$\beta_n / \beta \to 1.$$

**Proof.** We consider the quantile of the $\chi^2_d$-distribution. For all $\xi_1 > 0$, there exists a $n_0(\xi_1)$ such that

$$1 - \frac{1}{\tau n^{1+\rho}} = \int_{z_{1-E_n/n}}^{\infty} \frac{d^{d-1}}{2^{d/2} \Gamma(d/2)} \exp \left( -\frac{t}{2} \right) dt$$

$$\geq \int_{z_{1-E_n/n}}^{\infty} \exp \left( -\frac{t}{2 + \xi_1} \right) dt = \frac{\exp \left( -\frac{z_{1-E_n/n}}{2+\xi_1} \right)}{2 - \xi_1}$$

for all $n \geq n_0(\xi_1)$. This means

$$z_{1-E_n/n} \leq (2 + \xi_1)(1 + \rho) \log n$$

for all $n \geq n_0(\xi_1)$.

Analogously, for all $\xi_2 > 0$, there exists a $n_0(\xi_2)$ such that

$$z_{1-E_n/n} \geq (2 - \xi_2)(1 + \rho) \log n$$
3. Estimation method

for all \( n \geq n_0(\xi_2) \). Then we obtain
\[
\beta_n / \beta \to 1.
\]

In Section 3.9 we will discuss how to choose \( E_n \) in practice.

3.8. Correction method in the univariate and finite activity case

In this section we handle the practical issue that we cannot estimate the Lévy measure within the critical region \( B_n \). Our asymptotical results are not affected by this problem. However, in application it leads to biased estimators, because small jumps that lie in \( B_n \) are ignored. This leads to e.g. an underestimated jump intensity. We want to suggest a correction method for this issue. This correction method is based on the assumption that the Lévy measure behaves smooth on the boundary of \( B_n \). We restrict ourselves to the case of finite activity, because this case is interesting for application. Furthermore we only consider the univariate case. By construction, the set \( B_n \) is asymptotically getting small fast if the dimension is high, so the problem affects especially the univariate case.

We introduce the following notation for the upper bound of \( B_n \)
\[
\overline{B}_n := \sup \{ x \in B_n \} \in \left[ \sqrt{\beta'} C b_n, \sqrt{\beta''} b_n \right].
\]
That means \( B_n := [-\overline{B}_n, \overline{B}_n] \) in the univariate case. Of course, in general \( \overline{B}_n \) is stochastic, because \( B_n \) depends on \( \omega \). So all quantities below are defined pathwise.

We introduce an extrapolation technique for the function \( z \mapsto \hat{\nu}(z) \) in \( B_n \). In numerical mathematics such problems are often solved by using splines. Our approach is similar. We fit two polynomials \( p_1 \) and \( p_2 \) of degree 2, the first one in the negative part of \( B_n \) and the second one in the positive part of \( B_n \). The polynomials are defined by
\[
p_1(z) = a_{1,1}(z + \overline{B}_n) + a_{1,2}(z + \overline{B}_n)^2,
p_2(z) = a_{2,1}z + a_{2,2}z^2.
\]
The coefficients of \( p_1 \) and \( p_2 \) are defined in such a way that the following smoothness conditions in the knots \( -\overline{B}_n \) and \( \overline{B}_n \) hold true
\[
p_1'(-\overline{B}_n) = \frac{\hat{\nu}(-\overline{B}_n) - \hat{\nu}(-\overline{B}_n - \delta_n)}{\delta_n} =: \hat{\nu}'(-\overline{B}_n),
\]
\[
p_2''(-\overline{B}_n) = \frac{\hat{\nu}(-\overline{B}_n) - 2\hat{\nu}(-\overline{B}_n - \delta_n) + \hat{\nu}(-\overline{B}_n - 2\delta_n)}{\delta_n^2} =: \hat{\nu}''(-\overline{B}_n),
\]
\[
p_2'(\overline{B}_n) = \frac{\hat{\nu}(\overline{B}_n + \delta_n) - \hat{\nu}(\overline{B}_n)}{\delta_n} =: \hat{\nu}'(\overline{B}_n),
\]
\[
p_2''(\overline{B}_n) = \frac{\hat{\nu}(\overline{B}_n + \delta_n) - 2\hat{\nu}(\overline{B}_n - \delta_n) + \hat{\nu}(\overline{B}_n - 2\delta_n)}{\delta_n^2} =: \hat{\nu}''(\overline{B}_n),
\]
\[
p_1''(\overline{B}_n) = \frac{\hat{\nu}(\overline{B}_n + \delta_n) - 2\hat{\nu}(\overline{B}_n + \delta_n) + \hat{\nu}(\overline{B}_n + 2\delta_n)}{\delta_n^2} =: \hat{\nu}''(\overline{B}_n),
\]
\[
p_2''(\overline{B}_n) = \frac{\hat{\nu}(\overline{B}_n) - 2\hat{\nu}(\overline{B}_n + \delta_n) + \hat{\nu}(\overline{B}_n + 2\delta_n)}{\delta_n^2} =: \hat{\nu}''(\overline{B}_n),
\]
\[
p_1''(\overline{B}_n) = \frac{\hat{\nu}(\overline{B}_n) - 2\hat{\nu}(\overline{B}_n + \delta_n) + \hat{\nu}(\overline{B}_n + 2\delta_n)}{\delta_n^2} =: \hat{\nu}''(\overline{B}_n),
\]
\[
p_2''(\overline{B}_n) = \frac{\hat{\nu}(\overline{B}_n) - 2\hat{\nu}(\overline{B}_n + \delta_n) + \hat{\nu}(\overline{B}_n + 2\delta_n)}{\delta_n^2} =: \hat{\nu}''(\overline{B}_n),
\]
\[
p_1''(\overline{B}_n) = \frac{\hat{\nu}(\overline{B}_n) - 2\hat{\nu}(\overline{B}_n + \delta_n) + \hat{\nu}(\overline{B}_n + 2\delta_n)}{\delta_n^2} =: \hat{\nu}''(\overline{B}_n),
\]
3.8. Correction method in the univariate and finite activity case

\[ p_n''(B_n) = \frac{\nu'(B_n + 2\delta_n) - 2\nu'(B_n + \delta_n) + \nu(B_n)}{\delta_n^2} =: \nu''(B_n), \tag{3.24} \]

where \( \delta_n > 0 \). We will discuss below how to choose \( \delta_n \). The right hand sides can be considered as empirical derivatives. By solving the equation system we obtain the coefficients that are equal to

\[ \begin{align*}
  a_{1,1} &= \nu'(-B_n), \\
  a_{1,2} &= \nu''(-B_n)/2, \\
  a_{2,1} &= \nu'(B_n) - \nu''(B_n) B_n, \\
  a_{2,2} &= \nu''(B_n)/2.
\end{align*} \]

We do not require smoothness in the origin because it can happen that the structures of positive and negative jumps are completely different, e.g. if only negative jumps occur.

By constructing \( p_1 \) and \( p_2 \) this way it may happen that these polynomials decrease in \([-B_n, 0]\) and \([0, B_n]\), respectively. However, the real function \( z \mapsto \nu(z) \) cannot decrease because by definition it is a monotone increasing function. We solve this problem by partially setting \( p_1 \) and \( p_2 \), respectively, to a constant value if otherwise it would decrease. Therefore we define

\[ z_1^* := \arg \max_{z \in [-B_n, 0]} p_1(z) \quad \text{and} \quad z_2^* := \arg \min_{z \in [0, B_n]} p_2(z) \]

and compose the corrected estimator.

**Definition 3.43** (Corrected estimator)
The corrected estimator for the function \( z \mapsto \nu(z) \) is defined as follows

\[ \nu_c(z) := \begin{cases} 
  \nu(z) & \text{if } z < -B_n \\
  \nu(z) + p_1(z) & \text{if } -B_n \leq z < z_1^* \\
  \nu(z) + p_1(z_1^*) & \text{if } z_1^* \leq z \leq z_2^* \\
  \nu(z) + p_1(z_1^*) + p_2(z) - p_2(z_2^*) & \text{if } z_2^* < z \leq B_n \\
  \nu(z) + p_1(z_1^*) + p_2(B_n) - p_2(z_2^*) & \text{if } z > B_n
\end{cases} \]

This estimator is monotone increasing in \( z \). The next proposition gives explicit formulas for \( z_1^* \) and \( z_2^* \).

**Proposition 3.44**
We have

\[ z_1^* := \begin{cases} 
  -\frac{a_{1,1}}{2a_{1,2}} - B_n & \text{if } a_{1,2} \neq 0, \quad -\frac{a_{1,1}}{2a_{1,2}} - B_n \in [-B_n, 0] \\
  0 & \text{else}
\end{cases} \]

\[ z_2^* := \begin{cases} 
  -\frac{a_{2,1}}{2a_{2,2}} & \text{if } a_{2,2} \neq 0, \quad -\frac{a_{2,1}}{2a_{2,2}} \in [0, B_n] \\
  0 & \text{else}
\end{cases} \]

75
3. Estimation method

Proof. The polynomials $p_1$ and $p_2$ are of degree 2. This means that they have one local extremum at most. We start with $z_1^*$. The derivatives of $p_1$ are given by

$$p_1'(z) = a_{1,1} + 2a_{1,2}(z + B_n),$$
$$p_1''(z) = 2a_{1,2}. $$

By construction we always have $a_{1,1} \geq 0$.

If $a_{1,2} = 0$, by construction we have $p_1'(z) = a_{1,1} \geq 0$. So $p_1$ is linear and monotone non decreasing. That means

$$0 = \arg \max_{z \in [B_n,0]} p_1(z).$$

If $a_{1,2} \neq 0$, the extremum point of $p_1$ is at

$$z_{\text{ext}} = -\frac{a_{1,1}}{2a_{1,2}} - B_n.$$ 

If $a_{1,2} > 0$, we obtain $z_{\text{ext}} \leq -B_n$ and in case $a_{1,2} < 0$ we have $z_{\text{ext}} \geq -B_n$. Since $p_1''(z) = 2a_{1,2}$, $z_{\text{ext}}$ is a minimum point if $z_{\text{ext}} < -B_n$ and a maximum point if $z_{\text{ext}} > -B_n$. If $z_{\text{ext}} = -B_n$, we get $a_{1,1} = 0$. From the definition of $a_{1,1}$ follows that

$$\hat{\nu}(-B_n) = \hat{\nu}(-B_n - \delta_n).$$

Because we always have

$$\hat{\nu}(-B_n - \delta_n) \geq \hat{\nu}(-B_n - 2\delta_n),$$

it follows $p_1''(-B_n) = 2a_{1,2} < 0$, so $z_{\text{ext}}$ is also a maximum point in this case. In summary, we obtain, that $z \mapsto p_1(z)$ is monotone increasing in $[-B_n, z_1^*]$ and monotone decreasing in $(z_1^*, 0]$.

Similar arguments can be applied to show that $z \mapsto p_2(z)$ is monotone increasing in $(z_2^*, B_n]$ and monotone decreasing in $[0, z_2^*)$. 

In Theorem 3.45(ii) we show that the asymptotic properties of $\hat{\nu}_c(z)$ remain the same, if the sequence $\delta_n$ tends to zero with rate $n^{(\alpha - 1)/5}$. Due to practical reasons, that are explained below, we will use an additional random factor in the definition of $\delta_n$.

In this context it is interesting to investigate the empirical derivatives defined in (3.21), (3.22), (3.23) and (3.24) themselves. Theorem 3.45(i) shows that they converge to the real derivatives of the function $\nu(z)$. However, Theorem 3.45(ii) can be shown without Theorem 3.45(i).

Theorem 3.45
Let $1/2 < \alpha < 1$ in (3.1) and define

$$\delta_n := d_n n^{(\alpha - 1)/5},$$

where $\{d_n, n \in \mathbb{N}\}$ is an $\mathbb{R}$-valued random sequence satisfying $d_n \xrightarrow{P} d > 0$. 

76
3.8. Correction method in the univariate and finite activity case

(i) If \( \nu''(z) \) and \( \nu'''(z) \) exist and are bounded then, for \( r \in [0, (1 - \alpha)/10) \),

1. \( n^r \left| \hat{\nu}'(-B_n) - \nu'(-B_n) \right| \xrightarrow{P} 0, \)
2. \( n^r \left| \hat{\nu}''(-B_n) - \nu''(-B_n) \right| \xrightarrow{P} 0, \)
3. \( n^r \left| \hat{\nu}'(B_n) - \nu'(B_n) \right| \xrightarrow{P} 0, \)
4. \( n^r \left| \hat{\nu}''(B_n) - \nu''(B_n) \right| \xrightarrow{P} 0. \)

(ii) We have

\[
\sqrt{T} \sup_{z \in \mathbb{R}} |\hat{\nu}_c(z) - \hat{\nu}(z)| \xrightarrow{P} 0.
\]

Proof. We set w.l.o.g \( d_n \equiv 1. \)

Ad (i), 1.), by Theorem 3.25 we have

\[ n^r \delta_n^{-1} \left| \hat{\nu}(-B_n) - \nu(-B_n) \delta_n + \nu(-B_n) \delta_n \right| \leq 2 n^r \delta_n^{-1} \sup_{z \in \mathbb{R}} |\hat{\nu}(z) - \nu(z)| \xrightarrow{P} 0. \]

By Taylor’s theorem there exists a \( \xi_n \in (-B_n - \delta_n, B_n) \) such that

\[ n^r \left| \nu(-B_n) - \nu(-B_n) \delta_n + \nu(-B_n) \delta_n \right| \leq n^r \left| \nu''(\xi_n) \delta_n/2 \right| \xrightarrow{P} 0. \]

Ad (i), 2.), by Theorem 3.25 we have

\[ n^r \delta_n^{-2} \left| \hat{\nu}(-B_n) - 2 \nu(-B_n) - \nu(-B_n) \right| \leq 3 n^r \delta_n^{-2} \sup_{z \in \mathbb{R}} |\hat{\nu}(z) - \nu(z)| \xrightarrow{P} 0. \]

By Taylor’s theorem there exists a \( \xi_n \in (-B_n - \delta_n, B_n) \) and a \( \xi_n' \in (-B_n - 2\delta_n, B_n) \) such that

\[ n^r \left| \nu''(\xi_n) \delta_n/2! - 8 \nu'''(\xi_n') \delta_n/3! \right| \xrightarrow{P} 0. \]

The statements (i), 3.) and (i), 4.) can be shown analogously.

Ad (ii), we have

\[
\sqrt{T} \sup_{z \in \mathbb{R}} |\hat{\nu}_c(z) - \hat{\nu}(z)| \leq \sqrt{T} \left( p_1(z_1^*) + p_2(B_n) - p_2(z_2^*) \right)
\leq \sqrt{T} \left( |a_{1,1}| B_n + |a_{1,2}| B_n^2 + 2|a_{2,1}| B_n + 2|a_{2,2}| B_n^2 \right)
\leq \kappa \sqrt{T} \left( \hat{\nu}'(-B_n) B_n + \hat{\nu}''(-B_n) B_n^2 + \hat{\nu}'(B_n) B_n + \hat{\nu}''(B_n) B_n^2 \right), \quad (3.25)
\]
3. Estimation method

where κ is a positive constant. We start with the term \( \sqrt{T} \hat{\nu}(-B_n)B_n \). Similar to (i) we have

\[
\sqrt{T} B_n \delta_n^{-1} \left| \hat{\nu}(-B_n) - \hat{\nu}(-B_n - \delta_n) - \nu(-B_n) + \nu(-B_n - \delta_n) \right| \to 0.
\]

By Taylor’s theorem there exists a \( \xi_n \in (-B_n - \delta_n, -B_n) \) such that

\[
\sqrt{T} B_n (\nu(-B_n) - \nu(-B_n - \delta_n)) / \delta_n = \sqrt{T} B_n \nu'(\xi_n) \to 0,
\]

where \( \nu'(z) = f(z) \) is assumed to be bounded. Next, we consider the term \( \sqrt{T} B_n^2 \nu'' \) in (3.25). Similar to (i) we have

\[
\sqrt{T} B_n \delta_n^{-2} \left| \hat{\nu}(-B_n) - 2 \hat{\nu}(-B_n - \delta_n) + \hat{\nu}(-B_n - 2 \delta_n) - \nu(-B_n) + 2 \nu(-B_n - \delta_n) + \nu(-B_n - 2 \delta_n) \right| \to 0.
\]

By Taylor’s theorem there exists a \( \xi_n \in (-B_n - 2 \delta_n, -B_n) \) and a \( \xi'_n \in (-B_n - 2 \delta_n, -B_n) \) such that

\[
\sqrt{T} B_n^2 \left( \nu(-B_n) - 2 \nu(-B_n - \delta_n) + \nu(-B_n - 2 \delta_n) \right) / \delta_n^2
\]

\[
= \sqrt{T} B_n^2 (\nu' - \nu'(\xi'_n)) / \delta \to 0.
\]

Similar arguments can be applied for the remaining terms of (3.25).

**Remark 3.46**

If we use the iteration method, the asymptotic result of Theorem 3.45(ii) remains the same. This can be shown analogously to Theorem 3.36.

An important question is, how to choose the sequence \( d_n \) in practice. We use the following definition.

\[
d_n := \left\{ \begin{array}{ll}
\zeta \max \left\{ \frac{\nu(0)}{\nu(\infty)}, \frac{\nu(\infty)}{\nu(0)} \right\} \hat{\nu}(\infty)^{-1/5} & \hat{\nu}(\infty) > 0 \\
1 & \hat{\nu}(\infty) = 0
\end{array} \right.,
\]

where \( \zeta \) is a positive constant and

\[
\nu := \frac{1}{T} \sum_{k=1}^{n} X_{h,k} \mathbb{1}(X_{h,k} < -B_n), \quad \nu := \frac{1}{T} \sum_{k=1}^{n} X_{h,k} \mathbb{1}(X_{h,k} > B_n).
\]

Furthermore we use the convention \( 0/0 \) = 0.

That means

\[
d_n := \left\{ \begin{array}{ll}
\zeta \max \left\{ \frac{\nu(0)}{\nu(\infty)}; \frac{\nu(\infty)}{-\nu(0)} \right\} \left( \hat{\nu}(\infty) T \right)^{-1/5} & \hat{\nu}(\infty) > 0 \\
T^{-1/5} & \hat{\nu}(\infty) = 0
\end{array} \right..
\]

We have chosen the factor \( \hat{\nu}(\infty)^{-1/5} \) in the definition, since the quantity \( \hat{\nu}(\infty) T \) is equal to the observed jumps. The function \( \hat{\nu}(z) \) is a step function. To every step up point there exists an observed jump with an observed jump size that is equal to the step up point. That means the number of step up points of \( \hat{\nu}(z) \) is equal to \( \hat{\nu}(\infty) T \). Because the correction method is a numerical procedure, we
3.9. Simulations and real data application

think it is more natural to let $\delta_n$ depending on the number of step up points, instead of the time horizon.

We have chosen the quantities

$$m_1 := \frac{1}{T} \sum_{k=1}^{n} X_{h,k} \mathbb{1}(X_{h,k} < -B_n),$$

$$m_2 := \frac{1}{T} \sum_{k=1}^{n} X_{h,k} \mathbb{1}(X_{h,k} > B_n),$$

because we want to take the scaling of the jump size distribution into account. If the jumps sizes are larger, we have to use a larger $\delta_n$. Since $X_{h,k} = C X_{h,k} + J X_{h,k}$ we can use Lemma B.3(ii), Lemma B.4(iv), Proposition 2.32 and by the Strong Law of Large Numbers we obtain

$$m_1 \xrightarrow{P} \int_{-\infty}^{0} x \nu(dx) \quad \text{and} \quad m_2 \xrightarrow{P} \int_{0}^{\infty} x \nu(dx).$$

The expectation of the negative and positive jumps sizes, respectively, is given by

$$\frac{\int_{-\infty}^{0} x \nu(dx)}{\nu(0)} \quad \text{and} \quad \frac{\int_{0}^{\infty} x \nu(dx)}{\nu(\infty) - \nu(0)}.$$

This means $d_n$ depends on the maximum of the estimated expectations of the jump size distributions of the positive and negative jumps.

To ensure that $d_n$ is formally well defined, we have to exclude the case $\tilde{\nu}(\infty) = 0$. In this case we just set $d_n = 1$. The probability that this happens tends to zero if $\nu(\mathbb{R}) > 0$. If $\nu(\mathbb{R}) = 0$, the correction method has no effect anyway. Thus, it does not matter how we define $d_n$ in this case.

Next, we have to discuss how to define the constant $\zeta$. Under a practical point of view we have to ensure that for typical practical situations "enough" but not "too many" jumps occur that have a jump size in

$$[-B_n, -B_n - \delta_n], [-B_n, -B_n - 2\delta_n], [B_n, B_n + \delta_n] \quad \text{and} \quad [B_n, B_n + 2\delta_n].$$

We suggest to set $\zeta = 0.5$. That means if we can observe 20 jumps, we have

$$\delta_n = 0.5 \max \left\{ \frac{m_1}{\tilde{\nu}(0)}, \frac{m_2}{\tilde{\nu}(\infty) - \tilde{\nu}(0)} \right\} 20^{-1/5} = 0.27 \max \left\{ \frac{m_1}{\tilde{\nu}(0)}, \frac{m_2}{\tilde{\nu}(\infty) - \tilde{\nu}(0)} \right\}.$$

We think that this is a suitable value.

3.9. Simulations and real data application

In this section we check the finite sample properties of our estimators by a simulation study and apply our method to real historical time series. We restrict ourselves to finite activity processes. In Subsection 3.9.1 we start with the univariate case and perform simulations using the Merton model to generate random variables. Furthermore we use a historical time series of the DAX index. Subsection 3.9.2 deals with the multivariate case; simulations are performed by the bivariate Merton model. We consider exchange rate time series
3. Estimation method

of EUR/USD and EUR/JPY.

3.9.1. In the univariate case

To generate log returns we use the Merton model introduced in Chapter 2, Section 2.1.5. We start with a parameter constellation that is used in A"ıt-Sahalia (2004). Therefore we set

\begin{align*}
C &= 0.09, \\
\gamma[0] &= 0, \\
\beta &= 0, \\
\eta &= 0.6, \\
\lambda &= 5, \\
n &= 1000, \\
h &= \frac{1}{252},
\end{align*}

where \(\beta\) is the expectation and \(\eta\) the standard deviation of the jump size distribution. As usual in Finance, the parameters are scaled to one year, so \(h\) is one day and the overall time horizon \(T\) is about 4 years. For the evaluation 2000 simulation runs are performed. We estimate the following quantities

\[ \hat{C}, \hat{\gamma}[0] \text{ and } \hat{\lambda}_c := \hat{\nu}_c(\infty). \]

The uncorrected estimator is also considered. Instead of \(\hat{\nu}(\infty)\) we use

\[ \hat{\nu}_{z_0} := \hat{\nu}(\infty) - \hat{\nu}(z_0) + \hat{\nu}(-z_0), \quad z_0 := 0.079. \]

It is known that \(z \mapsto \hat{\nu}(z)\) is constant within the critical region, so the experimenter can use this information to consider the estimator at more suitable points. We choose \(z_0\) in such a way that the estimated critical region is contained in \((-z_0, z_0)\) for every simulation run. The true value is

\[ \nu(\infty) - \nu(z_0) + \nu(-z_0) = 4.48. \]

To compute the critical region we have to choose a suitable \(E_n\). If we set \(E_n\) too large, there are probably diffusion increments that lie outside of the critical region. This should be avoided in any case. However, if we choose \(E_n\) too small there are many jumps that lie within the critical region. So we have to find a compromise. We set \(E_n = 0.2\), so in the majority of cases no diffusion increment is erroneously considered as a jump. Furthermore in the correction method we set \(\zeta = 0.5\) as mentioned in Section 3.8.

Table 3.1 gives the result of the 2000 simulation runs. In Figure 3.1 the sample distribution is presented as well as the asymptotic approximation.

We realize that, for all estimators, the sample distributions and the asymptotic approximations fit well together. However, \(\hat{C}\) is a little bit biased. The estimators \(\hat{\gamma}[0], \hat{\nu}(\infty) - \hat{\nu}(z_0) + \hat{\nu}(-z_0)\) and \(\hat{\lambda}_c\), are almost unbiased.

Next we choose another parameter constellation.

\[ C = 0.09, \quad \gamma[0] = 0.035, \quad \beta = 0, \quad \eta = 0.1, \quad \lambda = 10. \]

Like in the scenario above, we set \(E_n = 0.2, \zeta = 0.5\) and \(z_0 = 0.079\). In this constellation the jump sizes are much smaller than in the last scenario, the variance of the Wiener process remains the same. So it is harder to distinguish between jumps and diffusion. Figure 3.2 shows an example of a stock price
3.9. Simulations and real data application

development driven by this parameters. We have set the initial value to 50.

Figure 3.3 shows the corresponding log returns and the estimated critical region. Increments that contain a jump are colored red. For this example, in Figure 3.4 the corrected estimator for the Lévy measure is compared to the real function \( \nu(z) \). We realize that there are a lot of small jumps, so the correction method has a great effect.

We perform 2000 simulation runs. For the uncorrected estimator we consider the same quantity as in the last scenario, which has the true value

\[
\nu(\infty) - \nu(z_0) + \nu(-z_0) = 4.28.
\]

Table 3.2 gives the simulation result and in Figure 3.5 the sample distribution is compared to the asymptotic approximation.

We realize that the bias and the variance of \( \hat{C} \) are greater than the values of the last scenario. A possible explanation therefor is that there are a lot of jumps within the critical region.

The sample distribution of \( \hat{\gamma}[0] \) and the asymptotic approximation fit well together. However, there is a small bias.

The estimator \( \hat{\nu}(\infty) - \hat{\nu}(z_0) + \hat{\nu}(-z_0) \) is almost unbiased and the sample distribution match well with the asymptotic approximation.

The bias of \( \hat{\lambda}_c \) is small. However, its variance is large. So the correction method works much better in the scenario above, where jumps are larger.

We have also investigated the number of iteration runs that are necessary to estimate the critical region. The values range form 2 to 7 and the mean is equal to 3.7.

Finally, we apply the estimation method to real data of the DAX index. We use daily closing prices in the time frame 28.07.2006 to 27.07.2011, so \( T = 5 \). The sample size is equal to \( n = 1270 \). We choose \( E_n = 0.2 \) and in the correction method we set \( \zeta = 0.5 \). The following results are obtained

\[
\hat{C} = 0.0379, \quad \hat{\gamma} = 0.0845, \quad \hat{\lambda} = 5.4, \quad \hat{\lambda}_c = 22.6.
\]

That means, for the diffusion part, we obtain the following economic features: The annual volatility is equal to \( \sqrt{\hat{C}} = 19.5\% \) and the drift rate is given by \( \hat{\gamma} + \frac{1}{2}\hat{C} = 10.3\% \). For the jump part, \( \hat{\lambda}_c = 22.6 \) means that the estimated number of jumps that occur p.a. is equal to 22.6. This result is in conformity with typical values that occur in Finance (see also Cont [2001]).
3. Estimation method

<table>
<thead>
<tr>
<th>$n = 1000$</th>
<th>$C$</th>
<th>$\hat{\gamma}_0$</th>
<th>$\hat{\nu}_0$</th>
<th>$\hat{\nu}(\infty)_c$</th>
</tr>
</thead>
<tbody>
<tr>
<td>true</td>
<td>0.0900</td>
<td>0.0000</td>
<td>4.48</td>
<td>5.00</td>
</tr>
<tr>
<td>Mean</td>
<td>0.0887</td>
<td>0.0000</td>
<td>4.42</td>
<td>5.03</td>
</tr>
<tr>
<td>Stdev</td>
<td>0.0043</td>
<td>0.1480</td>
<td>1.06</td>
<td>1.23</td>
</tr>
<tr>
<td>MSE</td>
<td>0.0000</td>
<td>0.0219</td>
<td>1.12</td>
<td>1.51</td>
</tr>
<tr>
<td>Min</td>
<td>0.0752</td>
<td>-0.5189</td>
<td>1.51</td>
<td>1.58</td>
</tr>
<tr>
<td>0.25-q.</td>
<td>0.0857</td>
<td>-0.0964</td>
<td>3.78</td>
<td>4.17</td>
</tr>
<tr>
<td>Median</td>
<td>0.0886</td>
<td>0.0027</td>
<td>4.28</td>
<td>4.97</td>
</tr>
<tr>
<td>0.75-q.</td>
<td>0.0914</td>
<td>0.0982</td>
<td>5.04</td>
<td>5.85</td>
</tr>
<tr>
<td>Max</td>
<td>0.1045</td>
<td>0.5715</td>
<td>8.57</td>
<td>9.91</td>
</tr>
</tbody>
</table>

Table 3.1.: Simulation result using the Merton model, parameters: $C = 0.09$, $\gamma[0] = 0$, $\beta = 0$, $\eta = 0.6$ and $\lambda = 5$.

Figure 3.1.: Simulation result using the Merton model, sample distribution and asymptotic approximation, parameters: $C = 0.09$, $\gamma[0] = 0$, $\beta = 0$, $\eta = 0.6$ and $\lambda = 5$. 
Figure 3.2.: Example of a stock price development driven by the Merton model, parameters: $C = 0.09$, $\gamma[0] = 0.035$, $\beta = 0$, $\eta = 0.1$ and $\lambda = 10$.

Figure 3.3.: Example of the log returns computed by the Merton model and the estimated threshold, parameters: $C = 0.09$, $\gamma[0] = 0.035$, $\beta = 0$, $\eta = 0.1$ and $\lambda = 10$. 
3. Estimation method

Figure 3.4.: Example of $\nu(z)$ and $\hat{\nu}(z)$ for the Merton model, parameters $C = 0.09$, $\gamma[0] = 0.035$, $\beta = 0$, $\eta = 0.1$ and $\lambda = 10$. 
3.9. Simulations and real data application

\[
\begin{array}{|c|c|c|c|c|}
\hline
\alpha = 1000 & C & \hat{\gamma}[0] & \hat{\nu}_2 & \lambda_c \\
\hline
\text{true} & 0.0900 & 0.0350 & 4.28 & 10.00 \\
\text{Mean} & 0.0948 & 0.0284 & 4.31 & 9.72 \\
\text{Stdev} & 0.0054 & 0.1504 & 1.04 & 3.38 \\
\text{MSE} & 0.0001 & 0.0227 & 1.08 & 11.51 \\
\text{Min} & 0.0786 & -0.5912 & 1.26 & 1.76 \\
0.25-q. & 0.0911 & -0.0747 & 3.53 & 7.32 \\
\text{Median} & 0.0945 & 0.0286 & 4.28 & 9.36 \\
0.75-q. & 0.0981 & 0.1276 & 5.04 & 11.68 \\
\text{Max} & 0.1146 & 0.5376 & 8.82 & 28.78 \\
\hline
\end{array}
\]

Table 3.2.: Simulation result using the Merton model, sample distribution and asymptotic approximation, parameters \( C = 0.09, \gamma[0] = 0.035, \beta = 0, \eta = 0.1 \) and \( \lambda = 10 \).

Figure 3.5.: Simulation result using the Merton model, sample distribution and asymptotic approximation, parameters \( C = 0.09, \gamma[0] = 0.035, \beta = 0, \eta = 0.1 \) and \( \lambda = 10 \).
3. Estimation method

3.9.2. In the multivariate case

We use the bivariate Merton model to generate bi-dimensional log returns. The setting is similar to the univariate case. We use \( n = 1000 \) and \( h = 1/252 \) and set the parameters of the diffusion process to

\[
C := \begin{pmatrix} 0.09 & 0.03 \\ 0.03 & 0.04 \end{pmatrix} \quad \text{and} \quad \gamma[0] := \begin{pmatrix} 0.035 \\ 0.025 \end{pmatrix}.
\]

Furthermore we set the covariance matrix \( \eta \) and the drift \( \beta \) of the bi-dimensional jump size distribution to

\[
\eta := \begin{pmatrix} 0.01 & -0.005 \\ -0.005 & 0.01 \end{pmatrix} \quad \text{and} \quad \beta := \begin{pmatrix} 0 \\ 0 \end{pmatrix}.
\]

The jump intensity is given by \( \lambda := 10 \).

That means that the parameter constellation of the first component of the Lévy process coincides with that of the second scenario of the previous subsection. To estimate the critical region we use \( E_n = 0.2 \). In Figure 3.6 an example of the bi-dimensional log returns driven by this parameter constellation as well as the critical regions that are computed in the single steps of the iteration method are given. The log returns that contain a jump are colored red.

We give estimators for \( C, \gamma[0], \lambda \) and for the term

\[
qcv := \int_{\mathbb{R}^2} x_1 x_2 \nu(dx) = \mu_{1,A}, \quad \mathbf{A} = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}.
\]

By Proposition 2.34 this quantity is equal to the expectation of the quadratic covariation of the jump part of the Lévy process in a time interval \([0,1]\). The true value in this scenario is 0.05.

We perform 2000 runs and give the result in Table 3.3.

We realize that the estimation of \( C_{1,1} \) and \( \gamma[0]_1 \) is much better than in Table 3.2, the values of the empirical bias and variance are much smaller.

The bias of \( \hat{C}_{2,2} \) is also very small and the empirical variance is almost as good as the asymptotic variance that is computed by Theorem 3.22.

The result for \( \hat{\gamma}[0]_2 \) is similar to \( \hat{\gamma}[0]_1 \), the bias is very small, but the empirical variance is greater than the asymptotic variance.

The estimator of \( C_{1,2} \) has a small bias. However, the empirical variance is much greater than the asymptotic variance. This result is obvious, since the jumps are negatively correlated, so jumps within the critical set disturb the estimation a lot.

The estimator of \( \lambda \) is as expected biased. However, the empirical variance is almost as good as the asymptotic variance.

The bias for \( qcv \) is very small and the empirical variance is a little bit greater.
3.9. Simulations and real data application

than the asymptotic variance.

Like in the last subsection, we apply the estimation method to real data. Therefore we consider a time series of daily data of the exchange rates of EUR/USD and EUR/JPY. The time frame is 28.07.2006 to 27.07.2011, so $T = 5$. The sample size is equal to $n = 1302$. We set $E_n = 0.2$ and obtain the following estimation result

$$
\hat{C} := \begin{pmatrix}
0.00964 & 0.00783 \\
0.00783 & 0.01670
\end{pmatrix}, \quad \gamma[0] := \begin{pmatrix}
0.0129 \\
-0.0396
\end{pmatrix}, \quad \lambda = 4, \quad \hat{qcv} = 0.00186.
$$

We realize that the values of $\hat{C}_{1,1}$, $\hat{C}_{2,2}$ and $\hat{\lambda}$ are smaller than the corresponding values of the estimators for the DAX index of the last subsection.
3. Estimation method

Figure 3.6.: Log returns and the critical regions that show up by the different steps of the iteration method

<table>
<thead>
<tr>
<th>n = 1000</th>
<th>$C_{1,1}$</th>
<th>$C_{2,2}$</th>
<th>$C_{1,2}$</th>
<th>$\hat{\gamma}[0]_1$</th>
<th>$\hat{\gamma}[0]_2$</th>
<th>$\lambda$</th>
<th>$\hat{q}_{cv}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>True</td>
<td>0.0900</td>
<td>0.0400</td>
<td>0.0300</td>
<td>0.0350</td>
<td>0.0250</td>
<td>10.00</td>
<td>-0.0500</td>
</tr>
<tr>
<td>Mean</td>
<td>0.0884</td>
<td>0.0394</td>
<td>0.0294</td>
<td>0.0329</td>
<td>0.0238</td>
<td>8.19</td>
<td>-0.0491</td>
</tr>
<tr>
<td>As. stdev</td>
<td>0.0040</td>
<td>0.0018</td>
<td>0.0013</td>
<td>0.1342</td>
<td>0.0894</td>
<td>1.41</td>
<td>0.0173</td>
</tr>
<tr>
<td>Stdev</td>
<td>0.0042</td>
<td>0.0019</td>
<td>0.0022</td>
<td>0.1476</td>
<td>0.0998</td>
<td>1.45</td>
<td>0.0203</td>
</tr>
<tr>
<td>MSE</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0218</td>
<td>0.0100</td>
<td>5.38</td>
<td>0.0004</td>
</tr>
<tr>
<td>Min</td>
<td>0.0757</td>
<td>0.0336</td>
<td>0.0227</td>
<td>-0.4427</td>
<td>-0.3107</td>
<td>4.28</td>
<td>-0.1376</td>
</tr>
<tr>
<td>0.25-q.</td>
<td>0.0854</td>
<td>0.0380</td>
<td>0.0279</td>
<td>-0.0660</td>
<td>-0.0428</td>
<td>7.06</td>
<td>-0.0617</td>
</tr>
<tr>
<td>Median</td>
<td>0.0883</td>
<td>0.0394</td>
<td>0.0293</td>
<td>0.0338</td>
<td>0.0240</td>
<td>8.06</td>
<td>-0.0477</td>
</tr>
<tr>
<td>0.75-q.</td>
<td>0.0912</td>
<td>0.0407</td>
<td>0.0309</td>
<td>0.1327</td>
<td>0.0884</td>
<td>9.07</td>
<td>-0.0344</td>
</tr>
<tr>
<td>Max</td>
<td>0.1045</td>
<td>0.0463</td>
<td>0.0386</td>
<td>0.5000</td>
<td>0.3623</td>
<td>13.61</td>
<td>0.0057</td>
</tr>
</tbody>
</table>

Table 3.3.: Simulation result using the bivariate Merton model
3.9. Simulations and real data application

3.9.3. Discussion

The basic requirement for the application of our method in practice is, that a sharp distinction between jumps and diffusion is existent. So an ideal situation is, if jumps are large and appear rarely. It has been our purpose to establish an estimation method that can deal with such situations, because they appear very often in Finance. By using a suitable critical region our approach can overcome the ill-posedness of such a problem.

Under a practical point of view, the estimation of the covariance matrix of the diffusion part is satisfying. The bias and the variances are very small.

The sample distribution of the estimator of the drift and the asymptotic approximation match very well. However, the variance is very large, this is a problem for many applications. This is a general problem in Finance, e.g. in the Black Scholes model we have exactly the same large variance. Of course we cannot obtain a better result than in the case where no jumps are existent.

The estimation of Lévy measure is satisfying as long as we stay away from zero. In the univariate case we have introduced a correction method for the interval within the critical region. This method depends on the assumption that the real Lévy measure behaves smooth on the boundary of the critical region. The correction method improves the estimation result depending on the question to which extent this assumption is fulfilled.

If we are in a higher dimension, it is less problematic to estimate the intensity without the correction method, because the bias is smaller than in the univariate case. The reason is that the volume of the multivariate critical region vanished asymptotically faster than in the multivariate case. Furthermore the critical region forms an ellipsoid which has a volume reducing effect, too. The MSE of the other estimators is also a little bit smaller than in the multivariate case.
4. Estimation in the Kou model

In this chapter we use our estimation method to derive estimators for the parameters of the Kou model. In Section 4.1 we introduce the estimation setting and define the estimators. Section 4.2 deals with pre-estimators that are based on the assumption, that we can observe the process continuously. The estimators that are defined by the critical region are considered in Section 4.3. In Section 4.4 a correction method is presented and in Section 4.5 simulations are performed.

4.1. Estimation setting

We have introduced and discussed the Kou model in Section 2.1.5. It is assumed that the logarithmic stock price follows a Jump diffusion process with asymmetric double exponential distributed jumps, i.e. the jump density is of the form

\[ f(x) := p \eta_1 \exp(-\eta_1 x) \mathbb{1}(x \geq 0) + (1-p) \eta_2 \exp(\eta_2 x) \mathbb{1}(x < 0), \]

where \( \eta_1 > 0, \eta_2 > 0 \) and \( p \in (0,1) \). The Lévy measure is given by

\[ \nu(dx) := \lambda f(x) dx, \]

where \( \lambda \) is the jump intensity. For given Lévy measure the parameters \( p, \eta_1 \) and \( \eta_2 \) can be computed by

\[ p = \frac{\nu(\infty) - \nu(0)}{\nu(\infty)}, \quad \eta_1 = \frac{\nu(\infty) - \nu(0)}{\int_0^\infty x \nu(dx)}, \quad \eta_2 = \frac{\nu(0)}{\int_{-\infty}^0 |x| \nu(dx)} \]

and have the following interpretation. The expectation of the distribution of the positive jumps is equal to \( 1/\eta_1 \) and the expectation of the distribution of the negative jumps is equal to \( 1/\eta_2 \). The parameter \( p \) controls the ratio of positive and negative jumps.

The asymptotic framework of this section coincides with Section 3.1, in particular we have high frequency data and the covariance matrix of the diffusion part is not singular.

By Chapter 3 we can estimate the diffusion parameters \( C \) and \( \gamma[0] \) and the Lévy measure \( \nu((-\infty, z]), z \in \mathbb{R} \). However, instead of estimating \( \nu \), it is more interesting to estimate directly the parameters \( p, \eta_1 \) and \( \eta_2 \). So in this chapter we present estimators for these parameters using the critical region.

As in Section 3.3 we start with pre-estimators that are based on the assumption
4. Estimation in the Kou model

that the exact jump times and sizes are known.

\textbf{Definition 4.1 (Pre-estimators)}

The pre-estimators are defined as follows

\[ \hat{p}_{\text{pre}} := \frac{\hat{\nu}_{\text{pre}}(\infty) - \hat{\nu}_{\text{pre}}(0)}{\hat{\nu}(\infty)}, \]

\[ \hat{\eta}_{1,\text{pre}} := \frac{\hat{\nu}_{\text{pre}}(\infty) - \hat{\nu}_{\text{pre}}(0)}{\frac{1}{T} \sum_{0 \leq t \leq T} \Delta L_t \mathbb{I}(\Delta L_t > 0)}, \]

\[ \hat{\eta}_{2,\text{pre}} := \frac{\hat{\nu}_{\text{pre}}(0)}{\frac{1}{T} \sum_{0 \leq t \leq T} \Delta L_t \mathbb{I}(\Delta L_t < 0)}. \]

Theorem 4.4 and 4.5 show that these estimators are asymptotically normally
distributed. Afterwards we define the estimators that are based on the critical
region. As in Section 3.8 we set

\[ \bar{B}_n := \sup \{ B_n \in [\sqrt{\beta} \sqrt{C} b_n, \sqrt{\beta}'' b_n] \}, \]

because we are in an univariate framework.

\textbf{Definition 4.2 (Estimators)}

The estimators are defined as follows

\[ \hat{p} := \frac{\hat{\nu}(\infty) - \hat{\nu}(0)}{\hat{\nu}(\infty)}, \]

\[ \hat{\eta}_1 := \frac{\hat{\nu}(\infty) - \hat{\nu}(0)}{\frac{1}{T} \sum_{k=1}^{n} X_{h,k} \mathbb{I}(X_{h,k} > \bar{B}_n)}, \]

\[ \hat{\eta}_2 := \frac{\hat{\nu}(0)}{\frac{1}{T} \sum_{k=1}^{n} X_{h,k} \mathbb{I}(X_{h,k} < -\bar{B}_n)}. \]

Theorem 4.6 and 4.7 prove that the difference between the estimator and the
pre-estimators converges to zero with rate \( \sqrt{T} \), so the asymptotic properties
remain the same.

The estimators of \( C \) and \( \gamma[0] \) coincide with those of the last chapter and are
not mentioned explicitly. Since \( \lambda = \nu(\mathbb{R}) \), the estimator is given by \( \hat{\lambda} := \hat{\nu}(\infty) \)
and we can apply Theorem 3.25 directly.

In Section 3.8 we have discussed the bias that occurs for a finite sample size. We
have introduced a correction method using extrapolation techniques. In the Kou
model we have exactly the same problem. The intensity \( \lambda \) is probably underesti-
mated and \( \eta_1 \) and \( \eta_2 \) that control the jump sizes are probably overestimated. In
Section 4.4 we introduce a correction method. This method is totally different
from the method that we have used in Section 3.8 Since small and large jumps
4.2. Pre-estimators

are controlled by the same parameters, we can derive the behavior of the small jumps from the behavior of the large jumps. We choose the following approach. We fix the critical region $B_n = B$ and let $n$ tend to infinity. Then, of course, the estimators are also asymptotically biased. The trick is to define a correction method in such a way that the estimators are even consistent, if the critical region is fixed. We extend the correction method to $\hat{C}$ and $\hat{\gamma}[0]$ the same way, because they are also biased by the small jumps that are within the critical region.

**Definition 4.3 (Correction method)**

The corrected estimators are defined as follows.

\[
\hat{\eta}_{1,c} := \left( \frac{1}{\hat{\eta}_1} - \bar{B}_n \right)^{-1},
\]

\[
\hat{\eta}_{2,c} := \left( \frac{1}{\hat{\eta}_2} - \bar{B}_n \right)^{-1},
\]

\[
\hat{p}_c := \left( \frac{1}{\hat{\eta} - \hat{C}_n} \exp((\hat{\eta}_{2,c} - \hat{\eta}_{1,c}) \bar{B}_n) + 1 \right)^{-1},
\]

\[
\hat{\lambda}_c := \frac{\hat{\lambda}}{\hat{p}_c \exp(-\hat{\eta}_{1,c} \bar{B}_n) + (1 - \hat{p}_c) \exp(-\hat{\eta}_{2,c} \bar{B}_n)},
\]

\[
\hat{\gamma}[0]_c := \hat{\gamma}[0] - \hat{\lambda}_c \left( \hat{p}_c \left( \frac{1}{\hat{\eta}_{1,c}} - \frac{e^{-\hat{\eta}_{1,c} \bar{B}_n}}{\bar{B}_n} \right) - (1 - \hat{p}_c) \left( \frac{1}{\hat{\eta}_{2,c}} - \frac{e^{-\hat{\eta}_{2,c} \bar{B}_n}}{\bar{B}_n} \right) \right),
\]

\[
\hat{C}_c := \hat{C} - \hat{\lambda}_c \left( \frac{2 - (\hat{\eta}_{1,c})^2 \bar{B}_n^2 e^{-\hat{\eta}_{1,c} \bar{B}_n} - 2 \hat{\eta}_{1,c} \bar{B}_n e^{-\hat{\eta}_{1,c} \bar{B}_n} - e^{-\hat{\eta}_{1,c} \bar{B}_n}}{(\hat{\eta}_{1,c})^2} \right)
\]

\[
- \frac{(1 - \hat{p}_c) \hat{\lambda}_c \left( 2 - (\hat{\eta}_{2,c})^2 \bar{B}_n^2 e^{-\hat{\eta}_{2,c} \bar{B}_n} - 2 \hat{\eta}_{2,c} \bar{B}_n e^{-\hat{\eta}_{2,c} \bar{B}_n} - e^{-\hat{\eta}_{2,c} \bar{B}_n}}{(\hat{\eta}_{2,c})^2} \right). \]

For all estimators defined in this chapter, the asymptotic properties remain the same, if we use the iteration method. This result can be shown analogously to Proposition 3.36 because the structure is the same.

**4.2. Pre-estimators**

In this section we start with the pre-estimators of $p$, $\eta_1$ and $\eta_2$ and in Theorem 4.4 and 4.5 we show that they are asymptotically normally distributed.
4. Estimation in the Kou model

**Theorem 4.4**
We have, as $T \to \infty$,
\[
\sqrt{T} (\hat{p}_{pre} - p) \xrightarrow{d} Z \sim N \left( 0, \frac{p(1-p)}{\lambda} \right)
\]

**Proof.** We have
\[
\sqrt{T} \left( \frac{\hat{\nu}_{pre}(\infty) - \hat{\nu}_{pre}(0) - p \hat{\nu}_{pre}(\infty)}{\hat{\nu}_{pre}(\infty)} \right) = \frac{1}{\hat{\nu}_{pre}(\infty)} \sqrt{T} (\hat{\nu}_{pre}(\infty) - \hat{\nu}_{pre}(0) - p \hat{\nu}_{pre}(\infty)).
\]

By Theorem 3.25 we obtain
\[
\hat{\nu}_{pre}(\infty) \xrightarrow{P} \lambda
\]
and by the definition of $\hat{\nu}_{pre}$ we have
\[
\hat{\nu}_{pre}(\infty) - \hat{\nu}_{pre}(0) - p \hat{\nu}_{pre}(\infty) = (1 - p) \frac{1}{T} N_T(0, \infty) - p \frac{1}{T} N_T(-\infty, 0)
\]
\[
= \frac{1}{T} \sum_{k=1}^{T} ((1 - p) N_{k-1,k}(0, \infty) - p N_{k-1,k}(-\infty, 0)),
\]
where $N_T$ is given in Definition 3.18. We want to apply the Central Limit Theorem to this term, so we compute
\[
\mathbb{E} [(1 - p) N_1(0, \infty) - p N_1(-\infty, 0)] = (1 - p) \nu ((0, \infty]) - p \nu ((-\infty, 0]) = 0
\]
and because $N_1(0, \infty)$ and $N_1(-\infty, 0)$ are independent we have
\[
\text{Var} [(1 - p) N_1(0, \infty) - p N_1(-\infty, 0)] = (1 - p)^2 \nu ((0, \infty]) + p^2 \nu ((-\infty, 0])
\]
\[
= \lambda ((1 - p)^2 p + p^2 (1 - p)).
\]

All together by the Central Limit Theorem and by Slutsky’s Theorem we have that
\[
\sqrt{T} (\hat{p} - p) \xrightarrow{d} Z \sim N \left( 0, \frac{p(1-p)}{\lambda} \right).
\]

**Theorem 4.5**
We have, as $T \to \infty$,
\[
\begin{align*}
(i) & \quad \sqrt{T} (\hat{\eta}_{1,pre} - \eta_1) \xrightarrow{d} Z \sim N \left( 0, \frac{\eta_1^2}{\lambda p} \right), \\
(ii) & \quad \sqrt{T} (\hat{\eta}_{2,pre} - \eta_2) \xrightarrow{d} Z \sim N \left( 0, \frac{\eta_2^2}{\lambda(1-p)} \right).
\end{align*}
\]
4.2. Pre-estimators

**Proof.** We just consider w.l.o.g. $\eta_1$. We have

\[
\sqrt{T} \left( \frac{\hat{\nu}_{pre}(\infty) - \hat{\nu}_{pre}(0)}{\frac{1}{T} \sum_{0<t\leq T} \Delta L_t \mathbb{1}(\Delta L_t > 0) - \eta_1} \right) = \frac{1}{\frac{1}{T} \sum_{0<t\leq T} \Delta L_t \mathbb{1}(\Delta L_t > 0)}.
\]

(4.1)

\[
\cdot \sqrt{T} \left( \frac{1}{T} N_T(0, \infty) - \eta_1 \frac{1}{T} \sum_{0 < t \leq T} \Delta L_t \mathbb{1}(\Delta L_t > 0) \right).
\]

(4.2)

Ad (4.1), by Proposition 2.32 we have

\[
\mathbb{E} \left[ \sum_{0 \leq t \leq 1} \Delta L_t \mathbb{1}(\Delta L_t > 0) \right] = \int_0^\infty x \nu(dx),
\]

\[
\text{Var} \left[ \sum_{0 \leq t \leq 1} \Delta L_t \mathbb{1}(\Delta L_t > 0) \right] = \int_0^\infty x^2 \nu(dx).
\]

So we obtain by the Strong Law of Large Numbers

\[
\frac{1}{T} \sum_{0 \leq t \leq T} \Delta L_t \mathbb{1}(\Delta L_t > 0) \xrightarrow{a.s.} \int_0^\infty x \nu(dx) = \frac{p \lambda}{\eta_1}.
\]

Ad (4.2) we have

\[
\frac{1}{T} N_T(0, \infty) - \eta_1 \frac{1}{T} \sum_{0 < t \leq T} \Delta L_t \mathbb{1}(\Delta L_t > 0)
\]

\[
= \frac{1}{T} \sum_{k=1}^T \left( N_1(0, \infty) - \eta_1 \sum_{0 < t \leq 1} \Delta L_t \mathbb{1}(\Delta L_t > 0) \right).
\]

The expectation of this term is equal to

\[
\mathbb{E} \left[ N_1(0, \infty) - \eta_1 \sum_{0 < t \leq 1} \Delta L_t \mathbb{1}(\Delta L_t > 0) \right] = \nu((0, \infty)) - \nu((0, \infty)) = 0
\]

and for the variance we have

\[
\text{Var} \left[ N_1(0, \infty) - \eta_1 \sum_{0 < t \leq 1} \Delta L_t \mathbb{1}(\Delta L_t > 0) \right]
\]

\[
= \text{Var} [N_1(0, \infty)] + \text{Var} \left[ \eta_1 \sum_{0 < t \leq 1} \Delta L_t \mathbb{1}(\Delta L_t > 0) \right]
\]

\[
+ 2 \text{Cov} \left[ N_1(0, \infty), \eta_1 \sum_{0 < t \leq 1} \Delta L_t \mathbb{1}(\Delta L_t > 0) \right]
\]

\[
= \nu((0, \infty)) + 2 \nu((0, \infty))
\]
4. Estimation in the Kou model

\(-2E \left[ N_1(0, \infty) \cdot \eta_1 \sum_{0<t\leq 1} \Delta L_t \mathbb{1}(\Delta L_t > 0) \right] + 2\nu((0, \infty))^2 \)

\(-2E \left[ \mathbb{E} \left[ N_1(0, \infty) \cdot \eta_1 \sum_{0<t\leq 1} \Delta L_t \mathbb{1}(\Delta L_t > 0) \middle| N_1(0, \infty) \right] \right] \)

\(-2E \left[ N_1(0, \infty) \cdot N_1(0, \infty) \frac{1}{\eta_1} \right] \)

\(\nu((0, \infty)) + 2\nu((0, \infty)) + 2\nu((0, \infty))^2 \)

where we have used the representation given in Definition 2.7 to compute the conditional expectation. Then the Central Limit Theorem can be applied.

4.3. Estimation method

In this section we consider the estimators for \(p, \eta_1\) and \(\eta_2\) that are based on the critical region. Theorem 4.6 and 4.7 show that the difference between these estimators and the pre-estimators converge in probability to zero with rate \(\sqrt{T}\). That means that the estimators are also asymptotically normally distributed with the same variances. The proofs are very short because we can directly use the results that are provided in the last chapter.

**Theorem 4.6**

Let \(1/2 < \alpha < 1\) in (3.1). Then, as \(n \to \infty\),

\(\sqrt{T}(\hat{p} - \hat{p}_{\text{pre}}) \xrightarrow{P} 0\) and \(\sqrt{T}(\hat{p} - p) \xrightarrow{d} Z \sim N\left(0, \frac{p(1-p)}{\lambda}\right)\).

**Proof.** By Theorem 3.25 we have

\(\sqrt{T} \left( \frac{\hat{\nu}(\infty) - \nu(0)}{\hat{\nu}(\infty)} - \frac{\nu_{\text{pre}}(\infty) - \nu_{\text{pre}}(0)}{\nu_{\text{pre}}(\infty)} \right) \xrightarrow{P} 0.\)

**Theorem 4.7**

Let \(1/2 < \alpha < 1\) in (3.1). Then, as \(n \to \infty\),

(i) \(\sqrt{T}(\hat{\eta}_1 - \hat{\eta}_{1,\text{pre}}) \xrightarrow{P} 0\) and \(\sqrt{T}(\hat{\eta}_1 - \eta_1) \xrightarrow{d} Z \sim N\left(0, \frac{\eta_1^2}{\lambda p}\right)\).

(ii) \(\sqrt{T}(\hat{\eta}_2 - \hat{\eta}_{2,\text{pre}}) \xrightarrow{P} 0\) and \(\sqrt{T}(\hat{\eta}_2 - \eta_2) \xrightarrow{d} Z \sim N\left(0, \frac{\eta_2^2}{\lambda (1-p)}\right)\).
4.4. Correction method

Proof. We consider w.l.o.g. \( \eta_1 \) and have by Lemma B.3(ii), B.4(iv) and Theorem 3.25

\[
\sqrt{T} \left( \frac{\hat{\nu}(\infty) - \hat{\nu}(0)}{\frac{1}{T} \sum_{k=1}^{n} X_{h,k} \mathbb{I}(X_{h,k} > B_n)} - \frac{\hat{\nu}_{\text{pre}}(\infty) - \hat{\nu}_{\text{pre}}(0)}{\frac{1}{T} \sum_{0 < t \leq T} \Delta L_t \mathbb{I}(\Delta L_t > 0)} \right) \overset{P}{\to} 0.
\]

\[\square\]

4.4. Correction method

This Section deals with the correction method. The theorems are arranged as follows. In the first part we fix the critical region \( B_n \equiv B \) in the definition of the estimator and give the limit of this quantity. Then, we correct the estimator in such a way that it is even consistent if the critical region is fixed. In the second part of the theorems we show that the corrected estimators still converge with the desired rates if we use the normal definition of the critical region \( B_n \). That means the correction method does not matter asymptotically.

Theorem 4.8

Let \( 1/2 < \alpha < 1 \) in (3.1).

(i) Let \( B > 0 \). Then, as \( n \to \infty \),

\[
\hat{\eta}_{1,B} := \left( \frac{\hat{\nu}(\infty) - \hat{\nu}(B)}{\frac{1}{T} \sum_{k=1}^{n} X_{h,k} \mathbb{I}(X_{h,k} > B)} \right) \overset{P}{\to} \left( \frac{1}{\eta_1 + B} \right)^{-1}
\]

and

\[
\hat{\eta}_{1,c,B} := \left( \frac{1}{\hat{\eta}_{1,B}} - B \right)^{-1} \overset{P}{\to} \eta_1.
\]

(ii) We have, as \( n \to \infty \),

\[
\sqrt{T} |\hat{\eta}_{1,c} - \hat{\eta}_1| \overset{P}{\to} 0.
\]

Proof. Ad (i), by Lemma B.3(i) and Lemma B.4(iv) we obtain

\[
\frac{1}{T} \sum_{k=1}^{n} X_{h,k} \mathbb{I}(X_{h,k} > B) - \frac{1}{T} \sum_{0 < t \leq T} \Delta L_t \mathbb{I}(\Delta L_t > B) \overset{P}{\to} 0.
\]

By Proposition 2.32 we have

\[
E \left[ \sum_{0 < t \leq 1} \Delta L_t \mathbb{I}(\Delta L_t > B) \right] = \int_B^{\infty} x \nu(dx)
\]

and

\[
\text{Var} \left[ \sum_{0 < t \leq 1} \Delta L_t \mathbb{I}(\Delta L_t > B) \right] = \int_B^{\infty} x^2 \nu(dx).
\]
4. Estimation in the Kou model

So we can apply Strong Law of Large Numbers and obtain
\[
\frac{1}{T} \sum_{0 < t \leq T} \Delta L_t \mathbb{1}(\Delta L_t > B) \xrightarrow{a.s.} \int_{\mathbb{B}} x \nu(dx).
\]

Finally, by Theorem 3.25 we have
\[
\hat{\nu}(\infty) - \hat{\nu}(B) \xrightarrow{P} \nu((\mathbb{B}, \infty))
\]
and thus,
\[
\frac{\hat{\nu}(\infty) - \hat{\nu}(B)}{\frac{1}{T} \sum_{k=1}^{n} X_{h,k} \mathbb{1}(X_{h,k} > B)} \xrightarrow{P} \left( \frac{1}{\eta_1} + B \right)^{-1}.
\]

The second statement follows directly.

Ad (ii), by \( B_n \in [\sqrt{2\beta^C n^{-\alpha} \log n}, \sqrt{2\beta^C n^{-\alpha} \log n}] \) we have
\[
\sqrt{T} |\hat{\eta}_{1,c} - \hat{\eta}_1| = \sqrt{T} \left| \frac{1}{1/\hat{\eta}_1 - B_n} - \hat{\eta}_1 \right| = \sqrt{T} \left| \frac{B_n}{1/\hat{\eta}_1 - B_n} \right| \xrightarrow{P} 0.
\]

\[\square\]

Theorem 4.9
Let \( 1/2 < \alpha < 1 \) in (3.1).

(i) Let \( B > 0 \). Then, as \( n \to \infty \),
\[
\hat{\eta}_{2,B} := \frac{\hat{\nu}(B) - \hat{\nu}(\infty)}{\frac{1}{T} \sum_{k=1}^{n} X_{h,k} \mathbb{1}(X_{h,k} < -B)} \xrightarrow{P} \left( \frac{1}{\eta_1} + B \right)^{-1}
\]
and
\[
\hat{\eta}_{2,c,B} := \left( \frac{1}{\hat{\eta}_{2,B}} - B \right)^{-1} \xrightarrow{P} \eta_2.
\]

(ii) We have, as \( n \to \infty \),
\[
\sqrt{T} |\hat{\eta}_{2,c} - \hat{\eta}_2| \xrightarrow{P} 0.
\]

Proof. The proof is obtained analogously to Theorem 4.6 \[\square\]

Theorem 4.10
Let \( 1/2 < \alpha < 1 \) in (3.1).

(i) Let \( B > 0 \). Then, as \( n \to \infty \),
\[
\hat{p}_{B} := \frac{\hat{\nu}(\infty) - \hat{\nu}(B) + \hat{\nu}(-B)}{\hat{\nu}(\infty) - \hat{\nu}(B) + \hat{\nu}(-B)} \xrightarrow{P} \left( 1 + \left( \frac{1}{p} - 1 \right) \exp((\eta_1 - \eta_2) B) \right)^{-1}
\]
and
\[
\hat{p}_{c,B} := \left( \frac{1}{p_{B}} - 1 \right) \exp((\hat{\eta}_{2,c,B} - \hat{\eta}_{1,c,B}) B) + 1 \xrightarrow{P} p.
\]
4.4. Correction method

(ii) We have, as \( n \to \infty \),
\[
\sqrt{T} \left| \hat{p}_c - \hat{p} \right| \xrightarrow{P} 0.
\]

Proof. Ad (i), by Theorem 3.25 we have
\[
\hat{p}_B \xrightarrow{P} \frac{\nu((B, \infty))}{\nu(\mathbb{R} \setminus (-B, B))} = \left( 1 + \left(\frac{1}{p} - 1\right) \exp((\eta_1 - \eta_2)B) \right)^{-1}
\]
The second statement follows directly.

Ad (ii), by the power series representation of the exponential function and \( \sqrt{T} B_n \to 0 \) we have
\[
\sqrt{T} \left| \hat{\lambda}_c - \hat{\lambda} \right| \leq \sqrt{T} \left| \left( 1 - \hat{p} \right) \left( 1 - \exp((\hat{\eta}_2, c - \hat{\eta}_1, c)B) \right) \right| \xrightarrow{P} 0.
\]

Theorem 4.11
Let \( 1/2 < \alpha < 1 \) in (3.1)

(i) Let \( B > 0 \), then, as \( n \to \infty \)
\[
\hat{\lambda}_B := \hat{\nu}(\infty) - \hat{\nu}(B) + \hat{\nu}(-B)
\]
and
\[
\hat{\lambda}_{c,B} := \frac{\hat{\lambda}_B}{\hat{p}_{c,B} \exp(-\hat{\eta}_{1,c,B}B) + (1 - \hat{p}_{c,B}) \exp(-\hat{\eta}_{2,c,B}B)} \xrightarrow{P} \lambda.
\]

(ii) We have, as \( n \to \infty \)
\[
\sqrt{T} \left| \hat{\lambda}_c - \hat{\lambda} \right| \xrightarrow{P} 0.
\]

Proof. Ad (i), by Theorem 3.25 we have
\[
\hat{\lambda}_B \xrightarrow{P} \nu(\mathbb{R} \setminus (-B, B)) = \lambda \left( p \exp(-\eta_1 B) + (1 - p) \exp(-\eta_2 B) \right).
\]
The second statement follows directly.

Ad (ii), by the power series representation of the exponential function and \( \sqrt{T} B_n \to 0 \) we have
\[
\sqrt{T} \left| \hat{\lambda}_c - \hat{\lambda} \right| \leq \sqrt{T} \left| \hat{\lambda} \left( 1 - \left( \hat{p}_c \exp(-\hat{\eta}_{1,c}B_n) + (1 - \hat{p}_c) \exp(-\hat{\eta}_{2,c}B_n) \right) \right) \right| \xrightarrow{P} 0.
\]
4. Estimation in the Kou model

**Theorem 4.12**

Let $1/2 < \alpha < 1$ in (3.1).

(i) Let $B_n > 0$, then, as $n \to \infty$

\[
\hat{\gamma}([0],B) := \frac{1}{T} \sum_{k=1}^{n} X_{h,k} \mathbb{1}(|X_{h,k}| \leq B)
\]

\[
P \to \gamma[0] + \lambda \left( \frac{1}{\eta_1} - \frac{e^{-\eta_1B}}{\eta_1} - e^{-\eta_1B} \right)
\]

and

\[
\hat{\gamma}([0],c,B) := \hat{\gamma}([0],B) - \lambda \left( \frac{1}{\eta_{1,c,B}} e^{-\eta_{1,c,B}B} + \frac{1}{\eta_{2,c,B}} e^{-\eta_{2,c,B}B} \right)
\]

\[
P \to \gamma[0].
\]

(ii) We have, as $n \to \infty$,

\[
\sqrt{T} |\hat{\gamma}([0],c) - \hat{\gamma}([0])| \xrightarrow{P} 0.
\]

**Proof.** At first, we have

\[
\sum_{0 < t < 1} \Delta L_t \mathbb{1}(|\Delta L_t| < B) = \sum_{0 \leq t \leq 1} \Delta L_t \mathbb{1}(\Delta L_t \in [0, B]) - \sum_{0 \leq t \leq 1} (-\Delta L_t) \mathbb{1}(\Delta L_t \in [-B, 0]).
\]

So we can apply Proposition 2.32 and obtain

\[
\mathbb{E} \left[ \sum_{0 < t \leq 1} \Delta L_t \mathbb{1}(|\Delta L_t| < B) \right] = \int_{-B}^{B} x \nu(dx)
\]

and

\[
\text{Var} \left[ \sum_{0 < t \leq 1} \Delta L_t \mathbb{1}(|\Delta L_t| < B) \right] = \int_{-B}^{B} x^2 \nu(dx).
\]

By the Strong Law of Large Numbers we have

\[
\frac{1}{T} \sum_{0 < t \leq 1} \Delta L_t \mathbb{1}(|\Delta L_t| < B) \xrightarrow{a.s.} \int_{-B}^{B} x \nu(dx).
\]

At second, we have

\[
\sum_{k=1}^{n} X_{h,k} \mathbb{1}(|X_{h,k}| \leq B)
\]
4.4. Correction method

\[= \sum_{k=1}^{n} X_{h,k} \mathbb{I}(|X_{h,k}| \leq B_n) + \sum_{k=1}^{n} X_{h,k} \mathbb{I}(|X_{h,k}| \leq B, |X_{h,k}| > B_n)\]

By Theorem 3.23 we have for the first term

\[\sum_{k=1}^{n} X_{h,k} \mathbb{I}(|X_{h,k}| \leq B_n) \xrightarrow{p} \gamma[0]\]

and for the second term by Lemma B.3(ii) and B.4(iv) we have

\[\sum_{k=1}^{n} X_{h,k} \mathbb{I}(|X_{h,k}| \leq B, |X_{h,k}| > B_n) - \sum_{0 \leq t \leq T} \Delta L_t \mathbb{I}(|\Delta L_t| < B) \xrightarrow{p} 0.\]

Finally we have by the definition of the parameters

\[\int_{-B}^{B} x \nu(dx) = \lambda \left( p \left( \frac{1}{\eta_1} - e^{-\eta_1 B} - e^{-\eta_1 B} \right) - (1 - p) \left( \frac{1}{\eta_2} - e^{-\eta_2 B} - e^{-\eta_2 B} \right) \right).\]

The second statement follows directly.

Ad (ii), by the power series representation of the exponential function and \[\sqrt{T} B_n \to 0\] we have

\[\sqrt{T} |\hat{\gamma}[0] - \hat{\gamma}[0]| \leq \sqrt{T} \left| \hat{\lambda}_c \left( \hat{\rho}_c \left( \frac{1}{\eta_1} - \frac{e^{-\eta_1 B_n}}{\hat{\eta}_1} - e^{-\eta_1 B_n} \right) \right) - (1 - \hat{\rho}_c) \left( \frac{1}{\eta_2} - \frac{e^{-\eta_2 B_n}}{\hat{\eta}_2} - e^{-\eta_2 B_n} \right) \right| \xrightarrow{p} 0.\]

\[\square\]

**Theorem 4.13**

Let \(1/2 < \alpha < 1\) in (3.1)

(i) Let \(B > 0\), then as \(n \to \infty\)

\[\hat{C}_B := \frac{1}{T} \sum_{k=1}^{n} X_{h,k}^2 \mathbb{I}(|X_{h,k}| \leq B) \xrightarrow{p} C + p \lambda \left( \frac{2 - \eta_1^2 B^2 e^{-\eta_1 B} - 2 \eta_1 B e^{-\eta_1 B} - 2 e^{-\eta_1 B}}{\eta_1^2} \right) + (1 - p) \lambda \left( \frac{2 - \eta_2^2 B^2 e^{-\eta_2 B} - 2 \eta_2 B e^{-\eta_2 B} - 2 e^{-\eta_2 B}}{\eta_2^2} \right)\]

101
4. Estimation in the Kou model

and

\[
\tilde{C}_{c,B} = \hat{C}_{c,B} - \frac{\hat{p}_{c,B} \hat{\lambda}_{c,B}}{(\hat{\eta}_{1,c,B})^2} \left( 2 - (\hat{\eta}_{1,c,B})^2 B e^{-\hat{\eta}_{1,c,B} B} - 2 \hat{\eta}_{1,c,B} B e^{-\hat{\eta}_{1,c,B} B} - 2 e^{-\hat{\eta}_{1,c,B} B} \right) \]

\[
- \frac{(1 - \hat{p}_{c,B}) \hat{\lambda}_{c,B}}{(\hat{\eta}_{2,c,B})^2} \left( 2 - (\hat{\eta}_{2,c,B})^2 B e^{-\hat{\eta}_{2,c,B} B} - 2 \hat{\eta}_{2,c,B} B e^{-\hat{\eta}_{2,c,B} B} - 2 e^{-\hat{\eta}_{2,c,B} B} \right).
\]

(ii) We have, as \( n \to \infty \),

\[
\sqrt{n} \left| \hat{C}_c - \hat{C} \right| \to P 0.
\]

**Proof.** Ad (i), By Theorem \[3.22\] and \[3.26\] we have

\[
\hat{C}_{c,B} = \frac{1}{T} \sum_{k=1}^{n} X_{h,k}^2 1(|X| \leq B_n) + \frac{1}{T} \sum_{k=1}^{n} X_{h,k}^2 1(|X| \leq B, |X| > B_n)
\]

\[\to P C + \int_{-B}^{B} x^2 \nu(dx)\]

and

\[
C + \int_{-B}^{B} x^2 \nu(dx)
\]

\[= C + p\lambda \left( \frac{2 - \eta_1^2 B^2 e^{-\eta_1 B} - 2 \eta_1 B e^{-\eta_1 B} - 2 e^{-\eta_1 B}}{\eta_1^2} \right)
\]

\[+ \left( 1 - p \right) \lambda \left( \frac{2 - \eta_2^2 B^2 e^{-\eta_2 B} - 2 \eta_2 B e^{-\eta_2 B} - 2 e^{-\eta_2 B}}{\eta_2^2} \right).
\]

The second statement follows directly. Ad (ii), by the power series representation of the exponential function and \( \sqrt{n} B_n^2 \to 0 \) we have

\[
\sqrt{n} \left| \hat{C}_c - \hat{C} \right|
\]

\[\leq \sqrt{n} \left| \hat{p}_{c,B} \hat{\lambda}_{c} \left( 2 - (\hat{\eta}_{1,c,B})^2 B_n^2 e^{-\hat{\eta}_{1,c,B} B_n} - 2 \hat{\eta}_{1,c,B} B_n e^{-\hat{\eta}_{1,c,B} B_n} - 2 e^{-\hat{\eta}_{1,c,B} B_n} \right) \hat{\eta}_{1,c,B} \right|
\]

\[+ \left( 1 - \hat{p}_{c,B} \right) \hat{\lambda}_{c} \left( 2 - (\hat{\eta}_{2,c,B})^2 B_n^2 e^{-\hat{\eta}_{2,c,B} B_n} - 2 \hat{\eta}_{2,c,B} B_n e^{-\hat{\eta}_{2,c,B} B_n} - 2 e^{-\hat{\eta}_{2,c,B} B_n} \right) \hat{\eta}_{2,c,B} \right| \]

\[\to P 0.
\]
4.5. Simulations and real data application

In this section we check the finite sample properties of our estimators and the correction method by a simulation study. We choose the following parameter constellation

\[ C = 0.09, \quad \gamma[0] = 0.035, \quad \lambda = 20, \quad p = 0.4, \quad \eta_1 = 12, \quad \eta_2 = 11. \]

Furthermore we consider daily data \((h = 1/252)\) and a sample size of \(n = 1000\). Figure 4.1 shows an example of a stock price development driven by this parameters. We have set the initial value to \(S_0 = 50\). The corresponding log returns and the estimated critical region are presented in Figure 4.1. The increments that contain a jump are colored red. As in Section 3.9 we have set \(E_n = 0.2\) to compute the critical region.

It is not possible to use Maximum Likelihood procedures for this estimation problem, because the density of the increments is not given in a closed form. It only can expressed as an improper convolution integral (cf. Riesner, 2006).

We perform 2000 simulation runs to check our estimation method. The result is shown in Table 4.1 and 4.2. In Figure 4.3 and 4.4 the sample distribution is presented as well as the asymptotic approximation.

We realize that uncorrected estimators for \(\lambda, \eta_1\) and \(\eta_2\) are very biased. Because the jumps are very small in this scenario, this result is not surprising. The correction method can reduce the bias a lot. However, the value of the sample variance of the corrected estimators is much larger. Nevertheless we obtain a

Figure 4.1.: Example of a stock price development driven by the Kou model, Parameters: \(C = 0.09, \gamma[0] = 0.035, \lambda = 20, p = 0.4, \eta_1 = 12, \eta_2 = 11.\)
4. Estimation in the Kou model

significant reduction of the MSE. Furthermore we realize that there are outliers, so the correction method seem not to be very robust.

The uncorrected estimator of $p$ is a little bit biased and the sample variance is greater than the variance of the asymptotic approximation. However, the bias is not reduced by the correction method, both estimators have comparable values. The sample variance of $\hat{p}_c$ is even larger.

The sample variance of $\hat{C}$ is greater than the variance of the asymptotic approximation and the estimation is a little bit biased. The corrected estimator is also biased and the sample variance is even larger. Whereas the mean of $\hat{C}$ is smaller than $C$, the mean of $\hat{C}_c$ is greater than $C$.

For $\gamma[0]$ the corrected estimator has a smaller bias than $\hat{\gamma}[0]$. However, its sample variance is larger and the MSE is larger, too.

In summary this evaluation shows that the correction method improves the estimation of $\lambda$, $\eta_1$ and $\eta_2$ significantly. For the other parameters we cannot empirically verify an improvement.

In [Gegler (2007)] and [Riesner (2006)] simulations are performed for the Kou-model, too. However, only two special cases are considered. In the first case it is assumed that $\eta_1$ and $\eta_2$ are equal and previously known. In the second case $\eta_1$ and $\eta_2$ are not previously known, but it is also required that $\eta_1 = \eta_2$.

Finally, we apply the method to real data. We choose the time series of the DAX index, that we have used in Section 3.9.1 The time frame is 28.07.2006 to 27.07.2011, so $T = 5$ and the sample size is equal to $n = 1270$. Furthermore we choose $E_n = 0.2$. We obtain the following results for the uncorrected estimators

$$\hat{C} = 0.0379, \quad \hat{\gamma}[0] = 0.0845, \quad \hat{\lambda} = 5.4, \quad \hat{p} = 0.407, \quad \hat{\eta}_1 = 14.43, \quad \hat{\eta}_2 = 17.11.$$  

Of course, the results of $\hat{C}$, $\hat{\gamma}$ and $\hat{\lambda}$ coincide with those of Section 3.9.1 because the estimators are the same.

For the corrected estimator we have

$$\hat{C}_c = 0.00244, \quad \hat{\gamma}[0]_c = 1.354, \quad \hat{\lambda}_c = 152.72, \quad \hat{p}_c = 0.106, \quad \hat{\eta}_1,c = 43.22, \quad \hat{\eta}_2,c = 81.33.$$  

The correction method has a big effect and we suspect that the values are too extreme. A possible explanation is that the method is not very robust. The corrected estimator for $\lambda$ in Section 3.9.1 is equal to 22.6.
Figure 4.2.: Example of the log returns computed by the Kou model and the estimated threshold, parameters: \( C = 0.09, \gamma[0] = 0.035, \lambda = 20, \rho = 0.4, \eta_1 = 12, \eta_2 = 11.\)
4. Estimation in the Kou model

<table>
<thead>
<tr>
<th>$n = 1000$</th>
<th>$\hat{C}$</th>
<th>$\hat{C}_c$</th>
<th>$\hat{\gamma}[0]$</th>
<th>$\hat{\gamma}[0]_c$</th>
<th>$\lambda$</th>
<th>$\lambda_c$</th>
</tr>
</thead>
<tbody>
<tr>
<td>true</td>
<td>0.0900</td>
<td>0.0900</td>
<td>0.0350</td>
<td>0.0350</td>
<td>20.00</td>
<td>20.00</td>
</tr>
<tr>
<td>Mean</td>
<td>0.0996</td>
<td>0.0821</td>
<td>-0.0272</td>
<td>0.0241</td>
<td>8.53</td>
<td>21.01</td>
</tr>
<tr>
<td>Stdev</td>
<td>0.0064</td>
<td>0.0074</td>
<td>0.1569</td>
<td>0.1950</td>
<td>1.48</td>
<td>5.16</td>
</tr>
<tr>
<td>MSE</td>
<td>0.0001</td>
<td>0.0001</td>
<td>0.0285</td>
<td>0.0382</td>
<td>133.63</td>
<td>27.65</td>
</tr>
<tr>
<td>Min</td>
<td>0.0797</td>
<td>0.0498</td>
<td>-0.5537</td>
<td>-0.7336</td>
<td>4.03</td>
<td>7.55</td>
</tr>
<tr>
<td>0.25-q.</td>
<td>0.0949</td>
<td>0.0773</td>
<td>-0.1282</td>
<td>-0.1006</td>
<td>7.56</td>
<td>17.49</td>
</tr>
<tr>
<td>Median</td>
<td>0.0993</td>
<td>0.0822</td>
<td>0.0287</td>
<td>8.57</td>
<td>20.43</td>
<td></td>
</tr>
<tr>
<td>0.75-q.</td>
<td>0.1038</td>
<td>0.0870</td>
<td>0.0751</td>
<td>9.58</td>
<td>23.82</td>
<td></td>
</tr>
<tr>
<td>Max</td>
<td>0.1272</td>
<td>0.1092</td>
<td>0.4884</td>
<td>14.62</td>
<td>54.69</td>
<td></td>
</tr>
</tbody>
</table>

Table 4.1.: Simulation result using the Kou model, parameters: $C$, $\gamma[0]$, $\lambda$.

<table>
<thead>
<tr>
<th>$n = 1000$</th>
<th>$\hat{p}$</th>
<th>$\hat{p}_c$</th>
<th>$\hat{\eta}_1$</th>
<th>$\hat{\eta}_1,c$</th>
<th>$\hat{\eta}_2$</th>
<th>$\hat{\eta}_2,c$</th>
</tr>
</thead>
<tbody>
<tr>
<td>true</td>
<td>0.4000</td>
<td>0.4000</td>
<td>12.00</td>
<td>12.00</td>
<td>11.00</td>
<td>11.00</td>
</tr>
<tr>
<td>Mean</td>
<td>0.3844</td>
<td>0.4149</td>
<td>6.45</td>
<td>12.86</td>
<td>6.05</td>
<td>11.18</td>
</tr>
<tr>
<td>Stdev</td>
<td>0.0826</td>
<td>0.1118</td>
<td>0.97</td>
<td>4.07</td>
<td>0.74</td>
<td>2.59</td>
</tr>
<tr>
<td>MSE</td>
<td>0.0071</td>
<td>0.0127</td>
<td>31.74</td>
<td>17.30</td>
<td>25.07</td>
<td>6.72</td>
</tr>
<tr>
<td>Min</td>
<td>0.0800</td>
<td>0.1298</td>
<td>3.60</td>
<td>4.83</td>
<td>3.50</td>
<td>4.66</td>
</tr>
<tr>
<td>0.25-q.</td>
<td>0.3250</td>
<td>0.3346</td>
<td>5.77</td>
<td>10.09</td>
<td>5.52</td>
<td>9.34</td>
</tr>
<tr>
<td>Median</td>
<td>0.3824</td>
<td>0.4101</td>
<td>6.42</td>
<td>12.13</td>
<td>6.02</td>
<td>10.85</td>
</tr>
<tr>
<td>0.75-q.</td>
<td>0.4413</td>
<td>0.4900</td>
<td>7.06</td>
<td>14.80</td>
<td>6.52</td>
<td>12.63</td>
</tr>
<tr>
<td>Max</td>
<td>0.6585</td>
<td>0.8258</td>
<td>10.57</td>
<td>44.44</td>
<td>8.89</td>
<td>26.46</td>
</tr>
</tbody>
</table>

Table 4.2.: Simulation result using the Kou model, parameters: $p$, $\eta_1$, $\eta_2$. 

106
4.5. Simulations and real data application

Figure 4.3.: Simulation result using the Kou model, sample distribution and asymptotic approximation, parameters: $C = 0.09$, $\gamma[0] = 0.035$ and $\lambda = 20$. 
4. Estimation in the Kou model

Figure 4.4.: Simulation result using the Kou model, sample distribution and asymptotic approximation, parameters: $p = 0.4$, $\eta_1 = 12$ and $\eta_2 = 11$. 
A. Some inequalities

Lemma A.1
Let \( r, s \in \mathbb{N}, n \in \mathbb{N}, k \in \{1, \cdots, n\} \) and \( i \in \{1, \cdots, n\} \). Then, there exists \( \kappa(r, s) > 0 \) such that, for all
\[
 a_{k,i,n} \in \mathbb{R}^+, \quad b_{k,i,n} \in \mathbb{R}^+, \quad c_{k,n} \in \mathbb{R}^+, \quad d_n \in \mathbb{R}^+.
\]
the following inequalities hold true.

\[(i) \quad \sum_{k=1}^{n} \left( c_{k,n} \left( \sum_{i_1,i_2=1}^{r} a_{k,i_1,n} b_{k,i_2,n} \right) \right)^s \leq \kappa(r, s) \sum_{i_1,i_2=1}^{r} c_{k,n} a_{k,i_1,n}^s b_{k,i_2,n}^s, \]
\[(ii) \quad \sum_{k=1}^{n} \left( c_{k,n} \left( \sum_{i_1,i_2=1}^{r} a_{k,i_1,n} a_{k,i_2,n} \right) \right)^s \leq \kappa(r, s) \sum_{i_1,i_2=1}^{r} c_{k,n} a_{k,i_1,n}^{2s}, \]
\[(iii) \quad \sum_{k=1}^{n} c_{k,n} \left( \sum_{i=1}^{r} a_{k,i,n} \right)^s - d_n \leq \sum_{k=1}^{n} c_{k,n} a_{k,1,n}^{s} - d_n + \kappa(r, s) \sum_{i=2}^{r} \sum_{m=1}^{s} c_{k,n} a_{k,1,n}^{2s-m} a_{k,i,n}^{m}. \]

Proof. Ad (i), because \( x \mapsto x^s \) is a convex function on \( \mathbb{R}^+ \), we have
\[
\left( \frac{\sum_{i_1,i_2=1}^{r} a_{k,i_1,n} b_{k,i_2,n}}{r^2} \right)^s \leq \frac{\sum_{i_1,i_2=1}^{r} a_{k,i_1,n}^s b_{k,i_2,n}^s}{r^{2s}}.
\]
Hence,
\[
\sum_{k=1}^{n} \left( c_{k,n} \left( \sum_{i_1,i_2=1}^{r} a_{k,i_1,n} b_{k,i_2,n} \right) \right)^s \leq \kappa(r, s) \sum_{i_1,i_2=1}^{r} c_{k,n} a_{k,i_1,n}^s b_{k,i_2,n}^s.
\]

Ad (ii), we have
\[
\sum_{k=1}^{n} \left( c_{k,n} \left( \sum_{i_1,i_2=1}^{r} a_{k,i_1,n} a_{k,i_2,n} \right) \right)^s \leq \sum_{k=1}^{n} \left( c_{k,n} r^s \max_{i=1,\cdots,r} \{ a_{k,i,n}^{2s} \} \right) ^s \leq r^s \sum_{k=1}^{n} c_{k,n} \sum_{i=1}^{r} a_{k,i,n}^{2s} = \kappa(r, s) \sum_{i=1}^{r} \sum_{k=1}^{n} c_{k,n} a_{k,i,n}^{2s}.
\]
A. Some inequalities

Ad (iii), we have

\[ \left| \sum_{k=1}^{n} c_{k,n} \left( \sum_{i=1}^{r} a_{k,i,n} \right)^{s} - d_{n} \right| \]

\[ = \left| \sum_{k=1}^{n} c_{k,n} \sum_{m=0}^{s} \binom{s}{m} a_{k,1,n}^{s-m} \left( \sum_{i=2}^{r} a_{k,i,n} \right)^{m} - d_{n} \right| \]

\[ \leq \left| \sum_{k=1}^{n} c_{k,n} a_{k,1,n}^{s} - d_{n} \right| + \left| \sum_{k=1}^{n} c_{k,n} \sum_{m=1}^{s} \binom{s}{m} a_{k,1,n}^{s-m} \left( \sum_{i=2}^{r} a_{k,i,n} \right)^{m} \right| \]

\[ \leq \left| \sum_{k=1}^{n} c_{k,n} a_{k,1,n}^{s} - d_{n} \right| + \kappa(r,s) \sum_{k=1}^{n} c_{k,n} \sum_{m=1}^{s} a_{k,1,n}^{s-m} (r-1)^{m} \max_{i=2,...,r} \{ a_{k,i,n}^{m} \} \]

\[ \leq \left| \sum_{k=1}^{n} c_{k,n} a_{k,1,n}^{s} - d_{n} \right| + \kappa(r,s) \sum_{k=1}^{n} c_{k,n} \sum_{m=1}^{s} a_{k,1,n}^{s-m} \sum_{i=2}^{r} a_{k,i,n}^{m} \]

\[ = \left| \sum_{k=1}^{n} c_{k,n} a_{k,1,n}^{s} - d_{n} \right| + \kappa(r,s) \sum_{i=2}^{r} \sum_{m=1}^{s} \sum_{k=1}^{n} c_{k,n} a_{k,1,n}^{s-m} a_{k,i,n}^{m} . \]

\[ \square \]

Lemma A.2
Let \( A \in \mathbb{R}^{d \times d} \), \( A = (a_{i,j})_{i,j=1}^{d} \) and \( x, y \in \mathbb{R}^{d} \). Then,

\[ x^{T} Ay \leq d! \max_{i,j=1,...,d} |a_{i,j}| |x| |y|. \]

Proof. We denote by \( \Pi \) the set of all permutations of \( \{1, \cdots, d\} \). Define \( \pi_{0} \in \Pi \) such that the following inequality is satisfied

\[ \sum_{i=1}^{d} |x_{i} y_{\pi(i)}| \leq \sum_{i=1}^{d} |x_{i} y_{\pi_{0}(i)}| \]

for all \( \pi \in \Pi \). Then we use Cauchy-Schwarz inequality and obtain

\[ x^{T} Ay = \sum_{i=1}^{d} x_{i} a_{i,j} y_{j} \leq \max_{i,j=1,...,d} |a_{i,j}| \sum_{i,j=1}^{d} |x_{i} y_{j}| = \max_{i,j=1,...,d} |a_{i,j}| \sum_{i \in \{1,...,d\}, \pi \in \Pi} |x_{i} y_{\pi(i)}| \]

\[ \leq d! \max_{i,j=1,...,d} |a_{i,j}| \sum_{i=1}^{d} |x_{i} y_{\pi_{0}(i)}| \leq \kappa |x| |y|. \]

\[ \square \]

Lemma A.3
Let \( \{ L_{t}, t \geq 0 \} \) be an \( \mathbb{R} \)-valued Lévy process and \( \{ C_{t} L_{t}', t \geq 0 \} \) an \( \mathbb{R} \)-valued
nonstandard Wiener process. Then, as $h \to 0$,

(i) $E[|L_h|] = \mathcal{O}(\sqrt{h})$,

(ii) $E[L_h^j] = h c_n(1) + \mathcal{O}(h^2) = \mathcal{O}(h)$, \hspace{1em} $j \in \mathbb{N}$, $j$ even,

(iii) $E[|L_h|^j] = \mathcal{O}(h)$, \hspace{1em} $j \in \mathbb{N}$, $j \geq 3$, $j$ odd,

(iv) $E[C L_h^j] = \mathcal{O}(h^{j/2})$, \hspace{1em} $j \in \mathbb{N}$, $j$ even,

where we have assumed that $E[L_1^{2(j+1)/2}] < \infty$ and have denoted by $c_j(1)$ the $j$-th cumulant of $L_1$.

**Proof.** The connection between cumulants and moments is given by (cf. Lukacs, 1970, p.27)

$$E[L_h^j] = c_j(h) + \sum_{k=1}^{n-1} \binom{n-1}{k-1} c_k(h) E[L_h^{j-k}].$$

So (ii) follows directly from Proposition 2.30. For (i) the Lyapunov inequality can be applied and for (iii) we have

$$E[|L_h|^j] \leq E[L_h^{j-1}] + E[L_h^{j+1}].$$

The cumulants $c_j(h)$ of a Wiener process are equal to zero for $j \geq 3$, so (iv) follows immediately. \hfill \square
B. Technical details of Section 3.4

Lemma B.1
Let $\alpha \in (1/2, 1]$ in (3.1) and $\kappa_1, \kappa_2 > 0$. Define

$$M_3^{(n)}(\chi, i) := \{\exists z \in \mathbb{R} : N_{T,i}(z - \kappa_1 b_n, z + \kappa_1 b_n) \geq \kappa_2 n^\chi\},$$

where $N_{T,i}$ is defined in Definition 3.18 and $b_n = \sqrt{2n \log n}$ (Definition 3.5).

Then, for all $\chi > 0 \lor 1 - 3/2\alpha$ and for all $i \in \{1, \ldots, d\}$ we have, as $n \to \infty$

$$\mathbb{P}(M_3^{(n)}(\chi, i)) \to 0.$$

Proof. We set w.l.o.g. $\kappa_1 = \kappa_2 = 1$ and define

$$A_m^{(n)} := \{N_{T,i}(2(m - 1) b_n, 2(m + 1) b_n) \geq n^\chi\}.$$

Obviously, we have

$$M_3^{(n)} \subset \bigcup_{m \in \mathbb{Z}} A_m^{(n)}.$$

We assume w.l.o.g $n^\chi \in \mathbb{N}$ (otherwise we consider $\lfloor n^\chi \rfloor$) and by the Poisson distribution we have

$$\mathbb{P}(A_m^{(n)}) = \sum_{k=n^\chi}^{\infty} \exp\left(-\tilde{\nu}_i((2(m - 1) b_n, 2(m + 1) b_n)) n^{1-\alpha}\right) \cdot$$

$$\frac{(\tilde{\nu}_i((2(m - 1) b_n, 2(m + 1) b_n)) n^{1-\alpha})^k}{k!} \leq \frac{(\tilde{\nu}_i((2(m - 1) b_n, 2(m + 1) b_n)) n^{1-\alpha})^{n^\chi}}{n^\chi!},$$

since by Tailor’s theorem there exists

$$\zeta_n \in (0, \tilde{\nu}_i((2(m - 1) b_n, 2(m + 1) b_n)) n^{1-\alpha})$$

such that

$$\sum_{k=n^\chi}^{\infty} \frac{(\tilde{\nu}_i((2(m - 1) b_n, 2(m + 1) b_n)) n^{1-\alpha})^k}{k!}$$

$$= \exp(\zeta_n) \frac{(\tilde{\nu}_i((2(m - 1) b_n, 2(m + 1) b_n)) n^{1-\alpha})^{n^\chi}}{n^\chi!} \leq \exp(\tilde{\nu}_i((2(m - 1) b_n, 2(m + 1) b_n)) n^{1-\alpha}) \cdot$$

$$\frac{(\tilde{\nu}_i((2(m - 1) b_n, 2(m + 1) b_n)) n^{1-\alpha})^{n^\chi}}{n^\chi!}.$$
B. Technical details of Section 3.4

Because the jump density \( f_i \) is bounded (Assumption 3.7), there exists a \( \kappa > 0 \) such that, for all \( m \in \mathbb{Z} \),

\[
\tilde{\nu}_i \left( (2(m-1)b_n, 2(m+1)b_n) \right) n^{1-\alpha} = n^{1-\alpha} \int_{2(m-1)b_n}^{2(m+1)b_n} f_i(x) \, dx \leq n^{1-\alpha} \max_{x \in [2(m-1)b_n, 2(m+1)b_n]} f_i(x) \cdot 4 b_n \leq \kappa n^{1-\alpha} b_n.
\]

Then,

\[
P \left( M_{3}^{(n)} \right) \leq \sum_{m \in \mathbb{Z}} P(A_m^{(n)}) \leq \sum_{m \in \mathbb{Z}} \left( \tilde{\nu}_i \left( (2(m-1)b_n, 2(m+1)b_n) \right) n^{1-\alpha} \right)^{\chi} \frac{n^{\chi}}{n!} \cdot \tilde{\nu}_i \left( (2(m-1)b_n, 2(m+1)b_n) \right) n^{1-\alpha} \leq \sum_{m \in \mathbb{Z}} \left( \kappa n^{1-\alpha} b_n \right)^{\chi} \frac{n^{\chi-1}}{n!} \tilde{\nu}_i \left( (2(m-1)b_n, 2(m+1)b_n) \right) n^{1-\alpha} \leq \left( \kappa n^{1-\alpha} b_n \right)^{\chi} \frac{n^{\chi-1}}{n!} n^{1-2} \tilde{\nu}(\mathbb{R}).
\]

This term tends to zero by Stirling’s formula.

\[\square\]

Lemma B.2

(i) Let \( 1/2 < \alpha \leq 1 \) in \([3.1]\) and \( i_1, i_2 \in \{1, \ldots, d\} \), then, as \( n \to \infty \),

\[
\sqrt{n} \frac{1}{T} \sum_{k=1}^{n} \left| ^C X_{h,k,i} \right| ^{CP} X_{h,k,i} \left| ^C X_{h,k,i} \right| \mathbb{1}(X_{h,k} \in B_n) \xrightarrow{P} 0.
\]

(ii) Let \( 1/2 < \alpha \leq 1 \) in \([3.1]\), \( i_1, i_2, i_3 \in \{1, \ldots, d\} \), \( r \in [0, \alpha/2) \) and \( j_1, j_2, j_3 \in \mathbb{N} \) satisfying

\[
j_1 + j_2 = 1, j_3 \geq 2 \quad \text{or} \quad j_1 + j_2 \geq 2, j_3 \geq 1.
\]

If \( \mathbb{E} \left[ L_{i,i}^{2j_1} \right] < \infty \) and \( \mathbb{E} \left[ L_{i,i}^{2j_2} \right] < \infty \), then, as \( n \to \infty \),

\[
n^{\alpha/2} \frac{1}{T} \sum_{k=1}^{n} \left| ^{CP} X_{h,k,i_1} \right| ^{j_1} \left| ^{CP} X_{h,k,i_2} \right| ^{j_2} \left| ^{C} X_{h,k,i_3} \right| ^{j_3} \xrightarrow{P} 0.
\]

Proof. Ad (i), we have

\[
n^{\frac{1}{2}-(1-\alpha)} \sum_{k=1}^{n} \left| ^{C} X_{h,k,i_1} \right| \left| ^{CP} X_{h,k,i_2} \right| \mathbb{1}(X_{h,k} \in B_n)
\]
\[
\begin{align*}
&= n^{1/2 - (1 - \alpha)}/2 \sum_{k=1}^{n} |C X_{h,k,i_1}| \left| |C P X_{h,k,i_2}| \mathbb{1}(X_{h,k} \in B_n, C X_{h,k} \notin B_n) \\
&\quad + n^{1/2 - (1 - \alpha)}/2 \sum_{k=1}^{n} |C X_{h,k,i_1}| \left| |C P X_{h,k,i_2}| \mathbb{1}(X_{h,k} \in B_n, C X_{h,k} \in B_n) \right|
\end{align*}
\]

The probability that the first summand is different from zero tends to zero by Proposition 3.6. Since
\[
C P X_{h,k} = X_{h,k} - |C X_{h,k}|
\]
an upper bound for the expectation of the second term is given by
\[
n^{1 + 1/2 - (1 - \alpha)} \mathbb{E} \left[ |C X_{h,k,i_1}| \left| |C P X_{h,k,i_2}| \mathbb{1}(C P X_{h,k} \in 2B_n) \right| \right] = n^{1/2 + \alpha} \mathbb{E} \left[ |C X_{h,k,i_1}| \left| |C P X_{h,k,i_2}| \mathbb{1}(|C P X_{h,k}| > 0) \mathbb{1}(C P X_{h,k} \in 2B_n) \right| \right] \leq n^{1/2 + \alpha} \mathbb{E} \left[ |C X_{h,k,i_1}| \left| |C P X_{h,k,i_2}| \mathbb{1}(|C P X_{h,k}| > 0) \mathbb{1}(C P X_{h,k} \in 2B_n(I, \beta^a)) \right| \right] \leq n^{1/2 + \alpha} \mathbb{E} \left[ |C X_{h,k,i_1}| 2^{\beta^n b_n} \mathbb{1}(|C P X_{h,k}| > 0) \right] = 2 n^{1/2 + \beta^n b_n} \mathbb{E} \left[ |C X_{h,k,i_1}| \right] \mathbb{P} \left( |C P X_{h,k}| > 0 \right),
\]
where we have used that \(B_n \subseteq B_n(I, \beta^a) = \{ x \in \mathbb{R}^d : x^T x \leq \beta^n b_n^2 \} \) and \(b_n = \sqrt{2h \log n} \) (Definition 3.5). This term tends to zero, since we have by the Poisson distribution
\[
\mathbb{P} \left( |C P X_{h,k}| > 0 \right) = \mathbb{P}(N_{h,i_2}(\mathbb{R}) > 0) = 1 - \exp (-\tilde{i}_{i_2}(\mathbb{R})h) = O(h)
\]
and by Lemma A.3 iv
\[
\mathbb{E} \left[ |C X_{h,k,i_1}| \right] = O(h).
\]
Ad (ii), we consider the expectation and apply Lemma A.3 The cases
\[
\begin{align*}
&j_1 + j_2 = 1, j_3 = 2, \quad j_1 \geq 2, j_2 = 0, j_3 \geq 1, \quad j_1 = 0, j_2 \geq 2, j_3 \geq 1 \quad \text{follow directly. For } j_1 \geq 1, j_2 \geq 1 j_3 \geq 1 \text{ we can use Hölder’s inequality and obtain}
\end{align*}
\]
\[
\mathbb{E} \left[ |C P X_{h,i_1}^j \cdot |C P X_{h,i_1}^{j_2}| \right] \leq \sqrt{\mathbb{E} \left[ |C P X_{h,i_1}^{j_1}| \right]} \mathbb{E} \left[ |C P X_{h,i_1}^{j_2}| \right] = O(h).
\]

\[\square\]

**Lemma B.3**

*Let \(1/2 < \alpha \leq 1\) in [3.1] and \(i \in \{1, \ldots, d\}\), then, as \(n \to \infty\)*

\[
\begin{align*}
(i) \quad &\sqrt{\frac{n}{T}} \sum_{k=1}^{n} C X_{h,i_1}^2 \mathbb{I}(X_{h,k} \notin B_n) \xrightarrow{P} 0, \\
(ii) \quad &\sqrt{T} \frac{1}{T} \sum_{k=1}^{n} C X_{h,i} \mathbb{I}(X_{h,k} \notin B_n) \xrightarrow{P} 0.
\end{align*}
\]

115
B. Technical details of Section 3.4

Proof. We set in (i) \( j = 2, r = 1/2 \) and in (ii) \( j = 1, r = (1 - \alpha)/2 \). Then

\[
\left| n^{r-(1-\alpha)} \sum_{k=1}^{n} C X_{h,k,i}^j 1(X_{h,k} \notin B_n) \right| \leq n^{r-(1-\alpha)} \sum_{k=1}^{n} |C X_{h,k,i}^j| 1(X_{h,k} \notin B_n)
\]

\[
= n^{r-(1-\alpha)} \sum_{k=1}^{n} |C X_{h,k,i}^j| 1(X_{h,k} \notin B_n, C X_{h,k} \in B_n)
\]

\[
+ n^{r-(1-\alpha)} \sum_{k=1}^{n} |C X_{h,k,i}^j| 1(X_{h,k} \notin B_n, C X_{h,k} \notin B_n)
\]

The probability that the latter term is different from zero tends to zero by Proposition 3.6. For the first term an upper bound is given by

\[
n^{r-(1-\alpha)} \sum_{k=1}^{n} |C X_{h,k,i}^j| 1(\|CP X_{h,k}\| > 0),
\]

since

\[CP X_{h,k} = X_{h,k} - C X_{h,k}.\]

We consider the expectation and obtain

\[
n^{1+r-(1-\alpha)} \mathbb{E} \left[ |C X_{h,k,i}^j| \right] \mathbb{P} (\|CP X_{h,k}\| > 0)
\]

\[
\leq n^{1+r-(1-\alpha)} \mathbb{E} \left[ |C X_{h,k,i}^j| \right] \left( 1 - \exp \left( -\nu(\mathbb{R}^d) h \right) \right) \rightarrow 0,
\]

where we have used the moments given in Lemma A.3(i).

Lemma B.4

(i) Let \( 1/2 < \alpha \leq 1 \) in (3.1) and \( i \in \{1, ..., d\} \). Then, as \( n \rightarrow \infty \),

\[
\sqrt{n} \frac{1}{T} \sum_{k=1}^{n} CP X_{h,k,i}^2 1(X_{h,k} \in B_n) \overset{p}{\rightarrow} 0.
\]

(ii) Let \( 1/2 < \alpha \leq 1 \) in (3.1) and \( i \in \{1, ..., d\} \). Then, as \( n \rightarrow \infty \),

\[
\sqrt{T} \frac{1}{T} \sum_{k=1}^{n} CP X_{h,k,i} 1(X_{h,k} \in B_n) \overset{p}{\rightarrow} 0.
\]

(iii) Let \( 1/2 < \alpha \leq 1 \) in (3.1) Then, as \( n \rightarrow \infty \),

\[
\sqrt{T} \sup_{z \in \mathbb{R}^d} \frac{1}{T} \sum_{k=1}^{n} 1(X_{h,k} \notin B_n, X_{h,k} \leq z)
\]

\[
- \frac{1}{T} \sum_{0 \leq t \leq T} 1(\Delta L_t \leq z, \Delta L_t \neq 0) \overset{p}{\rightarrow} 0.
\]
(iv) Let \( 1/2 < \alpha \leq 1 \) in (3.1), \( i \in \{1, \ldots, d\} \) and \( z \in \mathbb{R}^d \). Then, as \( n \to \infty \),
\[
\sqrt{T} \left( \frac{1}{T} \sum_{k=1}^{n} CPX_{h,k,i} I(X_{h,k} \not\in B_n, X_{h,k} \leq z) \right) - \frac{1}{T} \sum_{0 \leq t \leq T} \Delta L_t I(\Delta L_t \leq z) \xrightarrow{P} 0.
\]

(v) Let \( 1/2 < \alpha \leq 2/3 \) in (3.1), \( l \in \mathbb{N} \) and
\[
r \in (0, (1 - \alpha)/2] \quad \text{and} \quad p := \left\lfloor \frac{\log \alpha - 2r}{\alpha - 2r} + 1 \right\rfloor.
\]
We assume
\[
\mathbb{E} \left[ L_{1,i}^{2l(2r/p)} \right] < \infty, \quad \forall i = 1, \ldots, d.
\]
Then, as \( n \to \infty \),
\[
n^r \sup_{z \in \mathbb{R}^d} \left| \frac{1}{T} \sum_{k=1}^{n} \left( CPX_{h,k}^T A CPX_{h,k} \right)^l I(X_{h,k} \not\in B_n, X_{h,k} \leq z) \right. \left. - \frac{1}{T} \sum_{0 \leq t \leq T} (\Delta L_t^T A \Delta L_t)^l I(\Delta L_t \leq z) \right| \xrightarrow{P} 0.
\]

(vi) Let \( 2/3 < \alpha < 1 \) in (3.1), \( l \in \mathbb{N} \) and
\[
r \in (0, (1 - \alpha)/2]
\]
We assume
\[
\mathbb{E} \left[ L_{1,i}^{2l+2l(r=(1-\alpha)/2)} \right] < \infty, \quad \forall i = 1, \ldots, d.
\]
Then, as \( n \to \infty \),
\[
n^r \sup_{z \in \mathbb{R}^d} \left| \frac{1}{T} \sum_{k=1}^{n} \left( CPX_{h,k}^T A CPX_{h,k} \right)^l I(X_{h,k} \not\in B_n, X_{h,k} \leq z) \right. \left. - \frac{1}{T} \sum_{0 \leq t \leq T} (\Delta L_t^T A \Delta L_t)^l I(\Delta L_t \leq z) \right| \xrightarrow{P} 0.
\]

Proof. Recall, by Definition 3.3 we have
\[
B_n \subseteq B_n(I, \beta'') = \{ x \in \mathbb{R}^d : x^T x \leq \beta'' b_n^2 \}, \quad \text{where} \quad b_n = \sqrt{2h \log n}.
\]
We define \( \bar{b}_n := (1, \ldots, 1)^T b_n \) and have
\[
B_n(I, \beta'') \subseteq \left[ -\sqrt{\beta'' \bar{b}_n}, \sqrt{\beta'' \bar{b}_n} \right].
\]
B. Technical details of Section 3.4

We set \( u_k \in \mathbb{R}, k = 1, \cdots, n \) and \( v \in \mathbb{R} \), then

\[
\begin{align*}
&\left| \sum_{k=1}^{n} u_k \mathbbm{1}(X_{h,k} \notin B_n, X_{h,k} \leq z) - v \right| \\
\leq &\left| \sum_{k=1}^{n} u_k \mathbbm{1}(X_{h,k} \notin B_n, X_{h,k} \leq z, CPX_{h,k} \in B_n, \|CPX_{h,k}\| > 0) - v \right| \\
+ &\sum_{k=1}^{n} |u_k| \mathbbm{1}(X_{h,k} \notin B_n, X_{h,k} \leq z, CPX_{h,k} \notin B_n) \\
\leq &\left| \sum_{k=1}^{n} u_k \mathbbm{1}(X_{h,k} \notin B_n, X_{h,k} \leq z, CPX_{h,k} \in B_n, CPX_{h,k} \leq z, \|CPX_{h,k}\| > 0) - v \right| \\
+ &\sum_{k=1}^{n} |u_k| \mathbbm{1}(CPX_{h,k} \in (z, z + \sqrt{\beta''}b_n), \|CPX_{h,k}\| > 0) \\
+ &\sum_{k=1}^{n} |u_k| \mathbbm{1}(CPX_{h,k} \notin B_n) \\
\leq &\left| \sum_{k=1}^{n} u_k \mathbbm{1}(CPX_{h,k} \in B_n, CPX_{h,k} \leq z, \|CPX_{h,k}\| > 0) - v \right| \\
+ &\sum_{k=1}^{n} |u_k| \mathbbm{1}(CPX_{h,k} \notin B_n, CPX_{h,k} \leq z, \|CPX_{h,k}\| > 0) \\
+ &\sum_{k=1}^{n} |u_k| \mathbbm{1}(CPX_{h,k} \in (-2\sqrt{\beta''}b_n, 2\sqrt{\beta''}b_n), \|CPX_{h,k}\| > 0) \\
+ &\sum_{k=1}^{n} |u_k| \mathbbm{1}(CPX_{h,k} \in (z, z + \sqrt{\beta''}b_n), \|CPX_{h,k}\| > 0) \\
+ &\sum_{k=1}^{n} |u_k| \mathbbm{1}(CPX_{h,k} \notin B_n) \\
\leq &\left| \sum_{k=1}^{n} u_k \mathbbm{1}(CPX \leq z, \|CPX_{h,k}\| > 0) - v \right|
\end{align*}
\]
+ \sum_{k=1}^{n} |u_k| I(CP X_{h,k} \in (z - \sqrt{\beta' b_n}, z), \|CP X_{h,k}\| > 0)
+ \sum_{k=1}^{n} |u_k| I(CP X_{h,k} \in (-2\sqrt{\beta' b_n}, 2\sqrt{\beta' b_n}), \|CP X_{h,k}\| > 0)
+ \sum_{k=1}^{n} |u_k| \left( I(CP X_{h,k} \in (z, z + \sqrt{\beta' b_n}), \|CP X_{h,k}\| > 0) + 2I(C_{X,h,k} \notin B_n) \right)
\leq \sum_{k=1}^{n} u_k I(CP X \leq z, \|CP X_{h,k}\| > 0) - v \right)
\leq \sum_{k=1}^{n} u_k I(CP X \leq z, \|CP X_{h,k}\| > 0) - v \right)
(B.1)
+ \sum_{k=1}^{n} |u_k| \left( I(CP X_{h,k} \in (-\sqrt{\beta' b_n}, \sqrt{\beta' b_n}), \|CP X_{h,k}\| > 0) \right)
(B.2)
+ \sum_{k=1}^{n} |u_k| \left( I(CP X_{h,k} \in (-2\sqrt{\beta' b_n}, 2\sqrt{\beta' b_n}), \|CP X_{h,k}\| > 0) \right)
(B.3)
+ \sum_{k=1}^{n} |u_k| \left( 2I(C_{X,h,k} \notin B_n) \right).
(B.4)

We can insert this decomposition of the indicator function in (i)-(vi).

Ad (i), we set
\[ u_k := CP \sum_{i=1}^{n} CP X_{h,k}^2, \quad v := \sum_{k=1}^{n} CP X_{h,k}^2, \quad z = (\infty, \ldots, \infty)^T \]
which means that the terms in (B.1) and (B.2) vanish. So we have to control the remaining two terms. We start with the term in (B.3).
\[ \mathbb{P} \left( \sqrt{n} \frac{1}{T} \sum_{k=1}^{n} CP X_{h,k}^2 I(CP X_{h,k} \in (-2\sqrt{\beta' b_n}, 2\sqrt{\beta' b_n})) \geq \varepsilon \right) \]
\[ \leq \mathbb{P} \left( \sum_{k=1}^{n} \left( \sqrt{\beta' b_n} \right)^2 I(CP X_{h,k} \in (-2\sqrt{\beta' b_n}, 2\sqrt{\beta' b_n}), \|CP X_{h,k}\| > 0) \geq \varepsilon n^{1-\alpha-1/2} \right) \]
\[ \leq \mathbb{P} \left( \sum_{k=1}^{n} I(|CP X_{h,k,1}| \leq 2\sqrt{\beta' b_n}, \|CP X_{h,k,1}\| > 0) \geq n^{1-\alpha} \right). \]

So an upper bound is given by
\[ \mathbb{P}(M_2^{(n)} \cup M_3^{(n)}(1 - \alpha, i)). \]

Lemma [B.1] and Proposition [B.24] define the sets and prove that the probabilities tend to zero as \( n \) goes to infinity. Next, we have for the term in (B.4)
\[ \sqrt{n} \frac{1}{T} \sum_{k=1}^{n} CP X_{h,k}^2 I(C_{X,h,k} \notin B_n) \sim 0, \]
since the probability that this term is different from zero tends to zero by Proposition 3.6.

Ad (ii), we can prove the statement analogously to (i) by setting

\[ u_k := CP_{X_{h,k,i}}, \quad v := \sum_{k=1}^{n} CP_{X_{h,k,i}}, \quad z = (\infty, \cdots, \infty)^T. \]

Ad (iii), we set

\[ u_k := 1, \quad v := \sum_{0 \leq t \leq T} 1(\Delta L_t \leq z, \Delta L_t \neq 0). \]

By Proposition 3.24 the probability converges to zero that the term in (B.1) is different from zero. Next we consider the terms in (B.2) and (B.3). For all \( \varepsilon > 0 \) we have

\[
\begin{align*}
\mathbb{P}\left( \left. \frac{1}{\sqrt{T}} \sup_{z \in \mathbb{R}^d} \left| \frac{1}{T} \sum_{k=1}^{n} \left( 1(CP_{X_{h,k}} \in (z - \sqrt{b_n^\tau} \vec{b}, z + \sqrt{b_n^\tau} \vec{b}), \|CP_{X_{h,k}}\| > 0) \right) \right\} \right) > \varepsilon \\
\leq \mathbb{P}\left( \left. \frac{2}{\sqrt{T}} \sup_{z \in \mathbb{R}^d} \frac{1}{T} \sum_{k=1}^{n} \left( 1(CP_{X_{h,k}} \in \left( z - 2\sqrt{b_n^\tau} \vec{b}, z + 2\sqrt{b_n^\tau} \vec{b} \right), \|CP_{X_{h,k}}\| > 0) \right) \right\} \right) > \varepsilon \\
\leq \mathbb{P}\left( \left. \sup_{z \in \mathbb{R}^d} \sum_{k=1}^{n} \left( \frac{1}{d} \bigcup_{i=1}^{d} \left( CP_{X_{h,k,i}} \in \left( z^{(i)} - 2\sqrt{b_n^\tau} \vec{b}, z^{(i)} + 2\sqrt{b_n^\tau} \vec{b} \right), \|CP_{X_{h,k,i}}\| > 0 \right) \right) \right) > \varepsilon n^{(1-\alpha)/2}/2 \\
\leq \mathbb{P}\left( \left. \sup_{z \in \mathbb{R}^d} \sum_{k=1}^{n} \left( \exists i = 1, \cdots, d : CP_{X_{h,k,i}} \in \left( z^{(i)} - 2\sqrt{b_n^\tau} \vec{b}, z^{(i)} + 2\sqrt{b_n^\tau} \vec{b} \right), \|CP_{X_{h,k,i}}\| > 0 \right) \right) > \varepsilon n^{(1-\alpha)/2}/2 \right)
\end{align*}
\]
\[ |CP X_{h,k,i}| > 0 > \varepsilon n^{(1-\alpha)/2}/2 \].

An upper bound is given by
\[ P(M_2^{(n)}) + \sum_{i=1}^{d} P(M_3^{(n)}((1-\alpha)/2, i)) \to 0. \]

The term \((B.4)\) is considered in Proposition 3.6.

Ad (iv), we set
\[ u_k := CP X_{h,k,i}, \quad v := \frac{1}{T} \sum_{0 \leq t \leq T} \Delta L_t \mathbb{1}(\Delta L_t \leq z). \]

By Proposition 3.24 the probability converges to zero that the term in \((B.1)\) is different from zero. The term in \((B.2)\) vanishes if there exists an \(i \in \{1, \ldots, d\}\) such that \(z(i) = \infty\). If \(z \in \mathbb{R}^d\), we can find an upper bound for \(|u_k|\), because \(z\) is fixed. Thus \((B.2)\) tends to zero analogously to (iii). The term \((B.3)\) converges to zero analogously. The term \((B.4)\) is considered in Proposition 3.6.

Ad (v), the terms \((B.1)\) and \((B.4)\) converge to zero by Proposition 3.24 and 3.6. For the sum of \((B.2)\) and \((B.3)\) we use following upper bound
\[ 2n^r \sup_{z \in \mathbb{R}^d} \frac{1}{T} \sum_{k=1}^{n} (CP X_{h,k}^T A CP X_{h,k})^p \mathbb{1}(CP X_{h,k} \in (z - 2\sqrt{\beta''} b_n, z + 2\sqrt{\beta''} b_n), ||CP X_{h,k}|| > 0) \]
\[ \leq 2 \left( \frac{1}{T} \sum_{k=1}^{n} (CP X_{h,k}^T A CP X_{h,k})^p \right)^{1/p} \sup_{z \in \mathbb{R}^d} \left( n^r/(1-1/p) \right) \]
\[ \frac{1}{T} \sum_{k=1}^{n} \mathbb{1}(CP X_{h,k} \in (z - 2\sqrt{\beta''} b_n, z + 2\sqrt{\beta''} b_n), ||CP X_{h,k}|| > 0) \right)^{1-1/p}, \]

where we have used Hölder’s inequality and the given \(p\). The first term is stochastically bounded, since the expectation is uniform bounded by Lemma A.2 and Lemma A.3. For the second term we have analogously to (ii)
\[ P(M_2^{(n)}) + \sum_{i=1}^{d} P(M_3^{(n)}(1 - 3/2\alpha + \delta', i)), \]

where
\[ \delta := \frac{l \alpha}{\alpha - 2r} + 1 = \frac{\alpha}{\alpha - 2r} \quad \text{and} \quad \delta' := \frac{\delta (\alpha^2 - 4\alpha r + 4r^2)}{2r (-2\delta r + 2r + \delta \alpha)} > 0. \]

Ad (vi), we define
\[ p := \frac{l (1-\alpha)}{1-\alpha - r} + 1. \]
B. Technical details of Section 3.4

Then it can easily shown that

\[ 2l(2 \lor p) \leq 4l + 2l \left( r = \frac{1-a}{2} \right). \]

Thus, the statement can be analogously shown to (v) by

\[
\begin{align*}
\delta &:= p - \frac{l(1 - \alpha)}{1 - \alpha - r}, \\
\delta' &:= \frac{\delta \left( -2r + 1 - 2\alpha + 2\alpha r + \alpha^2 + r^2 \right)}{\delta - \delta \alpha + \delta r + r} > 0.
\end{align*}
\]
C. Technical details of Section 3.5

Definition C.1 (Different parts of the jump component)
Let \( \{L_t, t \geq 0\} \) be a Lévy process with triplet \((C, \gamma, \nu)\) and \(J\) the corresponding Poisson measure given in Proposition 2.23.

(i) The large jump part is defined by
\[
CP L_t(\omega)[\xi_1, \xi_2] := \int_0^t \int_{\xi_1 < ||x|| \leq \xi_2} x J((ds, dx), \omega)
\]
and contains all jumps with size in \((\xi_1, \xi_2]\), where \(\xi_1, \xi_2 \in [0, \infty]\) in case of \(\int_{||x|| \leq 1} ||x|| \nu(dx) < \infty\) and \(\xi_1, \xi_2 \in (0, \infty]\) in case \(\int_{||x|| \leq 1} ||x|| \nu(dx) = \infty\). Denote
\[
CP L_t[\xi] := CP L_t[\xi, \infty]
\]
and in case \(\int_{||x|| \leq 1} ||x|| \nu(dx) < \infty\),
\[
CP L_t := CP L_t[0].
\]

(ii) The small jump part is defined by
\[
sJ L_t(\omega)[\xi_1, \xi_2] := \int_0^t \int_{\xi_1 \leq ||x|| \leq \xi_2} (x J((ds, dx), \omega) - x \nu(dx))
\]
and contains all jumps with size in \((\xi_1, \xi_2]\) and the corresponding compensating drift term, where \(\xi_1, \xi_2 \in [0, \infty]\) in case of \(\int_{||x|| > 1} ||x|| \nu(dx) < \infty\) and \(\xi_1, \xi_2 \in (0, \infty]\) in case \(\int_{||x|| > 1} ||x|| \nu(dx) = \infty\). Define
\[
sJ L_t[\xi] := sJ L_t[0, \xi] \quad \text{and} \quad sJ L_t := sJ L_t[1].
\]

(iii) The compensating drift term is defined by
\[
\gamma_t[\xi_1, \xi_2] := t \int_{\xi_1 \leq ||x|| \leq \xi_2} x \nu(dx)
\]
and compensates the jumps with size in \((\xi_1, \xi_2]\), where \(\xi_1, \xi_2 \in [0, \infty]\) in case of \(\int_{||x|| \leq 1} ||x|| \nu(dx) < \infty, \int_{||x|| > 1} ||x|| \nu(dx) < \infty\) and \(\xi_1, \xi_2 \in [0, \infty]\) in case \(\int_{||x|| \leq 1} ||x|| \nu(dx) < \infty, \int_{||x|| > 1} ||x|| \nu(dx) = \infty\) and \(\xi_1, \xi_2 \in (0, \infty]\) in case \(\int_{||x|| \leq 1} ||x|| \nu(dx) = \infty, \int_{||x|| > 1} ||x|| \nu(dx) < \infty\) and \(\xi_1, \xi_2 \in (0, \infty]\) in case \(\int_{||x|| \leq 1} ||x|| \nu(dx) = \infty, \int_{||x|| > 1} ||x|| \nu(dx) = \infty\).
C. Technical details of Section 3.5

(iv) The large jump parts of the component processes are defined by

\[ CP_L_t(\omega)[\xi_1, \xi_2] := \int_0^t \int_{|x| \leq \xi_2} xJ(ds, dx), \quad i \in \{1, \cdots, d\}. \]

Please note the difference to the component of the large jump part that is denoted by \( CP_{L,i}[\xi_1, \xi_2] \). The quantities \( sJ_L_t[\omega][\xi_1, \xi_2] \) and \( \gamma_{i,t}[\xi_1, \xi_2] \) are defined analogously.

(v) The increments of the corresponding processes are denoted by

\[ \{CP_{X,h,k}[\xi_1, \xi_2], k = \{1, \cdots, d\}\}, \quad \{sJ_{X,h,k}[\xi_1, \xi_2], k = \{1, \cdots, d\}\}, \quad \{CP_{i,X,h,k}[\xi_1, \xi_2], k = \{1, \cdots, d\}\}, \quad \{sJ_{i,X,h,k}[\xi_1, \xi_2], k = \{1, \cdots, d\}\}. \]

Remark C.2

Like in the Lévy-Itô decomposition (Proposition 2.23) \( CP_L_t[\xi_1, \xi_2] \) or \( sJ_L_t[\xi_1, \xi_2] \) is independent to \( CP_{L,i}[\xi_1', \xi_2'] \) or \( sJ_{L,i}[\xi_1', \xi_2'] \) if \((\xi_1, \xi_2) \cap (\xi_1', \xi_2') = \emptyset\). See also Cont and Tankov (2004, Lemma 3.2).

Remark C.3

We keep in mind the difference between \( \gamma_{i,t}[\xi_1, \xi_2] \) and \( \gamma[\xi] \) that has been defined in Remark 2.15.

Lemma C.4

Let \( \alpha \in (1/2, 1] \) in (3.1) and \( \kappa_1, \kappa_2 > 0 \). We define

\[ M_2^{(n)}(\xi_n, \xi'_n, \chi, i) := \{ \exists z \in \mathbb{R} \setminus (-\xi_n, \xi_n) : N_{T,i}(z - \kappa_1 \xi'_n, z + \kappa_1 \xi'_n) \geq \kappa_2 n^\chi \}. \]

(i) Let

\[ 0 < \rho < \alpha/2 - 1/4 \quad \text{and} \quad \xi_n := n^{-\rho}. \]

Then there exists a sequence \( \xi_n \to 0 \) such that, as \( n \to \infty \),

\[ \frac{\xi_n}{\xi_n'} \to 0 \quad \text{and} \quad \sqrt{n}h \nu \left( \{ x \in \mathbb{R}^d : ||x|| \geq \xi_n \} \right) \to 0 \]

and

\[ P \left( M_2^{(n)}(\xi_n, \xi'_n, 1 - \alpha, i) \right) \to 0, \]

for all \( i \in \{1, \cdots, d\} \).

(ii) Let Assumption 3.14 be satisfied, choose

\[ r \in \begin{cases} [0 \vee 1/2 - 2\alpha/3, \alpha/3] & \text{if } \alpha \leq 3/5 \\ [0 \vee 1/2 - 2\alpha/3, (1 - \alpha)/2] & \text{if } \alpha > 3/5 \end{cases} \]

and define

\[ \xi_n := n^{(3r-\alpha)/6} \log(n) \quad \text{and} \quad \xi'_n := n^{(r-\alpha)/2} \log(n). \]
Let \( \theta \in [\gamma, \alpha/3] \), then, as \( n \to \infty \),
\[
\mathbb{P}\left( M^{(n)}_2(\xi_n, \xi'_n, \mathbf{1} - \alpha - \theta, i) \right) \to 0,
\]
for all \( i \in \{1, \ldots, d\} \).

**Proof.** We set w.l.o.g. \( \kappa_1 = \kappa_2 = 1 \) and assume \( n^\xi \in \mathbb{N} \).

Ad (i), for all positive zero sequences \( \xi_n \) and \( \xi'_n \) satisfying \( \xi'_n/\xi_n \to 0 \) we define
\[
A^{(n)}_m := \{ N_{T,i} (2(m-1)\xi'_n, 2(m+1)\xi'_n \setminus (-\xi_n/2, \xi_n/2)^c) \geq n^\xi \}
\]
and obtain
\[
M^{(n)}_2 \subset \bigcup_{m \in \mathbb{Z}} A^{(n)}_m.
\]

Analogously to Lemma B.4, we have
\[
\mathbb{P}\left( A^{(n)}_m \right) = \sum_{k=n^\xi}^{\infty} \exp\left( -\bar{\nu}_i \left( (2(m-1)\xi'_n, 2(m+1)\xi'_n \setminus (-\xi_n/2, \xi_n/2)^c) n^{1-\alpha} \right) \cdot \frac{(\bar{\nu}_i \left( (2(m-1)\xi'_n, 2(m+1)\xi'_n \setminus (-\xi_n/2, \xi_n/2)^c) n^{1-\alpha} \right)^k}{k!} \right)
\]
\[
\leq \frac{(\bar{\nu}_i \left( (2(m-1)\xi'_n, 2(m+1)\xi'_n \setminus (-\xi_n/2, \xi_n/2)^c) n^{1-\alpha} \right) n^\xi}{n^\xi!}.
\]

Since the jump density \( f_i \) is bounded in \( (-\xi_n/2, \xi_n/2)^c \) (Assumption 3.1), there exits a \( \kappa_{\xi_n} > 0 \) such that
\[
\bar{\nu}_i \left( (2(m-1)\xi'_n, 2(m+1)\xi'_n \setminus (-\xi_n/2, \xi_n/2)^c) n^{1-\alpha} \right)
\]
\[
= n^{1-\alpha} \int_{x \in (2(m-1)\xi'_n, 2(m+1)\xi'_n \setminus (-\xi_n/2, \xi_n/2)^c} f_i(x) \, dx
\]
\[
\leq 4 n^{1-\alpha} \max_{(2(m-1)\xi'_n, 2(m+1)\xi'_n \setminus (-\xi_n/2, \xi_n/2)^c} f_i(x) \cdot \xi'_n
\]
\[
=: \kappa_{\xi_n} n^{1-\alpha} \leq \kappa_{\xi_n} n^{1-\alpha - \rho}.
\]

Then,
\[
\mathbb{P}\left( M^{(n)}_2 \right) \leq \sum_{m \in \mathbb{Z}} \mathbb{P}(A^{(n)}_m)
\]
\[
\leq \sum_{m \in \mathbb{Z}} \frac{(\bar{\nu}_i \left( (2(m-1)\xi'_n, 2(m+1)\xi'_n \setminus (-\xi_n/2, \xi_n/2)^c) n^{1-\alpha} \right) n^\xi}{n^\xi!}
\]
\[
= \sum_{m \in \mathbb{Z}} \frac{(\bar{\nu}_i \left( (2(m-1)\xi'_n, 2(m+1)\xi'_n \setminus (-\xi_n/2, \xi_n/2)^c) n^{1-\alpha} \right) n^{\xi-1}}{n^\xi!}
\]
\[
\cdot \bar{\nu}_i \left( (2(m-1)\xi'_n, 2(m+1)\xi'_n \setminus (-\xi_n/2, \xi_n/2)^c) n^{1-\alpha} \right)
\]
\[
\leq \sum_{m \in \mathbb{Z}} \frac{(\kappa_{\xi_n} n^{1-\alpha - \rho} n^{\xi-1}}{n^\xi!} \bar{\nu}_i \left( (2(m-1)\xi'_n, 2(m+1)\xi'_n \setminus (-\xi_n/2, \xi_n/2)^c) n^{1-\alpha} \right)
\]
\[
125
\]
C. Technical details of Section 3.5

\[
\leq \frac{(\kappa \xi, n^{1-\alpha-\rho})^{n-1}_{n\chi!}}{n^{\chi}} n^{1-\alpha} 2 \tilde{v}_i((-\xi_n/2, \xi_n/2)^c).
\]

For each \(i \in \{1, \ldots, d\}\) we can find a sequence \(\xi_n\) such that the last term tends to zero by Stirling’s formula. Obviously we can also find a sequence \(\xi_n\) such that

\[
\sqrt{n} h \nu \left( \{ x : ||x|| \geq \xi_n \} \right) \xi_n^2 \to 0.
\]

Then we choose the slowest of all sequences.

Ad (ii), since Assumption 3.14, we have analogously to (i)

\[
\tilde{v}_i \left( (2(m-1) \xi_n', 2(m+1) \xi_n']) \setminus (-\xi_n/2, \xi_n/2)^c \right) n^{1-\alpha} \to 0.
\]

Then, analogously to (i),

\[
\mathbb{P} \left( M_{i_i}^{(n)} \leq \frac{n^{1-\alpha} 4 \xi_n'/(\xi_n/2)^3}{(n^{1-\alpha})!} n^{1-\alpha} 2 \tilde{v}_i((-\xi_n/2, \xi_n/2)^c) \right) \to 0.
\]

Then, analogously to (i),

\[
\mathbb{P} \left( M_{i_i}^{(n)} \right) \leq \frac{n^{1-\alpha} 4 \xi_n'/(\xi_n/2)^3}{(n^{1-\alpha})!} n^{1-\alpha} 2 \tilde{v}_i((-\xi_n/2, \xi_n/2)^c) \to 0.
\]

Remark C.5

It can be easily verified that the sequences \(\xi_n\) and \(\xi_n'\) defined in Lemma C.4(i) and (ii), respectively, also fulfill the conditions of Proposition 3.28(i) and (ii), respectively.

Lemma C.6

(i) Let \(1/2 < \alpha \leq 1\) in (3.4), \(i_1, i_2 \in \{1, \ldots, d\}\) and \(\mathbb{E} \left[ L_{i_1,i_2}^2 \right] < \infty\). Then, as \(n \to \infty\),

\[
\frac{1}{T} \sum_{k=1}^{n} |^j X_{h,k,i_1} |^C X_{h,k,i_2} |^1 (X_{h,k} \notin B_n) \to 0.
\]

(ii) Let \(1/2 < \alpha \leq 1\) in (3.4), \(i_1, i_2, i_3 \in \{1, \ldots, d\}\), \(r \in [0, \alpha/2]\) and \(j_1, j_2, j_3 \in \mathbb{N}\) satisfying

\[
 j_1 + j_2 = 1, j_3 \geq 2 \quad \text{or} \quad j_1 + j_2 = 2, j_3 \geq 1.
\]

If \(\mathbb{E} \left[ L_{i_1,i_1}^{2j_1} \right] < \infty\) and \(\mathbb{E} \left[ L_{i_2,i_2}^{2j_2} \right] < \infty\), then, as \(n \to \infty\),

\[
n^r \frac{1}{T} \sum_{k=1}^{n} |^j X_{h,k,i_1} |^{j_1} X_{h,k,i_2} |^{j_2} |^C X_{h,k,i_3} |^{j_3} \to 0.
\]
Proof. Ad (i), by Cauchy Schwartz inequality we obtain an upper bound
\[
\sqrt{\frac{1}{T} \sum_{k=1}^{n} C X_{h,k,i}^2 \mathbb{I}(X_{h,k} \notin B_n)} \sqrt{\frac{1}{T} \sum_{k=1}^{n} J X_{h,k,i}^2}.
\]
The first factor tends to zero in probability by Lemma C.7. The expectation of the second term is given by Lemma A.3 and is finite and independent of \( n \). Thus the second term is stochastically bounded.

Ad (ii), analogous to Lemma B.2(ii). \(\square\)

Lemma C.7
Let \( 1/2 < \alpha \leq 1 \) in (3.1) and \( i \in \{1, \cdots, d\} \), then, as \( n \to \infty \)
\[
\frac{1}{T} \sum_{k=1}^{n} C X_{h,k,i}^2 \mathbb{I}(X_{h,k} \notin B_n) \xrightarrow{P} 0.
\]

Proof. By Definition 3.5 we have \( B(C, \beta') \subseteq B_n \). We choose \( \bar{\beta}' \in (1, \beta') \), then
\[
\frac{1}{T} \sum_{k=1}^{n} C X_{h,k,i}^2 \mathbb{I}(X_{h,k} \notin B_n) = \frac{1}{T} \sum_{k=1}^{n} C X_{h,k,i}^2 \mathbb{I}(X_{h,k} \notin B_n, C X_{h,k} \in B_n(C, \bar{\beta}')) + \frac{1}{T} \sum_{k=1}^{n} C X_{h,k,i}^2 \mathbb{I}(X_{h,k} \notin B_n, C X_{h,k} \notin B_n(C, \bar{\beta}')).
\]
The probability that the latter term is different from zero tends to zero by Proposition 3.6. For the first term there exists \( \delta > 0 \) such that an upper bound is given by
\[
\frac{1}{T} \sum_{k=1}^{n} C X_{h,k,i}^2 \mathbb{I}(\| J X_{h,k} \| > \delta b_n) \cdot \left( \left( \sum_{i=1}^{d} \frac{\text{Var}(sJ X_{h,k,i})}{\delta^2 b_n^2} \right) + (1 - \exp (-\nu(|x| > 1)})h \right).
\]
The term converges to zero by Lemma A.3 (the second moment of \( sJ X_{1,1,i} \) does always exist) and by
\[
1 - \exp (-\nu(|x| > 1))h = O(h).
\]
C. Technical details of Section 3.5

**Lemma C.8**

Let $\frac{1}{2} < \alpha \leq 1$ in (3.1) and $i \in \{1, \cdots, d\}$, then, as $n \to \infty$

$$\frac{1}{T} \sum_{k=1}^{n} J^2_{\mathcal{X},h,k,i}(\mathcal{X}_{h,k} \in B_n) \xrightarrow{P} 0.$$  

**Proof.** We have

$$\frac{1}{T} \sum_{k=1}^{n} J^2_{\mathcal{X},h,k,i}(\mathcal{X}_{h,k} \in B_n)$$

$$\leq \frac{1}{T} \sum_{k=1}^{n} J^2_{\mathcal{X},h,k,i}(\mathcal{X}_{h,k} \in B_n, \mathcal{C}_{\mathcal{X},h,k} \notin B_n)$$

$$+ \frac{1}{T} \sum_{k=1}^{n} J^2_{\mathcal{X},h,k,i}(\mathcal{X}_{h,k} \in B_n, \mathcal{C}_{\mathcal{X},h,k} \in B_n)$$

The probability that the first term is different from zero converge to zero by Proposition 3.6. An upper bound for the second term is given by

$$\frac{1}{T} \sum_{k=1}^{n} sJ^2_{\mathcal{X},h,k,i}(sJ_{\mathcal{X},h,k} \in 2B_n)$$

$$\leq \frac{1}{T} \sum_{k=1}^{n} 4(\beta''b_n)^2 \mathbb{1}(\|\mathcal{C}_{\mathcal{X},h,k}[1]\| > 0).$$

The first term is considered in Lemma C.9 and the expectation of the second term is equal to

$$4 \frac{n}{T}(\beta''b_n)^2 \left(1 - \exp\left(-\frac{\nu(\{x : \|x\| > 1\})}{n^\alpha}\right)\right) \to 0.$$

**Lemma C.9**

Let $\frac{1}{2} < \alpha \leq 1$ in (3.1), $i \in \{1, \cdots, d\}$ and let $\delta_n$ be a positive zero sequence. Then, as $n \to \infty$,

$$\frac{1}{T} \sum_{k=1}^{n} sJ^2_{\mathcal{X},h,k,i}(sJ_{\mathcal{X},h,k} < \delta_n) \xrightarrow{L_1} 0.$$  

**Proof.** We can find a sequence $\delta'_n > 0$ satisfying

$$\delta'_n \to 0, \quad \frac{\delta'^2_n}{\delta'^3_n} \to 0 \quad \text{and} \quad h\left(\int_{\sqrt{\delta'_n} \leq \|x\| \leq 1} \|x\| \nu(dx)\right)^2 \to 0.$$  

128
Then the left hand side of i) is equal to
\[
\frac{1}{T} \sum_{k=1}^{n} s^j X^2_{h,k,i} \mathbb{1}(||s^j X_{h,k}|| < \delta_n, \sum_{(k-1)h<t\le kh} ||\Delta^s L_t||^2 \le \delta'_n) \quad \text{(C.1)}
\]
\[
+ \frac{1}{T} \sum_{k=1}^{n} s^j X^2_{h,k,i} \mathbb{1}(||s^j X_{h,k}|| > \delta_n, \sum_{(k-1)h<t\le kh} ||\Delta^s L_t||^2 > \delta'_n). \quad \text{(C.2)}
\]

For the expectation of \((C.1)\) an upper bound is given by
\[
\frac{n}{T} \mathbb{E} \left[ s^j X^2_{h,1,i} \mathbb{1}(||\Delta^s L_t|| \le \sqrt{\delta'_n}, \forall t \in ((k-1)h, kh]) \right]
\]
\[
\le \frac{2n}{T} \mathbb{E} \left[ s^j X^2_{h,1,i} [0, \sqrt{\delta'_n}] \right]
\]
\[
+ \frac{2n}{T} \mathbb{E} \left[ s^j X^2_{h,1,i} [\sqrt{\delta'_n}, 1] \mathbb{1}(||\Delta^s L_t|| \le \sqrt{\delta'_n}, \forall t \in ((k-1)h, kh]) \right].
\]

By Lemma A.3, the first term tends to zero, since
\[
\frac{n}{T} \mathbb{E} \left[ s^j X^2_{h,1,i} [0, \sqrt{\delta'_n}] \right] = \frac{n}{T} \mathbb{E} \left[ s^j X^2_{1,1,i} [0, \sqrt{\delta'_n}] \right]
\]
\[
= \frac{\int_{||x|| \le \sqrt{\delta'_n}} x^2 \nu(dx)}{\delta'_n} \to 0.
\]

The second term consists only of the compensating drift that is equal to
\[
\frac{n}{T} \mathbb{E} \left[ s^j X^2_{h,1,i} [\sqrt{\delta'_n}, 1] \mathbb{1}(||\Delta^s L_t|| \le \sqrt{\delta'_n}, \forall t \in ((k-1)h, kh]) \right] \to 0.
\]

For the expectation of \((C.2)\) an upper bound is given by
\[
\mathbb{E} \left[ \sum_{k=1}^{n} \delta_n^2 \mathbb{I} \left( \sum_{(k-1)h<t\le kh} ||\Delta^s L_t||^2 > \delta'_n \right) \right]
\]
\[
= \frac{n}{T} \delta_n^2 \mathbb{P} \left( \sum_{0<t\le h} ||\Delta^s L_t||^2 > \delta'_n \right) \le \frac{n}{T} \delta_n^2 \mathbb{E} \left[ \sum_{0<t\le h} ||\Delta^s L_t||^2 \right] \frac{1}{\delta'_n}
\]
\[
= \delta_n^2 \frac{n}{T} O(h),
\]

where we have used Markov’s Inequality and Proposition 2.32 and thus, the convergence follows from the assumption on \(\delta'_n\).

\[\square\]

**Lemma C.10**

(i) Let \(1/2 < \alpha \leq 1\) in \(\text{(3.1)}\), \(l \in \mathbb{N}\) and \(\mathbb{E} \left[ L^l_{1,i} \right] < \infty\) for all \(i \in \{1, \cdots, d\}\). Then, as \(n \to \infty\)

\[
\sup_{z \in \mathbb{R}^d} \left| \frac{1}{T} \sum_{k=1}^{n} \left( l^j X^T_{h,k} A^T X_{h,k} \right)^l \mathbb{1}(X_{h,k} \notin B_n, X_{h,k} \leq z) \right|
\]

is bounded by a constant independent of \(n\).
C. Technical details of Section 3.5

\[-\frac{1}{T} \sum_{0 \leq t \leq T} (\Delta L_t^T A \Delta L_t)^l \mathbb{1}(\Delta L_t \leq z) \xrightarrow{p} 0.\]

(ii) Let \(1/2 < \alpha \leq 1\) in (3.1), assume 3.14, choose

\[r \in \begin{cases} 0 \lor 1/2 - 2\alpha/3, \alpha/3 & \text{if } \alpha \leq 3/5 \smallskip \cr 0 \lor 1/2 - 2\alpha/3, (1 - \alpha)/2 & \text{if } \alpha > 3/5 \end{cases},\]

define

\[l := \left\lfloor \frac{3r}{\alpha - 3r} \right\rfloor + 2, \quad \text{and} \quad p := \frac{\log n + 1}{l},\]

and require

\[E \left[ L_{1,i}^{2(2/vp)} \right] < \infty, \quad i \in \{1, \ldots, d\}.\]

Then, as \(n \to \infty\),

\[n^r \sup_{z \in \mathbb{R}^d} \left| \frac{1}{T} \sum_{k=1}^n \left( J X_{h,k}^T A J X_{h,k} \right)^l \mathbb{1}(X_{h,k} \notin B_n, X_{h,k} \leq z) \right| \xrightarrow{p} 0.\]

Proof. We set \(r = 0\) in (i), in (ii) \(r\) is defined. We use the notation

\[u_k := \left( J X_{h,k}^T A J X_{h,k} \right)^l, \quad k = 1, \ldots, n\]

\[v := \sum_{0 \leq t \leq T} (\Delta L_t^T A \Delta L_t)^l \mathbb{1}(\Delta L_t \leq z).\]

We decompose the term in the indicator function and keep in mind

\[J X_{h,k} = C_P X_{h,k} [\xi_n] + \gamma_{c,h,k} [\xi, 1] + s J X_{h,k} [\xi_n]\]

and

\[B_n \subseteq B_n(I, \beta^*) = \{ x \in \mathbb{R}^d : x^T x \leq \beta^* b_n^2 \}, \quad \text{where} \quad b_n = \sqrt{2h \log n}.\]

Furthermore we define \(\tilde{b}_n := (1, \ldots, 1)^T b_n\).

In (i) and (ii), respectively, we use the sequences \(\{\xi_n, n \in \mathbb{N}\}\) and \(\{\xi'_n, n \in \mathbb{N}\}\) given in Lemma C.4(i) and (ii), respectively. It can easily verified that \(\{\xi_n, n \in \mathbb{N}\}\) also satisfies the condition of Proposition 3.28(i) and (ii), respectively.

The first step of the following decomposition is not shown in detail, since it is obtained analogously to Lemma B.4

\[\left| \sum_{k=1}^n u_k \mathbb{1}(X_{h,k} \notin B_n, X_{h,k} \leq z) - v \right| \leq \sum_{k=1}^n u_k \mathbb{1}(J X_{h,k} \leq z) - v\]
\[+ \sum_{k=1}^{n} |u_k| \mathbb{1} \left( JX_{h,k} \in (z - \sqrt{\beta_0 b_n}, z + \sqrt{\beta_0 b_n}) \right)\]

\[+ \sum_{k=1}^{n} |u_k| \mathbb{1} \left( JX_{h,k} \in (-2\sqrt{\beta_0 b_n}, \sqrt{\beta_0 b_n}) \right) + 2 \sum_{k=1}^{n} |u_k| \mathbb{1} \left( C X_{h,k} \notin B_n \right)\]

\[\leq \left| \sum_{k=1}^{n} u_k \mathbb{1} \left( JX_{h,k} \leq z, \| s^J X_{h,k}[\xi_n] \| < \xi_n'/2, \| \gamma_{h,k}[\xi_n, 1] \| < \xi_n'/2 \right) - v \right|\]

\[+ \sum_{k=1}^{n} |u_k| \mathbb{1} \left( JX_{h,k} \in (z - 2\sqrt{\beta_0 b_n}, z + 2\sqrt{\beta_0 b_n}), \| s^J X_{h,k}[\xi_n] \| < \xi_n'/2, \| \gamma_{h,k}[\xi_n, 1] \| < \xi_n'/2 \right)\]

\[+ \sum_{k=1}^{n} |u_k| \mathbb{1} \left( \| s^J X_{h,k}[\xi_n] \| \geq \xi_n'/2 \right) + \mathbb{1} \left( \| s^J X_{h,k}[\xi_n] \| \geq \xi_n'/2 \right)\]

\[+ 2 \sum_{k=1}^{n} |u_k| \mathbb{1} \left( C X_{h,k} \notin B_n \right)\]

\[\leq \left| \sum_{k=1}^{n} u_k \mathbb{1} \left( CP X_{h,k}[\xi_n] \leq z, JX_{h,k} \leq z, \| s^J X_{h,k}[\xi_n] \| < \xi_n'/2 \right) \right| - v\]

\[+ \sum_{k=1}^{n} |u_k| \mathbb{1} \left( CP X_{h,k}[\xi_n] > z, JX_{h,k} \leq z, \| s^J X_{h,k}[\xi_n] \| < \xi_n'/2, \| \gamma_{h,k}[\xi_n, 1] \| < \xi_n'/2 \right)\]

\[+ \sum_{k=1}^{n} |u_k| \mathbb{1} \left( CP X_{h,k}[\xi_n] \in (z - 2\xi_n', z + 2\xi_n') \right)\]

\[+ \sum_{k=1}^{n} |u_k| \mathbb{1} \left( \| CP X_{h,k}[\xi_n] \| \leq 2\xi_n' \right)\]

\[+ \sum_{k=1}^{n} |u_k| \mathbb{1} \left( \| \gamma_{h,k}[\xi_n, 1] \| \geq \xi_n'/2 \right) + \sum_{k=1}^{n} |u_k| \mathbb{1} \left( \| s^J X_{h,k}[\xi_n] \| \geq \xi_n'/2 \right)\]

\[+ 2 \sum_{k=1}^{n} |u_k| \mathbb{1} \left( C X_{h,k} \notin B_n \right)\]

\[\leq \left| \sum_{k=1}^{n} u_k \mathbb{1} \left( CP X_{h,k}[\xi_n] \leq z, \| s^J X_{h,k}[\xi_n] \| < \xi_n'/2, \| \gamma_{h,k}[\xi_n, 1] \| < \xi_n'/2 \right) \right|,\]
C. Technical details of Section 3.5

\[ \left| \left| \gamma_{h,k}^{c} [\xi_{n}, 1] \right| \right| < \xi_{n}'/2 - v \]

\[ + \sum_{k=1}^{n} |u_{k}| \mathbb{1} \left( CPX_{h,k}^{T}[\xi_{n}] \leq z, JX_{h,k} > z, \left| \left| s^{J}X_{h,k}^{T}[\xi_{n}] \right| \right| < \xi_{n}'/2, \right. \]

\[ \left. \left| \left| \gamma_{h,k}^{c} [\xi_{n}, 1] \right| \right| < \xi_{n}'/2 \right) \]

\[ + \sum_{k=1}^{n} |u_{k}| \mathbb{1} \left( CPX_{h,k}^{T}[\xi_{n}] \in (z, z + \xi_{n}') \right) \]

\[ + \sum_{k=1}^{n} |u_{k}| \mathbb{1} \left( CPX_{h,k}^{T}[\xi_{n}] \in (z - 2\xi_{n}', z + 2\xi_{n}') \right) \]

\[ + \sum_{k=1}^{n} |u_{k}| \mathbb{1} \left( \left| \left| \gamma_{h,k}^{c} [\xi_{n}, 1] \right| \right| \geq \xi_{n}'/2 \right) + \sum_{k=1}^{n} |u_{k}| \mathbb{1} \left( \left| \left| s^{J}X_{h,k}^{T}[\xi_{n}] \right| \right| \geq \xi_{n}'/2 \right) \]

\[ + 2 \sum_{k=1}^{n} |u_{k}| \mathbb{1} \left( C_{X_{h,k}} \notin B_{n} \right) \]

\[ \leq \sum_{k=1}^{n} |u_{k}| \mathbb{1} \left( CPX_{h,k}^{T}[\xi_{n}] \leq z \right) - v \]

\[ + \sum_{k=1}^{n} |u_{k}| \mathbb{1} \left( CPX_{h,k}^{T}[\xi_{n}] \in (z - 2\xi_{n}', z + 2\xi_{n}') \right) \]

\[ + \sum_{k=1}^{n} |u_{k}| \mathbb{1} \left( \left| \left| CPX_{h,k}^{T}[\xi_{n}] \right| \right| \leq 2\xi_{n}' \right) + 2 \sum_{k=1}^{n} |u_{k}| \mathbb{1} \left( \left| \left| \gamma_{h,k}^{c} [\xi_{n}, 1] \right| \right| \geq \xi_{n}'/2 \right) \]

\[ + 2 \sum_{k=1}^{n} |u_{k}| \mathbb{1} \left( \left| \left| s^{J}X_{h,k}^{T}[\xi_{n}] \right| \right| \geq \xi_{n}'/2 \right) + 2 \sum_{k=1}^{n} |u_{k}| \mathbb{1} \left( C_{X_{h,k}} \notin B_{n} \right) . \]

By inserting this decomposition in our statement it remains to show

1.) \( n^{r} \sup_{z \in \mathbb{R}^{d}} \left| \frac{1}{T} \sum_{k=1}^{n} \left( JX_{h,k}^{T}A^{J}X_{h,k} \right)^{l} \mathbb{1} \left( CPX_{h,k}^{T}[\xi_{n}] \leq z \right) \right| \]

\[ - \frac{1}{T} \sum_{0 \leq \ell \leq T} \left( \Delta L_{\ell}^{T}A^{J}L_{\ell} \right)^{l} \mathbb{1} \left( \Delta L_{\ell} \leq z \right) \rightarrow_{P} 0, \]

2.) \( n^{r} \sup_{z \in \mathbb{R}^{d}} \left| \frac{1}{T} \sum_{k=1}^{n} \left| JX_{h,k}^{T}A^{J}X_{h,k} \right|^{l} \left( 2 \mathbb{1} \left( CPX_{h,k}^{T}[\xi_{n}] \in (z - 2\xi_{n}', z + 2\xi_{n}') \right) \right)

\[ + \mathbb{1} \left( \left| \left| CPX_{h,k}^{T}[\xi_{n}] \right| \right| \leq 2\xi_{n}' \right) \right) \rightarrow_{P} 0, \]

3.) \( n^{r} \frac{1}{T} \sum_{k=1}^{n} \left| JX_{h,k}^{T}A^{J}X_{h,k} \right|^{l} \left( \left| \left| \gamma_{h,k}^{c} [\xi_{n}, 1] \right| \right| \geq \xi_{n}'/2 \right) \rightarrow_{P} 0, \)
4.) $n^r \frac{1}{T} \sum_{k=1}^{n} \langle J X_{h,k}^T A^j X_{h,k} \rangle^l \mathbb{1}(\|s^j X_{h,k}[\xi_n]\| \geq \xi_n/2) \xrightarrow{p} 0,$

5.) $n^r \frac{1}{T} \sum_{k=1}^{n} \langle J X_{h,k}^T A^j X_{h,k} \rangle^l \mathbb{1}(C X_{h,k} \notin B_n) \xrightarrow{p} 0.$

Ad 1.), we have

$$(J X_{h,k}^T A^j X_{h,k})^l = \left( C P X_{h,k}^T [\xi_n] A C P X_{h,k}[\xi_n] + s^j X_{h,k}^T [\xi_n] A s^j X_{h,k}[\xi_n] + \gamma_h^c[\xi_n, 1]^T A \gamma_h^c[\xi_n, 1] + C P X_{h,k}^T [\xi_n] A s^j X_{h,k}[\xi_n] + s^j X_{h,k}^T [\xi_n] A C P X_{h,k}[\xi_n] \right).$$

We use the decomposition in 1.) and use Lemma A.1(iii) and Lemma A.2. Then there exists a $\kappa > 0$ such that the following upper bound holds true.

$$n^r \sup_{z \in \mathbb{R}^d} \left| \frac{1}{T} \sum_{k=1}^{n} (C P X_{h,k}^T [\xi_n] A C P X_{h,k}[\xi_n])^l \mathbb{1}(C P X_{h,k}[\xi_n] \leq z) \right|$$

$$- \frac{1}{T} \sum_{0 \leq t \leq T} \left( \Delta L_t^T A \Delta L_t \right)^l \mathbb{1}(\Delta L_t \leq z) \right|$$

$$+ \kappa n^r \frac{1}{T} \sum_{m=1}^{l} \sum_{k=1}^{n} |C P X_{h,k}[\xi_n]|^{2l-2m}(\|s^j X_{h,k}[\xi_n]\|^{s^j X_{h,k}[\xi_n]}\|s^j X_{h,k}[\xi_n]\|)^m$$

$$+ \kappa n^r \frac{1}{T} \sum_{m=1}^{l} \sum_{k=1}^{n} |C P X_{h,k}[\xi_n]|^{2l-2m}(\|C P X_{h,k}[\xi_n]\|^{s^j X_{h,k}[\xi_n]}\|s^j X_{h,k}[\xi_n]\|)^m$$

$$+ \kappa n^r \frac{1}{T} \sum_{m=1}^{l} \sum_{k=1}^{n} |C P X_{h,k}[\xi_n]|^{2l-2m}(\|\gamma_h^c[\xi_n, 1]\|\gamma_h^c[\xi_n, 1])^m$$

$$+ \kappa n^r \frac{1}{T} \sum_{m=1}^{l} \sum_{k=1}^{n} |C P X_{h,k}[\xi_n]|^{2l-2m}(\|C P X_{h,k}[\xi_n]\|^{s^j X_{h,k}[\xi_n]}\|s^j X_{h,k}[\xi_n]\|)^m$$

$$+ \kappa n^r \frac{1}{T} \sum_{m=1}^{l} \sum_{k=1}^{n} |C P X_{h,k}[\xi_n]|^{2l-2m}(\|s^j X_{h,k}[\xi_n]\|\gamma_h^c[\xi_n, 1])^m.$$  

Ad (C.3) by Lemma A.2 there exists a $\kappa > 0$ such that

$$n^r E \left[ \frac{1}{T} \sum_{0 \leq t \leq T} (\Delta L_t A \Delta L_t)^l \mathbb{1}(\|\Delta L_t\| \leq \xi_n) \right]$$

$$\leq \kappa n^r E \left[ \frac{1}{T} \sum_{0 \leq t \leq T} |\Delta L_t|^{2l} \mathbb{1}(\|\Delta L_t\| \leq \xi_n) \right] = \kappa n^r \int_{|x| \leq \xi_n} |x|^{2l} \nu(dx)$$

133
C. Technical details of Section 3.5

\[ \leq \kappa n^r \sum_{i=1}^d \int_{|x_i| \leq \xi_n} |x_i|^{2l} \nu_i(dx_i), \]

where we have used Proposition 2.32. Thus, in (i) this term tends to zero, because \( r = 0 \) and in (ii) we use Assumption 3.14 and obtain an upper bound

\[ n^r \xi_n^2 \int_{-\xi_n}^{\xi_n} x^2 \nu(dx) \to 0. \]

Finally, by Proposition 3.28 we obtain, for all \( \varepsilon > 0 \),

\[ \mathbb{P} \left( n^r \sup_{z \in \mathbb{R}^d} \frac{1}{T} \sum_{k=1}^n (CP X_{h,k}^T \xi_n A^{CP} X_{h,k}^T \xi_n) \mathbf{1} (CP X_{h,k} \xi_n \leq z) \right) \]

\[ \leq \mathbb{P} \left( M^{(n)}_2 (\xi_n) \right) \to 0. \]

The terms \([C.4],[C.5],[C.6]\) and \([C.7]\) are considered in Lemma C.11. The term in \([C.8]\) is bounded by \([C.4]\) and \([C.6]\).

Ad 2.), obviously the following upper bound is given by

\[ 2 n^r \sup_{z \in \mathbb{R}^d} \left| \frac{1}{T} \sum_{k=1}^n |J_{X_{h,k}}^T A^J X_{h,k}|^l \mathbf{1} (CP X_{h,k} \xi_n \in (z - 2\xi_n, z + 2\xi_n)) \right| + 2 n^r \frac{1}{T} \sum_{k=1}^n |s^J X_{h,k}^T A s^J X_{h,k}|. \]

By Lemma A.2 for the second term, an upper bound is given by

\[ 2 n^r \frac{1}{T} \sum_{k=1}^n |s^J X_{h,k}|^{2l}. \]

This term is considered in Lemma C.11. For the first term we use Hölder’s inequality; in (i) we set \( p = 2 \) and in (ii) we use the given \( p \). Then we obtain

\[ \left( \frac{1}{T} \sum_{k=1}^n |J_{X_{h,k}}^T A^J X_{h,k}|^p \right)^{1/p} \sup_{z \in \mathbb{R}^d} \left( n^{r/(1-1/p)} \right)^{1-1/p}. \]

The first term is stochastically bounded, because we can use Lemma A.2, Lemma A.3 and the given moment condition to realize that the absolute moment is uniform bounded in \( n \). For the second term we define \( \theta := 0 \) in (i) and \( \theta := \alpha/3 - \delta' \) in (ii), where

\[ \delta := \frac{l_0}{\alpha - 3r} + 1 - \frac{\alpha}{\alpha - 3r} \quad \text{and} \quad \delta' := \frac{\alpha^2 - 6 \alpha r + 9 r^2}{3\delta \alpha - 9 \delta r + 9 r} > 0. \]
In (ii), by Assumption 3.14 there exists a $n$ such that

\[
The convergence follows from Lemma C.4 and 3.28.
\]

Thus, $\theta \in (r, \alpha/3)$. Then, for all $\varepsilon > 0$,

\[
\mathbb{P}
\left(
\sup_{z \in \mathbb{R}^d} \frac{1}{T} \sum_{k=1}^{n} \mathbb{1}(CPX_{h,k}[\xi_n] \in (z - 2\xi_n^{(i)}, z + 2\xi_n^{(i)}), CPX_{h,k}[\xi_n] \neq 0) > \varepsilon
\right)
\]

\[
\leq \mathbb{P}
\left(
\sup_{z \in \mathbb{R}^d} \frac{1}{T} \sum_{k=1}^{n} \mathbb{1}
\left(
\bigcap_{i=1}^{d} \left\{ CPX_{h,k,i}[\xi_n] \in (z(i) - 2\xi_n^{(i)}, z(i) + 2\xi_n^{(i)}), CPX_{h,k}[\xi_n] \neq 0 \right\}
\right) > \varepsilon
\right)
\]

\[
\leq \mathbb{P}
\left(
\sup_{z \in \mathbb{R}^d} \frac{1}{T} \sum_{k=1}^{n} \left( \sum_{i=1}^{d} \mathbb{1}(CPX_{h,k,i}[\xi_n] \in (z(i) - 2\xi_n^{(i)}, z(i) + 2\xi_n^{(i)}), CPX_{h,k}[\xi_n] \neq 0) \right) > \varepsilon
\right)
\]

\[
\leq \mathbb{P}
\left(
\sum_{i=1}^{d} \sup_{z(i) \in \mathbb{R}} \frac{1}{T} \sum_{k=1}^{n} \mathbb{1}(CPX_{h,k,i}[\xi_n] \in (z(i) - 2\xi_n^{(i)}, z(i) + 2\xi_n^{(i)}), CPX_{h,k}[\xi_n] \neq 0) > \varepsilon
\right)
\]

\[
\leq \mathbb{P}
\left(
\sup_{z \in \mathbb{R}^d} \frac{1}{T} \sum_{k=1}^{n} \left( \sum_{i=1}^{d} \mathbb{1}(CPX_{h,k,i}[\xi_n] \in (z(i) - 2\xi_n^{(i)}, z(i) + 2\xi_n^{(i)}), CPX_{h,k}[\xi_n] \neq 0) \right) > \varepsilon/d
\right)
\]

\[
\leq \mathbb{P}(M_3^{(n)}(\xi_n)) + \sum_{i=1}^{d} \mathbb{P}(M_2^{(n)}(\xi_n, \xi_n^{(i)} - \alpha, i)) \rightarrow 0.
\]

The convergence follows from Lemma C.4 and 3.28.

Ad 3.), in (i) the term is equal to zero for large $n$ because we have

\[
\frac{h \int_{\xi_n \leq ||x|| \leq 1} ||x|| \nu(dx)}{\xi_n^4/4} \leq \frac{h \nu((-\xi_n, \xi_n)^c)}{\xi_n^4/4} \leq \frac{\sqrt{n} \nu((-\xi_n, \xi_n)^c)}{\xi_n^2} \rightarrow 0.
\]

In (ii), by Assumption 3.14 there exists a $\kappa > 0$ such that

\[
\frac{h \int_{\xi_n \leq ||x|| \leq 1} ||x|| \nu(dx)}{\xi_n^4/4} \leq \sum_{i=1}^{d} \frac{h \int_{(-1, -\xi_n) \cup (\xi_n, 1)} ||x_i|| \nu_i(dx_i)}{\xi_n^4/4}
\]
C. Technical details of Section 3.5

\[
\sum_{i=1}^{d} \frac{\kappa h \int_{[-1,-\xi_n] \cup [\xi_n,1]} 1/x_i^2 dx_i}{\xi_n^l} \leq \frac{\kappa h}{\xi_n \xi_n'} \to 0.
\]

Ad 4.), we use the decomposition of 1.) and only have to show

\[
n^r T \sum_{k=1}^{n} \left| CPX_{h,k}^{T} [\xi_n] A^{CPX_{h,k}^{T}} [\xi_n] \right|^l \left( \sum_{i=1}^{d} I(||sJ X_{h,k,i}|| \geq \xi_n'/2) \right).
\]

We use Lemma A.2 and Lemma A.3 and have for the expectation

\[
n^r O(1) \cdot 4^d \sum_{i=1}^{d} \text{Var}(c_jh, X_{h,k}^{[\xi_n]})/\xi_n'^2 = n^r O(1) \cdot 4^d \sum_{i=1}^{d} \int_{-\xi_n}^{\xi_n} x_i^2 \nu(dx_i)/\xi_n'^2 = n^r h \cdot O(1)/\xi_n'^2 \to 0.
\]

Ad 5.), the probability that this term is different from zero is considered in Proposition 3.6. □

Lemma C.11

We use the setting given in Lemma C.10 (i) and (ii) and consider \(j_1, j_2 \in \{0, \ldots, 2l\}\) satisfying either \(j_1 > 0, j_2 \in \{1, \ldots, 2l\}\) or \(j_1 = 0, j_2 = 2l\). Then, as \(n \to \infty\),

1.) \[n^r T \sum_{k=1}^{n} \left| CPX_{h,k}^{T} [\xi_n] \right| j_1 \left| sJ X_{h,k}^{T} [\xi_n] \right| j_2 \xrightarrow{P} 0,
\]

2.) \[n^r T \sum_{k=1}^{n} \left| CPX_{h,k}^{T} [\xi_n] \right| j_1 \left| \gamma_{\tilde{h},k}^{T} [\xi_n,1] \right| j_2 \xrightarrow{P} 0.
\]

Proof. Ad 1., in the case \(j_1 > 0, j_2 \in \{1, \ldots, 2l\}\) we have

\[
\mathbb{P}\left( n^r T \sum_{k=1}^{n} \left| CPX_{h,k}^{T} [\xi_n] \right| j_1 \left| sJ X_{h,k}^{T} [\xi_n] \right| j_2 > \varepsilon \right) \leq n^r T \sum_{k=1}^{n} \left| CPX_{h,k}^{T} [1] \right| 2(j_1+1)/2 \]

\[
+ N_{(k-1)h,hk}((-\xi_n, \xi_n)^c) \left| sJ X_{h,k}^{T} [\xi_n] \right| j_2 > \varepsilon \quad \text{(C.9)}
\]

\[
+ \mathbb{P}\left( M^{(n)}_2 (\xi_n) \right) \quad \text{(C.10)}
\]

where Proposition 3.28 defines the set \(M^{(n)}_2 (\xi_n)\) and proves that \(M^{(n)}_2 (\xi_n)^c\) converges to zero. In \(\text{(C.9)}\) we have used that on \(\Omega \setminus M^{(n)}_2 (\xi_n)^c\) each \(CPX_{h,k}^{T} [\xi_n]\)
contains at most one jump, that is either larger or smaller than one. We consider the expectation of the term in (C.9)
\[
n^r \frac{n}{T} \mathbb{E} \left[ \left( |CP X_{h,k}[1]|^{2(j_1+1)/2} + N_{0,h,2} \left( -\xi_n, \xi_n \right) \right) \right] \mathbb{E} \left[ |s_j X_{h,1}[\xi_n]|^{j_2} \right]
\]
\[
\leq n^r \frac{n}{T} \mathbb{E} \left( \mathcal{O}(h) + h \nu \left( \{ x : ||x|| \geq \xi_n \} \right) \right) \sqrt{\mathbb{E} \left[ |s_j X_{h,1}^{2j_2}[\xi_n]|^{j_2} \right]}
\]
\[
\leq \mathcal{O}(1) \ n^{r-\alpha/2} \nu \left( \{ x : ||x|| \geq \xi_n \} \right) \left[ \sum_{i=1}^{d} \left( \int_{||x|| \leq \xi_n} x_i^{2j_2} \nu(dx) \right) \right],
\]
where we have used Lemma A.3. In the setting (i) of Lemma C.10 this term tends to zero by the conditions of Lemma C.4(i), for setting (ii) we can find an upper bound by using Assumption 3.14.

\[
\mathcal{O}(1) n^{r-\alpha/2} \sum_{i=1}^{d} \tilde{\nu}_i \left( -\xi_n, \xi_n \right)
\]
\[
\leq \mathcal{O}(1) n^{r-\alpha/2} \sum_{i=1}^{d} \int_{|x_i| \geq \xi_n} f_i(x_i) dx_i \sqrt{\sum_{i=1}^{d} \int_{||x|| \leq \xi_n} x_i^{2j_2} \nu(dx) }
\]
\[
\leq \mathcal{O}(1) n^{r-\alpha/2} \xi_n^{-2} \sum_{i=1}^{d} \int_{|x_i| \leq \xi_n} x_i^{2j_2} \nu(dx) \to 0.
\]

In the case \( j_1 = 0, j_2 = 2l \) there exists by Lemma A.1(ii), Lemma A.3(ii) and Proposition 2.30 a \( \kappa > 0 \) such that,

\[
n^r \mathbb{E} \left[ \frac{1}{T} \sum_{k=1}^{n} |s_j X_{h,k}[\xi_n]|^{2l} \right] = n^r \mathbb{E} \left[ \frac{1}{T} \sum_{k=1}^{n} \left( \sum_{i=1}^{d} s_j X_{h,1,i}^{2l}[\xi_n] \right) \right]
\]
\[
\leq \kappa n^r \frac{n}{T} \sum_{i=1}^{d} \mathbb{E} \left[ s_j X_{h,1,i}^{2l} \right] \leq n^r \frac{n h}{T} \sum_{i=1}^{d} \left( \int_{-\xi_n}^{\xi_n} x_i^{2l} \tilde{\nu}_i(dx_i) + \mathcal{O}(h) \right).
\]

This term tends to zero in setting (i), since

\[
n^0 \int_{-\xi_n}^{\xi_n} x_i^{2l} \tilde{\nu}_i(dx_i) \to 0
\]

and in setting (ii), because

\[
n^r \int_{-\xi_n}^{\xi_n} x_i^{2l} \tilde{\nu}_i(dx_i) \leq n^r \xi_n^{2l-2} \int_{-\xi_n}^{\xi_n} x_i^{2l} \tilde{\nu}_i(dx_i) \to 0.
\]

Ad 2., in setting (i) we have

\[
|\gamma_h^n[\xi_n, 1]| \leq h \int_{\xi_n \leq ||x|| \leq 1} ||x|| \nu(dx) \leq h \nu \left( \{ x : ||x|| \geq \xi_n \} \right)
\]
C. Technical details of Section 3.5

and in setting (ii) we have

$$|\gamma_h^c[\xi_n, 1]| \leq h \int_{\xi_n \leq ||x|| \leq 1} ||x|| \nu(dx) \leq \sum_{i=1}^{d} \int_{\xi_n \leq |x_i| \leq 1} |x_i| \tilde{\nu}_i(dx_i) \leq h/\xi_n.$$ 

So the term in 2.) converges analogously to 1.) to zero by using the conditions on $\xi_n$ given in Lemma C.4.
Bibliography


Bibliography


Bibliography


Bibliography


Bibliography


### List of Tables

3.1. Simulation result using the Merton model, parameters: $C = 0.09$, $\gamma[0] = 0$, $\beta = 0$, $\eta = 0.6$ and $\lambda = 5$. .................................................. 82

3.2. Simulation result using the Merton model, sample distribution and asymptotic approximation, parameters $C = 0.09$, $\gamma[0] = 0.035$, $\beta = 0$, $\eta = 0.1$ and $\lambda = 10$. .................................................. 85

3.3. Simulation result using the bivariate Merton model .................................................. 88

4.1. Simulation result using the Kou model, parameters: $C, \gamma[0], \lambda$. .............. 106

4.2. Simulation result using the Kou model, parameters: $p, \eta_1, \eta_2$. .............. 106
### List of Figures

1.1. Example of log returns and estimated threshold, generated by the Merton model ........................................ 4

3.1. Simulation result using the Merton model, sample distribution and asymptotic approximation, parameters: $C = 0.09$, $\gamma[0] = 0$, $\beta = 0$, $\eta = 0.6$ and $\lambda = 5$. ................................. 82

3.2. Example of a stock price development driven by the Merton model, parameters: $C = 0.09$, $\gamma[0] = 0.035$, $\beta = 0$, $\eta = 0.1$ and $\lambda = 10$. ................................. 83

3.3. Example of the log returns computed by the Merton model and the estimated threshold, parameters: $C = 0.09$, $\gamma[0] = 0.035$, $\beta = 0$, $\eta = 0.1$ and $\lambda = 10$. ................................. 83

3.4. Example of $\nu(z)$ and $\hat{\nu}_c(z)$ for the Merton model, parameters $C = 0.09$, $\gamma[0] = 0.035$, $\beta = 0$, $\eta = 0.1$ and $\lambda = 10$. ................................. 84

3.5. Simulation result using the Merton model, sample distribution and asymptotic approximation, parameters $C = 0.09$, $\gamma[0] = 0.035$, $\beta = 0$, $\eta = 0.1$ and $\lambda = 10$. ................................. 85

3.6. Log returns and the critical regions that show up by the different steps of the iteration method ............................ 88

4.1. Example of a stock price development driven by the Kou model, Parameters: $C = 0.09$, $\gamma[0] = 0.035$, $\lambda = 20$, $p = 0.4$, $\eta_1 = 12$, $\eta_2 = 11$. ................................. 103

4.2. Example of the log returns computed by the Kou model and the estimated threshold, parameters: $C = 0.09$, $\gamma[0] = 0.035$, $\lambda = 20$, $p = 0.4$, $\eta_1 = 12$, $\eta_2 = 11$. ................................. 105

4.3. Simulation result using the Kou model, sample distribution and asymptotic approximation, parameters: $C = 0.09$, $\gamma[0] = 0.035$ and $\lambda = 20$. ................................. 107

4.4. Simulation result using the Kou model, sample distribution and asymptotic approximation, parameters: $p = 0.4$, $\eta_1 = 12$ and $\eta_2 = 11$. ................................. 108
Zusammenfassung

Ziel dieser Dissertation ist es, ein neues Schätzverfahren für die Charakteristik eines Lévy-Prozesses vorzustellen. Lévy-Prozesse spielen eine bedeutende Rolle in der Finanzmathematik bei der Modellierung von Wertpapierkursverläufen. Sie sind definiert als eine Klasse von stochastischen Prozessen, die unabhängige und identisch verteilte Zuwächse haben, stochastisch stetig sind und bei null starten. Es lässt sich zeigen, dass sich jeder Lévy-Prozess aus einer Brown’schen Bewegung und einem reinen Sprungprozess zusammensetzt.


Zusammenfassung


Es existieren zwei verschiedene Typen von Lévy-Prozessen: Lévy-Prozesse mit endlicher Sprungaktivität und Lévy-Prozesse mit unendlicher Sprungaktivität. Unter endlicher Sprungaktivität ist zu verstehen, dass auf einem endlichen Zeitintervall fast sicher nur endlich viele Sprünge auftreten. Im Gegensatz dazu treten bei Prozessen mit unendlicher Sprungaktivität auf jedem kompakten Zeitintervall fast sicher unendlich viele Sprünge auf. Allerdings muss die Summe der aufaddierten quadrierten Sprunghöhen auf einem endlichen Zeitintervall trotzdem noch endlich sein. Da Lévy-Prozesse mit unendlicher Sprungaktivität schwieriger handhabbar sind, erhalten wir für diese beiden Typen unterschiedliche Resultate.

Im Falle der endlichen Sprungaktivität ist der in dieser Arbeit vorgeschlagene Schätzer von \( C \) asymptotisch normalverteilt; das Gleiche gilt für den Schätzer von \( \gamma \). Außerdem schätzen wir die Funktion \( z \mapsto \nu((\infty, z]), z \in \mathbb{R} \). Für feste Werte \( z \) erhalten wir asymptotische Normalität. Man kann den Schätzer auch als einen stochastischen Prozess in \( z \) auffassen. Wir zeigen, dass dieser schwach gegen einen Gauß-Prozess konvergiert.

Der Blumenthal-Getoor-Index ist eine Größe, die im Falle der unendlichen Sprungaktivität das Verhalten der unendlich vielen Sprünge misst. Wir entwickeln einen Schätzer und zeigen, dass dieser konsistent ist.


Obwohl die Schätzer asymptotisch unverzerrt sind, tritt bei der Anwendung häufig ein Bias auf. Intuitiv ist dies offensichtlich, da kleine Sprünge, die sich innerhalb der kritischen Region befinden, einfach ignoriert werden. Somit wird beispielsweise die Sprungintensität tendenziell eher unterschätzt. Für dieses Problem haben wir für den Fall, dass der Prozess eindimensional ist und endliche Sprungaktivität hat, einen Lösungsvorschlag entwickelt. Dieser basiert auf einem Extrapolationsansatz.


Oft benutzt man in der Finanzmathematik parametrische Lévy-Modelle, d.h. das Lévy-Maß wird vollständig durch Parameter beschrieben. Beliebt ist u.a. das Kou-Modell; bei diesem wird angenommen, dass die Sprünge doppelt asymmetrisch exponentiell verteilt sind. Somit lassen sich asymmetrische Sprungverteilungen gut modellieren. Wir verwenden unsere Methode, um Schätzer für diese
Zusammenfassung

Parameter abzuleiten und zeigen, dass diese asymptotisch normalverteilt sind. Aber auch hier gilt, dass die Schätzer bei der praktischen Anwendung häufig verzerrt sind. Deshalb entwickeln wir ein weiteres Korrekturverfahren speziell für das Kou-Modell, das die vollständige Parametrisierung des Prozesses ausnutzt. Wir testen die Schätzer und das Korrekturverfahren in Simulationen mit dem Ergebnis, dass sich die Verzerrung für die Sprungparameter durch die Korrektur deutlich reduzieren lässt. Es erhöht sich zwar die Schätzvarianz, der MSE allerdings verkleinert sich.
Acknowledgements

First of all, I would like thank my supervisor Prof. Dr. Ulrich Stadtmüller most sincerely. He has always been interested in my work and willing to discuss about research problems in minute detail. I learned a lot in our meetings and I adopted a lot of his methodology.

I also want to express my gratitude to my second adviser Prof. Dr. Rüdiger Kiesel. With his great expertise in mathematical Finance he perfectly complemented the supervision.

I really appreciate the scholarship which I received as a member of the Research Training Group 1100 from the Deutsche Forschungsgemeinschaft (DFG). The Research Training Group 1100 offers a pleasant and interesting research and work environment. I would particulary like to thank the head Prof. Dr. Karsten Urban for his great commitment.

Furthermore, I would like to thank Prof. Yacine Aït-Sahalia from Princeton University for hosting me at the Bendheim Center for Finance in fall 2010. It was a great experience to visit Princeton University.

I would like to extend my thanks to Prof. Cecilia Mancini from the University of Florence. She followed my invitation to give a talk in the seminary of the Research Training Group 1100. We had some similar ideas, so the discussion was very productive.

I also would like to thank my colleagues at the Research Training Group 1100 and at the Institute of Number Theory and Probability Theory. We had numerous discussions concerning our research work and had a good time together.

Finally, I would like to thank my family. Without their constant support and encouragement, this thesis would not have been possible.

Ulm, August 2011

Achim Gegler
Declaration

I hereby declare that this thesis was performed and written on my own and that references and resources used within this work have been explicitly indicated.

I am aware that making a false declaration may have serious consequences.

Ulm, August 17, 2011

(Signature)