TECHNISCHE UNIVERSITÄT WIEN
Vienna University of Technology

DIPLOMARBEIT

# On an Approach for Analyzing the Jump Activity Index of High Frequency Data

Ausgeführt am Institut für
Wirtschaftsmathematik
der Technischen Universität Wien

unter der Anleitung von
Ao.Univ.Prof. Dipl.-Ing. Dr.techn. Friedrich Hubalek

durch
Birgit Ondra
Siccardsburggasse 40/120
1100 Wien

Datum

Unterschrift

**Zusammenfassung**

Diese Diplomarbeit handelt vom 'jump activity index' $\beta$, welcher für beliebige Semimartingale definiert ist. Der Index $\beta$ ist eine Erweiterung des Blumenthal-Getoor Index, welcher für Lévy Prozesse berechnet werden kann. Durch den 'jump activity index' wird das Sprungverhalten von unendlich vielen kleinen Sprüngen eines Semimartingals beschrieben. Gibt es nur endlich viele Sprünge, so gilt $\beta = 0$. Dieser Index wurde von Aït-Sahalia und Jacod definiert und in ihrer Publikation 'Estimating the Degree of Activity of Jumps in High Frequency Data' [2], zeigen sie Resultate seines asymptotischen Verhaltens. In dieser Arbeit besprechen wir zwei ihrer zentralen Ergebnisse aus der soeben genannten Publikation, nachdem wir einen kurzen Einblick in die Entwicklung der Verwendung von high frequency data geben und einige notwendige mathematische Grundlagen beschreiben. Wir befassen uns hierbei mit Semimartingalen, mit der stabilen Konvergenz und mit stabilen Prozessen. Bei der Abhandlung der Beweise, versuchen wir jeden Schritt einzeln und klar verständlich darzustellen. Abschließend implementieren wir das Diskutierte in MATLAB um Ergebnisse auch graphisch darzustellen.

**Abstract**

This thesis deals with the topic of the jump activity index $\beta$, which is defined for any generic semimartingale. The index $\beta$ is an extension of the Blumenthal-Getoor index, which can be calculated for Lévy Processes only. The jump activity index $\beta$ characterizes the behavior of small jumps of semimartingales, in the case of infinitely many jumps being present. If the process only shows finitely many jumps, we get $\beta = 0$. The index is defined by Aït-Sahalia and Jacod and they show details about its asymptotic behavior in their paper 'Estimating the Degree of Activity of Jumps in High Frequency Data' [2]. We discuss two of the main results of this paper, after we give a brief overview of the development of the usage of high frequency data and the corresponding mathematical theory. By doing so, we provide some details on semimartingales, stable convergence and stable processes. In the discussion of the proofs we try to be as mindful as possible and to fragment them into steps which are easy to understand. Finally we provide some implementations in MATLAB, to allow for a visualization of the discussed.

# Contents

# Acknowledgments

First I want to express my gratitude to my supervisor Friedrich Hubalek, who supported me throughout the whole process of writing my thesis. He was always interested in my progress and I am very grateful for him providing more literature where needed.

Further, I want to thank my parents, Brigitte and Martin Ondra, for supporting me throughout my studies here at TU Wien and at Brunel University London, both morally and financially.

Last but not least I am grateful for all the helping words from my boyfriend, Michael Wallner. He was there for me in every phase of writing my thesis. Without him my thesis would not be of the same standard, as he always encouraged me to pursue my goals.

# Chapter 1

# Introduction

The increasing availability of high frequency data in combination with rising computer power has influenced the direction of academic interest with regard to the analysis of stochastic processes. The reason being that such changes allow for the implementation and usage of asymptotic results, which have rather built nice theory earlier. One such area which now combines theory and modeling of financial data, is formed by the analysis of jumps in high frequency data. As more models allowing for jumps seem to become popular, it is also essential to choose the right model for the empirical observations. When using Doob's words, we know that *a stochastic process is the mathematical abstraction of an empirical process whose development is governed by probabilistic law* [18]. This choice of the right abstraction and a suitable underlying law, is what can be supported by first applying certain estimators on the available high frequency data. Amongst other information which can be obtained, we can find out if jumps are present and, in the case of infinitely many being contained in the underlying process, we can categorize the jump behavior with an index $\beta \in (0, 2)$, which is not necessarily deterministic. By Aït-Sahalia and Jacod this index is called the 'jump activity index'. If there are only finitely many jumps, then $\beta = 0$. Contrarily, if we had $\beta = 2$ the paths would be continuous. In the latter case, one could imagine that there are so many arbitrarily small jumps, that they result in making the sample paths continuous. Hence, the index $\beta$ describes how the jumps of the underlying process behave, when categorizing them at some point in between these two extreme cases.

The jump activity index $\beta$ for a genuine semimartingale was defined by Aït-Sahalia and Jacod in 2008, in their paper 'Estimating the Degree of Activity of Jumps in High Frequency Data' [2]. In an earlier paper [1], they already analyze a Lévy process, being made up of the sum of a stable Lévy process and an independent other Lévy process. Therein they talk about characterizing the jump activity of the first. They also mention the Blumenthal-Getoor index, which coincides with the jump activity index, if the underlying process is Lévy. Other publications written on the topic of the jump behavior refer to the Blumenthal-Getoor index too, e.g. [33, 41]. In this thesis, however, we will focus on two main results of the paper [2] by

Aït-Sahalia and Jacod. We see that their choice of a stable process as in [1] is not maintained as such in [2], but we will see that the jump behavior close to zero is assumed to come from a Lévy process.

The jump activity index is defined as the infimum over the exponents $r > 0$ of the jumps, which still makes the sum over the jumps to the power of $r$ finite. Hence, we need to find out if the stated sum is finite or infinite, to determine the jump activity index. This becomes a more interesting question, when we only consider finitely many observations obtained from high frequency data, but want to conclude if the stated sum is infinite or not. This finite observation of the increments (as we cannot observe the jumps) is what we want to use to deduce if the sum over the jumps to the power of $r$ is finite or not. Since knowing this allows us to derive which value the jump activity index takes. The situation considered in [2] becomes even more challenging, as we allow the underlying semimartingale to consists of a continuous part too. This adds complexity, since the jump activity index is defined by the behavior of arbitrarily small jumps, as there are only finitely many big jumps present in our chosen setting. With the presence of a continuous part, these small jumps might not be seen in the increments however, because they are indistinguishable from the increments coming from the continuous part. To overcome this problem we will develop an estimator only considering the increments bigger than a certain threshold.

This thesis deals with the topic of the jump activity index, defined by Aït-Sahalia and Jacod. We start the next chapter by providing information on the general development of the usage of high frequency data. Then we describe some aspects of semimartingales and stable convergence, when stating various details on the framework often used in connection with high frequency data. Afterwards we focus on the work by Aït-Sahalia and Jacod by giving an overview of what they have done and by describing their approach. In Chapter 3 we solely focus on the aforementioned paper [2] and we discuss the proofs of their two main results within this paper. We will go through them as mindful as possible and we will point out any steps, which we do not manage to break down to the desired level of clarity. We will be doing this using footnotes and remarks. After exploiting the theory we dedicate Chapter 4 to the implementation of the discussed. We choose our implementations in such a way that we can draw a comparison to the results stated in the very same paper [2]. We dedicate the last section on carrying out some simulations not mentioned in the paper. After we conclude we provide all the self-programmed MATLAB code in the appendix. Whenever we make direct citations, we print the text in italics and state the corresponding page of the used source. Due to the depth of this topic we do not state all definitions, but assume them in the way given in [25], if not stated otherwise.

# Chapter 2

# General Aspects of Observing High Frequency Data

We use this chapter to provide a short overview of the development of how high frequency data is used for studying stochastic processes. Furthermore the mathematical environment necessary for such analysis is described. This is not done on an exhaustive level as it would lie beyond the scope of this thesis. Amongst the many sources available on this topic the ones mainly used for this thesis are [24, 25, 29, 37]. As this thesis will focus on one paper of Aït-Sahalia and Jacod, we also give a broader overview of the approach they use for the analysis of high frequency data.

## 2.1 Historical Overview

The increasing power of computers allowing for extensions of existing Monte Carlo methods, along with the more widespread availability of high frequency data, form two main reasons for the academic interest in the limit behavior of discretized processes [24]. The first allows to simulate even very complex functions and numerically approximate the desired expected values. The second can be seen in biology and finance, for example. In the area of biology, electrical or chemical activity can be measured at an ever higher frequency. In finance, prices are sometimes recorded every second or even more frequently, opening new areas like algorithmic trading or the herein mentioned analysis of the high frequency data.

The afore mentioned shows why this area is of interest now, but research in this field has started much earlier. Herein we have in mind the development of asymptotic results for stochastic processes. The connection lies in high frequency data allowing for the observation of ever smaller time steps, which can be formulated to happen for $n \to \infty$. Very relevant asymptotic results can be achieved when looking at the limit of sums of independent rvs and corresponding Central Limit Theorems (CLTs). The earliest CLTs have been formulated by

Bernstein in 1927[1] and by Lévy, who published his earliest results on this topic in 1935[2] [19]. This was even before the term *martingale was introduced into the modern probabilistic literature* in 1939 [19, p.1]. A detailed overview of how CLTs continued to develop till around 1980, can be found in the introduction to the chapter of CLTs in [19, Section 3.1]. These CLTs have often been formulated for martingales.

If we jump a bit further in time, Jacod started formulating CLTs for semimartingales [36]. Semimartingales are a generalization of martingales and Definition 2.2, in the next section, states their properties. In 1994 Jacod was the first to formulate a CLT for high frequency observations [36]. By doing so he set the starting point of theory being developed in connection with high frequency data, as many used his work to derive CLTs for certain specific situations [36]. His work from 1994 is called: 'Limit of random measures associated with the increments of a Brownian semimartingale.' but it was never published [34, 36]. The CLT he developed does not only hold in distribution however, but it converges in a slightly stronger way, which we describe in Definition 2.14. Such convergence is referred to as stable convergence and its definition builds on the work from Rényi [38]. Around this time Jacod was not the only one developing new methods for the analysis of high frequency data and a list naming more than a dozen papers can be found in [6, footnote 1]. There were many different areas developing in connection with high frequency data. How to deal with the problem of noise, was one of these areas. With regard to noise, established methods (e.g. realized variance) did not deliver acceptable results, when market frictions were present and time steps of length 1 minute or less were observed [10]. The topic of how to deal with blurred data sets is still prevailing. Aït-Sahalia and Jacod are dealing with such matter and show their new results on various *test statistics under the consideration when market microstructure noise is present* in their most recent joint paper [6, p.1011].

Many applications were found in finance, where such dense sets of data allowed for new approaches to study various issues related to the market's microstructure. Numerous papers have been dealing with the topics of *price discovery, competition among related markets, strategic behavior of market participants, and modeling of realtime market dynamics* [42]. Also, the volatility of asset returns was of interest. To perform the necessary analysis, the increments of the underlying processes are often observed. In this context, there are two basic types of sums frequently used for an observed process $X := (X_t)_{t\geq 0}$. We show their structure

[1] His main contribution came from the French article 'Sur l'extension du théoréme limite du calcul des probabilités aux sommes de quantités dépendantes' [12].

[2] The first such article by him was called 'Propriétés asymptotiques des sommes de variables aléatoires indépendants et enchaînées' [31].

by using the notation from [24, p.6]:

i) ‘normalized functionals’: $$V'^n(f,X)_t := \sum_{i=1}^{\lfloor t/\Delta_n \rfloor} f(\,(X_{i\Delta_n} - X_{(i-1)\Delta_n})\,/\sqrt{\Delta_n}\,),$$

ii) ‘non-normalized functionals’: $$V^n(f,X)_t := \sum_{i=1}^{\lfloor t/\Delta_n \rfloor} f(X_{i\Delta_n} - X_{(i-1)\Delta_n}).$$

In this display the step size is denoted by $\Delta_n$ and equally sized steps are assumed. The function $f$ may be random and it might also depend on the step size $\Delta_n$ and not only on the increment of the underlying process. The ‘normalized functional’ is showing germane results when the underlying process $X$ is continuous, whereas the ‘non-normalized functional’ is more relevant when looking at processes containing jumps [36]. We will later consider a certain type of the latter functional for our discussion of the jump behavior. In this representation however, we are interested in the asymptotic behavior when the the step size $\Delta_n \to 0$ for $n \to \infty$. Unfortunately, if we look at available data, we often get irregular observation points, when, for example, recording the traded prices, where trades can take place at any time. Allowing for irregular time steps makes the situation more difficult, as these observation times would occur randomly as well.

In the beginning, the research dealing with high frequency data built on models only allowing for continuous sample paths. We see the inclusion of jumps as the next step in the development of the analysis of such data [6]. This can be seen as a necessary step, as more and more models considering jumps became popular. An overview can be found in the book ‘Financial Modeling with Jump Processes’ [17], which provides some details about Lévy processes, Semimartingales and time inhomogeneous jump processes. To find the right model for an underlying process it is necessary to analyze some characteristic properties, such as testing for the presence of jumps and diffusion, the boundedness of variation or the behavior of the jumps [16]. Research is being developed in this area and Aït-Sahalia and Jacod are playing a major role in this movement. Amongst some others of relevance, Barndorff-Nielsen has contributed a lot. Many of his papers deal with realized power and bipower variations. The realized power variation is defined as a specific form of the ‘normalized functional’ $V'^n(f,X)_t$, namely when we choose $f(x) = |x|^p$ for $p > 0$ as shown in [36, Example 3.2], for example. The realized bipower variation is a generalisation thereof and details can be found in various of his papers, see for example [9, 11]. In the earlier he claims to be the first to introduce a method for separating the jump from the continuous part in the quadratic variation. He also points at the related work by Mancini, who considers truncated power variations in order to find an estimator for the quadratic variation coming from the jump component of the underlying process. In one of her papers, for example, she introduces a jump size estimation and one for the integrated infinitesimal variance under specific conditions [33].

The inclusion of jumps has shown to be a necessary extension for modeling financial processes and only allowing for continuous sample paths coming from a Brownian motion may be seen as an *ill fitting assumption* [11, p.2]. Hence, allowing for such a more general model can be considered as essential and, in combination with high frequency data, many properties of the jump process can be analyzed. Later in the thesis we will focus on 'estimating the degree of activity of jumps in high frequency data', which is also the title of the paper [2] by Aït-Sahalia and Jacod, which we will discuss. Before going into any details however, we provide a general overview of the mathematical framework used in connection with high frequency data in the following section.

## 2.2 Mathematical Definitions used for Dealing with High Frequency Data

For the analysis of any data over time, often an appropriate underlying process is assumed. We omit the detailed description of a stochastic basis $(\Omega, \mathscr{F}, \mathbf{F}, \mathbb{P})$, and if not stated otherwise an arbitrary stochastic basis is assumed. All random variables (rvs) and processes are assumed to take values in $\mathbb{R}$ throughout the whole thesis. Beside this probabilistic setting, we want to mention the space $\mathbb{D}(\mathbb{R})$ of all càdlàg[3] functions: $\mathbb{R}_+ \to \mathbb{R}$, which is referred to as the *Skorokhod space*. A detailed description of this space can be found in [25, Chapter VI], but here we only want to note the existence of the *Skorokhod topology*, which is described in Theorem VI.1.14 of the same book. The theorem states the existence of a metrizable topology, which makes $\mathbb{D}(\mathbb{R})$ a Polish space (i.e. it is a complete, separable metric space) and is characterized by the convergence of a sequence $(\alpha_n) \xrightarrow{Sk} \alpha$. A sequence converges in this way, iff for $n \to \infty$:

$$\sup_{0 \leq s} \left|\lambda_n(s) - s\right| \to 0 \qquad \text{and} \qquad \sup_{0 \leq s \leq N} \left|\alpha_n \circ \lambda_n(s) - \alpha(s)\right| \to 0 \quad \forall N \in \mathbb{N}, \tag{2.1}$$

for a sequence of strictly increasing functions $\lambda_n : \mathbb{R}_+ \to \mathbb{R}_+$ with $\lambda_n(0) = 0$ and $\lambda_n(t) \nearrow \infty$ for $t \nearrow \infty$.

In finance, where high frequency data tends to be available more often now, a common choice is to use a Brownian motion to model the observed stochastic process. The underlying process becomes more general, allowing for a better replication of the true world, if a Lévy process or a martingale is assumed, since both have the Brownian motion as a special form. A Lévy process generalizes it in the way of allowing for càdlàg almost surely (a.s.) instead of continuous a.s. sample paths (i.e. jumps are allowed) and it requires the increments to be independent and stationary only and not to be normally distributed. A martingale is connected to a Brownian motion over Lévy's martingale characterization, which says that every continuous martingale with the time as quadric variation, i.e. $[X, X]_t = t$ a.s. for all $t \geq 0$ (which means that $X^2 - [X, X]$ is a martingale), is a Brownian motion. When looking at the

[3] 'continue à droite, limite à gauche', i.e. right continuous with left limits existing.

connection between Lévy processes and martingales we can see the link in [37, Theorem 40] which is cited below.

***Theorem 2.1:*** Any Lévy process $X := (X_t)_{t\geq 0}$ can be shown as the sum of two Lévy processes $X = Y + Z$. Herein $Y := (Y_t)_{t\geq 0}$ is a martingale with bounded jumps fulfilling $Y_t \in L^p$ for all $p > 1$ and the sample paths of $Z := (Z_t)_{t\geq 0}$ have finite variation on all compacts.

If we want more freedom a relatively general approach, yet allowing for enough structure, is given by the set of processes of semimartingales. These processes are special since they have many 'nice' features, such as the space of semimartingales remaining stable under *stopping, localization, 'change of time', 'absolutely continuous change of probability measure' and 'changes of filtration'* [25, p.43]. Amongst all their properties one can even be used to characterize the set of processes: Semimartingales form the largest class of processes which can be used as stochastic integrators for all bounded predictable processes, when still maintaining the 'usual nice' properties (e.g. Lebesgue convergence theorem) [24, 25]. Main properties of semimartingales and the way this last feature can be used to characterize semimartingales are shown in several publications, e.g. [13, 37]. Furthermore semimartingales play an important role in mathematical finance, as the fundamental theorem of asset pricing (FTAP) *says that if no arbitrage is allowed, then the price process should at least be a semimartingale* [24, p.24]. As early as 1980 it was stated that prices must follow semimartingales in order to have no arbitrage opportunities [6, 20]. In this thesis however we shall stick to the traditional definition and notation as used in [25].

***Definition 2.2:*** *We call $X := (X_t)_{t\geq 0}$ a* **semimartingale** *if $X_0$ is a $\mathscr{F}_0$-measurable, finite-valued rv and, for a local martingale $M := (M_t)_{t\geq 0}$ which starts in 0 and a processes of finite variation $A := (A_t)_{t\geq 0}$, we have*

$$X = X_0 + M + A.$$

Note that the decomposition in Definition 2.2 can be taken one step further as for any local martingale M we have a unique decomposition into a continuous local martingale $M^c := (M^c_t)_{t\geq 0}$ and a purely discontinuous part $M^d := (M^d_t)_{t\geq 0}$ [25, Theorem I.4.18]. The term *purely discontinuous* means that $M^d_0 = 0$ and that for any continuous local Martingale $Z := (Z_t)_{t\geq 0}$ the product $ZM^d$ is a local martingale[4] [25, p.40]. Hence, any semimartingale can be written in the following way, as also stated in [25, Proposition I.4.27]:

$$X = X_0 + X^c + M^d + A. \tag{2.2}$$

[4] $ZM^d$ being a local martingale makes $Z$ and $M^d$ orthogonal. This term comes from their predictable quadratic covariation being zero, i.e. $\langle Z, M^d\rangle = 0$. This terminology can provide an explanation of *purely discontinuous* as being 'orthogonal' to 'continuous'.

Above we write $X^c$ for the *continuous martingale part* (instead of $M^c$), as we will refer to this part of the semimartingale by $X^c$ for the remainder of the thesis. For clarity we want to note that a purely discontinuous martingale $M$ can easily contain continuous parts and it is not necessarily made up of its jumps, since there is no guarantee that $\sum_{s\leq t} \Delta M_s$ even converges, where

$$\Delta M_s := M_s - M_{s-}$$

denotes the jump size at time $s \geq 0$. Even if the sum converges it can differ from $M$ as we can see in the example taken from [25, p.40] shown below:

***Example 2.3 (Poisson Process):*** Let $N := (N_t)_{t\geq 0}$ be a Poisson process with intensity function $a(t)$. Then it can be shown that $M_t := N_t - a(t)$ is a purely discontinuous local martingale. With the knowledge of $a(t)$ being continuous it can easily be seen that $\sum_{s\leq t} \Delta M_s = N_t \neq M_t$. Since this example forms a prototype of any purely discontinuous local martingale it becomes clear why $M$ is sometimes also referred to as the *compensated sum of jumps.* ■

We now want to take the decomposition shown in equation (2.2) further. As we are analyzing the jumps of semimartingales, we wish to represent the jumps in a more explicit way. Hence, we will introduce a *jump measure* and we will finally arrive at the *canonical representation* of a semimartingale used by Aït-Sahalia and Jacod in [2], which we discuss in the following chapter. The interested reader is referred to [25, Chapter 2] for more details, as we will proceed in a similar way. First, we give an adapted version (in line with this thesis we only state $\mathbb{R}$ as the image space, and not an arbitrary measurable space) of [25, Definition II.1.3].

***Definition 2.4:*** *The family $\mu := (\mu(\omega; dt, dx) : \omega \in \Omega)$ of nonnegative measures, satisfying $\mu(\omega; \{0\} \times \mathbb{R}) = 0$, is called a* **random measure** *on $\mathbb{R}_+ \times \mathbb{R}$.*

Having defined a random measure we now want to integrate with respect to this measure. For doing so, let $Y$ be an *optional function* on $\tilde{\Omega} := \Omega \times \mathbb{R}_+ \times \mathbb{R}$. A function $Y$ is called optional, if it is measurable with respect to the *optional $\sigma$-field* $\tilde{\mathscr{O}} := \mathscr{O} \times \mathcal{B}(\mathbb{R})$, which is the product of the $\sigma$-algebra generated by all càdlàg adapted processes on $\Omega \times \mathbb{R}_+$ (as in [25, Definition I.1.20], for example) and the Borel $\sigma$-algebra on $\mathbb{R}$. By seeing that $(t, x) \mapsto Y(\omega, t, x)$ is measurable for every $\omega \in \Omega$, we can provide the following definition of integrating with respect to $\mu$, in the way it can be found in [25, Equation (II.1.5)].

***Definition 2.5:*** *Let $Y$ be an optional function and $\mu$ be a random measure. Then we define* *the* **integral process** $Y \star \mu := (Y \star \mu_t)_{t \geq 0}$*, where we also allow* $t = \infty$*, by:*

$$Y \star \mu_t(\omega) := \begin{cases} \int_{[0,t]\times\mathbb{R}} Y(\omega,s,x)\mu(\omega;ds,dx) & \text{if } \int_{[0,t]\times\mathbb{R}} |Y(\omega,s,x)|\mu(\omega;ds,dx) < \infty, \\ \infty & \text{otherwise.} \end{cases}$$

We now use this framework to give the definition of the measure counting the jumps, as done in [25, Proposition II.1.16][5]. For this purpose, we also need to recall the *Dirac measure* at a point $a$, which we denote by $\varepsilon_a$.

***Definition 2.6:*** *For any adapted càdlàg* $\mathbb{R}$ *valued process* $X := (X_t)_{t \geq 0}$ *we define the random measure:*

$$\mu^X(\omega; dt, dx) := \sum_{s \geq 0} \mathbb{1}_{\{\Delta X_s(\omega) \neq 0\}} \varepsilon_{(s, \Delta X_s(\omega))}(dt, dx),$$

*on* $\mathbb{R}_+ \times \mathbb{R}$ *as the* **jump measure** *of* $X$*.*

We clearly see that $\mu^X([0,t], A)$, for $t \geq 0$ and for every Borel set $A \in \mathcal{B}(\mathbb{R})$, is a rv giving the number of jumps, which occur in the interval $[0,t]$ and have a size contained in the set $A$, i.e.:

$$\mu^X([0,t], A) = \mathbb{1}_{\{A\}} \star \mu_t^X = \sum_{s \leq t} \mathbb{1}_{\{A\}}(\Delta X_s).$$

Analogously to the situation of a Poisson process (as in Example 2.3), where the intensity $a(t)$ is also referred to as the compensator, making $M(t) = N(t) - a(t)$ a local martingale, we now proceed by giving the definition of the compensator of a random measure. For this, we first specify a function on $\tilde{\Omega}$ to be called *predictable* if it is measurable with respect to the *predictable* $\sigma$*-field* $\tilde{\mathscr{P}} := \mathscr{P} \times \mathcal{B}(\mathbb{R})$, which is the product of the $\sigma$-algebra generated by all càg[6] adapted processes on $\Omega \times \mathbb{R}_+$ (as, for example, given in [25, Definition I.2.1]) and the Borel $\sigma$-algebra on $\mathbb{R}$. This definition allows for the inclusion $\tilde{\mathscr{P}} \subset \tilde{\mathscr{O}}$, as $\mathscr{P} \subset \mathscr{O}$ which is shown in [25, Proposition II.1.24], meaning that every predictable process is optional. Let us introduce three more terms describing properties of random measures, as done in [25, Definition II.1.6].

***Definition 2.7:*** *With respect to the optional* $\sigma$*-field* $\tilde{\mathscr{O}}$ *and the predictable* $\sigma$*-field* $\tilde{\mathscr{P}}$ *we define the following for a random measure* $\mu$*:*

i) *We call* $\mu$ **optional**, *if the process* $Y \star \mu$ *is optional for every optional function* $Y$*.*

[5] The measure is not defined to be called *jump measure* in Proposition II.1.16, but an equivalent definition including the name *jump measure* can be found in, for example, [21, Theorem 11.15].

[6] 'continue à gauche', i.e. left continuous

ii) *We call* $\mu$ **predictable**, *if* $Y \star \mu$ *is predictable for every predictable function* $Y$.

iii) *A random measure as in* i) *is called* $\tilde{\mathscr{P}}$-$\sigma$-**finite**, *if we can find a strictly positive predictable function* $Y$, *such that* $Y \star \mu_\infty$ *is integrable.*

Now we are ready to generalize the concept compensators. In connection with processes, they are defined in such a way, that the difference of the process and its compensator is a local martingale. Below we extend this definition to introduce compensators of random measures, as it is done in [25, Theorem II.1.8].

***Definition 2.8:*** *For a* $\tilde{\mathscr{P}}$-$\sigma$-*finite random measure* $\mu$ *we get a unique (up to a* $\mathbb{P}$-*null set), predictable random measure* $\boldsymbol{\nu}$, *referred to as the* **compensator of** $\boldsymbol{\mu}$, *as soon as we can show one of the two equivalent properties:*

i) *For every nonnegative* $\tilde{\mathscr{P}}$-*measurable function* $Y$ *on* $\tilde{\Omega}$, *we have:*

$$\mathbb{E}[Y \star \nu_\infty] = \mathbb{E}[Y \star \mu_\infty].$$

ii) *For every* $\tilde{\mathscr{P}}$-*measurable function* $Y$ *which results in* $|Y| \star \mu$ *being locally integrable, we need to get that* $|Y| \star \nu$ *is locally integrable and that the process* $Y \star \nu$ *is the compensator of* $Y \star \mu$.

Note that another commonly used term for $\nu$ is the *dual predictable projection*, as used in [21] and that it is sometimes also referred to as *predictable compensator*. As we will see later, we are also interested in the integral $Y \star (\mu - \nu)$. We can define it in the logical way, to be

$$Y \star (\mu - \nu) := Y \star \mu - Y \star \nu,$$

if the process $|Y| \star \mu$ is locally integrable (as property ii) in Definition 2.8 lets us see that $|Y| \star \nu$ is also well-defined). But if we want to define the integral for any predictable function $Y$ another definition is required. A detailed explanation can be found in [25, Section II.1d] or [21, p.300ff.], for example. We shall just shortly quote the latter, by saying that such an integral can be defined for a predictable process $Y$ if $\int_{\mathbb{R}} |Y(t,x)|\, \nu(\{t\}, dx) < \infty$ for all $t \geq 0$ and if for $\tilde{Y}_t := \int_{\mathbb{R}} Y(t,x)\, \mu(\{t\}, dx) - \int_{\mathbb{R}} Y(t,x)\, \nu(\{t\}, dx)$ we have that $\sqrt{\sum_{s \leq \cdot} (\Delta \tilde{Y}_s)^2}$ is locally integrable. A slightly stronger, but easier to verify, condition, is stated in [24, equation (2.1.16)], namely, we can define the integral, if

$$(Y^2 \wedge |Y|) \star \nu_t < \infty \qquad \forall t > 0. \tag{2.3}$$

As stated in [21, Theorem 7.42], this setup implies the existence of a unique, purely discontinuous local martingale $M$, which fulfills

$$\Delta M = \Delta \tilde{Y}. \tag{2.4}$$

Consequently we denote $M = Y \star (\mu - \nu)$. We will need this compensated integral in order to display all small jumps contained in a semimartingale, as there might be infinitely many jumps of arbitrary small sizes. Before showing this connection, we want to introduce the *characteristics* of a semimartingale to uniquely identify a such a process. To understand the representation we want to mention a subclass of semimartingales, namely, we get a *special semimartingale* if we have the representation:

$$X = X_0 + M + A \qquad \text{as in Definition 2.2 with } A \text{ being predictable.}$$

Such a special semimartingale is always obtained, when the jumps of $X$ are bounded by a certain constant $a > 0$, as shown in [25, Lemma I.4.24]. Hence, we can split a semimartingale into the sum of its jumps bigger than a certain threshold and into a special semimartingale. It is common to use an arbitrary but fixed *truncation function* (i.e. a bounded function $h : \mathbb{R} \to \mathbb{R}$ satisfying $h(x) = x$ in a neighborhood of 0) to extract the big jumps by subtracting $\sum_{s\leq \cdot}(\Delta X_s - h(\Delta X_s))$ from the original process. In this thesis we will only allow

$$h_1(x) := x\mathbb{1}_{\{|x|\leq 1\}},$$

as this is the truncation function used in [2], and it is discussed in the next chapter. This choice of truncation function is the original one, but for various limit theorems a continuous version of $h(x)$ is preferred. The semimartingale without big jumps, can now be presented as:

$$X - \sum_{s\leq \cdot}(\Delta X_s - h_1(\Delta X_s)) = X_0 + M + B, \tag{2.5}$$

where we have a local martingale $M$ starting in 0, and a predictable process of finite variation $B$, which depends on the choice of the truncation function. With the help of this notation we are able to state [25, Definition 2.6] and get the following.

***Definition 2.9:*** *Let $X := (X_t)_{t\geq 0}$ be a semimartingale and choose a certain truncation function (here $x\mathbb{1}_{\{|x|\leq 1\}}$). Further we have*

i) *the predictable process of finite variation $B$ as stated in equation (2.5) above,*

ii) *the continuous process $C$ representing the quadratic variation of the continuous martingale part (i.e. $C = \langle X^c, X^c\rangle$ as will be defined in Theorem 2.11), and*

iii) *the predictable random measure $\nu$ being the compensator of the jump measure $\mu^X$.*

*Then the three predictable functions $(B, C, \nu)$ are called the* **characteristics** *of the semimartingale $X$.*

Note that the triplet $(B, C, \nu)$ needs to satisfy the required property of being predictable in order to uniquely represent a semimartingale [25, p.75]. Using the characteristics, we now introduce a notation, representing the jumps of the semimartingale explicitly. What we show is taken from [25, Theorem II.2.34] and [21, Theorem 11.24 and 11.25].

***Theorem 2.10:*** For a semimartingale $X$ with characteristics $(B, C, \nu)$ and the corresponding jump measure $\mu := \mu^X$ we can write

$$X = X_0 + B + X^c + (x\mathbb{1}_{\{|x|\leq 1\}}) \star (\mu - \nu) + (x\mathbb{1}_{\{|x|>1\}}) \star \mu, \tag{2.6}$$

and refer to it as the **canonical representation**.

*Proof:* Using what we stated in equation (2.5), representing the sum of the jumps bigger than one $\sum_{s\leq \cdot}(\Delta X_s - h_1(\Delta X_s))$ by its jump measure $\mu$ and splitting the martingale $M$ in two parts $X^c + M^d$, like done in equation (2.2), we get:

$$X = X_0 + B + X^c + M^d + (x\mathbb{1}_{\{|x|>1\}}) \star \mu$$

Hence, all we need to verify is the equality of $M^d = (x\mathbb{1}_{\{|x|\leq 1\}}) \star (\mu - \nu)$. First we will show that the right-hand side (r.h.s.) is well-defined. It is enough to show that equation (2.3) holds for the truncation function $h_1(x)$, which holds if:

$$\begin{aligned}(h_1^2 \wedge |h_1|) \star \nu_t &= \big((x^2 \wedge |x|)\mathbb{1}_{\{x\leq 1\}}\big) \star \nu_t \\ &= \big(x^2\mathbb{1}_{\{x\leq 1\}}\big) \star \nu_t \\ &< \big(x^2 \wedge 1)\big) \star \nu_t \overset{!}{<} \infty \qquad \forall t > 0.\end{aligned}$$

This last inequality is commonly known to hold, but it can also be derived by taking the following steps. From Definition 2.8, we know that the compensator $\nu$ needs to yield a locally integrable process $(x^2 \wedge 1) \star \nu_t$ (when using that $x^2 \wedge 1 = |x^2 \wedge 1|$ holds) as soon as $(x^2 \wedge 1) \star \mu_t = \sum_{s\leq t} |\Delta X_s|^2 \wedge 1$ is locally integrable. As we will state below in equation (2.7) the sum over the squared jumps is finite and so is the stated sum if we replace any jumps bigger than 1 by the smaller value 1. Doing this turns the process into a finite-valued one with bounded jumps, making it locally integrable. Hence, the desired property is derived. To show the equality, consider the special semimartingale

$$X' = X - X_0 - (x\mathbb{1}_{\{|x|>1\}}) \star \mu = B + \underbrace{X^c + M^d}_{M},$$

with the corresponding jump measure $\mu^{X'} = \mu\mathbb{1}_{\{|x|\leq 1\}}$ and the compensator $\nu^{X'} = \nu\mathbb{1}_{\{|x|\leq 1\}}$. Let $T$ be a predictable stopping time and since $B$ is predictable and $M$ is a martingale, we can perform the following transformation:

$$\begin{aligned}\Delta B_T\mathbb{1}_{\{T<\infty\}} &= \mathbb{E}\left[\Delta B_T\mathbb{1}_{\{T<\infty\}} \middle| \mathcal{F}_{T-}\right] \\ &= \mathbb{E}\left[\Delta X_T\mathbb{1}_{\{T<\infty\}} \middle| \mathcal{F}_{T-}\right] \\ &= \mathbb{E}\left[\int_{\mathbb{R}} x\mu^{X'}(\{T\}, dx)\mathbb{1}_{\{T<\infty\}} \middle| \mathcal{F}_{T-}\right]\end{aligned}$$

$$= \mathbb{E}\left[ \int_{\mathbb{R}} x\nu^{X'}(\{T\}, dx)\mathbb{1}_{\{T<\infty\}} \middle| \mathcal{F}_{T-} \right]$$
$$= \int_{\mathbb{R}} x\nu^{X'}(\{T\}, dx)\mathbb{1}_{\{T<\infty\}} \qquad \text{a.s.}$$

This implies that $\Delta B_t$ is indistinguishable from $\int_{\mathbb{R}} x\nu^{X'}(\{t\}, dx)$ and we obtain the following equality for all $t \geq 0$:

$$\begin{aligned} \int_{\mathbb{R}} x\mu^{X'}(\{t\}, dx) - \int_{\mathbb{R}} x\nu^{X'}(\{t\}, dx) &= \Delta X_t - \Delta B_t \\ &= \Delta M_t \\ &= \Delta M_t^d. \end{aligned}$$

Now we see the desired equality by the definition of the stochastic integral with respect to $(\mu^{X'} - \nu^{X'})$ as hinted at in equation (2.4). □

Next we want to emphasize *one of the main properties of semimartingales* $X$, [24, p.26], (used in the proof above), namely that the sum over the squared jumps is finite. Hence, the property can be written as:

$$\sum_{s \leq t} |\Delta X_s|^2 < \infty \qquad \forall t \geq 0. \tag{2.7}$$

This result is commonly known and can be found in [25, Section I.4c]. The proofs closely deal with the quadratic variation which forms part of the predictable quadratic variation[7] and coincides for continuous semimartingales. The result below is taken from [25, Theorem I.4.52].

***Theorem 2.11:*** First, for any locally square-integrable martingale $M := (M_t)_{t\geq 0}$ we denote by $\langle M, M\rangle$, the **predictable quadratic variation**, which is the predictable process fulfilling that $M^2 - \langle M, M\rangle$ is a local martingale. Second, for a semimartingale $X := (X_t)_{t\geq 0}$, let $[X, X] := X^2 - X_0^2 - 2X_- \cdot X$ be the **quadratic variation**, with $\cdot$ denoting the stochastic integration with respect to a semimartingale. For any semimartingale $X$ and its continuous part $X^c$ we then have:

$$[X, X]_t = \langle X^c, X^c\rangle_t + \sum_{s\leq t} |\Delta X_s|^2.$$

For the analysis of data, the jumps $\Delta X_s$ themselves cannot be observed. When monitoring a semimartingale $X$ on the time interval $[0, T]$ all we can record instead are the increments of

[7] The general definition would be for covariations but since this detail is not needed within the thesis it is omitted herein.

the whole process. When choosing the step size $\Delta_n$ we denote the corresponding increments by:

$$\Delta_i^n X := X_{i\Delta_n} - X_{(i-1)\Delta_n} \quad \text{for } i = 1, \dots, \lfloor T/\Delta_n \rfloor.$$

Using this definition we can state the following theorem wich is a simplified version of [25, Theorem I.4.47]. The proof makes use of a result stated in [25, Theorem I.4.31]. Namely, that for a semimartingale $X$ and a sequence $(H^n)_{n\in\mathbb{N}}$ of predictable processes converging pointwise to $H$, we have the following convergence: $H^n \cdot X_t \to H \cdot X_t$ in measure, uniformly on finite intervals, i.e. $\sup_{s\le t} |H^n \cdot X_t - H \cdot X_t| \xrightarrow{\mathbb{P}} 0$.

***Theorem 2.12:*** For a semimartingale $X$ and the step size $\Delta_n \to 0$ for $n \to \infty$ we see that the squared sum of increments with step size $\Delta_n$

$$H^n := \sum_{i=1}^{\lfloor T/\Delta_n \rfloor} |\Delta_i^n X|^2$$

converges to the quadratic variation $[X, X]$ in measure, uniformly on finite intervals, i.e. for all $\epsilon > 0$:

$$\lim_{n\to\infty} \mathbb{P}\left(\left\{\sup_{s\le t} \left|H_s^n - [X,X]_s\right| \ \ge \epsilon\right\}\right) = 0 \qquad \forall t \in [0,T].$$

*Proof:* Using the equality $(x-y)^2 = x^2 - y^2 - 2y(x-y)$ we can easily transform the sum over the increments in the following way:

$$\begin{aligned} H^n &= \sum_{i=1}^{\lfloor T/\Delta_n \rfloor} (X_{i\Delta_n} - X_{(i-1)\Delta_n})^2 \\ &= \sum_{i=1}^{\lfloor T/\Delta_n \rfloor} X_{i\Delta_n}^2 - X_{(i-1)\Delta_n}^2 - 2X_{(i-1)\Delta_n}\Delta_i^n X \\ &= X_{\lfloor T/\Delta_n \rfloor \Delta_n}^2 - X_0^2 - 2\sum_{i=1}^{\lfloor T/\Delta_n \rfloor} X_{(i-1)\Delta_n}\Delta_i^n X. \end{aligned}$$

The last sum above can be interpreted as stochastically integrating a predictable step process converging pointwise to the process $X$. Hence [25, Theorem I.4.31] mentioned above gives the desired result. □

The result just stated above is useful in applications, since it ensures that the sum over the squared increments converges. Within the observation of high frequency data the concept of *stable convergence in law* has also proved to be very useful. Sometimes it is just referred to as *stable convergence* and it is bridging the weaker form of convergence in law to the stronger convergence in probability. As commonly known, see for example [30, Definition 17.1], a sequence of rvs $X^n$ converges to $X$ in law (denoted by $X^n \xrightarrow{\mathcal{L}} X$) if its distribution function $F_n$ fulfills

$$\lim_{n\to\infty} F_n(x) = F(x)$$

for every continuity point of the distribution function of the limit (i.e. where $F_-(x) = F(x)$ is fulfilled). Be aware that we do not require the different $X^n$ to live on the same probability space, and denote the expected value with respect to $X^n$ by $\mathbb{E}_n$ just for this setting. We know that the definition above is hence equivalent to:

$$\mathbb{E}_n\left[f(X^n)\right] \to \mathbb{E}\left[f(X)\right] \qquad \forall f \text{ bounded, continuous,}$$

as can be found, for example, on [25, p.348]. The above representation hints at why this type of convergence is sometimes also referred to as *weak convergence*. The reason why convergence in law is sometimes less than one would desire in the analysis of high frequency data can be shown by the following example taken from [36, p.3].

***Example 2.13 (Mixed Normal rv):*** Within the framework of semimartingales we often consider *mixed normal* limits $X^n \xrightarrow{\mathcal{L}} VU$ with two independent rvs $U \sim N(0,1)$ and $V > 0$. We denote this by $X^n \xrightarrow{\mathcal{L}} MN(0,V^2)$, where $MN(0,V^2)$ is called a mixed normal distribution with random variance $V^2$. Since we are sometimes confronted with the situation of not knowing the distribution of $V$, we cannot calculate confidence intervals. The reason being that $X^n \xrightarrow{\mathcal{L}} MN(0,V^2)$ does not imply $(X^n, V^n) \xrightarrow{\mathcal{L}} (VU, V)$, which is essential for achieving $X^n/V^n \xrightarrow{\mathcal{L}} N(0,1)$. ■

Hence, a stronger type of convergence implying the joint convergence in law for any measurable rv $V$ is needed. As we will state below, this property is guaranteed by stable convergence and will be stated in Proposition 2.16. This type of convergence is also 'stronger' than convergence in law in the way, that it requires all the $X^n$ to be defined on the same probability space. This allows us to define the convergence on a subset, which does not necessarily make sense for the convergence in law [24, p.16]. When we show the main results of the paper [2] by Aït-Sahalia and Jacod in the following chapter, we will see that the main results are defined to hold in stable convergence on a subset only.

Following the literature by [7,22,25,35,38] we want to state that some CLTs are obtained using stability and that most known CLTs are stable. If they are not, there exists a subsequence of rvs along which the convergence is stable. The whole concept of stable convergence only started to develop in 1963, with Rényi introducing stable sequences of events $A_n \in \mathscr{F}$ iff for every $B \in \mathscr{F}$ the limit

$$\lim_{n\to\infty} \mathbb{P}(A_n \cap B)$$

exists. In line with this definition he calls a sequence of rvs $X^n$ stable if for every $B \in \mathscr{F}$ with $\mathbb{P}(B) > 0$ the conditional distribution on $B$ converges to a distribution function $F_B$:

$$\lim_{n\to\infty} \mathbb{P}(\{X^n < x\}|B) = F_B(x) \tag{2.8}$$

for every continuity point of $F_B(x)$. In 1978 Aldous and Eagleson first used the above to introduce stable convergence. Namely, if a sequence of rvs $X^n$ fulfills $X^n \xrightarrow{\mathcal{L}} X$ and equation (2.8) it is said to converge stably. Within literature there are several equivalent definitions of stable convergence. The interested reader can find the most detailed version in [25, Definition VIII.5.28], but for the scope of this thesis we use [35, Definition 1.7] to provide a better understanding of the general concept.

***Definition 2.14:*** *A sequence $X^n$ of rvs on $(\Omega, \mathscr{F}, \mathbb{P})$* **converges stably in law** *to a limit $X$ defined on an appropriate extension $(\Omega', \mathscr{F}', \mathbb{P}')$ of the aforementioned probability space if*

$$\lim_{n\to\infty} \mathbb{E}\left[Zf(X^n)\right] = \mathbb{E}[Zf(X)] \tag{2.9}$$

*for every $\mathscr{F}$-measurable, bounded rv $Z$ and any bounded, continuous function $f$. This convergence is denoted by:*

$$X^n \xrightarrow{\mathcal{L}-(s)} X.$$

The above definition can also be formulated for càdlàg processes, as it is done in [23, Section 2.1]. If all $X^n$ are càdlàlg then equation (2.9) must hold for all bounded continuous functions $f$ on he Skorokhod topology $\mathbb{D}(\mathbb{R})$, when $X^n \xrightarrow{\mathcal{L}-(s)} X$ shall hold. Further, with Definition 2.14 we can easily see an interesting property of the stable convergence. Indeed, the limit given in equation (2.9) does not depend on the distribution of the rv but on the rv itself. We shall take a closer look at a simple example, hinted at in [7].

***Example 2.15 (Stable Convergence):*** Let $Y$ and $Y'$ be two independent, identically distributed (i.i.d.) rvs with $\mathbb{P}(\{Y = 1\}) > 0$ and $\mathbb{E}[Y] \neq 1$. We then define a sequence in the following way:

$$X^n := \begin{cases} Y & n \text{ odd}, \\ Y' & n \text{ even}. \end{cases}$$

Obviously we have $X^n \xrightarrow{\mathcal{L}} Y$ being equivalent to $X^n \xrightarrow{\mathcal{L}} Y'$. But when we test the condition in equation (2.9) for $f(x) = x$ and $Z = \mathbb{1}_{\{Y=1\}}$ we get:

$$\mathbb{E}\left[Zf(X^n)\right] = \mathbb{E}[\mathbb{1}_{\{Y=1\}}X^n] = \begin{cases} \mathbb{P}(\{Y = 1\}) & n \text{ odd}, \\ \mathbb{P}(\{Y = 1\})\mathbb{E}[Y'] & n \text{ even}. \end{cases}$$

For $\mathbb{E}[Y] \neq 1$ as defined above we see that the sequence $\mathbb{E}\left[Zf(X^n)\right]$ does not converge and hence we do not have stable convergence. ■

As a last general result on stable convergence we cite two useful consequences as stated in [36, Lemma 2.3 and Proposition 2.5]. The first shows what is needed to have the weaker stable convergence be equivalent to convergence in probability. The second displays how the joint convergence for a mixed normal variable from Example 2.13 can be achieved.

***Proposition 2.16:*** Let $X^n$ be a sequence of rvs, stably converging to a rv $X$ defined on the same space as $X^n$.

i) We have the following equivalence:

$$X^n \xrightarrow{\mathcal{L}-(s)} X \Leftrightarrow X^n \xrightarrow{\mathbb{P}} X.$$

ii) If additionally the sequence $V^n$ and the rv $V$ is defined on the same space we also get the implication:

$$X^n \xrightarrow{\mathcal{L}-(s)} X, V^n \xrightarrow{\mathbb{P}} V \Rightarrow (X^n, V^n) \xrightarrow{\mathcal{L}-(s)} (X, V).$$

## 2.3 Approach used by Aït-Sahalia and Jacod

On the one hand Aït-Sahalia and Jacod are strongly involved in the development of the mathematical background for the analysis of semimartingales. This can be seen in various papers [2–5, etc.] which are written in connection with high frequency data, and also in the book Jacod published (together with Shiryaev) [25], where they analyze limit theorems of stochastic processes and do not miss any mathematical precision. On the other hand, in one of their papers [6], published just last year, they focus on the application of their theory. Therein they provide a short description of how to apply the statistics developed by them, and only refer to their earlier work, including the papers [2–5], for details about the required probability techniques. It can somehow be seen as a summary of their previous work and it is described how their mathematical framework can be used to analyze the log returns of asset prices, given over a finite interval of time $[0, T]$. Such a finite interval, is what we will assume for the remainder of the thesis too.

Within this set situation, they assume to have high frequency data as an input and wish to test which components of a semimartingale are present. They state that only *'regular' sampling schemes are considered* [6, p.1013], meaning that the time interval between two consecutive observations is fixed to be $\Delta_n$. The choice of a fixed step size is made, as arbitrary intervals would imply a much more complicated mathematical model. In their analysis of components of a semimartingale $X := (X_t)_{t\geq 0}$, they use a version of the canonical representation as in equation (2.6) of Theorem 2.10, and wish to find out if the continuous part, small jumps and big jumps are making up part of the underlying process or if they are not present.

To achieve this they introduce so called *truncated power variations*. In these sums (and all the ones of this sort to follow) they use the convention of $0^0 = 0$, so that we only consider real jumps by $|\Delta X_s|^r$ for any value of $r \geq 0$. The first variant of the truncated power variation only considers jumps smaller than certain thresholds $u_n$, which converge to 0, and this variant is denoted by:

$$B(p, u_n, \Delta_n) := \sum_{i=1}^{\lfloor T/\Delta_n \rfloor} |\Delta_i^n X|^p \mathbb{1}_{\{|\Delta_i^n X| \leq u_n\}}. \tag{2.10}$$

The second variant does the opposite by only taking account of the big jumps:

$$U(p, u_n, \Delta_n) := B(p, \infty, \Delta_n) - B(p, u_n, \Delta_n) = \sum_{i=1}^{\lfloor T/\Delta_n \rfloor} |\Delta_i^n X|^p \mathbb{1}_{\{|\Delta_i^n X| > u_n\}}. \tag{2.11}$$

These statistics show three degrees of freedom: the power $p$ to which the jumps are taken, the truncation levels $u_n$ and the step size $\Delta_n$. All three are then altered in appropriate ways to find out if there is a continuous part present or not and if jumps should be included in the model. In the case of jumps being found, they go one step further in wanting to analyze the behavior of jumps, namely, if they have finite or infinite activity. In the case of infinitely many jumps being present, the behavior of the jumps can be described more precisely with the help of the so called *jump activity index*. The analysis of this index is what Aït-Sahalia and Jacod deal with more precisely in their paper [2], in which the following definition can be found on p.2203.

***Definition 2.17:*** *For a generic semimartingale $X := (X_t)_{t\geq 0}$ we define the* **jump activity index** $\beta_t : \Omega \to [0, 2]$ *up to time $t \in [0, T]$ in the three steps:*

$$B(r)_t := \sum_{s\leq t} |\Delta X_s|^r, \qquad I_t := \{r \geq 0 : B(r)_t < \infty\} \qquad \text{and} \qquad \beta_t := \inf(I_t).$$

*This makes $\beta_t$ the infimum of the exponents r which make the sum over all jumps to the power of r finite.*

The jump activity index $t \mapsto \beta_t$ might grow in time as more jumps appear but it is bounded from above by 2 since we saw in equation (2.7) that for $r = 2$ the sum converges to the quadratic variation process. If the process $X$ has finite jump activity, then $\beta_t = 0$, since $B(0)_t$ counts the number of jumps up to time $t$. We cannot conclude the opposite however, i.e. $\beta_t = 0$ does not necessarily mean that the process only has finitely many jumps. An example is stated on [2, p.2206], giving the Gamma process, where the sum over the small jumps is only diverging slowly and hence resulting in $\beta_t$ being 0 in spite of infinitely many jumps occurring. Note also that from the definition we do not know if $\beta_t$ lies in the (random) interval $I_t$ or not. In the special case, when $X$ is a Lévy process, then the interval is non-random and $\beta_t$ does not depend on time. Hence, in this special case, we have $\beta_t \equiv \beta \in [0, 2]$.

This precise behavior of jumps, under the assumption of jumps being present, is what the remaining thesis deals with. If high frequency data is analyzed for its jump behavior, one must first assure that there are jumps present at all, as else the outcome might result in an estimate for the jump activity index being bigger than 2, as stated in [2, Remark 4], which obviously does not make sense. And also, a value of $\beta_t \in [0, 2]$ would not have any meaning. For the work in this thesis however, we will assume that the underlying process does contain jumps. The precise way of testing for jumps being present or not can be found in the paper [3] by Aït-Sahalia and Jacod but we do not include the analysis of this question in our thesis.

# Chapter 3

# Estimating the Jump Activity Index

The aim of this chapter is to discuss the results of the paper 'Estimating the Degree of Activity of Jumps in High Frequency Data' by Aït-Sahalia and Jacod [2]. In their paper they assume to observe the log-price $X$ of an asset at discrete time steps $\Delta_n$ over a time horizon $[0, T]$. As high frequency data is observed they make use of the assumption that $\Delta_n \to 0$ but they stick to only observing the sample paths over a fixed interval in time. By doing so they hope to find out more about the behavior of small jumps, i.e. the nature of a generalized Lévy measure $F_t$ near 0, where Definition 3.3 will state what such a measure needs to fulfill. They wish to construct estimators for the jump activity index $\beta_T$ as described in Definition 2.17 which are consistent for $\Delta_n \to 0$. Additionally they aim at providing rates of convergence and asymptotic distributions. In their work they show these desired results under some restrictions on the behavior of the Lévy measure, stated later in Assumption 3.12. When setting these few assumptions, they claim to stay as model-free as possible for achieving the desired outcome. In what follows we first state the required details of the chosen approach. Then we show the two main results of the paper [2] and finally we provide an extended version of the proofs, stating omitted details wherever possible. This means that we take the proofs from [2] but explain the steps using other sources or putting to paper our own thoughts where we feel that more details are needed and available. We do want to emphasize that this chapter is based on [2] if not stated differently.

## 3.1 Model Assumptions and Definitions

Within the paper the observed process is chosen to be a specific type of a semimartingale, namely an Itô semimartingale. We adopt this choice, as such semimartingales are behaving like Lévy processes, which is not necessarily true for a general semimartingale [24, p.34]. An

Itô semimartingale in the sense of Jacod is obtained when asking for a special property of the characteristics $(B, C, \nu)$, as seen in [24, Definition 2.1.1].

***Definition 3.1:*** *If the characteristics $(B, C, \nu)$ of a semimartingale $X := (X_t)_{t\geq 0}$ are absolutely continuous with respect to the Lebesgue measure, we call $X$ an* **Itô semimartingale**. *Hence, the following representation is possible:*

$$B_t = \int_0^t b_s\, ds, \qquad C_t = \int_0^t \sigma_s^2\, ds, \qquad \nu(dt, dx) = dt\, F_t(dx), \tag{3.1}$$

*with choosing the processes $b := (b_t)_{t\geq 0}$ and $c := (c_t)_{t\geq 0}$ to be predictable and the measure $F_t = F_t(\omega, dx)$ to be a predictable random measure.*

For the description of a predictable random measure, recall Definition 2.7.ii). If we wish, we can now display the canonical representation, as in Theorem 2.10, in a way including $C_t = \int_0^t \sigma_s^2\, ds$, and not only $B$ and $\nu$, like sometimes done by Aït-Sahalia and Jacod. By making use of the Martingale Representation Theorem, see, for example, [28, Section 3.4], one can extend the probability space in an appropriate way and then use the representation

$$X_t^c = \int_0^t \sigma_s\, dW_s, \tag{3.2}$$

with $W := (W_t)_{t\geq 0}$ being a Brownian motion on the extended probability space. Hence we can write:

$$X_t = X_0 + \int_0^t b_s\, ds + \int_0^t \sigma_s\, dW_s + (x\mathbb{1}_{\{|x|\leq 1\}}) \star (\mu - \nu)_t + (x\mathbb{1}_{\{|x|>1\}}) \star \mu_t.$$

The random measure $F_t$ introduced in equation (3.1) above, is referred to as *Lévy measure* by Jacod and Aït-Sahalia. This is because it fulfills what a traditional Lévy measure has as a property. A Lévy measure defined for a Lévy process $X' := (X_t')_{t\geq 0}$, as in [17, Definition 3.4], is the measure $F'$ on $\mathbb{R}$ representing the expected number of jumps in the unit interval (or equivalently in any interval of length one, as a Lévy process has stationary increments):

$$F'(A) := \mathbb{E}[\#\{t \in [0,1] : \Delta X_t \neq 0, \Delta X_t \in A\}], \qquad A \in \mathcal{B}(\mathbb{R}). \tag{3.3}$$

Since for a Lévy process we have $\mathbb{E}[\mu([0,t] \times A)] = \nu([0,t] \times A) = tF'(A)$ for all $A \in \mathcal{B}(\mathbb{R})$ (see [24, p.34], for example) we see that we have $F' = F$ for a Lévy process. The property referred to before can be found in [17, Proposition 3.7] and reads as follows.

***Theorem 3.2:*** A Lévy measure $F'$ fulfills:

$$F'(0) = 0 \qquad \text{and} \qquad \int_{\mathbb{R}} (x^2 \wedge 1) F'(dx) < \infty.$$

*Proof:* Since these properties are basic ones, we only shortly mention that $F'(0) = 0$ due to the definition in equation (3.3), as any jump is by its definition $\neq 0$. Further, the boundedness for the integral is derived from the sum over the squared jumps being finite, as stated in equation (2.7) for semimartingales. $\square$

Hence, we can also state this alternate version of a definition for a Lévy measure, as found on [17, p.27].

***Definition 3.3:*** *Let $F'$ be a Borel measure on $\mathbb{R}$. It is called a* **Lévy measure** *if*

$$F'(0) = 0 \qquad \text{and} \qquad \int_{\mathbb{R}} (x^2 \wedge 1) F'(dx) < \infty. \tag{3.4}$$

In the case of a semimartingale, the so called Lévy measure $F$ is a random measure, but we can choose a version of $F$, which satisfies the above equation (3.4) for each $(\omega, t)$, as the Lévy measure in Aït-Sahalia and Jacod's paper may depend on time and it may be random (in contrast to the Lévy measure $F'$ coming from a Lévy process). For the use within the proofs, we need one more property of the Lévy measure and hence provide one more definition.

***Definition 3.4:*** *With respect to any measure $G$ on $\mathbb{R}$ we introduce the notation*

$$\overline{G}(x) := G([-x, x]^c) \qquad \forall x \geq 0,$$

*representing the* **symmetrical tail function**.

When looking at the Lévy measure $F_t(\omega, dx)$ we see that $\overline{F}_t(\epsilon) < \infty$ is fulfilled (we can choose a version which fulfills this relation identically, see [24, p.35]) for all $\epsilon \in (0, 1]$ . We derive this, by using equation (3.4) and the consecutive simple transformation:

$$\begin{aligned}
\overline{F}_t(\epsilon) &= \int \mathbb{1}_{\{|x|>\epsilon\}} \, F_t(dx) \\
&= \frac{\epsilon^2}{\epsilon^2} \int \mathbb{1}_{\{|x|>\epsilon\}} \, F_t(dx) \\
&= \frac{1}{\epsilon^2} \int \mathbb{1}_{\{|x|>\epsilon\}} (x^2 \wedge \epsilon^2) \, F_t(dx) \\
&\leq \frac{1}{\epsilon^2} \int \mathbb{1}_{\{|x|>\epsilon\}} (x^2 \wedge 1) \, F_t(dx) \\
&\leq \frac{1}{\epsilon^2} \int (x^2 \wedge 1) \, F_t(dx) \overset{(3.4)}{<} \infty.
\end{aligned}$$

Now we state the first of the two assumptions as required for carrying out the proofs below. In the paper Aït-Sahalia and Jacod initially give a version only asking for local boundedness but by referring to a standard localization procedure, to be found in [23], and the fact that the main results are 'local' in time too, they finally arrive at the condition stated below.

***Assumption 3.5:*** *Let $L > 0$ be a constant. We assume that both $b$ and $\sigma$ of the characteristics of $X$ are bounded by $L$.*

The above assumption gives some weak restriction on two of the functions in the characteristics $(B, C, \nu)$. We also need to ask for more structure of the measure $\nu(dt, dx) = dt\, F_t(dx)$. Therefore we introduce two conditions. The first goes as follows.

***Definition 3.6:*** *Two measures $\mu_1$ and $\mu_2$ on the same space $(\Omega, \mathscr{F}, \mathbb{P})$ are called* **singular** *to one another (denoted by $\mu_1 \perp \mu_2$) if there exists a set $A \in \mathscr{F}$ with:*

$$\mu_1(A) = 1 \qquad \text{and} \qquad \mu_2(A^c) = 1.$$

Further, we need the definition of *stable processes* which form a subclass of Lévy processes. They are a generalization of a Brownian motion in the way that a Brownian motion $W := (W_t)_{t\geq 0}$ fulfills the condition of being *selfsimilar*, i.e. for all $a > 0$ we have that the following equality in distribution holds[1]:

$$(W_{at}/\sqrt{a})_{t\geq 0} \overset{d}{=} (W_t)_{t\geq 0}.$$

This property can be derived when the definition of a stable rv, stated below and taken from [17, Section 3.7], is transferred to a Lévy process [40, Definition 13.2].

***Definition 3.7:*** *A rv $U$ taking values in $\mathbb{R}$ with characteristic function $\varphi_U(z) := \mathbb{E}[e^{izU}]$, follows a* **stable** *distribution iff for all constants $a > 0$ there are two constants $b(a) > 0$ and $c(a) \in \mathbb{R}$ resulting in the equality for the characteristic function $\varphi_U$:*

$$\varphi_U(z)^a = \varphi_U(zb(a))e^{ic(a)z}, \qquad \forall z \in \mathbb{R}. \tag{3.5}$$

*If we have $c(a) = 0$ then the distribution is called* **strictly stable***.*

***Definition 3.8:*** *Let $X := (X_t)_{t\geq 0}$ be a Levy process with its distribution function at time 1 being the one of a stable rv, i.e. $X_1 \overset{d}{=} U$ from Definition 3.7. Then the process $X$ is called a* **stable process***.*

From the above definition we can easily show the selfsimilarity of a stable process. When we have a Lévy process $X$ we have, as can be found in [8, Theorem 1.3.3], its characteristic function complying with:

$$\varphi_{X_t} = e^{t/s \ln\left(\varphi_{X_s}\right)} \qquad \forall s, t > 0, \tag{3.6}$$

[1]The equality in distribution holds for the whole process and not only for every rv at time $t$.

i.e. when setting $s = 1$ we see that the dependence on time of a Lévy process is linear in the exponent of its characteristic function. Hence, we can derive that a Lévy process is fully characterized when knowing the law of $X_1$. Using this and the fact that a characteristic function fully defines the distribution, we can derive that a Lévy process $X_t$ is stable iff for all constants $a > 0$ there are two constants $b(a) > 0$ and $c(a) \in \mathbb{R}$ conforming with:

$$(X_{at})_{t\geq 0} \stackrel{d}{=} (b(a)X_t + c(a)t)_{t\geq 0}. \tag{3.7}$$

This property is sometimes referred to as *selfsimilarity up to a translation*. We show that the above is fulfilled by assuming $X_1$ to follow a stable distribution and using Definition 3.7 above. We assume the constants to be as in the stated definition and get:

$$\begin{aligned}
\varphi_{X_{at}}(z) &\stackrel{(3.6)}{=} \varphi_{X_1}(z)^{at} \\
&\stackrel{(3.5)}{=} \left(\varphi_{X_1}(zb(a))e^{ic(a)z}\right)^t \\
&\stackrel{(3.6)}{=} \varphi_{X_t}(zb(a))e^{ic(a)zt} \\
&= \mathbb{E}\left[e^{izb(a)X_t}e^{ic(a)zt}\right] \\
&= \mathbb{E}\left[e^{iz(b(a)X_t+c(a)t)}\right] \\
&= \varphi_{b(a)X_t+c(a)t}.
\end{aligned}$$

In this we see that a stable Lévy process is also selfsimilar (up to a translation). The proposition below, taken from [17, p.97] is even more detailed as it includes that $b(a)$ will always take the form $a^{1/\alpha}$ (which is proven in [39, Corollary 2.1.3]):

***Proposition 3.9:*** Let $X := (X_t)_{t\geq 0}$ be a stable Lévy process and $\alpha \in (0, 2]$. Then we call $X$ an **$\boldsymbol{\alpha}$-stable** Lévy process iff for all $a > 0$ there is a constant $c \in \mathbb{R}$ such that:

$$(X_{at})_{t\geq 0} \stackrel{d}{=} (a^{1/\alpha}X_t + ct)_{t\geq 0} \qquad \forall t \geq 0. \tag{3.8}$$

If we have $\alpha = 2$ then we get a Brownian motion with drift $(W_t + ct)_{t\geq 0}$. For $c = 0$ we refer to $X_t$ as being **strictly $\boldsymbol{\alpha}$-stable**.

To get an idea of the influence of parameter $\alpha$, we provide sample paths for the values $\alpha \in \{0.1, 1.0, 1.9\}$ in Figure 3.1. The simulation was carried out using the Matlab file `sPlotStable`, which can be found in the appendix. In the graphs we clearly see that the closer the value $\alpha$ is to 2, the more the process reminds us of a Brownian motion. But as long as $\alpha \neq 2$ jumps appear. Contrarily, when $\alpha$ is closer to 0, more of the very small jumps appear in the $\alpha$-stable process and big jumps become more likely too. When looking at this process at a big scale it looks similar to a Poisson process, despite being made up of infinitely many jumps. For any stable process several properties can be derived and some can be found, for example, in [8, 17, 39]. One being relevant to this thesis can be derived from [40, p.80] and

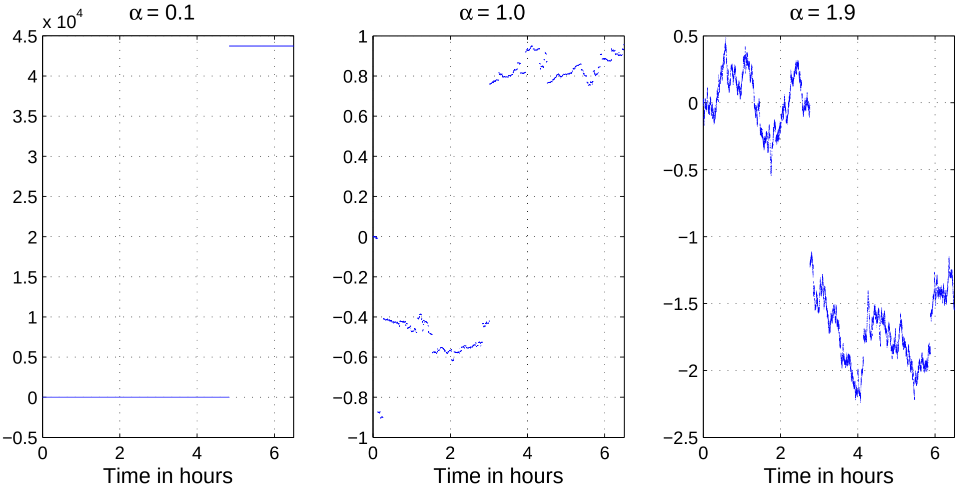


Figure 3.1: Sample path of an $\alpha$-stable process over 6.5 hours, with increments simulated every second.

is stated thereafter. It says that for $\alpha \in (0,2)$ the Lévy measure $F$ of an $\alpha$-stable, real-valued variable takes the form:

$$F(dx) = \frac{A}{|x|^{\alpha+1}} \mathbb{1}_{\{x<0\}}\,dx + \frac{B}{x^{\alpha+1}} \mathbb{1}_{\{x>0\}}\,dx, \tag{3.9}$$

for some $A, B > 0$. It is interesting to note that the density function of any $\alpha$-stable rv is generally not known in a closed form in terms of elementary functions. Only the Gaussian, the Cauchy and the Lévy distribution can be displayed in closed form and can be found on [17, p.98]. What we can see from the above representation in (3.9) is, that the jump activity index $\beta$ will coincide with the stability index $\alpha$. The reason being that we can carry out the following transformation when $X := (X_t)_{t\geq 0}$ is an $\alpha$-stable Lévy process and using that its Lévy measure is independent of time and not stochastic. Since we know that there are only finitely many big jumps (due to the càdlàg-property), it is enough to look at the jumps with absolute value being smaller than 1 only in order to derive the jump activity index $\beta$. First let us denote the sum of such jumps in a similar way to Definition 2.17:

$$B(r)'_t := \sum_{s\leq t} |\Delta X_s|^r \mathbb{1}_{\{|\Delta X_s|<1\}} = \int_0^t \int_{-1}^1 |x|^r \mu(ds, dx).$$

We now use the property of equal expected values from the jump measure $\mu$ and its compensator $\nu$. Further, we take advantage of the special form $\nu(ds, dx) = ds\, F(dx)$ (as $X$ is a Lévy process), with the Lévy measure $F$ being deterministic. Now we can derive:

$$\begin{aligned}\mathbb{E}[B(r)'_t] &= \mathbb{E}\left[\int_0^t ds \int_{-1}^1 |x|^r\, F(dx)\right] \\ &\overset{(3.9)}{=} t \int_{-1}^0 A|x|^{r-\alpha-1}\,dx + t\int_0^1 B|x|^{r-\alpha-1}\,dx\end{aligned}$$

$$= \begin{cases} \dfrac{t(A+B)}{r-\alpha} & \text{if } r \in (\alpha, \infty), \\ \infty & \text{if } r \leq \alpha. \end{cases}$$

So we have a finite expected value for $r$ bigger than $\alpha$. As the jump activity index is defined on the paths and not the expected value of $B(r)'_t$, we need to see the equivalence between the two. If we have a finite expected value, we easily see that the integrand needs to be finite a.s. (see, for example, [30, Lemma 9.26]). On the other hand, if the expected value is infinite, we need to show that $B(r)'_t = \infty$ a.s. The initial proof thereof comes from Blumenthal and Getoor and can be found in [14, Theorem 4.1]. These two authors were the eponym for the *Blumenthal-Getoor index*, which they define in Definition 2.1 in the same paper and they simply call it *index* $\beta$. The difference to the jump activity index $\beta$ in our context is, that the Blumenthal-Getoor index is only defined for Lévy processes. Coming back to showing the second direction, we will not state the proof of Blumenthal and Getoor however, but instead follow [24, Lemma A.2], which shows a general property of any Lévy process by not introducing any additional notation.

***Lemma 3.10:*** For any Lévy process $X := (X_t)_{t\geq 0}$ with Lévy measure $F$ and any $r \in [0, 2)$ the following implication holds:

$$\int_{-1}^{1} |x|^r \, F(dx) = \infty \quad \Rightarrow \quad B(r)'_t = \infty \text{ a.s.} \quad \forall t > 0.$$

*Proof:* To show the desired we make use of two standard results. Namely, the fact that the jump measure $\mu$ is a Poisson random measure (see, for example [15, Proposition 7.5.15]) and the form of the 'Laplace functional' of a Poisson random measure. This can be found, for example, in [15, Theorem 6.2.9], and allows for the following representation for any nonnegative measurable function $f$ on $\mathbb{R}_+ \times \mathbb{R}$:

$$\mathbb{E}\left[\exp\left(-\int f \, d\mu\right)\right] = \exp\left(-\int (1 - e^{-f}) \, d\nu\right).$$

We substitute $f(s, x) = |x|^r \mathbb{1}_{\{|x|<1\}} \mathbb{1}_{\{s \leq t\}}$ and derive $\int f \, d\mu = B(r)'_t$. Using the Laplace functional above we get:

$$\mathbb{E}\left[\exp\left(-B(r)'_t\right)\right] = \exp\left(-t \int_{-1}^{1} (1 - e^{-|x|^r}) \, F(dx)\right). \tag{3.10}$$

Now we know that $\int_{-1}^{1} |x|^r \, F(dx) = \infty$ and we see that the difference between the integrand on the r.h.s. above and herein is smaller than $K|x|^{2r}$ by writing:

$$(1 - e^{-|x|^r}) - |x|^r = 1 - \sum_{j=0}^{\infty} \frac{(-|x|^r)^j}{j!} - |x|^r = -\sum_{j=2}^{\infty} \frac{(-|x|^r)^j}{j!}.$$

By seeing that the term $|x|^r$ is hence in both integrands, this implies that the integral on the r.h.s. of (3.10) is infinite too, as $\int_{-1}^{1} |x|^r \, F(dx) = \infty$ holds. Therefore, the r.h.s. of (3.10) becomes equal to 0. As equation (3.10) must hold, we deduce that $B(r)'_t = \infty$ a.s., which concludes the proof. □

Summing up the above we have shown that only for $r \in (\alpha, \infty)$ the sample paths in $B(r)'_t$ are finite. As $B(r)_t$ is finite for the same values of $r$ we state the following:

***Lemma 3.11:*** For an $\alpha$-stable Lévy process $X := (X_t)_{t\geq 0}$ the value $\alpha \in (0, 2]$ coincides with the jump activity index $\beta$. Consequently we will refer to $X$ as a $\beta$-stable process in the remainder of the thesis.

We are now ready to formulate the assumption on the Lévy measure of the semimartingale $X$, which is split into two parts, denoted by $F_t = F'_t + F''_t$. Herein the first part is assumed to behave like the Lévy measure of a symmetric $\beta$-stable process near 0 and it is the one being responsible for the jump activity index $\beta$. The second part is assumed to be any type of Lévy measure with jump activity index $\beta'$ fulfilling $\beta' < \beta$. By Definition 2.17 of the jump activity index we see that $\beta$ of $F'_t$ is also the overall jump activity index of $X$ as only the bigger value of the two separate jump activity indices allows $B(\beta)_t < \infty$ (and we have $B(\beta')_t = \infty$) for the overall process $X$. For a formal definition we introduce some further notation below.

***Assumption 3.12:*** *We assume to have four constants $\beta \in (0, 2)$, $\beta' \in [0, \beta)$, $\gamma > 0$ and $L \in [1, \infty)$. The two parts of the Lévy measure $F'_t$ and $F''_t$ are singular, with:*

$$F_t = F'_t + F''_t.$$

*By $\Phi \subseteq \Omega \times (0, \infty) \times \mathbb{R}$ we denote the set showing that they are indeed singular. The set shows where the measures are equal to 0 in the way that*

$$F''_t(\omega, A) = 0 \qquad \forall A \notin \{x : (\omega, t, x) \in \Phi\},$$

*and for $F'_t$ the opposite holds*

$$F'_t(\omega, A) = 0 \qquad \forall A \notin \{x : (\omega, t, x) \in \Phi^c\}.$$

*For the parts of the Lévy measure we have:*

i) *For $F'_t$ there are predictable, nonnegative processes $a_t^{(+)}$, $a_t^{(-)}$, $z_t^{(+)}$, $z_t^{(-)}$ and a predictable function $f(\omega, t, x)$ which fulfill*

$$\text{a) } \frac{1}{L} \leq z_t^{(+)} \leq 1 \qquad \text{and} \qquad \frac{1}{L} \leq z_t^{(-)} \leq 1, \tag{3.11}$$

$$\text{b) } a_t^{(+)} + a_t^{(-)} \leq L, \tag{3.12}$$

$$\text{c) } 1 + |x| f(t, x) \geq 0, \tag{3.13}$$

$$\text{d) } |f(t, x)| \leq L, \tag{3.14}$$

*and allow for the representation:*

$$F'_t(dx) = \frac{1 + |x|^\gamma f(t, x)}{|x|^{1+\beta}} \left( a_t^{(+)} \mathbb{1}_{\left\{0 < x \leq z_t^{(+)}\right\}} + a_t^{(-)} \mathbb{1}_{\left\{-z_t^{(-)} < x \leq 0\right\}} \right) dx.$$

ii) *For $F''_t$ we only require that the following holds:*

$$\int_{\mathbb{R}} (|x|^{\beta'} \wedge 1) F''_t(dx) < L.$$

In order to see where $X$ has a jump activity index of $\beta$ we introduce the increasing, locally bounded processes:

$$A_t := \frac{a_t^{(+)} + a_t^{(-)}}{\beta} \qquad \text{and} \qquad \overline{A}_t := \int_0^t A_s \, ds, \tag{3.15}$$

allowing for the statement that we have $\beta_t = \beta$ on $\{\overline{A}_t > 0\}$ and else we obtain $\beta_t \leq \beta'$. Having stated all technical assumption we want to move on to describing the estimators now.

By choosing the observed process to follow a semimartingale we have a continuous martingale part present as well. When wanting to observe jumps this does not make the situation easier. As $\beta_T$ characterizes the behavior of the small jumps (as there are only finitely many bigger than any constant $\delta > 0$) it would be natural that small increments $\Delta X_i^n$ provide most information about the jumps. But in the used setting these small jumps also include the changes of the continuous part. Trying to get a good estimate for $\beta_T$ despite only obtaining this 'blurred picture' is what formed the challenge for Aït-Sahalia and Jacod in their paper. They chose to introduce an estimator only counting the jumps bigger than a certain threshold which is decreasing as the step size decreases. This is in line with their general approach of using truncated power variations as shortly described in Section 2.3. Here we define a special version of $U(p, u_n, \Delta_n)$, given in equation (2.11). In following their approach we set the first estimator in the way described below.

***Definition 3.13:*** *Let $\varpi \in (0, 1/2)$ and $\alpha > 0$ be two constants. For $t \in [0, T]$ and $n \in \mathbb{N}$ we set:*

$$U(\varpi, \alpha)_t^n := U(0, \alpha\Delta_n^{\varpi}, \Delta_n) = \sum_{i=1}^{\lfloor t/\Delta_n \rfloor} \mathbb{1}_{\{|\Delta_i^n X| > \alpha\Delta_n^{\varpi}\}},$$

*the sum over all increments bigger than the threshold $\alpha\Delta_n^{\varpi}$.*

The reason for defining the threshold in this way can be derived heuristically by having a closer look at the case when $X = W + Y$ with $W := (W_t)_{t\geq 0}$ being a Brownian motion and $Y := (Y_t)_{t\geq 0}$ a $\beta$-stable process without translation (i.e. $c = 0$ in equation (3.8)). From this same equation we get the selfsimilarity of $Y$ (and $W$ if we set $\beta = 2$) and hence the increments satisfy:

$$\begin{aligned}\Delta_i^n X &\stackrel{d}{=} W_{\Delta_n} + Y_{\Delta_n}\\ &\stackrel{d}{=} \Delta_n^{1/2} W_1 + \Delta_n^{1/\beta} Y_1,\end{aligned}$$

showing an overall order of magnitude $\Delta_n^{1/2}$ as we have $\beta \leq 2$. But as we are interested in the 'big jumps' created by $Y$, we only observe the increments which are 'big', as for the chosen $\varpi$ we have $\Delta_n^{1/2} \ll \Delta_n^{\varpi}$ asymptotically for $\Delta_n^{\varpi} \to 0$. Making this choice for the threshold we will later see in Proposition 3.30 that we have the following convergence in probability:

$$\Delta_n^{\varpi\beta} U(\varpi, \alpha)_t^n \xrightarrow{\mathbb{P}} \frac{\overline{A}_t}{\alpha^\beta}.$$

When wanting to derive $\beta$ from the r.h.s. it becomes logical that we calculate the statistic $U$ for another value of $\alpha$. For such a value $\alpha'$ the r.h.s. converges to $\overline{A}_t/\alpha'^\beta$ and, by dividing the two limits, $\overline{A}_t$ cancels out. This proceeding hints at defining an estimator for $\beta$ in the following way.

***Definition 3.14:*** *Let the two constants $\alpha$ and $\alpha'$ fulfill $0 < \alpha < \alpha'$. We define:*

$$\hat{\beta}_n(t, \varpi, \alpha, \alpha') := \frac{\ln(U(\varpi, \alpha)_t^n / U(\varpi, \alpha')_t^n)}{\ln(\alpha'/\alpha)},$$

*with the convention that we set $\hat{\beta}_n = 0$ as soon as one of the statistics $U$ is equal to zero.*

Deducing the CLT with regard to the value $\beta$ will include stable convergence (or more precisely the property stated in Proposition 2.16). We shall state here that the paper also mentions two other possible ways of deriving $\beta$ from the statistic $U$ by varying other degrees of freedom in the truncated power variation. One way is to have different step sizes, where the choice is made to double the step size in one of the statistics $U$ but to leave the values $\alpha$ to be the same. Another one suggested, is to alter the value of $\varpi$ in the two statistics $U$ but leave everything else equal. As we only concentrate on the estimator $\hat{\beta}_n$ in the proofs to follow, we solely mention that the paper declares this choice of $\hat{\beta}_n$ to be the best, as it has a smaller asymptotic variance than the other two possibilities mentioned above. Properties are also only given for this estimator and are denoted in the following section. Before moving on we want to provide a definition appearing in the first result, which provides information about the speed of convergence of a sequence of estimators. Further we include the description of a *tight* sequence, as it can be found in [25, Definition V.1.2].

***Definition 3.15:*** *A sequence of rvs $(Y_n)_{n\in\mathbb{N}}$, with $Y_n \xrightarrow{\mathbb{P}} Y$, is said to be* $\Delta_n^\theta$**-rate consistent** *if the sequence of variables*

$$\left(\frac{1}{\Delta_n^\theta}(Y_n - Y)\right)_{n\geq 1}$$

*is* **tight***, which means that it is impossible to have arbitrary large events taking place, i.e.:*

$$\lim_{N\to\infty} \limsup_{n} \mathbb{P}\left( \left| \frac{1}{\Delta_n^{\theta}} (Y_n - Y) \right| > N \right) = 0. \tag{3.16}$$

One may note that without the above setting, the general definition for tight, allows for different probabilities $\mathbb{P}_n$ for every element of the sequence.

## 3.2 Main Results

The main results in paper [2] are split into two groups. The ones which are formulated with as few restrictions on the process as possible, and the ones where a stable or a general Lévy process is assumed. In this thesis we only cite the general case. The two results to follow can only be achieved on the random set $\{\overline{A}_t > 0\}$ for a semimartingale, solely being restricted by Assumption 3.5 and Assumption 3.12 (below referred to as ‘the stated assumptions’). For stronger results a more restrictive setting must be chosen.

***Theorem 3.16:*** Under the stated assumptions and for $t > 0$, $\varpi \in (0, 1/2)$ and $0 < \alpha < \alpha'$ we get[2]:

$$\hat{\beta}_n(t, \varpi, \alpha, \alpha') \xrightarrow{\mathbb{P}} \beta \qquad \text{on } \{\overline{A}_t > 0\}.$$

And further $\hat{\beta}_n(t, \varpi, \alpha, \alpha')$ are $\Delta_n^{\chi-\epsilon}$-rate consistent on the set $\{\overline{A}_t > 0\}$ for any $\epsilon > 0$ and

$$\begin{aligned} \chi :=& \chi(\beta, \gamma, \beta', \varpi) \\ =& (\varpi\gamma) \wedge \frac{1 - \varpi\beta}{3} \wedge \frac{\varpi(\beta - \beta')}{1 + \beta'} \wedge \frac{1 - 2\varpi}{2} \wedge \frac{\varpi\beta}{2}. \end{aligned} \tag{3.17}$$

As the value $\chi$ only results in being positive but as it is not necessarily far away from zero, the aforementioned conditions are not strong enough to derive a distributional result. To achieve this we must ensure that $\chi = \varpi\beta/2$, i.e. $\varpi\beta/2$ needs to be the minimum of the values on the right. By undertaking four simple transformations we see the following:

$$\begin{array}{llllll} \text{i)} & (\varpi\gamma) > \dfrac{\varpi\beta}{2} & \Leftrightarrow & \gamma > \dfrac{\beta}{2}, & & \\ \text{ii)} & \dfrac{1 - \varpi\beta}{3} > \dfrac{\varpi\beta}{2} & \Leftrightarrow & 2 > 5\varpi\beta & \Leftrightarrow & \dfrac{2}{5\beta} > \varpi, \\ \text{iii)} & \dfrac{\varpi(\beta - \beta')}{1 + \beta'} > \dfrac{\varpi\beta}{2} & \Leftrightarrow & 0 > -\beta + \beta'(2 + \beta) & \Leftrightarrow & \dfrac{\beta}{2 + \beta} > \beta', \\ \text{iv)} & \dfrac{1 - 2\varpi}{2} > \dfrac{\varpi\beta}{2} & \Leftrightarrow & \dfrac{1}{2 + \beta} > \varpi. & & \end{array} \tag{3.18}$$

[2] In the paper the estimator $\hat{\beta}'_n$ (defined to be the one resulting out of statistics $U$ with different step sizes) was stated. But this clearly appears to be a typo.

These conditions will be included in the theorem below as they guarantee that the jump activity indices not being generated by $\beta$, but instead by $\beta'$ and $\beta-\gamma$, are not too close to the value of $\beta$. Having this we see the following stable convergence, also to a standardized statistic.

***Theorem 3.17:*** Let the stated assumptions hold and set $t > 0$ and $0 < \alpha < \alpha'$ as in Theorem 3.16. Additionally we assume $\gamma > \frac{\beta}{2}$, $\overline{\omega} \in \big(0, (2/(5\beta)) \wedge (1/(2+\beta))\big)$ and $\beta' \in \big[0, \beta/(2+\beta)\big)$. Then we get the following stable convergence in law on a subset and for a normally distributed variable independent of $X$:

$$\text{i)}\quad \frac{1}{\Delta_n^{\overline{\omega}\beta/2}}\left(\hat{\beta}_n(t,\overline{\omega},\alpha,\alpha')-\beta\right) \xrightarrow{\mathcal{L}-(s)} \mathcal{N}\left(0, \frac{\alpha'^{\beta}-\alpha^{\beta}}{\overline{A}_t(\ln(\alpha'/\alpha))^2}\right) \qquad \text{on } \{\overline{A}_t > 0\},$$

$$\text{ii)}\quad \frac{\ln(\alpha'/\alpha)}{\sqrt{1/U(\overline{\omega},\alpha)_t^n - 1/U(\overline{\omega},\alpha')_t^n}}\left(\hat{\beta}_n(t,\overline{\omega},\alpha,\alpha')-\beta\right) \xrightarrow{\mathcal{L}-(s)} \mathcal{N}(0,1) \qquad \text{on } \{\overline{A}_t > 0\}.$$

Obviously, when observing a process $X$ we do not initially know what jump activity index it has and hence we need to make the choice of $\overline{\omega}$ independent of the value of $\beta \in [0,2]$. With regard to the restrictions stated above, this leads to setting $\overline{\omega} := 1/5$ for any implementation. This choice of taking the upper boundary is made, as the convergence is faster for bigger values of $\overline{\omega}$. Before moving to any implementations, we discuss the proofs in the following section. The proofs of both theorems build on various lemmas, which we state in the order in which they are given in the paper. We provide the steps to prove Theorem 3.16, but with regard to the proof of Theorem 3.17 we do not show, but only cite, one lemma and one proposition (namely, the ones called 'Lemma 7' and 'Proposition 2' in the original paper [2]). The reason is twofold: On the one hand these two proofs strongly build on results taken from [25] and we feel that stating all details would lie beyond the scope of this thesis. On the other hand, we assume to have come across some inconsistency in the proof of a part of Lemma 3.28 (which is part of Lemma 9.b) in the original paper). The two results, where we do not go through the proofs, are used in combination with the outcome of this Lemma 3.28.ii.b) to imply the proof of Theorem 3.17. Seeing this we felt that it is a good choice to prove one theorem completely and to leave details of the other proof open for future work.

## 3.3 Detailed Discussion of the Proofs

Within this whole section $K$ always represents a constant with no fixed value and it changes for every condition as needed. Throughout the proofs we use different versions, like $K', K_1', K''$, etc., to allow for an easier way of following the steps. For ease of notation we also introduce $\mathbb{E}_{i-1}^n$ and $\mathbb{P}_{i-1}^n$ and also $\mathbb{E}_t^n$ and $\mathbb{P}_t^n$, to represent the expected value and probability conditioned on the sub-$\sigma$-algebra $\mathscr{F}_{(i-1)\Delta_n}$ and $\mathscr{F}_t$, respectively. We recall that the semimartingale

$X$ fulfills the stated assumptions throughout the whole section. As a first step, we give some general estimates, which will be useful within the proofs to follow.[3]

***Lemma 3.18:*** For all $u \in [0,1]$, $v \in (0,2]$ and $x, y \in (0,1]$ we obtain the following inequalities:

$$\text{i)} \ \overline{F}''_t(x) \le \frac{K}{x^{\beta'}}, \tag{3.19}$$

$$\text{ii)} \ \left| \overline{F}_t(x) - \frac{A_t}{x^\beta} \right| \le \frac{K}{x^{(\beta-\gamma)\vee\beta'}}, \tag{3.20}$$

$$\text{iii)} \ \overline{F}_t(x) \le \frac{K}{x^\beta}, \tag{3.21}$$

$$\text{iv)} \ \int_{\{|z|\le u\}} z^2 F_t(dz) \le K u^{2-\beta}, \tag{3.22}$$

$$\text{v)} \ \int_{\{|z|>u\}} (|z|^v \wedge 1) F_t(dz) \le \begin{cases} K_v & \text{if } v > \beta, \\ K_v \ln(1/u) & \text{if } v = \beta, \\ K_v u^{v-\beta} & \text{if } v < \beta, \end{cases} \tag{3.23}$$

$$\text{vi)} \ \overline{F}_t(x) - \overline{F}_t(x+y) \le \frac{K}{x^\beta} \left( 1 \wedge \frac{y}{x} + x^{\gamma\wedge(\beta-\beta')} \right). \tag{3.24}$$

*Proof:* Before showing the different estimates, we want to state that all of them hold almost surely, and hence we do not denote the dependence on $\omega$ in any of the proofs.
i) By making use of Definition 3.4 providing $\overline{F}''_t(x) = F''_t([-x,x]^c)$ and by $1 \le 1/x^{\beta'}$ we get

$$\begin{aligned} \overline{F}''_t(x) &= \int_{[-1,-x]\cup[x,1]} \frac{|z|^{\beta'}}{|z|^{\beta'}} F''_t(dz) + \int_{(-\infty,-1)\cup(1,\infty)} F''_t(dz) \\ &\le \frac{1}{|x|^{\beta'}} \left( \int_{[-1,-x]\cup[x,1]} |z|^{\beta'} F''_t(dz) + \int_{(-\infty,-1)\cup(1,\infty)} F''_t(dz) \right) \\ &= \frac{1}{x^{\beta'}} \int_{[-x,x]^c} (|z|^{\beta'} \wedge 1) F''_t(dz) \\ &\overset{(3.12)}{\le} \frac{L}{x^{\beta'}}. \end{aligned}$$

ii) To deal with the second inequality we first have a closer look at $\overline{F}'_t$ by using Assumption 3.12.i) and the definition of $A_t$ given in equation (3.15). For better readability we introduce the auxiliary functions $c(x)$ and the variable $K'_1$:

$$\overline{F}'_t(x) = a_t^{(+)} \int_{-z_t^{(-)}}^{-x} \frac{1+|z|^\gamma f(t,z)}{|z|^{1+\beta}}\, dz + a_t^{(-)} \int_x^{z_t^{(+)}} \frac{1+|z|^\gamma f(t,z)}{|z|^{1+\beta}}\, dz$$

[3] The estimates stated in Lemma 3.18 are not proven in the paper and so we are not following any ideas by the authors in this case.

$$= a_t^{(+)} \frac{1}{\beta} \frac{1}{|z|^\beta} \Big|_{-z_t^{(-)}}^{-x} - a_t^{(-)} \frac{1}{\beta} \frac{1}{z^\beta} \Big|_{x}^{z_t^{(+)}} + \underbrace{a_t^{(+)} \int_{-z_t^{(-)}}^{-x} \frac{|z|^\gamma f(t,z)}{|z|^{1+\beta}}\, dz + a_t^{(-)} \int_{x}^{z_t^{(+)}} \frac{|z|^\gamma f(t,z)}{|z|^{1+\beta}}\, dz}_{=:c(x)}$$

$$= \frac{1}{x^\beta} \underbrace{\frac{a_t^{(+)} + a_t^{(-)}}{\beta}}_{\overset{(3.15)}{=} A_t} - \underbrace{\frac{1}{\beta} \left( \frac{1}{|z_t^{(-)}|^\beta} a_t^{(-)} + \frac{1}{|z_t^{(+)}|^\beta} a_t^{(+)} \right)}_{:=K_1'} + c(x).$$

For the term $c(x)$ defined above, we can easily conduct the following estimate by using the boundedness stated in Assumption 3.12.i. (especially conditions a) and d)) and introducing a constant $K_2'$:

$$\begin{aligned} c(x) =& a_t^{(+)} \int_{-z_t^{(-)}}^{-x} \frac{|z|^\gamma f(t,z)}{|z|^{1+\beta}}\, dz + a_t^{(-)} \int_{x}^{z_t^{(+)}} \frac{|z|^\gamma f(t,z)}{|z|^{1+\beta}}\, dz \\ \leq& 2 \int_x^1 \frac{L}{|z|^{1+\beta-\gamma}}\, dz \\ \leq& \frac{K_2'}{x^{\beta-\gamma}}. \end{aligned}$$

Now we are ready to show the desired inequality by using the above:

$$\begin{aligned} \left| \overline{F}_t(x) - \frac{A_t}{x^\beta} \right| =& \left| \frac{A_t}{x^\beta} - K_1' + c(x) + \overline{F}_t''(x) - \frac{A_t}{x^\beta} \right| \\ =& \left| -K_1' + c(x) + \overline{F}_t''(x) \right| \\ \leq& K_1' + \frac{K_2'}{x^{\beta-\gamma}} + \frac{L}{x^{\beta'}} \\ \leq& \frac{K}{x^{(\beta-\gamma)\vee\beta'}}. \end{aligned}$$

iii) The third estimate follows directly from the second, as we have $\beta > (\beta - \gamma) \vee \beta'$ and hence (when keeping in mind that the constant $K$ differs from the above one) the above equations easily result in:

$$\begin{aligned} \overline{F}_t(x) =& \frac{A_t}{x^\beta} - K_1' + c(x) + \overline{F}_t''(x) \\ \leq& \frac{A_t}{x^\beta} - K_1' + \frac{K_2'}{x^{\beta-\gamma}} + \frac{L}{x^{\beta'}} \\ \leq& \frac{K}{x^\beta}. \end{aligned}$$

iv) In the following transformation we will make use of the above and the setting of $u < 1$ and $2 - \beta > 0$. We do not state the full density of $F_t'$, but represent parts of it by $(\ldots)$, which acts as a placeholder. This is done, for ease of notation. However, we do use its boundedness conditions as stated in Assumption 3.12. (especially a), b), d) from i) and ii)):

$$\int_{\{|z|\leq u\}} z^2 F_t(dz) = \int_{\{|z|\leq u\}} z^2 F_t'(dz) + \int_{\{|z|\leq u\}} z^2 F_t''(dz)$$

$$
\begin{aligned}
&= \int_{\{|z|\le u\}} \frac{1+|z|^\gamma f(t,z)}{|z|^{1+\beta-2}} (\ldots)\, dz + \int_{\{|z|\le u\}} |z|^{2-\beta}(|z|^\beta \wedge 1) F_t''(dz) \\
&\le \int_{\{|z|\le u\}} \frac{K_1}{|z|^{1+\beta-2}}\, dz + u^{2-\beta} \int_{\{|z|\le u\}} (|z|^\beta \wedge 1) F_t''(dz) \\
&\le 2K_1 u^{2-\beta} + u^{2-\beta} L \\
&\le K u^{2-\beta}.
\end{aligned}
$$

v) The first of the three cases $v > \beta$ is clear by the definition of the jump activity index $\beta$. For the other two estimates we make use of what we have shown in iii) above, by replacing the $\beta$ we used there with a $\tilde{\beta} := \beta - v$ (as the distribution function (having $x^{1+\beta}$ in the denominator) changes to having $x^{1+\beta-v}$ in the denominator, when multiplying it with $x^v$. Hence we deduce the following for the third case $v < \beta$:

$$
\begin{aligned}
\int_{\{|z|>u\}} (|z|^v \wedge 1) F_t(dz) &\le \int_{\{|z|>u\}} |z|^v F_t(dz) + K_v' \\
&\overset{\text{iii)}}{\le} \frac{K_v''}{u^{\beta-v}} + K_v' \\
&\le K_v u^{v-\beta}.
\end{aligned}
$$

In the same way we can receive the result for the second case, as we have $\int_u^1 z^{-1}\, dz = \ln(1/u)$.

vi) Again we make use of the boundedness and of the results above. We also display parts of the density of $F_t'$ by $(\ldots)$ once more, as we already did in the proof of part iv). For ease of notation we define the set $A := [-x-y, x+y] \backslash [-x, x]$ to be the union of the two intervals with length $y$ each.

$$
\begin{aligned}
\overline{F}_t(x) - \overline{F}_t(x+y) &= \int_A F_t'(dz) + \int_A F_t''(dz) \\
&= \int_A \frac{1}{|z|^{1+\beta}} (\ldots)\, dz + \underbrace{\int_A \frac{|z|^\gamma f(t,z)}{|z|^{1+\beta}} (\ldots)\, dz}_{\le K_1' x^{\gamma-\beta}} + \underbrace{\int_A F_t''(dz)}_{\le K'' x^{-\beta'}},
\end{aligned}
$$

wherein the last estimate holds because of $A \in [-x, x]^c$ and of inequality i). Further, we easily see that for the estimate of the two last summands we get $K_1' x^{(\gamma-\beta)} + K'' x^{-\beta'} \le (K_1' + K'') x^{(\gamma-\beta)\wedge -\beta'}$. Let us now take a closer look at the first summand, while using Assumption 3.12.i.a) and b):

$$
\begin{aligned}
\int_A \frac{1}{|z|^{1+\beta}} (\ldots)\, dz &= \int_A \frac{1}{|z|^{1+\beta}} \left( a_t^{(+)} \mathbb{1}_{\left\{0<z\le z_t^{(+)}\right\}} + a_t^{(-)} \mathbb{1}_{\left\{-z_t^{(-)}<z\le 0\right\}} \right) dz \\
&\le 2 \int_x^{(x+y)\wedge 1} \frac{1}{z^{1+\beta}} L\, dz \\
&= -2L \frac{1}{\beta z^\beta} \bigg|_x^{(x+y)\wedge 1}
\end{aligned}
$$

$$=K_2'\left(\frac{1}{x^\beta}-\frac{1}{((x+y)\wedge 1)^\beta}\right).$$

Summarizing what we have obtained so far, we see that we have:

$$\overline{F}_t(x)-\overline{F}_t(x+y)\leq\frac{K}{x^\beta}\left(1-\frac{x^\beta}{((x+y)\wedge 1)^\beta}+x^{\gamma\wedge(\beta-\beta')}\right),$$

and hence, what we still need to show to get the r.h.s. to being $\frac{K}{x^\beta}\left(1\wedge\frac{y}{x}+x^{\gamma\wedge(\beta-\beta')}\right)$, is:

$$1-\frac{x^\beta}{((x+y)\wedge 1)^\beta}\overset{!}{\leq}K'\left(1\wedge\frac{y}{x}\right). \tag{3.25}$$

If the minimum on the r.h.s. is 1 the inequality becomes obvious. Hence, we only need to consider $y/x<1\Leftrightarrow y<x$. On the left-hand side (l.h.s.) we also need to distinguish two cases, depending on the minimum of $(x+y)$ and 1. If we have $(x+y)\leq 1$ what we need to show is $1-(x/(x+y))^\beta\leq K'y/x$ and we do this by performing the following transformation:

$$\begin{aligned}1-\left(\frac{x}{x+y}\right)^\beta&\leq 1-\left(\frac{x}{x+y}\right)^2=\left(1-\frac{x}{x+y}\right)\left(1+\frac{x}{x+y}\right)=\frac{y}{x+y}\frac{2x+y}{x+y}\\&\leq\frac{y}{x+y}\frac{2(x+y)}{x+y}=\frac{2y}{x+y}\\&\leq 2\frac{y}{x}.\end{aligned}$$

If on the other hand we have $(x+y)>1$ (i.e. we can make use of $y>1-x$) then we need to show $1-x^\beta\leq K'y/x$, which is fulfilled if we show the stronger first estimate in:

$$\begin{aligned}&&1-x^\beta&\overset{!}{\leq}K'\frac{1-x}{x}\leq K'\frac{y}{x}\\ \Leftrightarrow&&0&\overset{!}{\leq}\underbrace{K'-(K'+1)x+x^{\beta+1}}_{=:f(x)}\qquad\text{for } x\in(0,1].\end{aligned}$$

By introducing $f(x)$, showing the desired inequality is equivalent to showing that $f(x)\geq 0$ for $x\in(0,1]$. We easily see this by finding that the values for the corner points of the interval are bigger than zero for 0 and equal to zero for 1:

$$f(0)=K'>0\qquad\text{and}\qquad f(1)=K'-(K'+1)+1=0$$

and by having the function only decrease on the interval $(0,1]$, which can be seen by its negative derivative

$$f'(x)=-(K'+1)+(\beta+1)x^{\beta-1}\overset{!}{\leq}0\Leftrightarrow(\beta+1)x^\beta\overset{!}{\leq}K'+1,$$

which is fulfilled independently of $\beta\in(0,2)$ as soon as $K'\geq 2$. $\square$

For the lemma to follow we consider the observed process $Y := (Y_t)_{t\geq 0}$ to be a symmetric strictly $\beta$-stable process which demands $A = B$ in the representation of the corresponding Lévy measure provided in equation (3.9) and yields in $F(dx) = A/|x|^{\beta+1}\mathbb{1}_{\{|x|>0\}}\,dx$. Before stating the associated lemma, we define the process without any jumps bigger than a certain threshold $\delta \in (0,1]$ to be denoted by:

$$Y(\delta)'_t := Y_t - \sum_{s\leq t} \Delta Y_s \mathbb{1}_{\{|\Delta Y_s|>\delta\}}.$$

***Lemma 3.19:*** For the symmetric strictly $\beta$-stable process $Y := (Y_t)_{t\geq 0}$ there exists a constant $K$ which depends on $(A, \beta)$, and fulfills:

$$\mathbb{P}\left(\{|Y(\delta)'_s| > \delta/2\}\right) \leq K\frac{s^{4/3}}{\delta^{4\beta/3}} \qquad \forall s > 0. \tag{3.26}$$

*Proof:* Before stating the idea of the proof we mention some properties of the stable process $Y$. From the structure of the Lévy measure $F$ we easily see that the symmetrical tail function can be represented as $\overline{F}(x) = A/x^\beta$. As $Y$ is a stable process we get from equation (3.8), that

$$(Y_t)_{t\geq 0} \stackrel{d}{=} (t^{1/\beta}Y_1)_{t\geq 0}. \tag{3.27}$$

We state that for the symmetrical tail function $\overline{G}(x) := \mathbb{P}(\{|Y_1| > x\})$ we have the following estimate

$$\overline{G}(x) = \frac{A}{x^\beta} + O\left(\frac{1}{x^{2\beta}}\right) \qquad \text{for } x \to \infty, \tag{3.28}$$

taken from [2, p.2212] and applied below to $x = \delta/(2s^{1/\beta})$ for very small values of $s$. Further we define a variable and the processes and sets required in the proof for fixed $s$ and $\delta$, such that the variable $\theta < 1/2$:

$$\begin{aligned}
\theta :=& s\overline{F}(\delta/2) = s\frac{A}{(\delta/2)^\beta}, && (\textit{used as a constant within the proof}),\\
D :=& \{|Y_s| > \delta/2\}, && (\textit{set where } |Y| \textit{ is} > \delta/2 \textit{ at } s),\\
Y' :=& Y(\delta)', && (\textit{process excluding any jumps} > \delta),\\
D' :=& \{|Y'_s| > \delta/2\}, && (\textit{set where } |Y'| \textit{ is} > \delta/2 \textit{ at } s),\\
Y'' :=& Y(\delta/2)', && (\textit{the process without any jumps} > \delta/2),\\
N(\delta/2)_s :=& \sum_{r\leq s} \mathbb{1}_{\{|\Delta Y_r|>\delta/2\}}, && (\textit{a Poisson process counting jumps} > \delta/2),\\
B :=& \{N(\delta/2)_s = 1\}, && (\textit{set where } |Y| \textit{ has one jump} > \delta/2 \textit{ till } s),\\
B' :=& \{N(\delta/2)_s = 0\}, && (\textit{set where } |Y| \textit{ has no jump} > \delta/2 \textit{ till } s).
\end{aligned}$$

Using this notation we now have to show the estimate for $\mathbb{P}(D')$. First, we use the basic Inclusion-Exclusion principle to get:

$$\mathbb{P}(D \cap B^c) = 1 - \mathbb{P}(D^c \cup B) = 1 - \mathbb{P}(D^c) - \mathbb{P}(B) + \mathbb{P}(D^c \cap B) = \mathbb{P}(D) - \mathbb{P}(B) + \mathbb{P}(D^c \cap B).$$

Further we see that $D \cap B' = D' \cap B'$, as the the processes $Y$ and $Y'$ are the same on the set $B'$, where there are no jumps bigger than $\delta/2$. We also use that $D'$ and $B'$ are independent, as stated in the proof in the paper, but this is not as clear and will be discussed in Remark 3.20. However, when assuming the just stated property and seeing that the Poisson process $N(\delta/2)_s$ has expected value $\theta$ and hence the probability of it being 0 is $\mathbb{P}(B') = e^{-\theta} \geq e^{-1/2}$, we derive:

$$\begin{aligned}\mathbb{P}(D') &= \frac{\mathbb{P}(D' \cap B')}{\mathbb{P}(B')} = \frac{\mathbb{P}(D \cap B')}{\mathbb{P}(B')} \leq \frac{\mathbb{P}(D \cap B^c)}{\mathbb{P}(B')} \\ &\leq K\mathbb{P}(D \cap B^c) = K\Big(\mathbb{P}(D) - \mathbb{P}(B) + \mathbb{P}(D^c \cap B)\Big) \\ &\leq K\Big(|\mathbb{P}(D) - \theta| + |\mathbb{P}(B) - \theta| + \underbrace{\mathbb{P}(D^c \cap B)}_{=\mathbb{P}(D^c|B)\mathbb{P}(B)}\Big)\end{aligned} \tag{3.29}$$

Having this transformation we need to show that all three summands are smaller than $K(s/\delta^\beta)^{4/3}$ or equivalently $< K\theta^{4/3}$, as we do not care about constants and $K$ can change for every estimate, as long as it does only depend on $A$ and $\beta$. Proceeding as stated, we make the first estimate by using the aforementioned properties:

$$\begin{aligned}\mathbb{P}(D) &= \mathbb{P}(\{|Y_s| > \delta/2\}) \overset{(3.27)}{=} \mathbb{P}(\{|Y_1| > s^{-1/\beta}\delta/2\}) = \overline{G}(s^{-1/\beta}\delta/2) \\ &\overset{(3.28)}{=} \frac{A}{(s^{-1/\beta}\delta/2)^\beta} + O\left(\frac{1}{(s^{-1/\beta}\delta/2)^{2\beta}}\right) = \theta + O\left(\theta^2\right),\end{aligned}$$

which implies

$$\big|\mathbb{P}(D) - \theta\big| = \big|\theta + O\left(\theta^2\right) - \theta\big| \leq K\theta^2 < K\theta^{4/3}.$$

For the estimate of the second summand, all we need to do is to use the series expansion of the exponential function, and we get:

$$\mathbb{P}(B) = \mathbb{P}(\{N(\delta/2) = 1\}) = \theta e^{-\theta} = \theta \sum_{n=0}^{\infty} \frac{(-\theta)^n}{n!} = \theta + O(\theta^2),$$

which easily results in the same estimate as for the first summand

$$\big|\mathbb{P}(B) - \theta\big| = \big|\theta + O\left(\theta^2\right) - \theta\big| \leq K\theta^2 < K\theta^{4/3}.$$

Getting that far, we only miss an estimate for the last summand. The analysis of this one is a bit more involved and we first give an estimate for the process $Y''$ without any jumps bigger than $\delta/2$, which hence has the Lévy measure $\mathbb{1}_{\{|x| \leq \delta/2\}} F(x)$ (which is deterministic and independent of time) and results in:

$$\begin{aligned}\mathbb{E}[(Y_s'')^2] &= s\int_{\{|x|\le\delta/2\}} x^2 F(dx)\\ &\overset{(3.18)}{<} sK'(\delta/2)^{2-\beta}\\ &= K''\theta\delta^2.\end{aligned} \tag{3.30}$$

Now we see that the set $D^c|B$ gives all paths where the absolute value of the process $Y$ at the time $s$ is $\le \delta/2$ conditioned on the event of exactly one jump bigger than $\delta/2$ happening till time $s$. Hence, on the set $B$ we can represent $Y_s$ as $Y_s''$ plus one jump of size bigger than $\delta/2$. This jump conditioned on $B$ is distributed by $\mathbb{1}_{\{|x|>\delta/2\}}F(x)/\theta$. The scaling factor $1/\theta$ is needed, as conditioned on the set $B$ we need to have $\mathbb{E}[N(\delta/2)_s|B] = 1$ and not $\mathbb{E}[N(\delta/2)_s] = \theta$. Hence, by using the above, and $e^{-\theta} < 1$ and

$$\{|Y_s''+x|\le\delta/2\}\subset\{|x|-|Y_s''|\le\delta/2\}=\{|Y_s''|\ge|x|-\delta/2\}, \tag{3.31}$$

we can proceed as follows:

$$\begin{aligned}\mathbb{P}(B\cap D^c) &= \mathbb{P}(B)\mathbb{P}(D^c|B)\\ &= \theta e^{-\theta}\ s\int_{\{|x|>\delta/2\}}\frac{1}{\theta}\mathbb{P}(\{|Y_s''+x|\le\delta/2\})\,F(dx)\\ &\le\ s\int_{\{|x|>\delta/2\}}\mathbb{P}(\{|Y_s''+x|\le\delta/2\})\,F(dx)\\ &\overset{(3.31)}{\le}\ s\int_{\{|x|>\delta/2\}}\mathbb{P}(\{|Y_s''|\ge|x|-\delta/2\})\,F(dx)\le\dots\end{aligned}$$

we would find an upper boundary by $s\overline{F}(\delta/2)\mathbb{P}(\{|Y_s''|\ge 0\}) = s\overline{F}(\delta/2) = \theta$ but this is too weak for the desired result. Hence we need too split the above in the event of the jump being smaller than a suitable threshold $\frac{\delta}{2}\theta^{1/3}$ and for the jump bigger than $\frac{\delta}{2}\theta^{1/3}$ we keep the probability $\mathbb{P}(\{|Y_s''|\ge\frac{\delta}{2}\theta^{1/3}\})$:

$$\begin{aligned}\dots\le{}&\underbrace{s\,F\left(\left\{\delta/2<|x|<\delta/2\left(1+\theta^{1/3}\right)\right\}\right)}_{=:S_1(s,\delta)}\\ &+\underbrace{s\,\overline{F}\left(\delta/2\left(1+\theta^{1/3}\right)\right)\mathbb{P}\left(\left\{|Y_s''|\ge\delta/2\left(\theta^{1/3}\right)\right\}\right)}_{=:S_2(s,\delta)}\end{aligned}$$

To provide an estimate for $S_1(s,\delta)$ we need to apply the estimate from Lemma 3.18.vi) to the stable process $Y$ (as $Y$ is stable we see that $f(t,x)\equiv 0$ and $F_t''(dx)\equiv 0$ in Assumption 3.12 and therefore the estimate in vi) is more precise not requiring the second summand) to obtain:

$$\begin{aligned}S_1(s,\delta) &= s\left(\overline{F}\,(\delta/2)-\overline{F}\left(\delta/2\left(1+\theta^{1/3}\right)\right)\right)\\ &\le sK'(2/\delta)^{\beta}\theta^{1/3} = K\theta^{4/3}.\end{aligned}$$

Finally we need to show that the summand $S_2(s,\delta)$ is smaller than $K\theta^{4/3}$. We do this by using Chebyshev's inequality (see, for example, [30, Implication 13.10]), and combine this with the estimate for $\mathbb{E}[(Y_s'')^2]$ from equation (3.30) above. Further we make use of Lemma 3.18.iii).

$$\begin{aligned} S_2(s,\delta) &\leq s\,\overline{F}\,(\delta/2)\,\mathbb{P}\left(\left\{|Y_s''| \geq \delta/2\left(\theta^{1/3}\right)\right\}\right) \\ &\leq s\,\overline{F}\,(\delta/2)\,\frac{\mathbb{E}[(Y_s'')^2]}{\delta^2\theta^{2/3}/4} \\ &\leq s\,\overline{F}\,(\delta/2)\,\frac{K''\theta\delta^2}{\delta^2\theta^{2/3}} \\ &\leq s\,K'\frac{1}{\delta^\beta}\;\;\frac{\theta^{1/3}}{1} = K\theta^{4/3}, \end{aligned}$$

where the last equality comes from $\theta = sA(2/\delta)^\beta$. Now we have the desired estimate for all summands, which completes the proof. □

We want to note, that Lemma 3.19 is just slightly stronger than Chebyshev's inequality. By using equation (3.30) for $Y'$ we get $\mathbb{E}[(Y_s'')^2] < 2^{2-\beta}K\theta\delta^2$. This implies the l.h.s. of equation (3.26) to be $< K(s/\delta^\beta)$ which is bigger than $K(s/\delta^\beta)^{4/3}$ for small values of $s$. We will be interested in small values when proving Lemma 3.25 and hence more steps had to be taken. Within these steps, we want to stress a step made in the proof, which needs further verification in our view.

***Remark 3.20:*** As commented on in our proof above, we feel that the statement by the authors saying that $D'$ and $B'$ are independent, is not obvious. We could not show this point here and want to mention, that it would also be enough to show that $\mathbb{P}(B'|D') > K' > 0$, for $K'$ independent of $\delta$ and $s$. The reason being that without independence we would need to replace the denominator in equation (3.29) by $\mathbb{P}(B'|D')$ and would still arrive at the same result when showing the existence of such a $K'$.

Next, we move our attention to the most general process assumed within this thesis, namely the semimartingale $X$. For the analysis thereof we will also make use of the *Burkholder-Davis-Gundy inequality (BDG inequality).* The standard version for any $p > 0$ only applies to continuous local martingales, see, for example, [27, Theorem 17.7]. As we want to apply the BDG inequality not only for continuous but for general local martingales, we need to use an extension, which is *possible only for exponents* $p \geq 1$, [27, p.524]. Hence, we cite the following from [27, Theorem 26.12] and introduce the notation for the supremum of a process $M := (M_t)_{t\geq 0}$ to be denoted by $M^*$, i.e.:

$$M^* := \sup_{t>0} |M_t|.$$

***Theorem 3.21 (BDG Inequality):*** For any local martingale $M := (M_t)_{t\geq 0}$ with $M_0 = 0$ and any $p \geq 1$, there is a constant $K_p > 0$ allowing for the BDG inequality:

$$K_p^{-1}\mathbb{E}\left[[M,M]_\infty^{p/2}\right] \leq \mathbb{E}\left[M^{*p}\right] \leq K_p\mathbb{E}\left[[M,M]_\infty^{p/2}\right].$$

To analyze the estimates for a semimartingale, we use the representation as given in equation (2.6) in Theorem 2.10 of the characteristics of a semimartingale. We are now interested in the process without the jumps bigger than a certain $\delta \in (0,1]$. First we define the sum of the steps bigger than $\delta$ as

$$X(\delta)''_t := \sum_{s\leq t} \Delta X_s \mathbb{1}_{\{|\Delta X_s|>\delta\}},$$

which allows us to define the altered process without any jumps bigger than $\delta$ by

$$\begin{aligned} X(\delta)'_t :=& X - X(\delta)''_t \\ =& X_0 + B + X^c + (x\mathbb{1}_{\{|x|\leq\delta\}}) \star (\mu - \nu) - B(\delta), \end{aligned} \tag{3.32}$$

when setting

$$b(\delta)_s := \int_{\{\delta<|x|<1\}} xF_s(dx) \qquad \text{and} \qquad B(\delta)_t := \int_0^t b(\delta)_s \, ds.$$

This means that $B(\delta)$ represents $(x\mathbb{1}_{\{\delta<|x|<1\}})\star\nu$. In a similar way we also define $B(\delta)'$, when we have $\beta' \leq 1$ as in this case Assumption 3.12.ii) guarantees the following integrals to be bounded[4]:

$$b(\delta)'_s := \int_{\{|x|<\delta\}} xF''_s(dx) \qquad \text{and} \qquad B'(\delta)_t := \int_0^t b(\delta)'_s \, ds. \tag{3.33}$$

We can deduce that $B(\delta)' = (x\mathbb{1}_{\{|x|<\delta\}}\mathbb{1}_\Phi) \star \nu$ and $(x\mathbb{1}_{\{|x|<\delta\}}\mathbb{1}_\Phi) \star \mu$ is of finite variation. The introduction of this notation will be useful in the consecutive lemma as we can easily distinguish between the cases $\beta' \leq 1$ and $\beta' > 1$. The following decomposition will be used:

$$X(\delta)' = X_0 + \hat{X} + \overline{X}(\delta)^a + \overline{X}(\delta)^b - B(\delta), \tag{3.34}$$

$$\text{where} \quad \text{i)} \quad \hat{X} := \begin{cases} B + X^c - B(\delta)' & \text{if } \beta' \leq 1, \\ B + X^c & \text{if } \beta' > 1, \end{cases} \tag{3.35}$$

$$\text{ii)} \quad \overline{X}(\delta)^a := \begin{cases} (x\mathbb{1}_{\{|x|<\delta\}}\mathbb{1}_\Phi) \star \mu & \text{if } \beta' \leq 1, \\ (x\mathbb{1}_{\{|x|<\delta\}}\mathbb{1}_\Phi) \star (\mu - \nu) & \text{if } \beta' > 1, \end{cases} \tag{3.36}$$

[4]The paper sets the integral in the definition of $b(\delta)'_s$ to be over the set $\{\delta < |x| < 1\}$, but as this is a typo we replace it with the set $\{|x| < \delta\}$ as is done by Aït-Sahalia and Jacod when they refer to $B(\delta')$ at a later stage.

$$\text{iii)} \quad \overline{X}(\delta)^b := (x \mathbb{1}_{\{|x|<\delta\}} \mathbb{1}_{\Phi^c}) \star (\mu - \nu). \tag{3.37}$$

Using the decomposition above we can now obtain various estimates on the newly introduced summands in the way stated below.

***Lemma 3.22:*** For all $\delta \in (0,1]$, $p \geq 2$, $s, t \geq 0$, and $s \leq 1$ in i)[5], there exist various constants $K$ which fulfill:

$$\text{i)} \qquad \mathbb{E}_t \left[ \left| \hat{X}_{t+s} - \hat{X}_t \right|^p \right] \leq K_p s^{p/2}, \tag{3.38}$$

$$\text{ii)} \qquad \mathbb{E}_t \left[ \left| \overline{X}(\delta)^a_{t+s} - \overline{X}(\delta)^a_t \right|^{\beta'} \right] \leq K s, \tag{3.39}$$

$$\text{iii)} \qquad \mathbb{E}_t \left[ \left| \overline{X}(\delta)^b_{t+s} - \overline{X}(\delta)^b_t \right|^2 \right] \leq K s \delta^{2-\beta}, \tag{3.40}$$

$$\text{iv)} \qquad |B(\delta)_{t+s} - B(\delta)_t| \leq \begin{cases} Ks & \text{if } \beta < 1, \\ Ks \ln(1/\delta) & \text{if } \beta = 1, \\ Ks\delta^{1-\beta} & \text{if } \beta > 1. \end{cases} \tag{3.41}$$

*Proof:* i) We use the representation of $\hat{X}_t$ in equation (3.35) above and analyze the increments of the summands, where $B(\delta)'$ is only present if $\beta' \leq 1$ and it is, together with the process $B$, of finite variation. Hence we find a $K'$ allowing for the following:

$$\mathbb{E}_t \left[ \left| \hat{X}_{t+s} - \hat{X}_t \right|^p \right] \leq \mathbb{E}_t \left[ \left| sK' + |X^c_{t+s} - X^c_t| \right|^p \right] \leq \ldots$$

For continuing the estimation, we derive an inequality from the convex function $g(x) := |x|^p$ by using that we have $g([x+y]/2) \leq [g(x)+g(y)]/2$ for any convex function, and we get:

$$\left( \frac{|x+y|}{2} \right)^p \leq \frac{|x|^p + |y|^p}{2}, \qquad \text{or equivalently} \qquad |x+y|^p \leq 2^{p-1}(|x|^p + |y|^p).$$

Further we see that $|X^c_{t+s} - X^c_t|$ is independent of $\mathscr{F}_t$, that it is distributed in the same way as $|X^c_s|$ and that $\sup_{r \leq s} |X^c_r|$ forms an upper boundary on which we can apply the BDG-inequality (as stated in Theorem 3.21). The quadratic variation of $X^c$ fulfills $[X^c, X^c]_s \leq K''s$ as we can represent $X^c$ as an integral with respect to a Brownian motion (see equation (3.2), where $\sigma_s$ is bounded). Now we can continue the above estimation by using the steps just mentioned:

$$\begin{aligned} \ldots \leq \mathbb{E}_t \left[ \left| sK' + |X^c_{t+s} - X^c_t| \right|^p \right] &\leq 2^{p-1}(|sK'|^p + \mathbb{E}_t[|X^c_{t+s} - X^c_t|^p] \\ &\leq K'_p(s^p + s^{p/2}) \leq K_p s^{p/2}. \end{aligned}$$

The last estimate only holds, as we assume to have $s \leq 1$ and $p \geq 2$.

[5] This additional restriction is not included in the paper, but is needed for the proof. However, only small values of $s$ are analyzed throughout the remainder of the paper, so it does not impose a problem.

ii) To prove this estimate we need to distinguish between the case in which $\beta' \leq 1$ and where $\beta' > 1$. One obvious reason being, that the BDG-inequality can only be applied for values of $p \geq 1$ (see Theorem 3.21), i.e. in the case when $\beta' > 1$.

In the first case, we make use of the inequality

$$\Big|\sum_m x_m\Big|^{\beta'} \leq \sum_m |x_m|^{\beta'}, \qquad \text{for } \beta' \leq 1, \tag{3.42}$$

which is obvious for $\beta' = 1$. We only need to show the estimate for $x_m > 0$ as with this assumption we can easily see that $|\sum_m a_m x_m|^{\beta'} \leq |\sum_m x_m|^{\beta'}$ for any $a_m \in \{-1, 1\}$ for all $m \in \mathbb{N}$. Hence, we can show the estimate for $x_m > 0$ and $\beta' \in (0, 1)$ by expanding the estimate $(x_1 + y)^{\beta'} \leq x_1^{\beta'} + y^{\beta'}$ for two summands to countable many summands. These details are not depicted here but we derive the estimate for two summands. The inequality can be transformed to $(x_1 + y)^{\beta'} - x_1^{\beta'} \leq y^{\beta'}$ and this is what we want to show, by making use of the function $x^{\beta'-1}$ being decreasing:

$$(x_1 + y)^{\beta'} - x_1^{\beta'} = x^{\beta'}\Big|_{x_1}^{x_1+y} = \int_{x_1}^{x_1+y} \beta' x^{\beta'-1}\, dx \leq \int_0^y \beta' x^{\beta'-1}\, dx = x^{\beta'}\Big|_0^y = y^{\beta'}.$$

One easily sees, that the stated estimate in equation (3.42) would not hold for $\beta' > 1$ (e.g. for $x_1 = x_2 = \delta$ and all other summands being zero, we have $l.h.s. = 2^{\beta'}\delta^{\beta'}$ but $r.h.s. = 2\delta^{\beta'}$, which is a counter example as $2^{\beta'} > 2$ for $\beta' > 1$). Seeing that $\overline{X}(\delta)^a$ is the sum over all jumps caused by the Lévy measure $F_t''$ and being smaller than $\delta$, we get[6]:

$$\big|\overline{X}(\delta)^a_{t+s} - \overline{X}(\delta)^a_t\big|^{\beta'} = \Big|\sum_{r \leq s} \Delta\overline{X}(\delta)^a_{t+r}\Big|^{\beta'} \leq \sum_{r \leq s} \big|\Delta\overline{X}(\delta)^a_{t+r}\big|^{\beta'}$$

$$\Rightarrow \mathbb{E}_t\left[\big|\overline{X}(\delta)^a_{t+s} - \overline{X}(\delta)^a_t\big|^{\beta'}\right] \leq \mathbb{E}_t\left[\int_t^{t+s} \int_{\{|x| \leq \delta\}} |x|^{\beta'}\, dr F_r''(dx)\right] \leq Ks.$$

To get the last estimate we make use of the required boundedness of $F_t''$ as stated in Assumption 3.12.ii).

In the second case, when we have $\beta' > 1$, we see that $\overline{X}(\delta)^a$ is a purely discontinuous, local martingale, due to its definition (compare equation (2.4) for the definition of the integral with respect to $(\mu - \nu)$). Hence we can apply the BDG-inequality (for $p = \beta' > 1$) on $\overline{X}(\delta)^a_{t+s} - \overline{X}(\delta)^a_t$ which has $\sum_{r \leq s} \big|\Delta\overline{X}(\delta)^a_{t+r}\big|^2$ as its quadratic variation (see, for example, [21, Definition 11.16]). By applying this, and also equation (3.42) for $\beta'/2 \leq 1$, we get:

[6] In the paper there was a typo having a 2 instead of $\beta'$ in the exponent of the first estimate and showing a spare $\beta'$ as another exponent.

$$\begin{aligned}
\mathbb{E}_t\left[\left|\overline{X}(\delta)^a_{t+s}-\overline{X}(\delta)^a_t\right|^{\beta'}\right] &\leq \mathbb{E}_t\left[\left(\sup_{r\leq s}\left|\overline{X}(\delta)^a_{t+r}-\overline{X}(\delta)^a_t\right|\right)^{\beta'}\right]\\
&\leq \mathbb{E}_t\left[\left(\sum_{r\leq s}\left|\Delta\overline{X}(\delta)^a_{t+r}\right|^2\right)^{\beta'/2}\right]\\
&\leq \mathbb{E}_t\left[\sum_{r\leq s}\left|\Delta\overline{X}(\delta)^a_{t+r}\right|^{\beta'}\right]\leq Ks.
\end{aligned}$$

In the last line the estimate is the same as when we had $\beta' \leq 1$ and hence we do not display the sum as an integral, but just state the estimate.

iii) Analogously to the above, for $\beta' > 1$ and $\overline{X}(\delta)^a$, we see that $\overline{X}(\delta)^b$ is a purely discontinuous local martingale, as it is an integral with respect to $(\mu-\nu)$ too. The quadratic variation is again the sum over the squared jumps. For proving the estimate, we make use of the increment of $\overline{X}(\delta)^b$ being taken to the power of 2 (and not $\beta' \in (0,2)$). This situation allows us to apply the Itô-Isometry. To be more precise we use an extension thereof, holding for any local martingale $Y := (Y_t)_{t\geq 0}$ and a predictable process $\xi$, which fulfills $E[\int_0^\infty \xi^2\, d[Y,Y]] < \infty$. This extended version is taken from [32, Theorem 5]. We simply put $\xi = \mathbb{1}_{(t,t+s]}$ in our case and by knowing that the sum over the squared jumps on the r.h.s. is finite, we can write:

$$\begin{aligned}
\mathbb{E}_t\left[\left|\overline{X}(\delta)^b_{t+s}-\overline{X}(\delta)^b_t\right|^2\right] &= \mathbb{E}_t\left[\sum_{r\leq s}\left|\Delta\overline{X}(\delta)^b_{t+r}\right|^2\right]\\
&= \mathbb{E}_t\left[\int_t^{t+s}\int_{\{|x|\leq\delta\}}|x|^2\,dr F'_r(dx)\right]\\
&\leq Ks\delta^{2-\beta},
\end{aligned}$$

wherein we get the last estimate from Lemma 3.18.iv).

iv) This final estimate easily follows from Lemma 3.18.v) when setting $v = 1$ and seeing that for the integrand in $B(\delta)_t = \int_0^t b(\delta)_s\, ds$ we have:

$$b(\delta)_s = \int_{\{\delta<|x|<1\}} xF_s(dx) < \int_{\{\delta<|x|\}} (|x|^v \wedge 1)F_s(dx). \qquad \square$$

Next we state a general result for a *counting process* $N := (N_t)_{t\geq 0}$. We recall that such a process has the property of being integer valued with $N_0 = 0$, right continuous and consists of jumps of size 1. Here we assume the process to be adapted to **F** and to have a predictable compensator, which is represented by $\Lambda_t = \int_0^t \lambda_s\, ds$. The feature we want to make use of is stated below.

***Lemma 3.23:*** Considering the counting process $N$, its compensator $\Lambda_t$ and assuming there is a constant $u > 0$ with $\lambda_s < u$ for all $s \in [0, T]$, we get the following[7] for any $t \geq 0$:

$$\left|\mathbb{P}_t(\{N_{t+s} - N_t = 1\}) - \mathbb{E}_t[\Lambda_{t+s} - \Lambda_t]\right| + \mathbb{P}_t(\{N_{t+s} - N_t \geq 2\}) \leq 2(us)^2.$$

*Proof:* We show the estimate by showing that each of the two summands is smaller than $(us)^2$. In both cases we denote by $T_1$ the jump time of the first jump after $t$ and respectively, by $T_2$ the second one. For the first summand we start the approximation by giving one equality

$$\begin{aligned}
\text{i) } \mathbb{P}_t(\{N_{t+s} - N_t = 1\}) &= \mathbb{P}_t(\{N_{(t+s)\wedge T_1} - N_t = 1\}) \\
&= \sum_{k=0}^{\infty} k\mathbb{P}_t(\{N_{(t+s)\wedge T_1} - N_t = k\}) \\
&= \mathbb{E}[N_{(t+s)\wedge T_1} - N_t] \\
&= \mathbb{E}[\Lambda_{(t+s)\wedge T_1} - \Lambda_t],
\end{aligned}$$

and one inequality

$$\begin{aligned}
\text{ii) } \mathbb{P}_t(\{N_{t+s} - N_t \geq 1\}) &= \sum_{k=1}^{\infty} \mathbb{P}_t(\{N_{t+s} - N_t = k\}) \\
&\leq \sum_{k=1}^{\infty} k\mathbb{P}_t(\{N_{t+s} - N_t = k\}) \\
&= \mathbb{E}_t[N_{t+s} - N_t] \\
&= \mathbb{E}_t[\Lambda_{t+s} - \Lambda_t] \\
&\leq us,
\end{aligned}$$

which we combine to obtain

$$\begin{aligned}
\left|\mathbb{E}_t[\Lambda_{t+s} - \Lambda_t] - \mathbb{P}_t(\{N_{t+s} - N_t = 1\})\right| &\overset{\text{i)}}{=} \mathbb{E}_t[\Lambda_{t+s} - \Lambda_{(t+s)\wedge T_1}] \\
&= \mathbb{E}_t[(\Lambda_{t+s} - \Lambda_{(t+s)\wedge T_1})\mathbb{1}_{\{T_1 < t+s\}}] \\
&\leq us\mathbb{P}_t(\{N_{t+s} - N_t \geq 1\}) \\
&\overset{\text{ii)}}{\leq} (us)^2.
\end{aligned}$$

Having this, we move to the estimate for the second summand. We use the transformation

$$\begin{aligned}
\text{iii) } \mathbb{E}[\mathbb{1}_{\{T_2 < t+s\}}|\mathscr{F}_{T_1}]] &= \mathbb{E}[N_{(t+s)\wedge T_2} - 1|\mathscr{F}_{T_1}] \\
&= \mathbb{E}[N_{(t+s)\wedge T_2} - N_{T_1}|\mathscr{F}_{T_1}] \\
&= \mathbb{E}[\Lambda_{(t+s)\wedge T_2} - \Lambda_{T_1}|\mathscr{F}_{T_1}]
\end{aligned}$$

[7] There is no constant 2 on the r.h.s. of the estimate but from what is shown in the proof of the paper, this constant needs to be added.

$$= \mathbb{E}\bigg[ \int_{T_1}^{(t+s)\wedge T_2} \lambda_r \, dr \bigg| \mathscr{F}_{T_1} \bigg],$$

and by also applying the tower property and using the boundedness $\lambda_r < u$, we get:

$$\begin{aligned}
\mathbb{P}_t(\{N_{t+s} - N_t \geq 2\}) &= \mathbb{P}_t(\{T_2 \leq t+s\}) \\
&= \mathbb{E}_t[\mathbb{1}_{\{T_1 < t+s\}} \mathbb{1}_{\{T_2 < t+s\}}] \\
&= \mathbb{E}_t[\mathbb{1}_{\{T_1 < t+s\}} \mathbb{E}[\mathbb{1}_{\{T_2 < t+s\}} | \mathscr{F}_{T_1}]] \\
&\overset{\text{iii)}}{=} \mathbb{E}_t\Big[\mathbb{1}_{\{T_1 < t+s\}} \mathbb{E}\bigg[ \int_{T_1}^{(t+s)\wedge T_2} \lambda_r \, dr \bigg| \mathscr{F}_{T_1} \bigg]\Big] \\
&\leq us\mathbb{E}_t[\mathbb{1}_{\{T_1 < t+s\}}] \\
&= us\mathbb{P}_t(\{N_{t+s} - N_t \geq 1\}) \\
&\overset{\text{ii)}}{\leq} (us)^2.
\end{aligned}$$

Hence we see that the sum of the two parts is smaller than $2(us)^2$, which is what we wanted to show. □

As our next step, we define a specific counting process (which appeared for a stable process in the proof of Lemma 3.19). Namely, we count the number of jumps of the semimartingale $X$, which are bigger than a certain $\delta \in (0, 1]$, and use the notation:

$$N(\delta)_t := \sum_{s \leq t} \mathbb{1}_{\{|\Delta X_s| > \delta\}} \qquad \forall t \in [0, T]. \tag{3.43}$$

Now we will give an estimate of the likeliness of the event that, over a timespan of length $s$, at least one jump bigger than $\delta$ occurs (i.e. $\{N(\delta)_{t+s} - N(\delta)_t \geq 1\}$) while, at the same time, the process containing only jumps smaller than $\delta$, changes by more than $\delta\zeta$ (i.e. $\{|X(\delta)'_{t+s} - X(\delta)'_t| > \delta\zeta\}$). This will be useful in the proof of Lemma 3.25 when we have an estimate for $X(\delta)'$ but want to extend it to the whole process $X$ to arrive at the desired result. The idea will go along the lines that if the whole process changes by less than an occurring jump in the interval, then the process $X(\delta)'$ must have changed by at least the remaining difference. This is why we have $X(\delta)'$ changing more than a certain threshold in the probability depicted below.

***Lemma 3.24:*** For all $\delta \in (0, 1]$, $\zeta \in (0, 1/2)$ and $p \geq 2$ and for $N(\delta)_t$, the following holds[8]:

$$\mathbb{P}_t\left(\{N(\delta)_{t+s} - N(\delta)_t \geq 1\} \cap \left\{\left|X(\delta)'_{t+s} - X(\delta)'_t\right| > \delta\zeta\right\}\right) \leq K_{s,p}\frac{s^{p/2}}{\zeta^p \delta^p} + K\frac{s^2}{\zeta^2 \delta^{2\beta}}.$$

*Proof:* We make use of the representation as in equation (3.32), i.e.

$$X(\delta)' = X_0 + B + X^c + (x\mathbb{1}_{\{|x| \leq \delta\}}) \star (\mu - \nu) - B(\delta).$$

[8]Despite not stated in the Lemma nor the proof, the paper only mentions a constant $K_p$, which we denote by $K_{s,p}$ as it depends on the value of $s$ too.

We define $M(\delta) := (x\mathbb{1}_{\{|x|\leq\delta\}}) \star (\mu - \nu)$. Further we do not include $B(\delta)'$, as introduced in equation (3.33), and hence, within this proof we denote $\hat{X} = B + X^c$ regardless of the value of $\beta'$. The estimate of Lemma 3.22.i) still holds (as the omitted $B(\delta)'$ in the case $\beta' \leq 1$ has finite variation) and, by first using a direct implication of Chebyshev's inequality (see, for example, [15, Remark 13.11]), we derive the following estimate for one of the increments of two of the summands of $X(\delta)'$:

$$\begin{aligned}\mathbb{P}_t\left(\left\{\left|\hat{X}_{t+s} - \hat{X}_t\right| \geq \delta\zeta/3\right\}\right) &\leq 3^p \frac{\mathbb{E}_t\left[\left|\hat{X}_{t+s} - \hat{X}_t\right|^p\right]}{\zeta^p\delta^p} \\ &\leq K_p' \frac{s^{p/2}}{\zeta^p\delta^p}.\end{aligned}$$

The choice of the threshold being $\delta\zeta/3$ is given by how we want to approach this proof. The idea of showing the inequality is that for any three rvs $Y_1, Y_2, Y_3$ we have the inclusion $\{|Y_1+Y_2+Y_3| > \delta\zeta\} \subseteq \{|Y_1| > \delta\zeta/3\} \cup \{|Y_2| > \delta\zeta/3\} \cup \{|Y_3| > \delta\zeta/3\}$. Further the probability gets bigger if we leave out the intersection with another set. This gives us:

$$\begin{aligned}&\mathbb{P}_t\left(\{N(\delta)_{t+s} - N(\delta)_t \geq 1\} \cap \left\{\left|X(\delta)'_{t+s} - X(\delta)'_t\right| > \delta\zeta\right\}\right) \\ &\quad\leq \mathbb{P}_t\Big(\{N(\delta)_{t+s} - N(\delta)_t \geq 1\} \cap \left\{\left|M(\delta)_{t+s} - M(\delta)_t\right| > \delta\zeta/3\right\}\Big) \\ &\qquad + \underbrace{\mathbb{P}_t\left(\left\{\left|\hat{X}_{t+s} - \hat{X}_t\right| \geq \delta\zeta/3\right\}\right)}_{\leq K_p' \frac{s^{p/2}}{\zeta^p\delta^p}} + \mathbb{P}_t\left(\left\{\left|B(\delta)_{t+s} - B(\delta)_t\right| \geq \delta\zeta/3\right\}\right),\end{aligned}$$

where we already state the estimate for the second summand from above. When looking at the summand including $B(\delta)$ we use the result of Lemma 3.22.iv) and formulate it to be independent of $\beta$ for $\delta \in (0,1]$:

$$|B(\delta)_{t+s} - B(\delta)_t| \leq \left\{\begin{matrix} Ks & \text{if } \beta < 1, \\ Ks\ln(1/\delta) & \text{if } \beta = 1, \\ Ks\delta^{1-\beta} & \text{if } \beta > 1, \end{matrix}\right\} \leq Ks/\delta, \quad \text{as} \begin{cases} 1 \leq 1/\delta, \\ \ln(1/\delta) \leq 1/\delta \Leftrightarrow 1/\delta \leq e^{1/\delta}, \\ \delta/\delta^\beta \leq 1/\delta \Leftrightarrow \delta^2 \leq \delta^\beta. \end{cases}$$

Hence, for small values of $s$ we can demand that $|B(\delta)_{t+s} - B(\delta)_t| \leq \delta\zeta/3$ (when we have $Ks/\delta \leq \delta\zeta/3$). Else we can choose $K_{s,p}$ in such a way, that the upper boundary fulfills $K_{s,p}\frac{s^{p/2}}{\zeta^p\delta^p} \geq 1$ (in which case the estimate for a probability obviously holds). The choice for $K_{s,p}$ needs to satisfy:

$$K_{s,p}\Big(\frac{s^{1/2}}{\zeta\delta}\Big)^p \geq 1 \;\Leftrightarrow\; K_{s,p}\frac{s^{1/2}}{\zeta\delta} \geq 1 \;\Leftrightarrow\; K_{s,p}s^{1/2} > 1/2 > \delta\zeta,$$

where $1/2$ is the upper boundary for $\delta\zeta$ due to their allowed range of values and hence we see that $K_{s,p}$ is independent of $\delta$ and $\zeta$. Having this, we only need to show that the first summand is bounded by $K\frac{s^2}{\zeta^2\delta^{2\beta}}$ in order to achieve the desired result. Therefore we introduce the shorthand notation for the increments:

$$N_s := N(\delta)_{t+s} - N(\delta)_t \qquad \text{and} \qquad M_s := M(\delta)_{t+s} - M(\delta)_t.$$

To get the desired, we perform the following basic transformation first:

$$\begin{aligned}
\mathbb{P}_t(N_s \geq 1, M_s \geq \zeta\delta/3) &= \mathbb{E}_t\left[\mathbb{1}_{\{N_s\geq 1\}}\,\mathbb{1}_{\{9M_s^2/(\zeta\delta)^2\geq 1\}}\right] \\
&\leq \mathbb{E}_t\left[N_s\mathbb{1}_{\{N_s\geq 1\}}\,\frac{9}{(\zeta\delta)^2}M_s^2\mathbb{1}_{\{9M_s^2/(\zeta\delta)^2\geq 1\}}\right] \\
&\leq \frac{9}{(\zeta\delta)^2}\mathbb{E}_t\left[N_s M_s^2\right].
\end{aligned}$$

Herein the last estimate is only allowed for $M_s^2$ (and would not work for $M_s$) as we would not know if negative values make the expected value smaller when considering the whole probability space. From here we move on by applying the integration by parts for semimartingales (as stated in [21, Theorem 9.33], for example) twice and obtain:

$$\begin{aligned}
N_s M_s^2 &= \int_0^s N_{r-}\,dM_r^2 \int_0^s M_{r-}^2\,dN_r + [N, M^2]_s \\
&= 2\int_0^s N_{r-}M_{r-}\,dM_r + \int_0^s N_{r-}\,d[M,M]_r + \int_0^s M_{r-}^2\,dN_r + [N, M^2]_s \\
&= 2\int_0^s N_{r-}M_{r-}\,dM_r + \int_0^s M_{r-}^2\,dN_r + \sum_{r\leq s} N_{r-}|\Delta M_r|^2,
\end{aligned}$$

where the last equality comes from $N$ (counting jumps bigger than $\delta$) and $M$ (adding the compensated jumps smaller or equal to $\delta$) having no common jumps and by the definition of the quadratic variation. We proceed by seeing that for $\lambda_r := \overline{F}_r(\delta)$, the process $\Lambda_s := \int_0^s \lambda_r dr$ is the compensator of the Poisson process $N$. From Lemma 3.18.iii) we see that $\lambda_r \leq K\delta^{-\beta}$. By iv) of the same lemma, we see that the compensator $\Lambda_s' := \int_0^s \lambda_r' dr$ of the quadratic variation of $M$ is bounded in a similar way, namely $\lambda_r' := \int_{\{|z|\leq\delta\}} z^2 F_r(dz) \leq K\delta^{2-\beta}$. Using these compensators we can now transform the expected value $\mathbb{E}_t[N_s M_s^2]$ to obtain the desired result. Obviously, the first summand vanishes, as it is a martingale, else we get:

$$\begin{aligned}
\mathbb{E}_t\left[N_s M_s^2\right] &= \mathbb{E}_t\bigg[\int_0^s M_r^2\,dN_r + \sum_{r\leq s} N_{r-}|\Delta M_r|^2\bigg] \\
&= \mathbb{E}_t\left[\int_0^s M_r^2\,d\Lambda_r + \int_0^s N_r\,d\Lambda_r'\right] \\
&\leq K\delta^{-\beta}\mathbb{E}_t\left[\int_0^s M_r^2\,dr + \int_0^s \delta^2 N_r\,dr\right] \\
&\leq K\delta^{-\beta}\mathbb{E}_t\left[\int_0^s \Lambda_r'\,dr + \int_0^s \delta^2\Lambda_r\,dr\right] \\
&\leq K\delta^{-2\beta}\mathbb{E}_t\left[\int_0^s r\delta^2\,dr + \int_0^s \delta^2 r\,dr\right] \leq K\delta^{2(1-\beta)}s^2.
\end{aligned}$$

□

The proof of the following three estimates uses all the results shown in the lemmas before. It can also be seen as relevant, as many further results will build on it. In this thesis, however, we do not go through the proof in detail, but rather describe what the authors did in the paper, where this statement is referred to as Lemma 6. The easy step therein is to take the boundaries found in Lemma 3.22. It is more challenging to find an estimate for the summand $\overline{X}(\delta)^b$, accounting for the compensated small jumps caused by $F_t'$, as the upper bound stated there is not good enough for achieving this result. To find a more precise value limiting the growth of this part, the authors construct a stable process, on which they apply Lemma 3.19, which was also shown in many steps in order give a strong enough estimate.

***Lemma 3.25:*** For $\alpha > 0$, $\varpi \in (0, 1/2)$ and $\eta \in (0, 1/2 - \varpi)$ we set

$$\rho := \eta \wedge (\varpi(\beta - \beta') - \beta'\eta) \wedge (\varpi\gamma) \wedge (1 - \varpi\beta - 2\eta), \tag{3.44}$$

and obtain a constant $K$ depending on the values $\alpha$, $\varpi$, $\eta$ and on the characteristics of $X$, which gives the following estimates for all $s \in (0, \Delta_n]$ and $t \in [0, T]$:

$$\text{i)} \qquad \left| \mathbb{P}_t(|X_{t+s} - X_t| > \alpha\Delta_n^{\varpi}) - \mathbb{E}_t\left[\int_t^{t+s} \overline{F}_r(\alpha\Delta_n^{\varpi})\, dr\right]\right| \le K\Delta_n^{1-\varpi\beta+\rho},$$

$$\text{ii)} \qquad \mathbb{P}_t(\alpha\Delta_n^{\varpi} < |X_{t+s} - X_t| \le \alpha\Delta_n^{\varpi}(1 + \Delta_n^{\varpi})) \le K\Delta_n^{1-\varpi\beta+\rho},$$

$$\text{iii)} \qquad \mathbb{P}_t(|X_{t+s} - X_t| > \alpha\Delta_n^{\varpi}) \le Ks\Delta_n^{1-\varpi\beta}.$$

*Proof:* Instead of showing the details of the proof, we want to describe the general approach used by the authors. As we will only be interested in small values of $\Delta_n$, we can always assume to have $\Delta_n$ as small as needed and we set $\delta_n = \alpha\Delta_n^{\varpi}$ for this whole proof. The value $\rho$ is chosen to be the minimum of four values. This can be seen as the natural way, as we often get upper boundaries in the form: $\Delta_n^z$, which increases for smaller values of $z > 0$ as $\Delta_n \in (0, 1)$. Hence, as we are combining various upper bounds with different values for $z$ we get a collective upper boundary, by taking the minimum of all appearing values of $z$, resulting in the value $\rho$ in our case. When starting the estimates, we proceed in the way we did in the proof of Lemma 3.24 as we make estimates for the three summands on the r.h.s. of[9]

$$X(\delta_n)' - \overline{X}(\delta_n)^b - X_0 = \hat{X} + \overline{X}(\delta_n)^a - B(\delta_n),$$

which is obtained when transforming equation (3.34). By using Chebyshev's inequality and the results of Lemma 3.22 a first estimate for the increments of this part of the semimartingale is given (i.e. for the increments of $X(\delta_n)' - \overline{X}(\delta_n)^b$, which represents the semimartingale $X$ without the big jumps and without the compensated small jumps resulting out of the Lévy

[9] Here there appears to be a typo in equation (67) in the paper, as the increments of $X(\delta_n)' + \overline{X}(\delta_n)^b$ are denoted, wherein the plus should be replaced by a minus.

measure $F_t'$). To adapt the estimate in order to be applicable to $X(\delta_n)'$ (where only the big jumps are missing) the probability space is extended and for every $n$ a stable Lévy process

$$Y(\delta_n)' := \overline{X}(\delta_n)^b + X'$$

can be defined, where $X'$ is an auxiliary semimartingale defined to make $Y(\delta_n)'$ a stable process. Using the result of Lemma 3.19, an estimate for the increments of $\overline{X}(\delta_n)^b$ is obtained, which is combined to provide an estimate for the increments of $X(\delta_n)'$. Finally the processes $N(\delta)$, counting jumps bigger than a certain threshold, as introduced in equation (3.43), are recalled. By considering these Poisson processes and their compensators, for the values

$$\delta_n \qquad \text{and} \qquad \delta_n' = \alpha\Delta_n^{\overline{\omega}}(1+\Delta_n^{\eta}),$$

the situation of Lemma 3.24 is resembled. The application thereof, together with various transformations, allows for the extension of the estimates derived for $X(\delta_n)'$ to the whole semimartingale $X$. □

The following lemma is the one referred to earlier, namely 'Lemma 7' in the original paper. As mentioned, we do not give any details on how it is proved. The outcome only forms part of the proof of Proposition 3.31 (which is the second result we will only cite). The proof of this lemma can be found on p.2237ff. of [2].

***Lemma 3.26:*** When being in the same situation as in Lemma 3.25 (i.e. $s \leq \Delta_n$, etc.), and for any bounded martingale $M := (M_t)_{t\geq 0}$, we see that there is a $K_M$, allowing for the estimate:

$$\begin{aligned}&\left|\mathbb{E}_t[(M_{t+s}-M_t)\mathbb{1}_{\{|X_{t+s}-X_t|>\alpha\Delta_n^{\overline{\omega}}\}}]\right| \\ &\qquad \leq K\Delta_n^{1-\overline{\omega}\beta+\rho} + K\Delta_n^{1-(\overline{\omega}+\eta)\beta}\mathbb{E}_t[|M_{t+s}-M_t|] + K\Delta_n^{(1-(\overline{\omega}+\eta)\beta)/2}\sqrt{\mathbb{E}_t[|M_{t+s}-M_t|^2]}.\end{aligned}$$

Next, we show two lemmas, which are auxiliary limit theorems. The stated assumptions still apply and the process $\overline{A}_t$ is as defined in equation (3.15). We will use the first of the two for proving the second, which is then used directly for verifying Proposition 3.30.

***Lemma 3.27:*** For $\overline{\omega} \in (0, 1/2)$, $\alpha > 0$ and $\rho' := 1/2 \wedge (\overline{\omega}(\beta-(\beta-\gamma)\vee\beta'))$ we see that for all $t > 0$ the sequence

$$\left(\Delta_n^{-\rho'}\left|\Delta_n^{\overline{\omega}\beta}\sum_{i=1}^{\lfloor t/\Delta_n\rfloor}\mathbb{E}_{i-1}^n\left[\int_{(i-1)\Delta_n}^{i\Delta_n}\overline{F}_s(\alpha\Delta_n^{\overline{\omega}})\,ds\right]-\frac{\overline{A}_t}{\alpha^\beta}\right|\right)_{n\geq 1}$$

is tight.

*Proof:* To increase readability we first introduce the following notation:

$$\theta_i^n := \int_{(i-1)\Delta_n}^{i\Delta_n} \overline{F}_s(\alpha\Delta_n^{\overline{\omega}})\, ds \qquad \text{and} \qquad \eta_i^n := \int_{(i-1)\Delta_n}^{i\Delta_n} A_s\, ds. \tag{3.45}$$

By recalling Definition 3.15, which states in equation (3.16) what a tight sequence needs to fulfill, we see that we need to show

$$\lim_{N\to\infty} \limsup_n \ \mathbb{P}\left(\Delta_n^{-\rho'} \left| \sum_{i=1}^{\lfloor t/\Delta_n \rfloor} \mathbb{E}_{i-1}^n \left[\Delta_n^{\overline{\omega}\beta}\theta_i^n\right] - \frac{\overline{A}_t}{\alpha^\beta} \right| > N \right) = 0.$$

In the process of showing this, we will use the well known fact, that for any rv $Y$ we have:

$$\mathbb{E}[|Y|] < \infty \quad \Rightarrow \quad \mathbb{P}(|Y| > N) \to 0. \tag{3.46}$$

One way of seeing this is by the inequality

$$\infty > \mathbb{E}[|Y|] > \mathbb{E}[\mathbb{1}_{\{|Y|\le N\}}|Y|] + N\mathbb{P}(|Y| > N) \qquad \forall N \in \mathbb{N}. \tag{3.47}$$

As $N$ can become arbitrary large, we see that the convergence to 0 must hold to maintain the inequality. To derive tightness we need to be more precise. Let us assume we want to verify that the sequence $Y^n$ of rvs is tight. We show that boundedness in $\mathcal{L}^1$:

$$\mathbb{E}[|Y^n|] < K \qquad \forall n \in \mathbb{N}, \tag{3.48}$$

i.e. the boundedness by a $K > 0$ (independent of $n$) for all $Y^n$, is enough to imply tightness. Let us hence assume we have the above boundedness. Following the idea of equation (3.47), we get the equivalence:

$$\begin{aligned} & K > \mathbb{E}[|Y^n|] > \mathbb{E}[\mathbb{1}_{\{|Y^n|\le N\}}|Y^n|] + N\mathbb{P}(|Y^n| > N) \\ \Leftrightarrow \qquad & \frac{K}{N} > \frac{K - \mathbb{E}[\mathbb{1}_{\{|Y^n|\le N\}}|Y^n|]}{N} > \mathbb{P}(|Y^n| > N). \end{aligned}$$

Seeing this, we can carry out the following estimate for the criterion of tightness:

$$\lim_{N\to\infty} \limsup_n \ \mathbb{P}(|Y^n| > N) < \lim_{N\to\infty} \limsup_n \frac{K}{N} = \lim_{N\to\infty} \frac{K}{N} = 0.$$

From all this we see that we only need to show the boundedness of the sequence in order to achieve tightness. But before we make use thereof we again proceed like in Lemma 3.24 where we saw that the sum of three rvs being bigger than $N$ has a smaller probability than each of the rvs being bigger than $N/3$. By including the summands $\mathbb{E}_{i-1}^n[\eta_i^n/\alpha^\beta]$ and $\eta_i^n/\alpha^\beta$ we can write:

$$\mathbb{P}\left(\Delta_n^{-\rho'} \left| \sum_{i=1}^{\lfloor t/\Delta_n \rfloor} \mathbb{E}_{i-1}^n \left[\Delta_n^{\overline{\omega}\beta}\theta_i^n\right] - \frac{\overline{A}_t}{\alpha^\beta} \right| > N \right)$$

$$
\begin{aligned}
&= \mathbb{P}\left(\Delta_n^{-\rho'}\left|\sum_{i=1}^{\lfloor t/\Delta_n\rfloor}\left(\mathbb{E}_{i-1}^n\left[\Delta_n^{\overline{\omega}\beta}\theta_i^n\right]+\frac{\mathbb{E}_{i-1}^n\left[\eta_i^n\right]-\mathbb{E}_{i-1}^n\left[\eta_i^n\right]+\eta_i^n-\eta_i^n}{\alpha^\beta}\right)-\frac{\overline{A}_t}{\alpha^\beta}\right|>N\right)\\
&= \mathbb{P}\left(\Delta_n^{-\rho'}\left|\sum_{i=1}^{\lfloor t/\Delta_n\rfloor}\mathbb{E}_{i-1}^n\left[\Delta_n^{\overline{\omega}\beta}\theta_i^n-\frac{\eta_i^n}{\alpha^\beta}\right]\right.\right.\\
&\qquad\qquad\left.\left.+\frac{1}{\alpha^\beta}\sum_{i=1}^{\lfloor t/\Delta_n\rfloor}\left(\mathbb{E}_{i-1}^n\left[\eta_i^n\right]-\eta_i^n\right)+\frac{1}{\alpha^\beta}\left(\sum_{i=1}^{\lfloor t/\Delta_n\rfloor}\eta_i^n-\overline{A}_t\right)\right|>N\right)\\
&\le \mathbb{P}\Bigg(\underbrace{\Delta_n^{-\rho'}\left|\sum_{i=1}^{\lfloor t/\Delta_n\rfloor}\mathbb{E}_{i-1}^n\left[\Delta_n^{\overline{\omega}\beta}\theta_i^n-\frac{\eta_i^n}{\alpha^\beta}\right]\right|}_{=:Y_1^n}>\frac{N}{3}\Bigg)\\
&\qquad\qquad+\mathbb{P}\Bigg(\underbrace{\left|\sum_{i=1}^{\lfloor t/\Delta_n\rfloor}\Delta_n^{-\rho'}\left(\eta_i^n-\mathbb{E}_{i-1}^n\left[\eta_i^n\right]\right)\right|}_{=:Y_2^n}>\frac{\alpha^\beta N}{3}\Bigg)\\
&\qquad\qquad\qquad+\mathbb{P}\Bigg(\underbrace{\Delta_n^{-\rho'}\left|\left(\sum_{i=1}^{\lfloor t/\Delta_n\rfloor}\eta_i^n-\overline{A}_t\right)\right|}_{=:Y_3^n}>\frac{\alpha^\beta N}{3}\Bigg).
\end{aligned}
$$

In a straight forward way, we first analyze the third summand with respect to $Y_3^n$. From the definition of $\eta_i^n$ we see that $\sum_{i=1}^{\lfloor t/\Delta_n\rfloor}\eta_i^n=\int_0^{\Delta_n\lfloor t/\Delta_n\rfloor}A_s\,ds$ and therefore we obtain the below from the boundedness of $A_s$, as stated in Assumption 3.12.i.b), and from $\rho'\le 1/2$:

$$|Y_3^n|\le\Delta_n^{-\rho'}\Delta_nK=\Delta_n^{1-\rho'}K\le K.$$

This boundedness for all $n$ of the rv, clearly implies that the sequence is tight. Now we look at the other two summands, where we show that equation (3.48) holds and hence get tightness. This means, we show that there is a common upper boundary for the expected value. We start with $Y_1^n$, where we give an estimate for the summands, which we get from applying Lemma 3.18.ii) and using the definition of $\rho'$:

$$
\begin{aligned}
\left|\Delta_n^{\overline{\omega}\beta}\theta_i^n-\frac{\eta_i^n}{\alpha^\beta}\right| &=\Delta_n^{\overline{\omega}\beta}\left|\int_{(i-1)\Delta_n}^{i\Delta_n}\overline{F}_s(\alpha\Delta_n^{\overline{\omega}})-\frac{A_s}{(\alpha\Delta_n^{\overline{\omega}})^\beta}\,ds\right|\\
&\le\Delta_n^{\overline{\omega}\beta}\Delta_n\frac{K''}{(\alpha\Delta_n^{\overline{\omega}})^{(\beta-\gamma)\vee\beta'}}\le K'\Delta_n^{1+\rho'}.
\end{aligned}
$$

Now we look at the expected value of $Y_1^n$ and can easily see the following when using the above estimate:

$$\mathbb{E}[|Y_1^n|]\le\mathbb{E}\Big[\Delta_n^{-\rho'}\sum_{i=1}^{\lfloor t/\Delta_n\rfloor}K'\Delta_n^{1+\rho'}\Big]\le\Delta_n^{-\rho'}(t\Delta_n^{-1})K'\Delta_n^{1+\rho'}\le K.$$

To achieve the tightness of the whole sequence we still need to show that $\mathbb{E}[|Y_2^n|] < K$. As we see that the summands in $Y_2^n$ are martingale increments, we make use of this property. Let us denote such an increment by $\zeta_i^n := \Delta_n^{-\rho'}\left(\eta_i^n - \mathbb{E}_{i-1}^n\left[\eta_i^n\right]\right)$. By looking at any martingale $M := (M_s)_{s\geq 0}$, with $M_0 = 0$, and any $k \in \mathbb{N}$, we see the following equality:

$$\mathbb{E}\Big[(M_{k\Delta_n})^2\Big] = \mathbb{E}\Big[\sum_{i=1}^{k}\left((M_{i\Delta_n})^2 - (M_{(i-1)\Delta_n})^2\right)\Big] = \mathbb{E}\Big[\sum_{i=1}^{k}\left(M_{i\Delta_n} - M_{(i-1)\Delta_n}\right)^2\Big].$$

Further, we know that (on a probability space) the fact $\mathbb{E}[|Y|^2] < \infty$ implies $\mathbb{E}[|Y|] \leq 1 + \mathbb{E}[|Y|^2] < \infty$, as can be found in [29, Theorem 4.19]. Hence, for all $\mathbb{E}[|Y_2^n|]$ to have a common upper bound, is implied by $\mathbb{E}[|Y_2^n|^2]$ having a common upper bound. Due to $Y_2^n$ being a martingale, and when replacing $(M_{i\Delta_n} - M_{(i-1)\Delta_n})$ by $\zeta_i^n$ in the above equation, we see that:

$$\mathbb{E}[|Y_2^n|^2] = \mathbb{E}\Big[\sum_{i=1}^{\lfloor t/\Delta_n\rfloor}|\zeta_i^n|^2\Big] < K \qquad \forall n \in \mathbb{N}, \tag{3.49}$$

implies the desired boundedness. The statement above follows easily, since $A_s$ is bounded and we can write

$$\begin{aligned}\left|\eta_i^n - \mathbb{E}_{i-1}^n\left[\eta_i^n\right]\right| &= \left|\int_{(i-1)\Delta_n}^{i\Delta_n} A_s\,ds - \mathbb{E}_{i-1}^n\Big[\int_{(i-1)\Delta_n}^{i\Delta_n} A_s\,ds\Big]\right| \leq K'\Delta_n \\ \Rightarrow \qquad |\zeta_i^n|^2 &= \Delta_n^{-2\rho'}\left|\eta_i^n - \mathbb{E}_{i-1}^n\left[\eta_i^n\right]\right|^2 \leq K'\Delta_n^{2-2\rho'},\end{aligned}$$

which allows us to complete the proof by using $1-2\rho' > 0$, and making the following estimate:

$$\mathbb{E}\Big[\sum_{i=1}^{\lfloor t/\Delta_n\rfloor}|\zeta_i^n|^2\Big] \leq (t\Delta_n^{-1})K'\Delta_n^{2-2\rho'} < K. \qquad \square$$

As stated before, we can use the outcome of the next lemma to show Proposition 3.30. To be precise, we use part i) for doing so. In the second part, we do not show the outcome of ii.b). The reason being that there appears to be an inconsistency. See the remark after the following proof for more details.

***Lemma 3.28:*** i) For $\overline{\omega} \in (0, 1/2)$, $\alpha > 0$ and

$$\chi' := (\overline{\omega}\gamma) \wedge \frac{1-\overline{\omega}\beta}{3} \wedge \frac{\overline{\omega}(\beta-\beta')}{1+\beta'} \wedge \frac{1-2\overline{\omega}}{2},$$

and for all $\epsilon > 0$, $t \in [0,T]$, the sequence

$$\left(\Delta_n^{\epsilon-\chi'}\left|\Delta_n^{\overline{\omega}\beta}\sum_{i=1}^{\lfloor t/\Delta_n\rfloor}\mathbb{P}_{i-1}^n\left(|\Delta_i^n X| > \alpha\Delta_n^{\overline{\omega}}\right) - \frac{\overline{A}_t}{\alpha^\beta}\right|\right)_{n\geq 1}$$

is tight. Specifically, what we obtain is:

$$\Delta_n^{\overline{\omega}\beta} \sum_{i=1}^{\lfloor t/\Delta_n \rfloor} \mathbb{P}_{i-1}^n \left( |\Delta_i^n X| > \alpha \Delta_n^{\overline{\omega}} \right) \xrightarrow{\mathbb{P}} \frac{\overline{A}_t}{\alpha^\beta}.$$

ii) Under the condition of $\gamma > \frac{\beta}{2}$, $\overline{\omega} \in \big(0, (2/(5\beta)) \wedge (1/(2+\beta))\big)$ and $\beta' \in \big[0, \beta/(2+\beta)\big)$, and for a bounded continuous martingale $M := (M_t)_{t \geq 0}$, we have for all $t \in [0,T]$ and $\alpha > 0$:

$$\text{a) } \Delta_n^{-\overline{\omega}\beta/2} \left| \Delta_n^{\overline{\omega}\beta} \sum_{i=1}^{\lfloor t/\Delta_n \rfloor} \mathbb{P}_{i-1}^n \left( |\Delta_i^n X| > \alpha \Delta_n^{\overline{\omega}} \right) - \frac{\overline{A}_t}{\alpha^\beta} \right| \xrightarrow{\mathbb{P}} 0,$$

$$\text{b) } \Delta_n^{\overline{\omega}\beta/2} \sum_{i=1}^{\lfloor t/\Delta_n \rfloor} \left| \mathbb{E}_{i-1}^n [\Delta_i^n M \mathbb{1}_{\{|\Delta_i^n X| > \alpha \Delta_n^{\overline{\omega}}\}}] \right| \xrightarrow{\mathbb{P}} 0.$$

*Proof:* i) To show this part, we see immediately that we only have to prove tightness, as the convergence in probability follows directly (as $\Delta_n^{\epsilon - \chi'} \to \infty$). The tightness is derived using the same approach as in the proof of Lemma 3.27 as we first include the summands $\Delta_n^{\overline{\omega}\beta} \mathbb{E}_{i-1}^n[\theta_i^n] - \Delta_n^{\overline{\omega}\beta} \mathbb{E}_{i-1}^n[\theta_i^n]$, where $\theta_i^n$ is defined as in equation (3.45). Hence, we get that the sequence in i) is tight if both sequences

$$\left( \Delta_n^{\epsilon - \chi'} \left| \Delta_n^{\overline{\omega}\beta} \sum_{i=1}^{\lfloor t/\Delta_n \rfloor} \underbrace{\left( \mathbb{P}_{i-1}^n \left( |\Delta_i^n X| > \alpha \Delta_n^{\overline{\omega}} \right) - \mathbb{E}_{i-1}^n [\theta_i^n] \right)}_{\leq K' \Delta_n^{1 - \overline{\omega}\beta + \rho}} \right| \right)_{n \geq 1}$$

and

$$\left( \Delta_n^{\epsilon - \chi'} \left| \Delta_n^{\overline{\omega}\beta} \sum_{i=1}^{\lfloor t/\Delta_n \rfloor} \mathbb{E}_{i-1}^n [\theta_i^n] - \frac{\overline{A}_t}{\alpha^\beta} \right| \right)_{n \geq 1}$$

are tight. To show this property of the first sequence, we use the estimate given for each summand in Lemma 3.25.i) as shown above. Therein $\rho$ is defined as in equation (3.44) of the same lemma. When adding all $\lfloor t/\Delta_n \rfloor$ summands, we see that each element of the sequence is smaller than $K\Delta^{\epsilon - \chi' + \rho}$. Hence, the first sequence is tight if we have $\chi' - \epsilon \leq \rho$. For the second sequence we only need to use the result of Lemma 3.27, to see that it is tight if we have $\chi' - \epsilon \leq \rho'$, where $\rho'$ is defined in the same lemma. When we combine the two conditions we get the strongest result, when setting

$$\chi' - \epsilon = \rho \wedge \rho'.$$

The result still holds but gets weaker, when the exponent is smaller than $\rho \wedge \rho'$, and so we can immediately see, that what we want to show, allows for an arbitrary large value of $\epsilon$. The reason for having 0 as the lower boundary of $\epsilon$ but still needing it, is that the definition of $\rho$ in equation (3.44) includes an $\eta \in (0, 1/2 - \overline{\omega})$, giving an open interval as the range of allowed values. As $\rho$ can take the value $\eta$ we show that the exponent of this lemma could take this value by allowing $\chi' = 1/2 - \varpi$ resulting in $\chi' - \epsilon = \rho = \eta < 1/2 - \varpi$ with the right choice of $\epsilon > 0$. To derive the definition of $\chi'$ we now look at

$$\rho \wedge \rho' = \Big(\eta \wedge (\varpi(\beta - \beta') - \beta'\eta) \wedge (\varpi\gamma) \wedge (1 - \varpi\beta - 2\eta)\Big) \wedge \Big(1/2 \wedge (\varpi(\beta - (\beta - \gamma) \vee \beta'))\Big),$$

where we only used the definition of the two values. We stated the possible range for $\eta$ above, and from Lemma 3.25 we also have $\varpi \in (0, 1/2)$. Let us first consider the case, when the minimum comes from a term containing $\eta$. If it is $\eta$ itself, we only have the existing upper bound of $< 1/2 - \varpi$. Else the term needs to be equal to $\eta$ or smaller, to account for the minimum. Considering this gives us the following upper bounds coming from the

$$\begin{aligned} &2^{\text{nd}} \text{ term: if } \quad \varpi(\beta - \beta') - \beta'\eta = \eta \qquad &&\Rightarrow \qquad \eta = \frac{\varpi(\beta - \beta')}{1 + \beta'}, \\ &4^{\text{th}} \text{ term: if } \quad 1 - \varpi\beta - 2\eta = \eta \qquad &&\Rightarrow \qquad \eta = \frac{1 - \varpi\beta}{3}. \end{aligned}$$

To arrive at the definition of $\chi'$ we need to show that the last two terms are redundant. For $1/2$ this is obvious, as it is always bigger than $\eta$. For the last term we distinguish the two cases:

$$\begin{aligned} &1)\ \beta - \gamma > \beta': \quad && \varpi(\beta - (\beta - \gamma)) = \varpi\gamma && \text{equal to } 3^{\text{rd}} \text{ term}, \\ &2)\ \beta - \gamma \leq \beta': \quad && \varpi(\beta - \beta') > \frac{\varpi(\beta - \beta')}{1 + \beta'} && \text{equal to boundary from } 2^{\text{rd}} \text{ term}. \end{aligned}$$

Putting all the bounds together we arrive at the definition of $\chi'$ and i) is proven.

ii) The given restrictions on $\beta'$, $\gamma$ and $\varpi$ are the same as in Theorem 3.17. These were derived from the transformations in equation (3.18), which demand all components of $\chi$ to be bigger than $\varpi\beta/2$, which immediately implies $\chi' > \varpi\beta/2$ as we have $\chi = \chi' \wedge \varpi\beta/2$. With $\varpi\beta/2$ being smaller, we immediately get the estimate in ii.a) from i) as $\epsilon$ can be chosen in the way that $\chi' - \epsilon = \varpi\beta/2$. The proof of part ii.b) is omitted here but commented on in the remark below. □

***Remark 3.29:*** Within the proof of Lemma 3.28.ii.b) the authors claim to use the result of Lemma 3.26 but we find that this result is too weak to show the desired. The problem lies in the first estimate displayed on p.2241, where the factor in the last summand is denoted by $\Delta_n^{(1-\eta\beta)/2}$ but to obtain the shown estimates, we feel that it would need to be $\Delta_n^{1-(\eta\beta)/2}$ instead. This replacement is assumed to be possible when looking at the last estimate of the proof of Lemma 3.26 on p.2239, where this factor is introduced. If we assume that the position of the bracket is just a typo, then the estimates displayed in stand-alone equations in the proof of Lemma 3.28 do hold. There is another typo in this first estimate displayed on p.2241, which just misses the power 2, which is then included in the more general estimate to follow. Further the proof states that $1-2\eta\beta > \varpi\beta$ (which is equivalent to $\eta < (1-\varpi\beta)/(2\beta)$) holds, because of the upper boundary $(1-\varpi\beta)/3$, shown for $\eta$ in part i) of the proof. For $\beta$ bigger than 1.5 however, this is not necessarily true. Even more so, the proof does not account for the dependence of $\rho$ on $\eta$, but it is defined in this way in equation (3.44). As we need $\rho > \varpi\beta/2$ (to make one of the summands finite) we clearly have to have $\eta > \varpi\beta/2$ as $\rho$ is the minimum over $\eta$ and some other values. Hence we would need that

$$\frac{1-\varpi\beta}{2\beta} > \frac{\varpi\beta}{2},$$

or more precisely, when using what we obtained in part i) of the same proof, we require that

$$\frac{1-\varpi\beta}{2\beta} > \eta = \left(\frac{1-\varpi\beta}{3} \wedge \frac{\varpi(\beta-\beta')}{1+\beta'} \wedge \frac{1-2\varpi}{2}\right) - \epsilon > \frac{\varpi\beta}{2}, \tag{3.50}$$

holds for the restricted values of $\varpi$ and any value of $\beta \in (0,2)$. If we choose, for example, $\beta = 1.75$ and $\varpi = 0.22$, which lies within the stated boundaries, the above inequality does not hold. Hence, the first inconsistency within the proof becomes irrelevant, as this second part of the proof does not hold for all values within the claimed boundaries. Seeing this, we feel that there must either be some other way of showing that the result holds for the stated boundaries, or we use this way and adapt the boundaries for $\varpi$ which make equation (3.50) hold. This would result in only allowing some lower values for $\varpi$. If such a restriction was to be needed, this would also influence the allowed range of $\varpi$ in Theorem 3.17 as the outcome of Lemma 3.28.ii.b) is used to proof the stated theorem. For sure, we do not want to doubt the stated boundaries in Theorem 3.17 and we do not go into any details of formulating new boundaries which would allow for the stated reasoning. What we will do, is carry out some simulation tests for lower values of $\varpi$ complying with equation (3.50). We will discuss the outcome of the implementations in Section 4.2.

Now we come to the last step in the process of proving Theorem 3.16. In the proposition below we formulate the result obtained for the conditioned probabilities $\mathbb{P}^n_{i-1}\left(|\Delta^n_i X| > \alpha\Delta^{\varpi}_n\right)$ in Lemma 3.28.i), for the idicator functions of these sets. By doing so, we get the desired result for the estimator $U(\varpi,\alpha)^n_t$, which we have already mentioned before introducing the estimator $\hat{\beta}_n$ in Definition 3.14.

***Proposition 3.30:*** Let $\chi$ be as defined in equation (3.17), then we get for every $t \in [0,T]$ and every $\epsilon > 0$, that the sequence

$$\left(\Delta_n^{\epsilon-\chi}\left|\Delta_n^{\varpi\beta}U(\varpi,\alpha)^n_t - \frac{\overline{A}_t}{\alpha^\beta}\right|\right)_{n\geq 1}$$

is tight, and especially:

$$\Delta_n^{\varpi\beta}U(\varpi,\alpha)^n_t \xrightarrow{\mathbb{P}} \frac{\overline{A}_t}{\alpha^\beta}.$$

*Proof:* We recall Definition 3.13 of $U(\varpi,\alpha)_t^n$ and denote by $D_i^n := \{|\Delta_i^n X| > \alpha\Delta_n^{\varpi}\}$ the set defining the indicator function in each summand (for arbitrary but fixed $\varpi$ and $\alpha$). Now we show the tightness in the way we did in Lemma 3.27. First we include the summands $\Delta_n^{\varpi\beta}\mathbb{E}_{i-1}^n[\mathbb{1}_{D_i^n}] - \Delta_n^{\varpi\beta}\mathbb{E}_{i-1}^n[\mathbb{1}_{D_i^n}]$. For ease of notation we define $\zeta_i^n := \Delta_n^{\varpi\beta/2}(\mathbb{1}_{D_i^n} - \mathbb{E}_{i-1}^n[\mathbb{1}_{D_i^n}])$ and see that it is a martingale increment. (Hence, the same notation as in the proof of Lemma 3.27.) To show the desired, we state the following estimate, using what we just announced:

$$
\begin{aligned}
&\mathbb{P}\left(\Delta_n^{\epsilon-\chi}\left|\Delta_n^{\varpi\beta}U(\varpi,\alpha)_t^n - \frac{\overline{A}_t}{\alpha^\beta}\right| > N\right)\\
&\quad = \mathbb{P}\left(\Delta_n^{\epsilon-\chi}\left|\Delta_n^{\varpi\beta}\sum_{i=1}^{\lfloor t/\Delta_n\rfloor}\big(\mathbb{1}_{D_i^n} - \mathbb{E}_{i-1}^n[\mathbb{1}_{D_i^n}] + \mathbb{E}_{i-1}^n[\mathbb{1}_{D_i^n}]\big) - \frac{\overline{A}_t}{\alpha^\beta}\right| > N\right)\\
&\quad \le \mathbb{P}\left(\Delta_n^{\epsilon-\chi}\left|\Delta_n^{\varpi\beta/2}\sum_{i=1}^{\lfloor t/\Delta_n\rfloor}\zeta_i^n\right| > \frac{N}{2}\right) + \mathbb{P}\left(\Delta_n^{\epsilon-\chi}\left|\Delta_n^{\varpi\beta}\sum_{i=1}^{\lfloor t/\Delta_n\rfloor}\mathbb{P}_{i-1}^n(D_i^n) - \frac{\overline{A}_t}{\alpha^\beta}\right| > \frac{N}{2}\right)
\end{aligned}
$$

The second probability above contains exactly the same summands as in Lemma 3.28.i) and since we already saw in the proof of Lemma 3.28.ii) that $\chi = \chi' \wedge (\varpi\beta/2)$, we can chose $\epsilon$ in an appropriate way to see that the sequence is tight. For the analysis of the first probability above, we note that $\epsilon - \chi + \varpi\beta/2 \ge \epsilon > 0$ which means that we do not need to consider the factor coming from $\Delta_n$, which is not already included in $\zeta_i^n$. For the remaining term we use the fact that we are dealing with martingale increments $\zeta_i^n$. As described in the proof of Lemma 3.27, in this situation we only need to show that equation (3.49) holds. So all we need to complete the proof, is the following inequality, where we take the estimate $\mathbb{P}_{i-1}^n(D_i^n) \le K'\Delta_n^{1-\varpi\beta}$ from Lemma 3.25.iii):

$$
\begin{aligned}
\mathbb{E}\left[|\zeta_i^n|^2\right] &= \Delta_n^{\varpi\beta}\mathbb{E}\left[\left|\mathbb{1}_{D_i^n} - \mathbb{P}_{i-1}^n(D_i^n)\right|^2\right]\\
&= \Delta_n^{\varpi\beta}\mathbb{E}\Big[\mathbb{1}_{D_i^n} - 2\mathbb{1}_{D_i^n}\mathbb{P}_{i-1}^n(D_i^n) + \mathbb{P}_{i-1}^n(D_i^n)^2\Big]\\
&\le \Delta_n^{\varpi\beta}\Big(\mathbb{P}(D_i^n) + 2\mathbb{P}(D_i^n)K'\Delta_n^{1-\varpi\beta} + (K')^2\Delta_n^{2(1-\varpi\beta)}\Big)\\
&\le K''\Delta_n
\end{aligned}
$$

$$
\Rightarrow \mathbb{E}\left[\sum_{i=1}^{\lfloor t/\Delta_n\rfloor}|\zeta_i^n|^2\right] \le (t\Delta_n^{-1})K''\Delta_n \le K. \qquad \square
$$

We are now at the final step required for showing the proofs. The proposition below forms the last part of showing Theorem 3.17. As stated earlier, we will only cite its outcome as it strongly builds on results from [25], namely Thorem IX.7.28. Further it uses the result of Lemma 3.28.ii.b), where we pointed out an inconsistency in Remark 3.29. In paper [2] the proof is given on p.2242.

***Proposition 3.31:*** For every $\gamma > \frac{\beta}{2}$, $\overline{\omega} \in \big(0, (2/(5\beta)) \wedge (1/(2+\beta))\big)$ and $\beta' \in \big[0, \beta/(2+\beta)\big)$ as in ii) of Lemma 3.28, and $\alpha' > \alpha$, the pair of processes

$$\Delta_n^{-\overline{\omega}\beta/2}\left(\Delta_n^{\overline{\omega}\beta} U(\overline{\omega},\alpha)_t^n \; - \frac{\overline{A}_t}{\alpha^\beta}, \Delta_n^{\overline{\omega}\beta} U(\overline{\omega},\alpha')_t^n \; - \frac{\overline{A}_t}{\alpha'^\beta}\right)_{n\geq 1}$$

converges stably in law to a continuous Gaussian martingale $(\overline{W}, \overline{W}') := (\overline{W}_t, \overline{W}'_t)_{t\geq 0}$ independent of $\mathscr{F}$ and fulfilling:

$$\mathbb{E}[\overline{W}_t^2] = \frac{\overline{A}_t}{\alpha^\beta}, \qquad \mathbb{E}[\overline{W}_t'^2] = \frac{\overline{A}_t}{\alpha'^\beta} \qquad \text{and} \qquad \mathbb{E}[\overline{W}_t \overline{W}'_t] = \frac{\overline{A}_t}{\alpha'^\beta}.$$

Having stated all needed building blocks for the proofs we are now ready to verify Theorem 3.16. As the shown results had been interwoven and building onto each other, all we need for proving the first of the main results, is the outcome of Proposition 3.30.

*Proof of Theorem 3.16:* For better readability we introduce the shorthand notation, which we will use when transforming the estimator $\hat{\beta}_n$:

$$V_n := \Delta_n^{\epsilon-\chi}\left(\Delta_n^{\overline{\omega}\beta} U(\overline{\omega},\alpha)_t^n \; - \frac{\overline{A}_t}{\alpha^\beta}\right) \qquad \text{and} \qquad V'_n := \Delta_n^{\epsilon-\chi}\left(\Delta_n^{\overline{\omega}\beta} U(\overline{\omega},\alpha')_t^n \; - \frac{\overline{A}_t}{\alpha'^\beta}\right).$$

As we want to show that $\hat{\beta}_n(t,\overline{\omega},\alpha,\alpha') \xrightarrow{\mathbb{P}} \beta$ on $\{\overline{A}_t > 0\}$, we analyze the difference of the sequence and its desired limit. On this set we obviously do not have a problem with dividing by $\overline{A}_t \neq 0$. Using the definition of the estimator $\hat{\beta}_n(t,\overline{\omega},\alpha,\alpha')$, we can write (by including the factor $\ln(\alpha'/\alpha)$ for ease of notation):

$$\begin{aligned}
\ln\big(\alpha'/\alpha\big)\big(\hat{\beta}_n(t,\overline{\omega},\alpha,\alpha') - \beta\big) &= \ln\big(U(\overline{\omega},\alpha)_t^n / U(\overline{\omega},\alpha')_t^n\big) - \beta \ln\big(\alpha'/\alpha\big) \\
&= \ln\left(\frac{\alpha^\beta U(\overline{\omega},\alpha)_t^n}{\alpha'^\beta U(\overline{\omega},\alpha')_t^n}\, \frac{\Delta_n^{\overline{\omega}\beta}/\overline{A}_t}{\Delta_n^{\overline{\omega}\beta}/\overline{A}_t}\right) \\
&= \ln\left(\frac{1 + \alpha^\beta\big(\Delta_n^{\overline{\omega}\beta} U(\overline{\omega},\alpha)_t^n - \overline{A}_t/\alpha^\beta\big)/\overline{A}_t}{1 + \alpha'^\beta\big(\Delta_n^{\overline{\omega}\beta} U(\overline{\omega},\alpha')_t^n - \overline{A}_t/\alpha'^\beta\big)/\overline{A}_t}\right) \ldots
\end{aligned}$$

By using that $\Delta_n^{\chi-\epsilon} V_n = (\Delta_n^{\overline{\omega}\beta} U(\overline{\omega},\alpha)_t^n - \overline{A}_t/\alpha^\beta) \xrightarrow{\mathbb{P}} 0$ by Proposition 3.30 and the approximation $\ln(x+1) \approx x$ for $x \to 0$ (which is easily seen when applying l'Hospital's rule to $\lim_{x\to 0} \ln(1+x)/x = \lim_{x\to 0} 1/(1+x) = 1$), we further get:

$$\ldots \overset{\mathbb{P}}{=} \frac{\alpha^\beta \Delta_n^{\chi-\epsilon} V_n - \alpha'^\beta \Delta_n^{\chi-\epsilon} V'_n}{\overline{A}_t} \xrightarrow{\mathbb{P}} 0.$$

So we have shown the desired convergence in probability. Further we point out that from the above we can also see the $\Delta_n^{\chi-\epsilon}$-rate consistency. The above is equivalent to:

$$\Delta_n^{\epsilon-\chi}\big(\hat{\beta}_n(t,\overline{\omega},\alpha,\alpha') - \beta\big) \overset{\mathbb{P}}{=} \frac{\alpha^\beta V_n - \alpha'^\beta V'_n}{\ln(\alpha'/\alpha)\overline{A}_t}, \tag{3.51}$$

which is tight, as $V_n$ and $V_n'$ are tight by what we showed in Proposition 3.30. This completes the proof. □

Using what we have just shown and replacing $\Delta_n^{\chi-\epsilon}$ with $\Delta_n^{\varpi\beta/2}$ we get the proof of Theorem 3.17. Also, we need to make use of the outcome of Proposition 3.31, giving us the desired property of stable convergence.

*Proof of Theorem 3.17:* To show the first part, we arrive at equation (3.51) with the exponent $-\varpi\beta/2$ instead of $\epsilon-\chi$ in exactly the same way as above. This change of exponent is achieved by defining $V_n$ and $V_n'$ in a slightly different way, namely:

$$V_n := \Delta_n^{-\varpi\beta/2}\left(\Delta_n^{\varpi\beta}U(\varpi,\alpha)_t^n - \frac{\overline{A}_t}{\alpha^\beta}\right) \qquad \text{and} \qquad V_n' := \Delta_n^{-\varpi\beta/2}\left(\Delta_n^{\varpi\beta}U(\varpi,\alpha')_t^n - \frac{\overline{A}_t}{\alpha'^\beta}\right).$$

Now we use the outcome of Proposition 3.31, which describes the stable convergence of $(V_n, V_n')$ to $(\overline{W}, \overline{W}')$. By using the stated expected values in Proposition 3.31, we easily see that

$$\mathbb{E}\left[\left(\frac{\alpha^\beta}{\ln(\alpha'/\alpha)\overline{A}_t}\overline{W}\right)^2\right] = \frac{\alpha^\beta}{\ln(\alpha'/\alpha)^2\overline{A}_t}.$$

By seeing the same result for $\overline{W}'$ we get that the r.h.s. of equation (3.51) converges stably to a normally distributed variable with variance $(\alpha'^\beta - \alpha^\beta)/(\ln(\alpha'/\alpha)^2\overline{A}_t)$. Hence, we have shown the proof of part i).

To show part ii) we just state that it directly follows from the convergence in probability of the statistic $U(\varpi,\alpha)_t^n$, which we derived in Proposition 3.30, in combination with Proposition 2.16.ii). The latter is giving a general result on combining stable convergence with convergence in probability. Hence, dividing the l.h.s. of part i) by the standard deviation of the r.h.s. gives the desired outcome. □

The proof of the distributional result above was now made under the assumption of using all outcomes of the lemmas and propositions, in the way they are stated in paper [2]. When we were discussing them however, we felt that there might be some inconsistencies with regard to the parameter $\varpi$. We will carry out some simulations with regard to this assumption in Section 4.2. Before, we will include simulations rebuilding tables of the stated paper. These implementations can be found in the next chapter and herewith we end the discussion of the theory.

# Chapter 4

# Implementation of the Jump Activity Index Estimator

In contrast to the previous chapters, this one does not deal with the theory as such, but rather with the implementation thereof. We assume to have a certain process made up of a jump process and a continuous part, as defined in equation (4.1). By the jump process $Y$ being $\beta$-stable we fulfill Assumption 3.12, with $\beta' = 0$ and with $f(x,t) \equiv 0$, which allows us to take $\gamma$ arbitrarily large (and obviously we can assume $\gamma > \beta/2$ which is required in Theorem 3.17). We follow the framework by Aït-Sahalia and Jacod in Section 6 of their paper [2] and visualize the main results for the estimator $\hat{\beta}$. Hereby we show results for the mean and standard deviation and include histograms for the normalized estimator converging to a normally distributed rv. Then we show some results for different parameters, which can not be found in the stated paper. All simulations are carried out using code written in MATLAB. We shortly describe how the code works within this chapter and we provide the commented code in the appendix.

## 4.1 Monte Carlo Simulation visualizing the Main Results

The aim of this section is to implement the estimator $\hat{\beta}$ and to depict its convergence for different values of $\beta$, different probabilities for big jumps and different time steps $\Delta_n$. Doing this, we wish to recreate Table 1 of [2]. This allows us to somehow 'test' our simulation by comparing our results to the stated table. A difficulty in this comparison is, that the authors apply a so-called 'small sample bias correction', which they describe in [2, Section 5]. This correction does not change the asymptotic results, as the correction term they subtract from the estimator $\hat{\beta}$ is of order $\Delta_n^{1-2\varpi}$. Despite not explicitly stated in the paper how their estimates are affected by this correction, we can tell from its structure that it has a bigger influence for larger values of $\beta$ and larger values of $\Delta_n$. So the comparison we make is between our 'raw' outcomes and the authors' 'corrected' estimates. Our approach will show, where the application of the 'raw' estimates might not deliver the desired results for certain parameters. We will point out such situations when comparing the mean values of the estimators.

The aforementioned table shows different values of $\beta$ in the rows. Different time steps in the discrete observation and different tail probabilities are considered in the columns. These tail probabilities give the likeliness of the underlying process having increments bigger than the truncation level $\alpha\Delta_n^{\overline{\omega}}$. This probability (set to be 0.25%, 0.5%, 1.0% or 2.5%) gives a rough percentage of how many of the observed increments will actually be counted by the statistic $U(\overline{\omega},\alpha)_t^n$, given in Definition 3.13. Despite the stated assumptions allowing for a quite general stochastic process, we chose to carry out our Monte Carlo simulations using an underlying stochastic process of the following structure:

$$\begin{aligned} dX_t =& \sigma_t\, dW_t + \theta\, dY_t, \qquad (4.1)\\ \text{with} \quad \sigma_t = v_t^{1/2}, \quad dv_t =& \kappa(\eta - v_t)\, dt + \gamma v_t^{1/2}\, dB_t \quad \text{and} \quad \mathbb{E}[dW_t dB_t] = \rho\, dt. \end{aligned}$$

Herein the process $Y$ is a $\beta$-stable jump process. Within the paper the case $\beta = 0$ is included in the analysis, but we discuss it separately and do not include its analysis just now. Both, $B$ and $W$ are Brownian motions which are correlated in the stated way. We note that the stochastic volatility follows a Cox-Ingersoll-Ross model with mean reversion towards the value $\eta$. In the choice of the parameters, we adopt the authors' and set them to: $\eta = 1/16$, $\gamma = 1/2$ (which does not stay in any relation to the $\gamma$ of Assumption 3.12), $\kappa = 5$, $\rho = -1/2$, and the scale-parameter $\theta$ will be chosen in such a way, that the tail probability $\mathbb{P}(\theta|\Delta_i^n Y_t| > \alpha\Delta_n^{\overline{\omega}})$ has a set value. The parameters to define the truncation level are chosen to be $\alpha = 5\eta$, $\alpha' = 10\eta$ and $\overline{\omega} = 1/5$. To get an idea of how such a process could look like, we provide Figure 4.1. Therein we display the continuous part and the jump part separately and also include the sum of the two. We chose the value $\beta = 1$, step size of 1 second and the tail probability (defining the scale parameter $\theta$) to be 1%.

We see that this choice of the scale parameter $\theta$ makes the sum of the two processes reassemble the pure jump process. This is obviously caused by looking at the sum at a big scale, where we do not see the small changes caused by the continuous part. The code for generating this sample path is provided in the appendix. Here we describe the approach for implementing the stated model in some more detail. To simulate a discrete version of the process, we write a function for the continuous part, namely `simulateContDis(dt,T,eta,gamma,kappa,rho)`, and a second one, which we call `simulateStable(dt,T,beta,theta)`, for generating the jump process. Therein the Greek letters are placeholders for the parameters mentioned above. The value `dt` represents the time steps $\Delta_n$ in seconds and `T` represents the chosen amount of observed trading days. We assume one trading day to be made up of 6.5 hours, resulting in $60^2 \cdot 6.5 = 23,400$ seconds. As our results hold for small time steps, we convert our time step to be in the unit of days and hence transform the input `dt` by dividing it by $23,400$. Despite

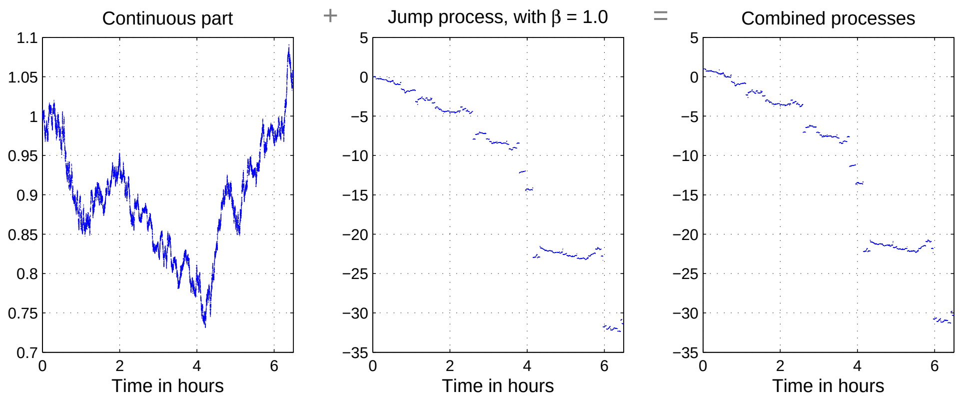


Figure 4.1: Sample path of the model stated in equation (4.1) with $\beta = 1$, $\Delta_n = 1$ second and tail probability of 1%.

our code allowing for more freedom, we choose to observe 1 trading day in what follows. In simulating the continuous part we follow an Euler scheme (see for example, [24, Section 5.6.3]) and discretize the stochastic differential equation for $v_t$ for increments of size $\Delta_n$. We use the inbuilt MATLAB-function `mvnrnd` to get two correlated Brownian motions. Then we obtain the continuous part by summing up over the simulated increments. For the simulation of a stable process $Y$ we follow [17, Algorithm 6.6], which combines a uniformly distributed rv $U$ on $(-\pi/2, \pi/2)$ and an independent standard exponential rv $W$ in such a way, that we obtain an increment $\Delta_i^n Y$ of a stable process. The definition is the following

$$\Delta_i^n Y = (\Delta_n)^{1/\beta} \frac{\sin(\beta U)}{\cos(U)^{1/\beta}} \left( \frac{\cos((1-\beta)U)}{W} \right)^{(1-\beta)/\beta},$$

and the interested reader is referred to the aforementioned source of the algorithm for more details. Now we have what we need to simulate a sample path of the underlying process.

The next step in recreating Table 1 is to calculate the scale parameter $\theta$. To set the tail probability for big jumps to a certain value, we need to define $\theta$ such that $\mathbb{P}(\theta|\Delta_i^n Y_t| \geq \alpha \Delta_n^{\varpi})$ is the desired tail probability. We do not program our own code for getting the inverse distribution function, but instead use the toolbox Veillette provides on the Mathworks file exchange.[1] His toolbox on stable distributions is mainly based on the book [39]. Further, Veillette is writing his PhD with Taqqu, one of the authors of this book. What he does to obtain the inverse distribution function is to use a combination of Newton's method and the bisection method. He implements this in the code `stblinv`, which also uses two other functions of the same toolbox, namely `stblcdf` and `stblpdf`.

[1]The toolbox can be downloaded from the official Mathworks website: `http://www.mathworks.com/matlabcentral/fileexchange/37514-stbl-alpha-stable-distributions-for-matlab` [Accessed 09/05/13].

Our code `calcSampleMeanStd(dt,tail_pr,beta,N,T,om,alph,eta,gamma,kappa,rho)` makes direct use of `stblinv`. Therein we calculate the entries replicating Table 1, i.e. the mean and standard deviation of estimator $\hat{\beta}$. As an input for this function we choose the amount of simulation runs to be $N = 5,000$, which is in line with the choice of Aït-Sahalia and Jacod. The code is written in three main loops. The outer loop goes through the different values of $\beta$, so through the rows. The next goes through the tail probabilities in combination with the size of the time steps. Herein we calculate $\theta$ for each entry in the table. Then the innermost loop follows, and accounts for the number of simulations $N$, creating the sample paths. We first apply the two aforementioned codes to get a sample path of the underlying process. Then we calculate the statistics $U(\varpi, \alpha)_t^n$ and $U(\varpi, \alpha')_t^n$ (in a loop for only two iterations). Having this we calculate $\hat{\beta}$ in accordance with Definition 3.14, also using the convention that we set $\hat{\beta} = 0$ if either of the statistics $U$ is 0. Before moving to the next position in the table, we store the mean and the standard deviation of all $N$ sample paths into the output matrix. We run this function using the script file `sBasicTable`.

| **Step size $\Delta_n$**<br>**Tail probability** | **1 sec**<br>**0.25%** | **1 sec**<br>**0.5%** | **1 sec**<br>**1.0%** | **1 sec**<br>**2.5%** | **5 sec**<br>**1.0%** |
|---|---|---|---|---|---|
| $\beta = 1.50$ | 1.54<br>(1.52) | 1.53<br>(1.51) | 1.55<br>(1.50) | 1.60<br>(1.52) | 1.57<br>(1.53) |
| $\beta = 1.25$ | 1.27<br>(1.27) | 1.26<br>(1.26) | 1.26<br>(1.25) | 1.28<br>(1.26) | 1.29<br>(1.27) |
| $\beta = 1.00$ | 1.01<br>(1.01) | 1.01<br>(1.01) | 1.00<br>(1.00) | 1.00<br>(1.00) | 1.02<br>(1.01) |
| $\beta = 0.75$ | 0.76<br>(0.76) | 0.75<br>(0.76) | 0.75<br>(0.75) | 0.75<br>(0.75) | 0.76<br>(0.76) |
| $\beta = 0.50$ | 0.51<br>(0.51) | 0.50<br>(0.50) | 0.50<br>(0.50) | 0.50<br>(0.50) | 0.50<br>(0.50) |
| $\beta = 0.25$ | 0.25<br>(0.25) | 0.25<br>(0.25) | 0.25<br>(0.25) | 0.25<br>(0.25) | 0.25<br>(0.25) |

Table 4.1: Sample mean of estimator $\hat{\beta}$ resulting out of a Monte Carlo simulation with $5,000$ runs (in comparison with the outcome displayed in [2]).

The obtained output of mean values can be seen in Table 4.1. Additionally to our calculated values we place the ones from Table 1 in [2] in brackets, just below our values. If the two are equal we shade the cell in gray. Hence, we easily see that we obtain the same estimators for most values with $\beta \leq 1$. For $\beta = 1.50$ our estimates are always poorer than the ones depicted in the paper. Further we see that in the last column, where the time step is 5 seconds, and not only 1 second, the results are still further away from the actual value of $\beta$. A possible reason could be the influence of the bias correction which is applied to the original table. This would be in line with what we stated earlier, as the correction term, which is subtracted as part of the small sample bias correction, has more influence for bigger values of $\beta$ and bigger time steps. We see this by the structure of the correction term given in [2, equation (42)], which is

$$\frac{1}{\ln(\alpha/\alpha')}\left[\frac{\beta(\beta+1)\sigma^2}{2}\left(\frac{1}{\alpha^2}-\frac{1}{\alpha'^2}\right)\Delta_n^{1-2\varpi}+\frac{d_\beta\theta^\beta}{2c_\beta}\left(\frac{1}{\alpha^\beta}-\frac{1}{\alpha'^\beta}\right)\Delta_n^{1-2\beta}\right]$$

$$\text{with} \qquad c_\beta = \frac{\Gamma(\beta+1)}{2\pi}\sin(\pi\beta/2) \qquad \text{and} \qquad d_\beta = -\frac{\Gamma(2\beta+1)}{8\pi}\sin(\pi\beta).$$

Herein $\sigma$ was assumed to come from a constant volatility and $\Gamma$ represents the Gamma function. For values of $\beta > 1$ we see that both summands are being positive. Hence the estimator without the bias correction can be assumed to be too big for large values of $\beta$. This assumption would be in line with what we see in our table above, but it remains an assumptions.

We point out that not restricting our results onto a subset is fine in our situation, as we have $\{\overline{A}_t > 0\}$ being the whole set. Hence we are in line with the outcome of the main results. Further we have Theorem 3.17 providing information on the distribution. Just now we only have a closer look at the standard deviation, which should be smaller for smaller values of $\Delta_n$ and $\beta$ (as $(\alpha'\beta - \alpha\beta)/\overline{A}_t$ with $\overline{A}_t$ being proportional to $1/\beta$, is decreasing for smaller values of $\beta$). We can illustrate this by using the aforementioned script file `sBasicTable` to give out the values of the standard deviations. We display them in Table 4.2. Therein we use the same logic as in the table before and hence shade the values being equal to the outcome in paper [2] in gray. In every column we see that the values decrease for smaller values of $\beta$, which is what we would hope to get. Further we see that the biggest difference between our simulation and the authors' is present for $\beta = 0.25$ and for the bigger time step of 5 seconds. We do not see the reason for this coming from the bias correction. In both these cases the convergence is said to be slower, either because of the larger observation intervals, or because $\beta$ is closer to $\beta' = 0$, which also influences the rate of convergence, as we saw in the definition of $\chi$ in equation (3.17). Maybe this slower speed of convergence plays a role in seeing different values in these cases.

In contrast to the authors' approach we chose not to include a row for the value $\beta = 0$ in both tables, as this value is not considered by the mathematical analysis as stated in Assumption 3.12. Nevertheless we also included the implementation for $\beta = 0$ in our code `simulateStable(dt,T,beta,theta)`. In this case the parameter `theta` represents the ar-

| **Step size $\Delta_n$**<br>**Tail probability** | **1 sec**<br>**0.25%** | **1 sec**<br>**0.5%** | **1 sec**<br>**1.0%** | **1 sec**<br>**2.5%** | **5 sec**<br>**1.0%** |
|---|---|---|---|---|---|
| $\beta = 1.50$ | 0.26<br>(0.26) | 0.19<br>(0.18) | 0.13<br>(0.13) | 0.08<br>(0.08) | 0.30<br>(0.24) |
| $\beta = 1.25$ | 0.23<br>(0.23) | 0.16<br>(0.16) | 0.11<br>(0.11) | 0.07<br>(0.07) | 0.26<br>(0.19) |
| $\beta = 1.00$ | 0.19<br>(0.19) | 0.14<br>(0.14) | 0.10<br>(0.10) | 0.06<br>(0.06) | 0.22<br>(0.14) |
| $\beta = 0.75$ | 0.16<br>(0.16) | 0.11<br>(0.11) | 0.08<br>(0.08) | 0.05<br>(0.05) | 0.18<br>(0.11) |
| $\beta = 0.50$ | 0.12<br>(0.13) | 0.09<br>(0.09) | 0.06<br>(0.06) | 0.04<br>(0.04) | 0.14<br>(0.08) |
| $\beta = 0.25$ | 0.11<br>(0.09) | 0.09<br>(0.06) | 0.09<br>(0.04) | 0.07<br>(0.04) | 0.10<br>(0.05) |

Table 4.2: Sample standard deviation of estimator $\hat{\beta}$ resulting out of a Monte Carlo simulation with $5,000$ runs (in comparison with the outcome displayed in [2]).

rival rate of a Poisson process. If we choose the model only containing finitely many jumps in the same way as in the Monte Carlo simulations in [2], then we do not assume random jump sizes.[2] We instead fix the jump size to 0.1 and choose the arrival rate `theta` such that $\mathbb{P}(\theta|\Delta_i^n Y| \geq \alpha\Delta_n^{\varpi})$ is equal to one of the set tail probabilities. Maintaining all other parameters as we had them before, results in the Monte Carlo simulation of the mean value of $\hat{\beta}$ showing values greater than 2. This value does not result in the desired, but it is no contradiction. As soon as we choose the step size $\Delta_n$ small enough (in relation to the jump size) we arrive at the mean of $\hat{\beta}$ being equal to 0. This is logical, as at one point the statistics $U(\varpi, \alpha)_t^n$ and $U(\varpi, \alpha')_t^n$ will both contain all jumps coming from the Poisson process. However, the observation of $\beta = 0$ lies without the considered scenarios of the main results as we only allow for $\beta > 0$ in Assumption 3.12. Still, the estimator results in the desired for small enough step sizes.

[2]The text on p.2220 explains to fix the jump size, but Table 1 showing the results, states that we are considering a compound Poisson process, which might hint at random jump sizes. Also, the results in Table 1 are all close to 0 and hence very different to our outcome for $\beta = 0$, which reaches values up to 6.4 in some simulations.

What we have shown in the two tables so far, can be seen as visualizing the outcome of Theorem 3.16. The behavior described in Theorem 3.17 is a stronger result and we cannot see it out of the analysis above. To see the distributional result we instead need to observe all sample paths. We do this in the function `makeHistNorm`, where we normalize $\hat{\beta}$, resulting out of one sample path, in the way shown in equation ii) of Theorem 3.17. We use these normalized values to create a histogram, and overlay the normal distribution function to allow for a visual comparison. To get a feeling for the convergence towards a normally distributed rv we choose bigger time steps than before first, namely 30 and 10 are computed in addition to the step size of 1 second. Again, we vary the step size by column. To get an overview of different values of $\beta$ we chose three values, namely 1.5, 1.0 and 0.5 shown in the rows. Using the script file `sBasicHistNorm` we create the described simulations depicted in Figure 4.2. Within the simulations we fixed the tail probability to 1% and left the other parameters as they had been before.

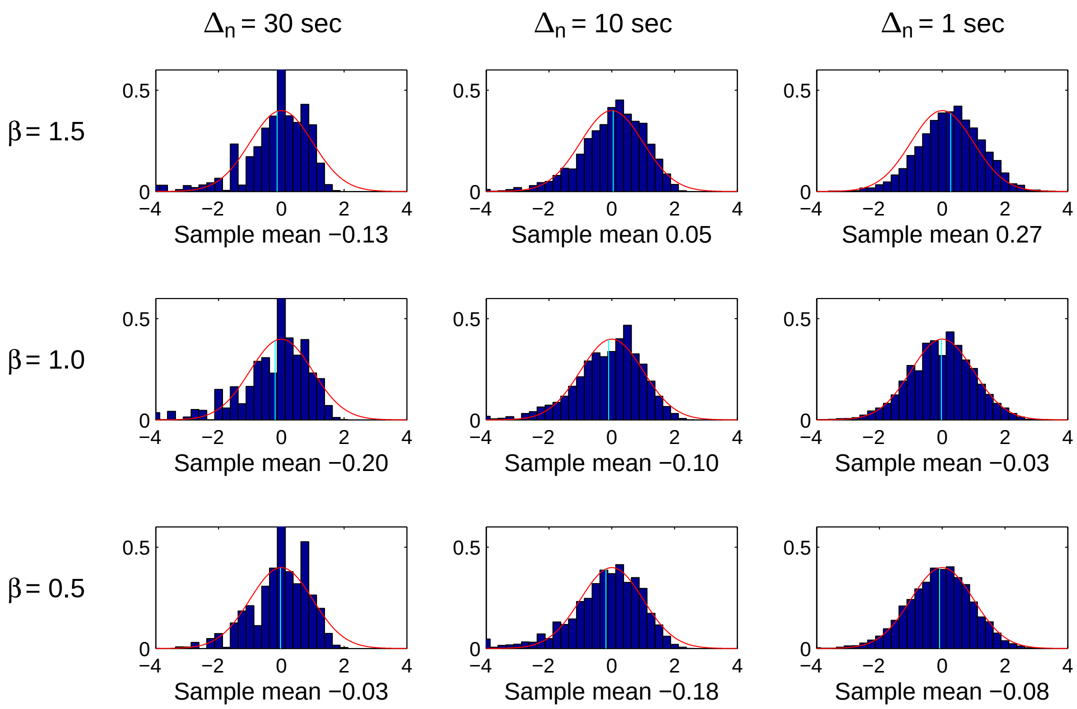


Figure 4.2: Distribution of the normalized estimator $\hat{\beta}$ after 5,000 simulations, highlighting its mean and its similarity to a normal rv.

We clearly see that we get closer to a normally distributed rv for smaller time steps. According to the theory we should have the sample mean of the normalized $\hat{\beta}$ converging to zero. We feel comfortable with seeing this for the smaller values of $\beta$ in the second and third row, but we are not yet able to observe this tendency for $\beta = 1.5$. In the figures it looks as if the mean of $\hat{\beta}$ might converge to a bigger value, as the normalized estimator is obtained by subtracting the actual value of $\beta$ from the estimator $\hat{\beta}$ and scaling it. Despite not depicted, we get the same distributions for the other values of Table 4.1, and also for $\beta = 1.75$, with this one being shifted to the right as well. If we try to see this convergence for smaller or bigger values of $\beta$, this should be possible by the constraints given in Theorem 3.17, but it is not what our simulations verify when choosing a time step of 1 second. For values smaller than $\beta = 0.25$ and bigger than $\beta = 1.5$ we cannot make a comparison to any results of paper [2], as these are not included therein. We provide a short discussion for values lying outside these bounds in the following section.

## 4.2 Discussion of certain Simulation Results

The analysis of any results obtained from the Monte Carlo simulation for values of $\beta$ outside the interval $[0.25, 1.5]$ was not included in paper [2]. When we use our code to compute the mean of the estimator $\hat{\beta}$ and the corresponding standard deviations, as done in Table 4.1 and Table 4.2, we do not see such good fittings for values outside the stated interval. For example, when looking at the mean we still get good estimators in the situation of $\beta = 0.2$ (the estimators lie in the interval $[0.19, 0.21]$ for different tail probabilities). But when the actual value of $\beta$ is even smaller, the outcome gets worse (e.g. for $\beta = 0.1$ the corresponding results of Table 4.1 would lie in $[0.05, 0.17]$). When looking at values of $\beta > 1.5$, we again do not see good fittings. For example, we tested the value $\beta = 1.95$ and the corresponding means of the estimators lie in the interval $[2.40, 4.15]$, which per se does not make sense. The reasons could be manifold. The exact analysis of this outcome will be left for future research, but here we want to mention some possible reasons we found. We will summarize what our thoughts were and how we have tested them.

We assumed that one reason could be that the time steps are still too big to see the desired convergence. Hence, we computed the mean of the estimator $\hat{\beta}$ and of the normalized estimator, which is defined in Theorem 3.17, for smaller observation intervals. We include the code we programmed for getting these and the following simulations in the appendix in Section A.2. When decreasing the size of the time steps, we first calculated the value $\theta$ for every scenario, but as $\theta$ does not only depend on the tail probability, but also on the step size, changing $\theta$ implies to change the underlying process. This would mean, that at smaller time steps the jump part would become bigger, as the underlying process was defined to be $dX_t = \sigma_t \, dW_t + \theta \, dY_t$ in equation (4.1) and $\theta$ increases for smaller time steps[3]. Obviously we want to have the same underlying process for shorter observation intervals. Therefore, we

[3]The dependence of $\theta$ on the size of the time step is clear by its definition and we tested the stated relation by computing the values of $\theta$ in dependence on the time steps and $\beta$.

will only refer to the computations we carried out with a fixed level of $\theta$, when talking about the convergence we wanted to observe from choosing a smaller step size.

As we stated earlier, we still wanted to test if the convergence to a normally distributed variable as stated in Theorem 3.17 is only visible for smaller values of $\overline{\omega}$ when looking at big values of $\beta$. This restriction might come from what we stated in equation (3.50) in Remark 3.29. From what we saw from our simulations first, we thought that for a lower value $\overline{\omega}$, which is in line with the restrictions, the normalized estimator $\hat{\beta}$ does converge and that it does not converge for the given value $\overline{\omega} = 0.2$. For example, when considering $\beta = 1.75$ we had the mean of the standardized estimator growing to 0.33, when using $\overline{\omega} = 0.2$, and for $\overline{\omega} = 0.12$ we only had $-0.03$ and no growing tendency. This was for a time step of 1/8 seconds and 5,000 simulation runs. When we wanted to strengthen our assumption by looking at a smaller step size however, we found that the convergence was visible for $\overline{\omega} = 0.2$ as well. For instance, we got that the mean of the normalized estimator is 0.01 for $\beta = 1.95$ and a step size of only 1/128. Seeing this we could say that in our empirical observations the convergence is happening in a slower way, when looking at bigger values of $\beta$. One reason could be that the standard deviation is bigger in such a case, as we described when commenting on the simulated standard deviations, which were depicted in Table 4.2.

For small values of $\beta < 0.25$ we have neither obtained good estimators, when choosing time steps of 1 second. We again tried to decrease the step size to achieve better values. When, for example, reducing the step size to 1/128 or even 1/256 the estimator $\hat{\beta}$ reached the values 0.081 and 0.087 respectively, when the underlying $\beta = 0.1$. This could still be in line with a slower convergence, which could come from a lower value of $\chi$, which is defined in Theorem 3.16 and describes the convergence rate. The outcome of the distributional result in Theorem 3.17 is not yet visible at this step size however. We plotted histograms for the above scenarios and they were both dominated by one bar around zero. We feel that the jump sizes might be so small that there is a computational inaccuracy or fault showing up or that the statistics used to compute $\hat{\beta}$ dismiss too many values and are hence too small to reach the actual value of $\beta$. These are just assumptions, however, and the actual reason could be manifold.

# Chapter 5

# Conclusion and Future Scope

Within this thesis we had a closer look at the work of Aït-Sahalia and Jacod in connection with the analysis of high frequency data. We first introduce the reader to some mathematical background, often used when dealing with such data. Then we focused on one of the papers by Aït-Sahalia and Jacod, namely 'Estimating the Degree of Activity of Jumps in High Frequency Data' [2]. Therein they introduced the jump activity index $\beta$. Within this paper they go into mathematical depth when showing their results. It was the aim of this thesis, to discuss two of the main results and to break down the proofs into steps, which are easy to understand. We have tried to go into as much detail as possible and feel to have thereby allowed for an easier understanding. By discussing the various lemmas, which form part of these two proofs, we have come across some minor typos, which we stated in the footnotes. Further, we used Remark 3.20 and Remark 3.29 to point out two issues, which we could not clarify in detail. We leave them open for future work and hope to have contributed to the understanding of the proofs by the steps presented in this thesis.

To round up our discussion of the proofs, we have also implemented some simulations. These showed the desired outcome when we were comparing our results to the ones displayed in the aforementioned paper. Using our implementations, we found that the convergence for values of $\beta > 1.5$ is slower than the one seen for the values of $\beta \in [0.25, 1.5]$, which were considered in the simulations by Aït-Sahalia and Jacod too. When running Monte Carlo simulations for $\beta < 0.25$ we soon find that the outcome is not as precise. With regard to the results obtained by us, it is interesting to mention one paper, implementing the jump activity index and extending it further [26]. This publication is written by Jing, Kong, Liu and Mykland. Therein the estimator $\hat{\beta}$ is modified in a way, which makes it consider more of the increments, instead of ignoring all increments smaller than a certain threshold. It is shown that this modification results in an improved estimator. Further, this paper [26] states various tables, in which the new estimator is compared to the original one. Table 1 of paper [26] also depicts values for $\beta = 1.75$ (else the same values as in the analysis of Aït-Sahalia and Jacod are displayed), wherein we see that the error of the original estimator is most significant for this high value of $\beta = 1.75$. This does not provide us with an answer, of why our values differ from the actual value for $\beta = 1.75$ more than in other cases, but it is at least in line with our observations.

The paper we discussed throughout this thesis can be seen as being one small part of the work by Aït-Sahalia and Jacod. In their most recent joint paper [6], they describe various estimators they have developed with regard to high frequency data. Further, they state that others have already started to develop refinements of some of their estimators (e.g. in paper [26] stated above). By reading their work, it also becomes clear that there are many areas left open for further research. These include the establishment of estimators with lower standard deviations when looking at the same model assumptions. When looking at real data observations another challenge comes in, namely the one of market microstructure noise blurring the data. As we feel that current literature is in the opinion of jumps being present in financial data, we assume that there will be more research happening in this area. It will be interesting to see, how things develop and what kind of new statistics will be introduced in the near future.

# Appendix A

# Description of the Matlab Code

For creating the figures within the thesis we used Matlab self-programmed functions consisting of some existing Matlab-functions too. We utilized the software Matlab version 7.11.0.584 (R2010b). The existing Matlab-functions we included in our code are the inbuilt ones for generating uniformly distributed rvs and multivariate normally distributed rvs. Further, we made use of a function calculating the inverse of the stable cumulative distribution function, which is part of the toolbox of alpha-stable distributions programmed by Mark Veillette and available for free on the Mathworks file exchange.[1] The code for the self-programmed functions is provided below.

## A.1 Code for Monte Carlo Simulations

This section displays all code used within Section 4.1. The script files are included as well and they can be identified by their name starting with an `s` followed by a capital letter. Within the name of the following script files we find the word `Basic` meaning that the simulations are carried out without considering any bias correction, which is mentioned in section [2, 5. Small sample bias correction]. The order of the code below is that we first give the two functions generating the continuous and generating the jump part of the sample paths. The function `simulateStable(dt,T,beta,theta)` produces a sample path of a $\beta$-stable process, for values of $\beta \in (0, 2)$. When we choose the input parameter $\beta = 0$ a Poisson process with arrival rate $\theta$ is created. Else the variable $\theta$ is optional and representing the scaling factor introduced in equation 4.1. We commented on the output when assuming a Poisson process but did not include any figures. After stating these two functions, we provide the code for producing Tables 4.1 and 4.2. This is one function calculating the mean and standard deviation and one script file executing the code. Afterwards we include the code generating Figure 4.2 displaying the histograms of the estimator $\hat{\beta}$. One function generating the histograms and one script file is again included.

[1] The toolbox can be downloaded from the official Mathworks website: `http://www.mathworks.com/matlabcentral/fileexchange/37514-stbl-alpha-stable-distributions-for-matlab` [Accessed 09/05/13].

### A.1.1 `simulateStable`

Matlab help:

```
This function generates one increment of a beta-stable process for 0 < beta < 2. The
increment is taken over the time span dt. The increments are then added up to give
the process. We proceed like in [1], Algorithm 6.6 and also make use of the inverse
transform method, see Remark 8.17 [3], for example.  For beta = 0 we generate a
Poisson process with arrival rate theta and fixed jump size of 0.1. The assumptions
regarding time (e.g. 1 day has 6.5 trading hours) are taken from [2], p.2220.

input: dt ...... time interval in seconds
       T ....... no. of 6.5 hour trading days observed
       beta .... beta, making the process beta-stable
       theta ... OPTIONAL parameter: with two different functions:
                 1) if 0 < beta < 2, then we multiply the process with
                    this value, i.e. we return a multiple of the stable
                    process
                 2) if beta = 0, this gives the arrival rate of a Poisson
                    process with fixed jump size of 0.1
                 Default value: 1

output: Y ...... a vector of all values of the process starting at 0 and at
                 all multiples of the time interval till T

syntax: Y = simulateStable(dt,T,beta,theta)

references:
[1] Cont, R., and Tankov, P. Financial modelling with Jump Processes.
    Chapman & Hall / CRC Press, 2003
[2] Ait-Sahalia, Y., and Jacod, J. Estimating the Degree of Activity of
    Jumps in High Frequency Data. Ann. Statist. (37), 2009
[3] Kusolitsch, N. Maß- und Wahrscheinlichkeitstheorie: Eine Einführung.
    Springer-Verlag Wien, 2011
```

Matlab code:

```
function  Y = simulateStable(dt,T,beta,theta)

% theta is an OPTIONAL parameter and only needed in case of a Poisson process
if nargin < 4
    if beta == 0
    error('error:input:value', ['For beta = 0 a Poisson process is' ...
```

```
            'simulated. We need the input theta to define the arrival rate.'])
    end
    theta = 1;
end

% adjusting time to be counted in days
dt = dt/(6.5*60^2);

% no. of increments needed
n = floor(T/dt);

% initializing variables
Y = zeros(1,n+1);
if beta ~=0;
    % initializing variables
    increments = zeros(1,n);
    gamma = zeros(1,n);
    W = zeros(1,n);

    % generating n uniformly distributed rvs on (-pi/2,pi/2)
    gamma = pi*(rand(1,n)-1/2);

    % generating n independent standard exponential rvs
    W = -log(rand(1,n));

    % generating increments of the beta-stable process using Algorithm 6.6;
    % where in order to simulate the increments we use the stable property:
    %    Y_t / t^(1/beta) = Y_1
    increments = dt^(1/beta)*sin(beta*gamma)./(cos(gamma).^(1/beta))...
                 .*(cos((1-beta).*gamma)./W).^((1-beta)/beta);

    % and summing them up to get the process scaled by the factor theta
    Y = theta*[0,cumsum(increments)];

else
    % initializing variables
    time = zeros(1,2);
    j = 0;

    % generating jump times (exponentially distributed inter-arrival times)
    % and assigning the constant value between the previous and this jump
    while time(1) < T
        time(2) = -log(rand(1))/theta + time(1);
        Y(floor(time(1)*(6.5*60^2))+1:min(ceil(time(2)*(6.5*60^2)),...
            n+1))=j*0.1;
```

```
        j=j+1;
        time(1) = time(2);
    end
end
end
```

### A.1.2 simulateCont

MATLAB help:

```
This function simulates one path of X:=int_0^T sigma_t dW_t as given in [1] on p.2220
by using discretization. The process sigma_t fulfills (where we do not include a
jump process dJ_t):
   sigma_t = v_t^(1/2),
   dv_t = kappa(eta - v_t) dt + gamma v_t^(1/2) dB_t,
with:
   E[dW_t*dB_t] = rho dt.

input: dt ...... time interval in seconds
       T ....... no. of 6.5 hour trading days observed
       eta ..... parameter as in the equation above
       gamma ... parameter as in the equation above
       kappa ... parameter as in the equation above
       rho ..... parameter as in the equation above

output: X .... a vector of all values of the process starting at 0 and at
               all multiples of the time interval till T

syntax: X = simulateCont(dt,T,eta,gamma,kappa,rho)

references:
[1] Ait-Sahalia, Y., and Jacod, J. Estimating the Degree of Activity of
    Jumps in High Frequency Data. Ann. Statist. (37), 2009
```

MATLAB code:

```
function  X = simulateCont(dt,T,eta,gamma,kappa,rho)

% adjusting time to be counted in days
dt = dt/(6.5*60^2);

% no. of simulations needed
```

```
n = floor(T/dt);

% initializing variables
X = ones(1,n+1);
v = ones(1,n+1);
v(1) = eta;

% creating the correlated random increments, satisfying E[dW_t*dB_t]=rho dt
dW_dB = mvnrnd(zeros(1,2),dt.*[1,rho;rho,1],n).';

% stepwise simulation of stochastic variance and the process X
for j=1:n
    v(j+1) = v(j) + kappa*(eta-v(j))*dt + gamma*sqrt(v(j))*dW_dB(2,j);
    X(j+1) = sqrt(v(j))*dW_dB(1,j) + X(j);
end
end
```

### A.1.3 calcSampleMeanStd

MATLAB help:

```
This function calculates the mean and standard deviation of the estimator beta hat.
We proceed as described in [1] in the theoretical part (i.e. we do not include a
bias correction). We allow for n different combinations of time intervals and tail
probabilities, and for m different values of beta. The output will be two m times n
matrices.

input: dt ........ n time intervals stating the frequency of observations
                   in seconds
       tail_pr ... n set tail probabilities
       beta ...... m values for the jump activity index beta
       N ......... no. of simulation runs
       T ......... no. of 6.5 hour trading days observed
       om ........ value for scaling the truncation level
       alph ...... value for scaling the truncation level
       eta ....... parameter for simulation of the continuous part
       gamma ..... parameter for simulation of the continuous part
       kappa ..... parameter for simulation of the continuous part
       rho ....... parameter for simulation of the continuous part

output: OUTPUT_beta_hat ....... m times n matrix giving the mean of
                                the estimation of beta hat
        OUTPUT_beta_hat_std ... m times n matrix of corresponding
                                standard deviations
```

```
syntax: [OUTPUT_beta_hat, OUTPUT_beta_hat_std] = ...
         calcSampleMeanStd(dt,tail_pr,beta,N,T,om,alph,eta,gamma,kappa,rho)

references:
[1] Ait-Sahalia, Y., and Jacod, J. Estimating the Degree of Activity of
    Jumps in High Frequency Data. Ann. Statist. (37), 2009
```

---

MATLAB code:

---

```
function [OUTPUT_beta_hat, OUTPUT_beta_hat_std] = calcSampleMeanStd(dt,...
                                tail_pr,beta,N,T,om,alph,eta,gamma,kappa,rho)

% adjusting time to be counted in days
dt_days = dt/(6.5*60^2);

% defining the length
alph_n = length(alph);
tail_pr_n = length(tail_pr);
if tail_pr_n ~= length(dt)
    error('error:input:lenght', ['You need to provide two vectors of' ...
          'the same length for dt and tail_pr.'])
end
beta_m = length(beta);

% initializing variables
OUTPUT_beta_hat = zeros(beta_m,tail_pr_n);
OUTPUT_beta_hat_std = zeros(beta_m,tail_pr_n);
cut_off = alph*dt_days.^om; % an alpha_n times tail_pr_n matrix
theta = 0;
U = zeros(alph_n,1);

for m=1:beta_m                      % first main loop: for rows
    for n=1:tail_pr_n               % second main loop: for columns
        % theta is adapted to fit the tail probability - see [1] p. 2220
        if beta ~= 0
           % for a stable process: reach 2*stblcdf(-cut_off(1,n)/theta,
           % beta(m),0,dt_days(n)^(1/beta(m)),0) = taip_pr, is required
           theta = cut_off(1,n)/stblinv(1-tail_pr(n)/2,beta(m),0,...
                   dt_days(n)^(1/beta(m)),0);
        else
           % for a poisson process adjust the arrival rate to obtain P(next
           % arrival < dt)= theta
           theta = -log(1-tail_pr(n))/dt_days(n);
```

```
        end

        % set the estimator back to zero
        beta_hat = zeros(1,N);

        for j=1:N                   % third main loop: for simulation runs
            jump_Y = simulateStable(dt(n),T,beta(m),theta);
            Y = simulateCont(dt(n),T,eta,gamma,kappa,rho)+jump_Y;
            diff_Y = diff(Y);

            % assign the statistics U only considering big enough jumps
            for k=1:alph_n
                U(k) = sum(abs(diff_Y)>cut_off(k,n));
            end

            % by convention the estimator is set to be zero, if either of
            % the statistics U is zero (see p.2211 of [1]).
            if (U(1)~=0)*(U(2)~=0)
                beta_hat(j) = log(U(1)/U(2))/log(alph(2)/alph(1));
            end
        end
        % store the mean and standard deviation
        OUTPUT_beta_hat(m,n) = mean(beta_hat);
        OUTPUT_beta_hat_std(m,n) = std(beta_hat);
    end
end
end
```

### A.1.4 sBasicTable

MATLAB help:

```
Script to create own version of table 1 on p.2221 of [1]. We do not implement a bias
correction here, as we stick to the the mathematical results only.
```

MATLAB script:

```
% if output shall be put in a set excel file set print=1, else print=0
print=1;

% fixing the step sizes and tail probabilities (per column)
dt =[1,1,1,1,5];
```

```
tail_pr = [.0025,.005,.01,.025,.01];

% fixing the assumed values of beta (per row)
beta = [1.5,1.25,1,.75,.5,.25];

% simulation runs
N = 5000;

% input parameters of the model
T = 1;
om = 0.2;
eta = 1/16;
alph = [5;10].*eta;
gamma = 1/2;
kappa = 5;
rho = -1/2;

% calculating the mean and standard deviation of the estimator beta hat
[OUTPUT_beta_hat, OUTPUT_beta_hat_std] = calcSampleMeanStd(dt,tail_pr,...
                                       beta,N,T,om,alph,eta,gamma,kappa,rho)

% if print==1, then the output is written into the stated excel file
if print
    filename = '00_Trial_Output';
    xlswrite(filename,OUTPUT_beta_hat,'MATLAB',['C4:G9']);
    xlswrite(filename,OUTPUT_beta_hat_std,'MATLAB',['K4:O9']);
end
```

### A.1.5 makeHistNorm

MATLAB help:

```
This function calculates the possible values of beta hat in simulations and plots a
histogram to compare the empirical outcome to the theoretical normal distribution.

input: dt ........ time interval stating the frequency of observations
                   in seconds
       tail_pr ... set tail probability
       beta ...... jump activity index beta
       N ......... no. of simulation runs
       T ......... no. of 6.5 hour trading days observed
       om ........ value for scaling the truncation level
       alph ...... value for scaling the truncation level
       eta ....... parameter for simulation of the continuous part
```

```
       gamma ..... parameter for simulation of the continuous part
       kappa ..... parameter for simulation of the continuous part
       rho ....... parameter for simulation of the continuous part

syntax: makeHistNorm(dt,tail_pr,beta,N,T,om,alph,eta,gamma,kappa,rho)

references:
[1] Ait-Sahalia, Y., and Jacod, J. Estimating the Degree of Activity of
    Jumps in High Frequency Data. Ann. Statist. (37), 2009
```

---

MATLAB code:

---

```
function makeHistNorm(dt,tail_pr,beta,N,T,om,alph,eta,gamma,kappa,rho)

% adjusting time to be counted in days
dt_days = dt/(6.5*60^2);

% test for scalar input
if 1-(length(tail_pr) == length(dt) == length(beta) == 1)
    error('error:input:lenght', ['You should only provide scalar input' ...
          'for dt, tail_pr and beta.'])
end

% initializing variables
cut_off = alph*dt_days^om;
theta = cut_off(1)/stblinv(1-tail_pr/2,beta,0,dt_days^(1/beta),0);
U = zeros(2,1);
beta_hat = zeros(1,N);
beta_hat_norm = zeros(1,N);

% run simulations for beta hat
for j=1:N                  % one main loop: for simulation runs
    jump_Y = simulateStable(dt,T,beta,theta);
    Y = simulateCont(dt,T,eta,gamma,kappa,rho)+jump_Y;
    diff_Y = diff(Y);

    % assign the statistics U only considering big enough jumps
    for k=1:2
        U(k) = sum(abs(diff_Y)>cut_off(k));
    end

    % by convention the estimator is set to be zero, if either of the
    % statistics U is zero (see p.2211 of [1]), further for not dividing by
    % zero in beta_hat_norm, we need to exclude the case U(1)==U(2)
```

```
    % making sqrt(1/U(2)-1/U(1))=0 (in this case beta_hat = 0, so we do
    % not need to set it to zero)
    if (U(1)~=0)*(U(2)~=0)*(U(1)~=U(2))
        beta_hat(j) = log(U(1)/U(2))/log(alph(2)/alph(1));
        beta_hat_norm(j) = log(cut_off(2)/cut_off(1))/...
                                sqrt(1/U(2)-1/U(1))*(beta_hat(j)-beta);
    end
end

% get the sample mean of the standardized variables
mu = mean(beta_hat_norm);

% plot the histogram resulting from the normalized beta hat
bin = 0.25;
[h, x] = hist(beta_hat_norm,[-4:bin:4]);
h = h/(bin*N);      % change the values from absolute to percentage
bar(x,h,'hist');
xlim([-4 4])
ylim([0 0.6])
str0=sprintf('Sample mean %.2f\n', mu);
xlabel(str0,'Fontsize',12);

% add the density function of a normal rv and include a vertical line for
% the mean of the normalized samples
hold on
plot([-4:0.01:4],normpdf([-4:0.01:4]),'Color','red')
plot([mu mu], [0 normpdf(mu)], 'Color', 'cyan')
hold off
end
```

### A.1.6 sBasicHistNorm

Matlab help:

```
Script to create own version of figure 3 on p.2224 of [1]. We do not implement a
bias correction here, as we stick to the the mathematical results only.
```

Matlab script:

```
% fixing arbitrary amount of sampling times and one tail probability
dt = [30,10,1];
dt_n = length(dt);
```

```
tail_pr = 0.01;

% fixing the assumed values of beta
beta = [1.5,1,0.5];
beta_n = length(beta);

% simulation runs
N = 5000;

% input parameters of the model
T = 1;
om = 0.2;
eta = 1/16;
alph = [5;10].*eta;
gamma = 1/2;
kappa = 5;
rho = -1/2;

% calculate the possible values of beta hat in simulations and plot histograms
for k=1:beta_n
    for j=1:dt_n
        subplot(beta_n,dt_n,(k-1)*dt_n+j)
        makeHistNorm(dt(j),tail_pr,beta(k),N,T,om,alph,eta,gamma,kappa,rho);
    end
end
```

## A.2 Code for Discussion of certain Simulation Results

For completeness we include the code we used to simulate the mean of $\hat{\beta}$ and of its normalized version, for a fixed $\beta$ but different values of $\varpi$ and $\theta$. We described the outcome in Section 4.2 but did not include any tables. Such tables could be created using the following function `calcTestFixedBeta(dt,tail_pr,beta,N,T,om,alph,eta,gamma,kappa,rho,print)` and a corresponding script file `sTextFixedBeta`. For running the simulations we always changed the values of `beta`, `dt` and `om` accordingly, wherein the latter represents $\varpi$. It works in such a way that one sample path of the continuous and the jump part is simulated once and that these are used to create two versions of an underlying process. These versions only differ by a different factor $\theta$. Onto these versions we then apply the estimator $\hat{\beta}$. We again use two versions. This time with two different values of $\varpi$. Proceeding in this way, gives out four versions of the estimator corresponding to two different underlying processes. Any thoughts we had with regard to these simulations can be found in the aforementioned section.

### A.2.1 calcTestFixedBeta

MATLAB help:

```
This function considers four scenarios for one set level of beta. In the first omega
is 0.2 and the value theta is fixed. Comment one part of the code to either make
theta = 1 or to being the value resulting out of dt=1 and omega=0.2. In the second
we have the same fixed theta but omega is the value given as input. In the third
we again have omega=0.2 but the value of theta is always calculated using the tail
probability given as input and in the way it is desribed in the paper [1] on p.2220.
The fourth varies theta but leaves omega the set input value. In all cases the mean
of beta hat and the mean of the normalized beta hat is calculated and written into
an excel file if the variable print=1.

input: dt ........ n time intervals stating the frequency of
                   observations in seconds
       tail_pr ... one set tail probabilities
       beta ...... one value for the jump activity index beta
       N ......... no. of simulation runs
       T ......... no. of 6.5 hour trading days observed
       om ........ a second value for omega, which will be compared to
                   the value 0.2 from the paper
       alph ...... value for scaling the truncation level
       eta ....... parameter for simulation of the continuous part
       gamma ..... parameter for simulation of the continuous part
       kappa ..... parameter for simulation of the continuous part
       rho ....... parameter for simulation of the continuous part
       print ..... if set to 1, then the means of the estimators will be
                   written into an excel file - this assumed n=6, i.e. 6
                   diffterent sizes of time steps; if 0 nothin happens

syntax: calcTestFixedBeta(dt,tail_pr,beta,N,T,om,alph,eta,gamma,kappa,...
                                                              rho,print)

references:
[1] Ait-Sahalia, Y., and Jacod, J. Estimating the Degree of Activity of
    Jumps in High Frequency Data. Ann. Statist. (37), 2009
    Chapman & Hall / CRC Press, 2003
```

MATLAB code:

```
function calcTestFixedBeta(dt,tail_pr,beta,N,T,om,alph,eta,gamma,kappa,...
```

```
                                                                   rho,print)

% use the input to define the second value, which omega should take; the
% first one is as in the paper om = 0.2
om = [0.2, om];

% initialize output wrtiten into excel file, first line gives mean of
% normalized beta hat, second line gives mean of beta hat:
n = length(dt);
% 1st: fixed theta, omega (1)
mean_thetaFix_om1 = zeros(2,n);
% 2nd: fixed theta, omega (2)
mean_thetaFix_om2 = zeros(2,n);
% 3rd: changing theta, omega (1)
mean_thetaVar_om1 = zeros(2,n);
% 4th: changing theta, omega (2)
mean_thetaVar_om2 = zeros(2,n);

% calculate fixed value for theta (which comes from dt=1 and om=0.2), in
% the first two entries of theta
theta = zeros(1,4);
dt_days = 1/(6.5*60^2);
cut_off = alph*dt_days.^0.2;
theta([1,2]) = cut_off(1)/stblinv(1-tail_pr/2,beta,0,dt_days^(1/beta),0)...
                                                                  *ones(1,2);
% or set it to be one
% theta([1,2])=ones(1,2)

U = zeros(2,1);

% for varying values of theta we need these parameters
dt_days = dt./(6.5*60^2);
% put the values coming from the second value of omega behind the n entries
% coming from the first, i.e. 2x2n matirx
cut_off = [alph*dt_days.^om(1), alph*dt_days.^om(2)];

for j=1:n
    % second two entries of theta depend on the time and the value of omega
    theta(3) = cut_off(1,j)/stblinv(1-tail_pr/2,beta,0,dt_days(j)^(1/beta),0);
    theta(4) = cut_off(1,n+j)/stblinv(1-tail_pr/2,beta,0,dt_days(j)^(1/beta),0);

    % clear values of beta
     beta_hat = zeros(4,N);
     beta_hat_norm = zeros(4,N);
```

```
    for k=1:N
        jump_Y = simulateStable(dt(n),T,beta);
        cont_X = simulateCont(dt(n),T,eta,gamma,kappa,rho);

        for l=1:4
            Y = cont_X + theta(l)*jump_Y;
            diff_Y = diff(Y);

            % assign the statistics U only considering big enough jumps,
            % with the cutoff level coming from the right omega
            for m=1:2
                U(m) = sum(abs(diff_Y)>cut_off(m,j+(mod(l+1,2))*n));
            end

            % by convention the estimator is set to be zero, if either of the
            % statistics U is zero (see p.2211 of [1]), further for not dividing
            % by zero in beta_hat_norm, we need to exclude the case U(1)==U(2)
            % making sqrt(1/U(2)-1/U(1))=0 (in this case beta_hat = 0, so we do
            % not need to set it to zero)
            if (U(1)~=0)*(U(2)~=0)*(U(1)~=U(2))
                beta_hat(l,k) = log(U(1)/U(2))/log(alph(2)/alph(1));
                beta_hat_norm(l,k) = log(cut_off(2,j+(mod(l+1,2)))/...
                            cut_off(1,j+(mod(l+1,2))))/sqrt(1/U(2)-1/U(1))...
                            *(beta_hat(l,k)-beta);
            end
        end
    end
    % 1st: fixed theta, omega (1)
    mean_thetaFix_om1(1,j) = mean(beta_hat(1,:));
    mean_thetaFix_om1(2,j) = mean(beta_hat_norm(1,:))
    % 2nd: fixed theta, omega (2)
    mean_thetaFix_om2(1,j) = mean(beta_hat(2,:));
    mean_thetaFix_om2(2,j) = mean(beta_hat_norm(2,:))
    % 3rd: changing theta, omega (1)
    mean_thetaVar_om1(1,j) = mean(beta_hat(3,:));
    mean_thetaVar_om1(2,j) = mean(beta_hat_norm(3,:))
    % 4th: changing theta, omega (2)
    mean_thetaVar_om2(1,j) = mean(beta_hat(4,:));
    mean_thetaVar_om2(2,j) = mean(beta_hat_norm(4,:))
end

% if print==1, then the output is written into the stated excel file
if print
    filename = '00_Trial_Output';
    xlswrite(filename,mean_thetaFix_om1,['testBeta=', ...
```

```
                                              sprintf('%.2f', beta)],['C4:H5']);
    xlswrite(filename,mean_thetaFix_om2,['testBeta=', ...
                                              sprintf('%.2f', beta)],['C7:H8']);
    xlswrite(filename,mean_thetaVar_om1,['testBeta=', ...
                                            sprintf('%.2f', beta)],['C10:H11']);
    xlswrite(filename,mean_thetaVar_om2,['testBeta=', ...
                                            sprintf('%.2f', beta)],['C13:H14']);
end
end
```

### A.2.2 sTestFixedBeta

Matlab help:

```
Script to test convergence of fixed beta with changing or fixed value of theta and
for different omega. For testing different values, change the level of beta or 'the
second value' omega should take, as the first is set to be 0.2.
```

Matlab script:

```
% fixing the assumed value of beta
beta = 1.75;

% we compare the outcome for two different values in omega, one will be
% left to be om = 0.2 as in the paper and the other one can be set here
om = 0.12;

% if output shall beta put in excel file set print=1, else print=0
print=1;

% fixing the sampling times
dt =[10,5,1,1/2,1/4,1/8];

% simulation runs
N = 5000;

% for calculating theta, we set the tail probability to
tail_pr = .01;

% input parameters of the model
T = 1;
eta = 1/16;
```

```
alph = [5;10].*eta;
gamma = 1/2;
kappa = 5;
rho = -1/2;

% store the values in an excel file
calcTestFixedBeta(dt,tail_pr,beta,N,T,om,alph,eta,gamma,kappa,rho,print)
```

## A.3 Code for other Figures

In this last section we display the code for creating the other figures within the thesis. We first provide two functions plotting graphs. The function `plotStable(dt,T,beta,theta)` produces a figure depicting one sample path of a $\beta$-stable process. In the same manner `plotContDis(dt,T,eta,gamma,kappa,rho)` creates the continuous part of the model stated in equation (4.1). Using these two functions we created Figure 3.1, showing sample paths of $\beta$-stable processes, and Figure 4.1, showing one sample path of the considered model within the implementations. The corresponding script files are included below.

### A.3.1 `plotStable`

MATLAB help:

```
This function plots one sample path created by simulateStable.  The assumptions
regarding time (e.g. 1 day has 6.5 trading hours) are taken from [1], p.2220.

input: dt ...... time interval in seconds
       T ....... no. of 6.5 hour trading days observed
       beta .... beta, making the process beta-stable
       theta ... OPTIONAL parameter: with two different functions:
                 1) if 0 < beta < 2, then we multiply the process with
                    this value, i.e. we return a multiplr of the stable
                    process
                 2) if beta = 0, this gives the
                    arrival rate of a Poisson process with fixed jump
                    size of 0.1

output: Y ...... a vector of all values of the process starting at 0 and at
                 all multiples of the time interval till T

syntax: Y = plotStable(dt,T,beta,theta)

references:
```

[1] Ait-Sahalia, Y., and Jacod, J. Estimating the Degree of Activity of
    Jumps in High Frequency Data. Ann. Statist. (37), 2009

---

MATLAB code:

---

```
function  Y = plotStable(dt,T,beta,theta)

% calculating no. of required simulations
n = floor(T*(6.5*60^2)/dt);

% initializing variables
Y = zeros(1,n+1);
time = [0:dt/(60^2):n*dt/(60^2)];  % time in hours

% simulating the beta-stable process
if nargin < 4
    Y = simulateStable(dt,T,beta);
else
    Y = simulateStable(dt,T,beta,theta);
end

% plotting the sample path
plot(time,Y,'.','MarkerSize',1);
xlabel('Time in hours','Fontsize',12);
xlim([time(1),time(n+1)]);
str1 = sprintf('\\alpha','interpreter','latex');
str2 = sprintf(' = %.1f', beta);
str = strcat(str1,str2);
title(str,'Fontsize',12);
grid on
end
```

---

### A.3.2 plotCont

MATLAB help:

---

```
This function plots one sample path of X:=int_0^T sigma_t dW_t as generated by sim-
ulateStable. The assumptions regarding time (e.g. 1 day has 6.5 trading hours) are
taken from [1], p.2220.

input: dt ...... time interval in seconds
       T ....... no. of 6.5 hour trading days observed
```

```
       eta ..... parameter as in the equation above
       gamma ... parameter as in the equation above
       kappa ... parameter as in the equation above
       rho ..... parameter as in the equation above


output: X .... a vector of all vales of the process starting at 0 and at
               all multiples of the time intervall till T

syntax: X = plotContDis(dt,T,eta,gamma,kappa,rho)

references:
[1] Ait-Sahalia, Y., and Jacod, J. Estimating the Degree of Activity of
    Jumps in High Frequency Data. Ann. Statist. (37), 2009
```

MATLAB code:

```
function  X = plotCont(dt,T,eta,gamma,kappa,rho)

% calculating no. of required simulations
n = floor(T*(6.5*60^2)/dt);

% initializing variables
X = zeros(1,n+1);
time = [0:dt/(60^2):n*dt/(60^2)];  % time in hours

% simulating the beta-stable process
X = simulateContDis(dt,T,eta,gamma,kappa,rho);

% plotting the sample path
plot(time,X,'.','MarkerSize',.5);
xlim([time(1),time(n+1)]);
xlabel('Time in hours','Fontsize',12);
str=sprintf('Continuous part');
title(str,'Fontsize',12);
grid on
end
```

### A.3.3 sPlotStable

MATLAB help:

```
Script to plot one sample path of a stable process for three different values of
beta.
```

MATLAB script:

```
% initializing variables
beta = [0.1,1,1.9];
dt=1;
T=1;

% creating the three plots next to each other
for j=1:3
    subplot(1,3,j)
    Y=plotStable(dt,T,beta(j));
end
```

### A.3.4 sContJumpSum

MATLAB help:

```
Scrit to generate one sample path of the model stated in [1], p.2220, by fist sim-
ulating the continuous part, then the jump part and then adding them together. We
generate 3 plots.
```

MATLAB script:

```
% input parameters of the model
beta = 1;
dt=1;
T=1;
eta = 1/16;
gamma = 1/2;
kappa = 5;
rho = -1/2;

% calculating the scale parameter theta
tail_pr = 0.01;
om = 0.2;
alph = 5*eta;
cut_off = alph*(dt/(6.5*60^2))^om;
theta = cut_off/stblinv(1-tail_pr/2,beta,0,(dt/(6.5*60^2))^(1/beta),0);
```

```
% needed values for the third figure
n = floor(T*(6.5*60^2)/dt);
time = [0:dt/(60^2):n*dt/(60^2)];  % time in hours

% generating 3 subplots: 1) continuous part
subplot(1,3,1)
X = plotCont(dt,T,eta,gamma,kappa,rho);

% 2) jump part
subplot(1,3,2)
Y = plotStable(dt,T,beta,theta);
% overwrite title
str0 = sprintf('Jump process, with ');
str1 = sprintf(' \\ beta','interpreter','latex');
str2 = sprintf(' = %.1f', beta);
str = strcat(str0,str1,str2);
title(str,'Fontsize',12);

% 3)sum of the two processes
subplot(1,3,3)
plot(time,(X+Y),'.','MarkerSize',1);
xlabel('Time in hours','Fontsize',12);
xlim([time(1),time(end)]);
title('Combined processes','Fontsize',12);
grid on
```

# Bibliography

[1] Aït-Sahalia, Y., and Jacod, J. Fisher's information for discretely sampled Lévy processes. *Econometrica 76*, 4 (2008), 727–761.

[2] Aït-Sahalia, Y., and Jacod, J. Estimating the Degree of Activity of Jumps in High Frequency Data. *Ann. Statist. 37*, 5A (2009), 2202–2244.

[3] Aït-Sahalia, Y., and Jacod, J. Testing for jumps in a discretely observed process. *Ann. Statist. 37*, 1 (2009), 184–222.

[4] Aït-Sahalia, Y., and Jacod, J. Is Brownian motion necessary to model high-frequency data? *Ann. Statist. 38*, 5 (2010), 3093–3128.

[5] Aït-Sahalia, Y., and Jacod, J. Testing whether jumps have finite or infinite activity. *Ann. Statist. 39*, 3 (2011), 1689–1719.

[6] Aït-Sahalia, Y., and Jacod, J. Analyzing the Spectrum of Asset Returns: Jump and Volatility Components in High Frequency Data. *J. Econ. Lit. 50*, 4 (2012), 1007–1050.

[7] Aldous, D. J., and Eagleson, G. K. On mixing and stability of limit theorems. *Ann. Probability 6*, 2 (1978), 325–331.

[8] Applebaum, D. *Lévy Processes and Stochastic Calculus.* Cambridge Studies in Advanced Mathematics.

[9] Barndorff-Nielsen, O. E. Power and Bipower Variation with Stochastic Volatility and Jumps (with discussion). *J. Financ. Economet. 2*, 1 (2004), 1–48.

[10] Barndorff-Nielsen, O. E., Hansen, P. R., Lunde, A., and Shephard, N. Designing realized kernels to measure the ex post variation of equity prices in the presence of noise. *Econometrica 76*, 6 (2008), 1481–1536.

[11] Barndorff-Nielsen, O. E., and Shephard, N. Econometrics of Testing for Jumps in Financial Economics Using Bipower Variation.*J. Financ. Economet. 4*, 1 (2006), 1–30.

[12] Bernstein, S. Sur l'extension du théoréme limite du calcul des probabilités aux sommes de quantités dépendantes. *Math. Ann. 97*, 1 (1927), 1–59.

[13] Bichteler, K. Stochastic integration and $L^p$-theory of semimartingales. *Ann. Probab. 9*, 1 (1981), 49–89.

[14] Blumenthal, R. M., and Getoor, R. K. Sample functions of stochastic processes with stationary independent increments. *J. Math. Mech. 10*, 3 (1961), 493–516.

[15] Çinlar, E. *Probability and Stochastics.* Springer, New York, 2011.

[16] Cont, R., and Mancini, C. Nonparametric tests for pathwise properties of semimartingales. *Bernoulli 17*, 2 (2011), 781–813.

[17] Cont, R., and Tankov, P. *Financial modelling with jump processes.* Chapman & Hall/CRC Financial Mathematics Series. Chapman & Hall/CRC, Boca Raton, FL, 2004.

[18] Doob, J. L. *Stochastic processes.* Wiley Classics Library. John Wiley & Sons Inc., New York, 1990. Reprint of the 1953 original, A Wiley-Interscience Publication.

[19] Hall, P., and Heyde, C. C. *Martingale limit theory and its application.* Academic Press Inc. [Harcourt Brace Jovanovich Publishers], New York, 1980. Probability and Mathematical Statistics.

[20] Harrison, J. M., and Pliska, S. R. Martingales and stochastic integrals in the theory of continuous trading. *Stochastic Process. Appl. 11*, 3 (1981), 215–260.

[21] He, S. W., and Wang, J. G. and Yan, J. A. *Semimartingale theory and stochastic calculus.* Kexue Chubanshe (Science Press), Beijing, 1992.

[22] Jacod, J. On continuous conditional Gaussian martingales and stable convergence in law. In *Séminaire de Probabilités, XXXI*, vol. 1655 of *Lecture Notes in Math.* Springer, Berlin, 1997, pp. 232–246.

[23] Jacod, J. Asymptotic properties of realized power variations and related functionals of semimartingales. *Stochastic Process. Appl. 118*, 4 (2008), 517–559.

[24] Jacod, J., and Protter, P. *Discretization of processes*, vol. 67 of *Stochastic Modelling and Applied Probability.* Springer, Heidelberg, 2012.

[25] Jacod, J., and Shiryaev, A. N. *Limit theorems for stochastic processes*, second ed., vol. 288. Springer-Verlag, Berlin, 2003.

[26] Jing, B.-Y., Kong, X.-B., Liu, Z., and Mykland, P. On the jump activity index for semimartingales. *J. Econometrics 166*, 2 (2012), 213–223.

[27] Kallenberg, O. *Foundations of modern probability*, second ed. Springer-Verlag, New York, 2002.

[28] Karatzas, I., and Shreve, S. E. *Brownian Motion and Stochastic Calculus*, second ed. Springer-Verlag, 1991.

[29] Klenke, A. *Wahrscheinlichkeitstheorie*, second ed. Springer-Verlag, Berlin, 2006.

[30] Kusolitsch, N. *Maß- und Wahrscheinlichkeitstheorie: Eine Einführung*, first ed. Springer-Verlag, Wien, 2011.

[31] Lévy, P. Propriétés asymptotiques des sommes de variables aléatoires indépendants et enchaînées. *J. Math. Pures et Appliquées 14* (1935), 347–402.

[32] Lowther, G. *Quadratic Variations and the Ito Isometry.* Alomst Sure, March 2010. Blog: `http://almostsure.wordpress.com/2010/03/29/quadratic-variations-and-the-ito-isometry/` [Accessed 30/10/13].

[33] Mancini, C. Non-parametric threshold estimation for models with stochastic diffusion coefficient and jumps. *Scand. J. Stat. 36*, 2 (2009), 270–296.

[34] Mykland, P. A., and Zhang, L. The econometrics of high frequency data. In *Statistical Methods for Stochastic Differential Equations*, M. Kessler, A. Lindner, and M. Sorensen, Eds. Chapman & Hall/CRC, 2012, pp. 109–190.

[35] Podolskij, M. *New Theory on Estimation of Integrated Volatility with Applications.* PhD thesis, Ruhr-Universität Bochum, 2006.

[36] Podolskij, M., and Vetter, M. Understanding limit theorems for semimartingales: a short survey. *Stat. Neerl. 64*, 3 (2010), 329–351.

[37] Protter, P. E. *Stochastic integration and differential equations*, second ed. Springer-Verlag, Berlin, 2005.

[38] Rényi, A. On stable sequences of events. *Sankhyā Ser. A 25* (1963), 293–302.

[39] Samorodnitsky, G., and Taqqu, M. S. *Stable non-Guassian Random Processes, Stochastic Models with Infinite Variance.* Stochastic Modeling. Chapman & Hall/CRC, Boca Raton, Florida, 2000.

[40] Sato, K. *Lévy Processes and Infinitely Divisible Distributions.* Cambridge University Press, 1999.

[41] Woerner, J. H. C. Analyzing the fine structure of continuous time stochastic processes. In *Seminar on Stochastic Analysis, Random Fields and Applications VI*, vol. 63 of *Progr. Probab.* Birkhäuser/Springer Basel AG, Basel, 2011, pp. 473–492.

[42] Zivot, E., and Yan, B. Analysis of High-Frequency Financial Data with S-Plus. Online: `http://faculty.washington.edu/ezivot/research/hfanalysis.pdf` [Accessed 21/11/13].