Bernoulli **20**(2), 2014, 919–957
DOI: 10.3150/13-BEJ510

Efficient maximum likelihood estimation for Lévy-driven Ornstein–Uhlenbeck processes

HILMAR MAI

Weierstrass Institute, Mohrenstr. 39, 10117 Berlin, Germany. E-mail: mai@wias-berlin.de

We consider the problem of efficient estimation of the drift parameter of an Ornstein–Uhlenbeck type process driven by a Lévy process when high-frequency observations are given. The estimator is constructed from the time-continuous likelihood function that leads to an explicit maximum likelihood estimator and requires knowledge of the continuous martingale part. We use a thresholding technique to approximate the continuous part of the process. Under suitable conditions, we prove asymptotic normality and efficiency in the Hájek–Le Cam sense for the resulting drift estimator. Finally, we investigate the finite sample behavior of the method and compare our approach to least squares estimation.

Keywords: discrete time observations; efficient drift estimation; jump filtering; Lévy process; maximum likelihood estimation; Ornstein–Uhlenbeck process

1. Introduction

Let $(L_t, t \geq 0)$ be a Lévy process on a filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t), P)$ adapted to the filtration $(\mathcal{F}_t)_{t\geq 0}$. Denote by (b, σ^2, μ) the Lévy–Khintchine triplet of L. We call for every $a \in \mathbb{R}$ a strong solution X to the stochastic differential equation

$$\mathrm{d}X_t = -aX_t\,\mathrm{d}t + \mathrm{d}L_t, \qquad t \in \mathbb{R}_+, \qquad X_0 = \tilde{X}, \tag{1}$$

a Lévy-driven Ornstein–Uhlenbeck (OU) process or Ornstein–Uhlenbeck type process. The initial condition $\tilde{X}$ is assumed to be independent of L. We consider the problem of estimating the mean reversion parameter a when observations $X_{t_1}, \dots, X_{t_n}$ on an interval $[0, T_n]$ are given. It is well known that the drift of X is identifiable only in the limit $T_n \to \infty$, even when time-continuous observations are given. Therefore, we work under the asymptotic scheme $T_n \to \infty$ and $\Delta_n = \max_{1\leq i\leq n-1}\{|t_{i+1} - t_i|\} \downarrow 0$ as $n \to \infty$.

The OU process serves us here as a toy model to understand the interplay of jumps and continuous component of X in this estimation problem. This interplay is fundamental also for drift estimation in more general models (cf. [18]).

Ornstein–Uhlenbeck type processes have important applications in various fields. In mathematical finance they are well know as a main building block of the Barndorff–Nielsen–Shephard stochastic volatility model (cf. [3]). But also in neuroscience they are popular for the description of the membrane potential of a neuron (cf. [11] and [17]).

Estimation of Lévy-driven Ornstein–Uhlenbeck processes has been considered by several authors (see [22] and the references therein) mostly when the driving Lévy process is a subordinator. Some examples are [13] on nonparametric estimation of the Lévy density of L, in [4] the

1350-7265

Davis–McCormick estimator was applied in the OU context and parametric estimation based on a cumulant M-estimator was studied in [12]. In [8], least squares estimation of the drift parameter for an α-stable driver is discussed, when no Gaussian component is present. Reference [22] found that the rate of convergence of the least absolute deviation estimator is either the standard parametric rate, when L has a Gaussian component, or is faster than the standard rate, when L is a pure jump process and depends on the activity of the jumps. In [25], joint parametric estimation of the drift and the Lévy measure was treated via estimating functions. Unfortunately, none of these methods lead to an efficient estimator of the drift when L is a general Lévy process.

To construct an efficient estimator, our starting point will be the continuous time likelihood function. From this likelihood function, an explicit maximum likelihood estimator can be derived, which is efficient in the sense of the Hájek–Le Cam convolution theorem. In the likelihood function the continuous martingale part of X under the dominating measure appears, which is not directly observed in our setting. For discrete observations, we approximate the continuous part of X by neglecting increments that are larger than a certain threshold that has to be chosen appropriately. We will call this thresholding technique a jump filter. For this discretized likelihood estimator with jump filtering, we prove asymptotic normality and efficiency by showing that it attains the same asymptotic distribution as the benchmark estimator based on time-continuous observations.

This leads to the main mathematical question underlying this estimation problem. Can we recover the continuous part of X in the high-frequency limit via jump filtering? If L has only compound Poisson jumps it is intuitively clear that the answer is yes. But when L has infinitely many small jumps in every finite interval this is a much more challenging question. It turns out that even in this situation jump filtering works under mild assumptions on the behavior of the Lévy measure around zero. The main condition here is that the Blumenthal–Getoor index of the jump part is strictly less than two, which is apparently a necessary condition in this context. In this setting, jump filtering becomes possible since on a small time scale increments of the continuous part and of the jump part exhibit a different order of magnitude such that they can be distinguished via thresholding. To control the small jumps of L under thresholding, we derive an estimate for the Markov generator of a thresholded pure jump Lévy process. Estimates of this type without thresholding were given, among others, in [7] and [10].

The problem of separation between continuous and jump part of a process appears naturally in many situations. For example, in estimation of the integrated volatility of a jump diffusion process via realized volatility the quadratic variation of the jump component has to be removed. This problem has been solved by thresholding in [19] for Poisson jumps and in [20] for more general jump behavior. Efficiency questions in this context in a simple parametric model have been addressed in [1]. When properties of the jump component are of interest thresholding techniques are equally useful as demonstrated, among others, in [2]. In contrast to our discussion all these references consider the separation problem for a finite and fixed observation horizon $T_n = T < \infty$.

We also demonstrate in a simulation example that jump filtering leads to a major improvement of the drift estimate for finite sample size. Let us also mention that implementation of the drift estimator is straightforward and computation time is not an issue even for large data sets.

The paper is organized as follows: in Section 2 we derive the maximum likelihood estimator based on time-continuous observations, give its asymptotic properties and obtain the efficient limiting distribution for this estimation problem. Section 3 deals with estimating the drift parameter from discrete observations when L has finite jump activity. In Section 4 we build on the results from Sections 2 and 3 to prove efficiency also for possibly infinite jump activity. The finite sample behavior of the estimator is investigated in Section 5 based on simulated data together with an analysis of the impact of the jump filter on the performance of the estimator.

2. Maximum likelihood estimation

Let us summarize some important facts on Ornstein–Uhlenbeck type processes. It follows from Itô's formula that an explicit solution of (1) is given by

$$X_t = \mathrm{e}^{-at} X_0 + \int_0^t \mathrm{e}^{-a(t-s)} \, \mathrm{d}L_s, \qquad t \in \mathbb{R}_+. \tag{2}$$

The integral in (2) can by partial integration be defined path-wise as a Riemann–Stieltjes integral, since the integrand is of finite variation (see [6], e.g.). This solution to equation (1) is unique up to indistinguishability.

When L is non-deterministic it was shown in [9] that equation (1) admits a causal stationary solution (cf. also [21] and [24]) if and only if

$$\int_{|x|>1} \log|x| \mu(\mathrm{d}x) < \infty \quad \text{and} \quad a > 0. \tag{3}$$

Under these conditions X has a unique invariant distribution G and $X_t \xrightarrow{\mathcal{D}} X_\infty \sim G$ as $t \to \infty$.

The Ornstein–Uhlenbeck process X exhibits a modification with càdlàg paths and hence it induces a measure P^a on the space $D[0, \infty)$ of càdlàg functions on the interval $[0, \infty)$. Denote by P_t^a the restriction of P^a to the σ-field $\mathcal{F}_t$. If $\sigma^2 > 0$ then these induced measure are locally equivalent (cf. [26]) and the corresponding Radon–Nikodym derivative or likelihood function is given by

$$\frac{\mathrm{d}P_t^a}{\mathrm{d}P_t^0} = \exp\left(-\frac{a}{\sigma^2} \int_0^T X_s \, \mathrm{d}X_s^c - \frac{a^2}{2\sigma^2} \int_0^T X_s^2 \, \mathrm{d}s \right),$$

where X^c denotes the continuous P^0-martingale part of X. This leads to the explicit maximum likelihood estimator

$$\hat{a}_T = -\frac{\int_0^T X_s \, \mathrm{d}X_s^c}{\int_0^T X_s^2 \, \mathrm{d}s} \tag{4}$$

for a when the process is fully observed on $[0, T]$. The estimator $\hat{a}_T$ cannot be applied in this form, since time-continuous observations are usually not available in most applications. Therefore, we will develop in the next section a discrete version of $\hat{a}_T$ and prove its efficiency. The main challenge there will be that the continuous part X^c is not directly observed and hence has to be approximated from discrete observations of X via jump filtering.

However, for us $\hat{a}_T$ will serve as a benchmark for the estimation problem with discrete observations. Asymptotic normality and efficiency in the Hájek–Le Cam sense of $\hat{a}_T$ follow easily from general results for exponential families of stochastic processes (cf. [16] and [18]). These results provide an efficiency bound for the case of discrete observations in the next section. Let us summarize them in the following theorem.

Theorem 2.1. (i) *Under the condition $\sigma^2 > 0$ the estimator $\hat{a}_T$ exists uniquely and is strongly consistent under P^a, that is,*

$$\hat{a}_T \xrightarrow{a.s.} a$$

under P^a as $T \to \infty$.

(ii) *Suppose that additionally* (3) *holds and that X has bounded second moments such that the invariant distribution satisfies $E_a[X_\infty^2] < \infty$. Then under P^a*

$$\sqrt{T}(\hat{a}_T - a) \xrightarrow{\mathcal{D}} N\left(0, \frac{\sigma^2}{E_a[X_\infty^2]}\right)$$

and

$$\sigma^{-1} S_T^{1/2}(\hat{a}_T - a) \xrightarrow{\mathcal{D}} N(0, 1) \tag{5}$$

as $T \to \infty$, where $S_T = \int_0^T X_s^2 \, \mathrm{d}s$.

(iii) *The statistical experiment $\{P^a, a \in \mathbb{R}_+\}$ is locally asymptotically normal.*

(iv) *The estimator $\hat{a}_T$ is asymptotically efficient in the sense of Hájek–Le Cam.*

For a proof we refer to [18], Section 4.2.

Remark 2.2. When σ^2 is known or a consistent estimator is at hand, we can use (5) to construct confidence intervals for a.

3. Discrete observations: Finite activity

In this section, we consider the estimation of a for discrete observations. The maximum likelihood estimator for the drift given in (4) involves the continuous martingale part that is unknown when only discrete observations are given. Hence, we will approximate the continuous part of the process by removing observations that most likely contain jumps. We restrict our attention in this section to the case that the driving Lévy process has jumps of finite activity. The jump filtering technique provides us in the high-frequency limit an asymptotically normal and efficient estimator. Based on these results, we will treat the general case of an infinitely active jump component in Section 4.

3.1. Estimator and observation scheme

Let X be an Ornstein–Uhlenbeck process defined by (1) and suppose we observe X at discrete time points $0 = t_1 < t_2 < \cdots < t_n = T_n$ such that $T_n \to \infty$ as well as $\Delta_n = \max_{1 \le i \le n-1}\{|t_{i+1} - t_i|\} \downarrow 0$ and $n\Delta_n T_n^{-1} = \mathrm{O}(1)$ as $n \to \infty$. The last condition assures that the number of observations n does not grow faster than $T_n \Delta_n^{-1}$. It can always be fulfilled by neglecting observations and will simplify the formulation of the proof considerably. Denote by (b, σ^2, μ) the Lévy–Khintchine triplet of L. Assume throughout this section that $\lambda = \mu(\mathbb{R}) < \infty$ for the Lévy measure μ.

By deleting increments that are larger than a threshold $v_n > 0$ we filter increments that most likely contain jumps and thus approximate the continuous part with the remaining increments. Applied to the time-continuous likelihood estimator (4) this method leads to the following estimator:

$$\bar{a}_n := -\frac{\sum_{i=0}^{n-1} X_{t_i} \Delta_i X \mathbf{1}_{\{|\Delta_i X| \le v_n\}}}{\sum_{i=0}^{n-1} X_{t_i}^2 (t_{i+1} - t_i)}. \tag{6}$$

Here $v_n > 0, n \in \mathbb{N}$, is a cut-off sequence that will be chosen as a function of the maximal distance between observations Δ_n and $\Delta_i X = X_{t_{i+1}} - X_{t_i}$.

In the finite activity case, the jump part J of L can be written as a compound Poisson process

$$J_t = \sum_{i=1}^{N_t} Z_i,$$

where N is a Poisson process with intensity λ and the jump heights $Z_1, Z_2, \ldots$ are i.i.d. with distribution F.

3.2. Asymptotic normality and efficiency

The indicator function that appears in $\bar{a}_n$ deletes increments that are larger than v_n. In [19], it was shown that increments of the continuous part of X over an interval of length Δ_n are with high probability smaller than $\Delta_n^{1/2}$. Hence, we set $v_n = \Delta_n^{\beta}$ for $\beta \in (0, 1/2)$ to keep the continuous part in the limit unaffected by the threshold. In order to be able to choose v_n such that $X_n^c = \sum_{i=0}^{n-1} \Delta_i X \mathbf{1}_{\{|\Delta_i X| \le v_n\}}$ approximates the continuous martingale part in the limit, we make the following assumptions on the jumps of L and the observation scheme.

Assumption 3.1. (i) *Suppose that μ and a satisfy* (3), *the drift $b = 0$, the process X has bounded second moments*,

(ii) *the distribution F of the jump heights is such that*

$$F\big((-2\Delta_n^{\beta}, 2\Delta_n^{\beta})\big) = \mathrm{o}\big(T_n^{-1}\big),$$

(iii) *and there exists $\beta \in (0, 1/2)$ such that the maximal distance between observations satisfies* $T_n \Delta_n^{(1-2\beta)\wedge(1/2)} = \mathrm{o}(1)$.

Remark 3.2. Suppose that F has a bounded Lebesgue density f. Then

$$F\big((-\Delta_n^\beta, \Delta_n^\beta)\big) = \mathrm{O}\big(\Delta_n^\beta\big)$$

and Assumption 3.1(ii) becomes $\Delta_n^\beta T_n = \mathrm{o}(1)$. Together with Assumption 3.1(i), we obtain that $\beta = 1/3$ leads to an optimal compromise between Assumption 3.1(i) and (ii).

Remark 3.3. Assumption 3.1(iii) means here that for given $T_n \to \infty$ we require $\Delta_n \downarrow 0$ fast enough such that there exists $\beta \in (0, 1/2)$:

$$T_n \Delta_n^{1-2\beta} = \mathrm{o}(1) \quad \text{and} \quad T_n \Delta_n^{1/2} = \mathrm{o}(1).$$

Of course one of these two conditions will be dominating and determine the order of Δ_n.

Remark 3.4. Assumption 3.1(ii) gives a lower bound for the choice of the threshold β. At the same time Assumption 3.1(iii) limits the range of possible β's from above, since the available frequency of observations, that is, the order of Δ_n, may be limited in specific applications. Hence, the distribution F, the observation length T_n and frequency Δ_n fix a range for the choice of β. At this point the question of a data driven method to choose β arises, but this will not be considered in this work. The condition $b = 0$ is necessary in this context, since otherwise there is no hope to recover the continuous martingale part via jump filtering.

The following theorem gives as the main result of this section a central limit theorem for the discretized MLE with jump filter.

Theorem 3.5. *Suppose that Assumption 3.1 holds and $\sigma^2 > 0$. Set $v_n = \Delta_n^\beta$ for $\beta \in (0, 1/2)$, then*

$$T_n^{1/2}(\bar{a}_n - a) \xrightarrow{\mathcal{D}} N\left(0, \frac{\sigma^2}{E_a[X_\infty^2]}\right) \qquad \text{as } n \to \infty.$$

The estimator $\bar{a}_n$ is asymptotically efficient.

Remark 3.6. Asymptotic efficiency follows immediately from Theorem 2.1 and the first statement of Theorem 3.5.

3.3. Proofs

We divide the proof of the theorem into several lemmas. First of all, we need a probability bound for the event that the continuous component of X exceeds a certain threshold.

By the Lévy–Itô decomposition (cf. [23]) and since $b = 0$ in our setting the driving Lévy process can be decomposed as $L = W + J$, where W is a standard Wiener process and J is a pure jump Lévy process independent of W. Denote by D the drift component of X, that is,

$$D_t = -a \int_0^t X_s \, \mathrm{d}s.$$

Lemma 3.7. *Let* $\sup_{s\geq 0}\{E[|X_s|^l]\}<\infty$ *for some* $l\geq 1$. *For any* $\delta\in(0,1/2)$ *and* $i\in\{1,\dots,n-1\}$, *we have*

$$P\bigl(|\Delta_i W+\Delta_i D|>\Delta_n^{1/2-\delta}\bigr)=\mathrm{O}\bigl(\Delta_n^{l(1/2+\delta)}\bigr)\qquad \textit{as } n\to\infty.$$

Proof. In the first step, we separate $\Delta_i W$ and $\Delta_i D$:

$$P\bigl(|\Delta_i W+\Delta_i D|>\Delta_n^{1/2-\delta}\bigr)\leq P\left(|\Delta_i W|>\frac{\Delta_n^{1/2-\delta}}{2}\right)+P\left(|\Delta_i D|>\frac{k\Delta_n^{1/2}}{2}\right).$$

By Lemma 22.2 in [15],

$$P\left(|\Delta_i W|>\frac{\Delta_n^{1/2-\delta}}{2}\right)\leq 2\Delta_n^{\delta}\mathrm{e}^{-1/(8\Delta_n^{\delta})}.$$

It follows from Jensen's inequality that

$$\left|\int_{t_i}^{t_{i+1}}X_s\,\mathrm{d}s\right|^l\leq \Delta_n^{l-1}\int_{t_i}^{t_{i+1}}|X_s|^l\,\mathrm{d}s.$$

This leads to

$$E\bigl[|\Delta_i D|^l\bigr]\leq \Delta_n^{l-1}a^l\int_{t_i}^{t_{i+1}}E\bigl[|X_s|^l\bigr]\,\mathrm{d}s\leq \Delta_n^l a^l\sup_{s\in[t_i,t_{i+1}]}\bigl\{E\bigl[|X_s|^l\bigr]\bigr\}.$$

Finally, Markov's inequality yields

$$P\left(|\Delta_i D|>\frac{\Delta_n^{1/2-\delta}}{2}\right)\leq a^l\sup_{s\in[t_i,t_{i+1}]}\bigl\{E\bigl[|X_s|^l\bigr]\bigr\}\frac{2^l\Delta_n^l}{\Delta_n^{l(1/2-\delta)}}=\mathrm{O}\bigl(\Delta_n^{l(1/2+\delta)}\bigr).$$

□

3.3.1. *Jump filtering*

First, we will investigate how to choose the cut-off sequence v_n in order to filter the jumps. Define for $n\in\mathbb{N}$ and $i\in\{1,\dots,n\}$ the following sequence of events

$$A_n^i=\bigl\{\omega\in\Omega:\ \mathbf{1}_{\{|\Delta_i X|\leq v_n\}}(\omega)=\mathbf{1}_{\{\Delta_i N=0\}}(\omega)\bigr\}.$$

Here N denotes the counting measure that counts the jumps of L.

Lemma 3.8. *Suppose that Assumption* 3.1 *holds and set* $v_n=\Delta_n^{\beta}$, $\beta\in(0,1/2)$, *then it follows that for* $A_n=\bigcap_{i=1}^n A_n^i$ *we have*

$$P(A_n)\to 1\qquad \textit{as } n\to\infty.$$

Proof. Observe that

$$P\bigl(A_n^c\bigr)=P\left(\bigcup_{i=1}^n\bigl(A_n^i\bigr)^c\right)\leq\sum_{i=1}^n P\bigl(\bigl(A_n^i\bigr)^c\bigr).$$

By setting

$$K_n^i = \{|\Delta_i X| \le v_n\},$$
$$M_n^i = \{\Delta_i N = 0\},$$

we can rewrite $(A_n^i)^c$ as

$$(A_n^i)^c = \{\mathbf{1}_{K_n^i} \ne \mathbf{1}_{M_n^i}\} = (K_n^i \setminus M_n^i) \cup (M_n^i \setminus K_n^i).$$

Here the events $K_n^i \setminus M_n^i$ and $M_n^i \setminus K_n^i$ correspond to the two types of errors that can occur when we search for jumps. In the first case, we miss a jump and in the second case we neglect an increment although it does not contain any jumps. Next, we are going to bound the probability of both errors:

$$P((A_n^i)^c) = P(K_n^i \setminus M_n^i) + P(M_n^i \setminus K_n^i). \tag{7}$$

Set $\Delta_i = t_{i+1} - t_i$. For the first type of error, we obtain

$$\begin{aligned} P(K_n^i \setminus M_n^i) &= P(|\Delta_i X| \le v_n, \Delta_i N > 0) \\ &= \sum_{j=1}^{\infty} \mathrm{e}^{-\lambda\Delta_i} \frac{(\lambda\Delta_i)^j}{j!} P(|\Delta_i X| \le v_n | \Delta_i N = j) \\ &\le P(\Delta_i N = 1) P(|\Delta_i X| \le v_n | \Delta_i N = 1) + \mathrm{O}(\Delta_n^2) \end{aligned} \tag{8}$$

and

$$\begin{aligned} P(|\Delta_i X| \le v_n | \Delta_i N = 1) \le{}& P(|\Delta_i X| \le v_n, |\Delta_i J| > 2v_n | \Delta_i N = 1) \\ &+ P(|\Delta_i X| \le v_n, |\Delta_i J| \le 2v_n | \Delta_i N = 1). \end{aligned} \tag{9}$$

The first term on the right-hand side is bounded by

$$\begin{aligned} &P(|\Delta_i X| \le v_n, |\Delta_i J| > 2v_n | \Delta_i N = 1) \\ &\quad = P(|\Delta_i W + \Delta_i J + \Delta_i D| \le v_n, |\Delta_i J| > 2v_n | \Delta_i N = 1) \\ &\quad \le P(|\Delta_i W + \Delta_i D| > v_n, \Delta_i N = 1) P(\Delta_i N = 1)^{-1} \\ &\quad \le P(|\Delta_i W + \Delta_i D| > v_n) P(\Delta_i N = 1)^{-1} = P(\Delta_i N = 1)^{-1} \mathrm{O}(\Delta_n^{2-2\beta}), \end{aligned} \tag{10}$$

where we used Lemma 3.7 with $l = 2$. Denote by F the distribution of the jump heights of J. Then we obtain for the second term on the right-hand side of (9)

$$P(|\Delta_i X| \le v_n, |\Delta_i J| \le 2v_n | \Delta_i N = 1) \le P(|\Delta_i J| \le 2v_n | \Delta_i N = 1) = F((-2v_n, 2v_n)).$$

For the second addend in (7) it follows by independence of W and J that

$$\begin{aligned} P\big(M_n^i \setminus K_n^i\big) &= P\big(|\Delta_i X| > v_n, \Delta_i N = 0\big) \\ &\le P\big(|\Delta_i W + \Delta_i D| > v_n\big). \end{aligned}$$

Lemma 3.7 yields

$$P\big(|\Delta_i W + \Delta_i D| > v_n\big) = \mathrm{O}\big(\Delta_n^{2-2\beta}\big). \tag{11}$$

Finally, (8), (10) and (11) lead to

$$P\big(\big(A_n^i\big)^c\big) \le F\big(\big(-2\Delta_n^\beta, 2\Delta_n^\beta\big)\big)\Delta_n + \mathrm{O}\big(\Delta_n^{2-2\beta}\big)$$

such that the statement follows, since we have shown that

$$P\big(A_n^c\big) \le \sum_{i=1}^n P\big(\big(A_n^i\big)^c\big) \le \mathrm{O}(T_n) F\big(\big(-2\Delta_n^\beta, 2\Delta_n^\beta\big)\big) + \mathrm{O}\big(T_n \Delta_n^{1-2\beta}\big).$$

□

3.3.2. *Approximation of the continuous martingale part*

Lemma 3.9. *Under Assumption 3.1, we obtain*

$$\left|\sum_{i=0}^{n-1} X_{t_i}\big(\Delta_i X \mathbf{1}_{\{|\Delta_i X| \le v_n\}} - \Delta_i X^c\big)\right| = \mathrm{O}_p\big(T_n \Delta_n^{1/2}\big)$$

as $n \to \infty$.

Proof. On A_n from Lemma 3.8, we have

$$\sum_{i=0}^{n-1} X_{t_i}\big(\Delta_i X \mathbf{1}_{\{|\Delta_i X| \le v_n\}} - \Delta_i X^c\big) = \sum_{i=0}^{n-1} X_{t_i}\big(\Delta_i X \mathbf{1}_{\{\Delta_i N = 0\}} - \Delta_i X^c\big). \tag{12}$$

By Lemma 3.8, we have $P(A_n) \to 1$ as $n \to \infty$. Observe now that the difference of the increments on the right-hand side of (12) is unequal to zero only if a jump occurred in that interval, that is,

$$\Delta_i X \mathbf{1}_{\{\Delta_i N = 0\}} - \Delta_i X^c = \begin{cases} -\Delta_i X^c; & \Delta_i N > 0, \\ 0; & \Delta_i N = 0. \end{cases}$$

Define $C_i^n = \{\Delta_i N > 0\}$ and observe that

$$E\left|\mathbf{1}_{A_n} \sum_{i=0}^{n-1} X_{t_i}\big(\Delta_i X \mathbf{1}_{\{\Delta_i N = 0\}} - \Delta_i X^c\big)\right| = E\left|\sum_{i=0}^{n-1} X_{t_i} \Delta_i X^c \mathbf{1}_{A_n \cap C_i^n}\right|.$$

The ith increment of X^c can be written as $\Delta_i X^c = \Delta_i W + \Delta_i D$. Therefore,

$$E\left|\sum_{i=0}^{n-1} X_{t_i}\Delta_i X^c \mathbf{1}_{A_n\cap C_i^n}\right| \le \sum_{i=0}^{n-1} E\big[\big|X_{t_i}(\Delta_i W + \Delta_i D)\big|\mathbf{1}_{A_n\cap C_i^n}\big]$$
$$\le \sum_{i=0}^{n-1} E\big[\big(|X_{t_i}\Delta_i W| + |X_{t_i}\Delta_i D|\big)\mathbf{1}_{C_i^n}\big].$$

The number of jumps of J follows a Poisson process with intensity λ such that $P(C_i^n) \le \Delta_n\lambda$. The independence of $N \perp W$ and $\Delta_i N \perp X_{t_i}$ yields

$$\sum_{i=0}^{n-1} E\big[|X_{t_i}\Delta_i W|\mathbf{1}_{C_i^n}\big] = \sum_{i=0}^{n-1} E\big[|X_{t_i}|\big]E\big[|\Delta_i W|\big]P\big(C_i^n\big) \le \mathrm{O}\big(T_n\Delta_n^{1/2}\big).$$

Finally, by Hölder's inequality

$$\sum_{i=0}^{n-1} E\big[|X_{t_i}\Delta_i D|\mathbf{1}_{C_i^n}\big] \le \sum_{i=0}^{n-1} E\big[X_{t_i}^2(\Delta_i D)^2\big]^{1/2} P\big(C_i^n\big)^{1/2} = \mathrm{O}\big(T_n\Delta_n^{1/2}\big).$$ □

3.3.3. *Central limit theorem for the discretized estimator*

To prove Theorem 3.5, we show next that when we discretize the time-continuous estimator $\hat{a}_T$ as

$$\hat{a}_n = -\frac{\sum_{i=1}^n X_{t_i}\Delta_i X^c}{\sum_{i=1}^n X_{t_i}^2(t_{i+1}-t_i)},$$

then $\hat{a}_n$ attains the same asymptotic distribution as $\hat{a}_T$ itself. In the last step, we will then show that the discretized MLE and the estimator with jump filter show the same limiting behavior.

Lemma 3.10. *If Assumption 3.1 is fulfilled, then the convergence*

$$T_n^{1/2}(\hat{a}_n - a) \xrightarrow{\mathcal{D}} N\big(a, \sigma^2 E_a\big[X_\infty^2\big]^{-1}\big) \qquad \text{as } n\to\infty$$

holds under P^a.

Proof. Let W denotes a P^a-Wiener process. The continuous P^0-martingale part can be written as

$$X_t^c = \sigma W_t - a\int_0^t X_s\,\mathrm{d}s.$$

This leads to the decomposition

$$T_n^{1/2}(\breve{a}_n - a) = T_n^{1/2}a\left(\frac{\sum_{i=0}^{n-1} X_{t_i}\int_{t_i}^{t_{i+1}} X_s\,\mathrm{d}s}{\sum_{i=0}^{n-1} X_{t_i}^2\Delta_i} - 1\right) - T_n^{1/2}\frac{\sigma\sum_{i=0}^{n-1} X_{t_i}\Delta_i W}{\sum_{i=0}^{n-1} X_{t_i}^2\Delta_i} = S_n^1 - S_n^2.$$

We will show now that $S_n^1 \xrightarrow{p} 0$ and $S_n^2 \xrightarrow{\mathcal{D}} N(0, \sigma^2 E_a[X_\infty^2]^{-1})$ as $n \to \infty$ such that the statement of the proposition follows. Define $\lfloor t \rfloor_n = \max_{i \le n}\{t_i | t_i \le t\}$. Let us first consider convergence of S_n^1. Observe that

$$S_n^1/a = \frac{T_n^{-1/2}(\sum_{i=0}^{n-1} X_{t_i} \int_{t_i}^{t_{i+1}} X_s \,\mathrm{d}s - \sum_{i=0}^{n-1} X_{t_i}^2 \Delta_i)}{T_n^{-1} \sum_{i=0}^{n-1} X_{t_i}^2 \Delta_i}. \tag{13}$$

For the numerator, we obtain

$$\begin{aligned} T_n^{-1/2} E\left[\left|\sum_{i=0}^{n-1} X_{t_i} \int_{t_i}^{t_{i+1}} X_s \,\mathrm{d}s - \sum_{i=0}^{n-1} X_{t_i}^2 \Delta_i\right|\right] &\le T_n^{-1/2} \int_0^{T_n} E_a\big[\big|X_{\lfloor t \rfloor_n} X_t - X_{\lfloor t \rfloor_n}^2\big|\big] \,\mathrm{d}t \\ &= \mathrm{O}\big(T_n^{1/2} \Delta_n^{1/2}\big) \end{aligned} \tag{14}$$

such that the numerator converges to zero in L^1. A similar estimate for the denominator yields

$$T_n^{-1} E_a\left[\left|\int_0^{T_n} X_t^2 \,\mathrm{d}t - \sum_{i=0}^{n-1} X_{t_i}^2 \Delta_i\right|\right] = \mathrm{O}\big(\Delta_n^{1/2}\big),$$

and since the ergodic theorem implies that $T_n^{-1} \int_0^{T_n} X_t^2 \,\mathrm{d}t \xrightarrow{p} E_a[X_\infty^2]$ as $n \to \infty$, we conclude

$$T_n^{-1} \sum_{i=0}^{n-1} X_{t_i}^2 \Delta_i \xrightarrow{p} E_a\big[X_\infty^2\big] \tag{15}$$

as $n \to \infty$. This convergence together with (13) and the estimate (14) imply that $S_n^1 \xrightarrow{p} 0$ as $n \to \infty$.

It remains to prove convergence of S_n^2. From Itô's isometry and stationarity of X, we obtain for the numerator of S_n^1 that

$$\begin{aligned} T_n^{-1} E_a\left[\left(\int_0^{T_n} X_t \,\mathrm{d}W_t - \sum_{i=0}^{n-1} X_{t_i} \Delta_i W\right)^2\right] &= T_n^{-1} E_a\left[\left(\int_0^{T_n} (X_t - X_{\lfloor t \rfloor_n}) \,\mathrm{d}W_t\right)^2\right] \\ &= T_n^{-1} E_a\left[\int_0^{T_n} (X_t - X_{\lfloor t \rfloor_n})^2 \,\mathrm{d}t\right] \\ &= T_n^{-1} \int_0^{T_n} E_a\big[(X_t - X_{\lfloor t \rfloor_n})^2\big] \,\mathrm{d}t = \mathrm{O}(\Delta_n). \end{aligned}$$

The numerator of S_n^2 is a continuous martingale and its quadratic variation converges due to the ergodic theorem to the second moment of X. The martingale central limit theorem implies now

$$T_n^{-1/2} \sigma \int_0^{T_n} X_t \,\mathrm{d}W_t \xrightarrow{\mathcal{D}} N\big(0, \sigma^2 E_a\big(X_\infty^2\big)\big)$$

such that also

$$T_n^{-1/2}\sigma\sum_{i=0}^{n-1}X_{t_i}\Delta_i W \xrightarrow{\mathcal{D}} N\big(0,\sigma^2 E_a\big(X_\infty^2\big)\big)$$

as $n\to\infty$. This convergence together with (15) and Slutsky's lemma lead to

$$S_n^1 \xrightarrow{\mathcal{D}} N\big(0,\sigma^2 E_a\big[X_\infty^2\big]^{-1}\big)$$

as $n\to\infty$. This completes the proof. □

Proof of Theorem 3.5. By Lemma 3.10 $T_n^{1/2}(\hat{a}_n - a)\xrightarrow{\mathcal{D}} N(0,\frac{\sigma^2}{E_a[X_\infty^2]})$ as $n\to\infty$. By Slutsky's lemma, it remains to show

$$T_n^{1/2}(\bar{a}_n-\hat{a}_n)\xrightarrow{p} 0 \qquad \text{as } n\to\infty. \tag{16}$$

Observe that

$$T_n^{1/2}(\bar{a}_n-\hat{a}_n)=T_n^{1/2}\left(\frac{\sum_{i=1}^n X_{t_i}\Delta_i X\mathbf{1}_{\{|\Delta_i X|\le v_n\}}-\sum_{i=1}^n X_{t_i}\Delta_i X^c}{\sum_{i=1}^n X_{t_i}^2\Delta_i}\right).$$

By Lemma 3.9 we obtain under P_a that

$$T_n^{-1/2}\left(\sum_{i=1}^n X_{t_i}\Delta_i X\mathbf{1}_{\{|\Delta_i X|\le v_n\}}-\sum_{i=1}^n X_{t_i}\Delta_i X^c\right)\xrightarrow{p} 0 \qquad \text{as } n\to\infty$$

and

$$T_n^{-1}\sum_{i=1}^n X_{t_i}^2\Delta_i \xrightarrow{p} E_a\big[X_\infty^2\big],$$

such that (16) follows. □

4. Discrete observations: Infinite activity

In this section, we generalize the results from Section 3 to the case that the jump part of the driving Lévy process can be of infinite activity. We give conditions on the Lévy measure and suitable rates for the cut-off sequence that ensure separation in the high-frequency limit between jump part and continuous part. Under these conditions, we will then prove asymptotic normality and efficiency of the drift estimator $\bar{a}_n$ given in (6).

The observation scheme considered here will be like in Section 3.1, that is, $0=t_1<t_2<\cdots<t_n=T_n$ such that $T_n\to\infty$ as well as $\Delta_n=\max_{1\le i\le n-1}\{|t_{i+1}-t_i|\}\downarrow 0$ and $n\Delta_n T_n^{-1}=\mathrm{O}(1)$ as $n\to\infty$.

4.1. Asymptotic normality and efficiency

In this section, we state as the main result of this paper a CLT for the estimation error of $\bar{a}_n$. The limiting distribution will imply asymptotic efficiency of $\bar{a}_n$. But before we can formulate the theorem, we introduce some notation and mild assumptions on the jump part of L that enable us to separate the jump part and continuous part via jump filtering.

Let N denote the Poisson random measure associated to the jump part of L. The jump component J of X, the components M of jumps smaller than one and U of jumps larger than one and the drift D are given by

$$\begin{aligned} J_t &= \int_0^t \int_{-\infty}^{\infty} x\big(N(\mathrm{d}x, \mathrm{d}s) - \mu(\mathrm{d}x)\lambda(\mathrm{d}s)\big), \\ M_t &= \int_0^t \int_{-1}^{1} x\big(N(\mathrm{d}x, \mathrm{d}s) - \mu(\mathrm{d}x)\lambda(\mathrm{d}s)\big), \\ U_t &= J_t - M_t, \\ D_t &= -a \int_0^t X_s \,\mathrm{d}s, \end{aligned} \tag{17}$$

respectively. Owing to this decomposition of X we can apply the results from Chapter 3 to D, W and U and thus can focus on M. To control the small jumps of M, we impose the following assumption on the Lévy measure μ.

Assumption 4.1. (i) *Suppose that* (3) *holds*, $b = 0$ *and* X *has bounded second moments.*

(ii) *There exists an* $\alpha \in (0, 2)$ *such that as* $v \downarrow 0$

$$\int_{-v}^{v} x^2 \mu(\mathrm{d}x) = \mathrm{O}\big(v^{2-\alpha}\big). \tag{18}$$

(iii) *There exists* $\eta > 0$ *and* $n_0 \in \mathbb{N}$ *such that for all* $\varepsilon \le \eta$ *and* $n \ge n_0$

$$E[\Delta_i M \mathbf{1}_{\{|\Delta_i M| \le \varepsilon\}}] = 0 \qquad \textit{for all } i \in \{1, \dots, n-1\}.$$

Remark 4.2. Assumption 4.1(ii) controls the intensity of small jumps, which is determined by the mass of μ around the origin. When γ denotes the Blumenthal–Getoor index of L defined by

$$\gamma = \inf_{c \ge 0} \left\{ \int_{|x| \le 1} |x|^c \mu(\mathrm{d}x) < \infty \right\} \le 2,$$

then $\alpha = \gamma$ satisfies (18), that is, Assumption 4.1(i) states that the Blumenthal–Getoor index is less than two. This is a natural condition in the context of jump filtering (see, e.g., [20] in the context of volatility estimation).

Remark 4.3. To compare the finite to the infinite activity setting let us contrast Assumption 3.1(ii) with Assumption 4.1(ii). Both assumptions control the behavior of small jumps. When the Lévy measure is finite such that $\mu(\mathrm{d}x) = \lambda F(\mathrm{d}x)$ for some probability distribution F we can rewrite (18) as

$$\lambda \int_{-\Delta_n^\beta}^{\Delta_n^\beta} x^2 F(\mathrm{d}x) = \mathrm{O}\big(\Delta_n^\beta\big),$$

since $\alpha = 0$ in this case. At the same time Assumption 3.1(ii) dominates Assumption 4.1(ii) in the sense that

$$\int_{-\Delta_n^\beta}^{\Delta_n^\beta} x^2 \mu(\mathrm{d}x) \le \mu\big(\big(-\Delta_n^\beta, \Delta_n^\beta\big)\big) = \lambda F\big(\big(-\Delta_n^\beta, \Delta_n^\beta\big)\big).$$

Hence, if $F((-\Delta_n^\beta, \Delta_n^\beta)) = \mathrm{O}(\Delta_n^\beta)$ then Assumption 3.1(ii) implies Assumption 4.1(ii) such that a direct comparison becomes possible when F has a bounded Lebesgue density as in Example 4.7 below.

Lemma 4.4. *If the Lévy measure $\mu_{|[-1,1]}$ of M is symmetric around zero, then Assumption* 4.1(iii) *holds.*

Proof. Assumption 4.1(iii) is equivalent to the symmetry of the distribution of M restricted to $(-\delta, \delta)$ such that the statement follows, since an infinitely divisible distribution is symmetric if and only if its Lévy–Khintchine triplet is of the form $(0, \sigma^2, \mu)$ with μ being symmetric (cf. [23]). □

Remark 4.5. Assumption 4.1(iii) is a symmetry condition on the distribution $P^{\Delta_i M}_{|(-\varepsilon,\varepsilon)}$ of the increments of M restricted to $(-\varepsilon, \varepsilon)$ for every $0 < \varepsilon \le \delta$. A sufficient condition for Assumption 4.1(iii) in terms of the Lévy measure of L was given in Lemma 4.4. But it is easy to see that this is not a necessary condition. The main point here is that $P^{\Delta_i M}_{|(-\varepsilon,\varepsilon)}$ is not infinitely divisible anymore.

Since our method does not depend on the choice of the maximal jump size in the definition of M in (17), it follows that the condition given in Lemma 4.4 can be relaxed to $\mu_{|(-\varepsilon,\varepsilon)}$ being symmetric for some $\varepsilon > 0$ by redefining M to have a Lévy measure supported on $(-\varepsilon, \varepsilon)$.

The main result of this chapter is the following central limit theorem for the drift estimator with jump filter.

Theorem 4.6. *Suppose that Assumption* 4.1 *holds and $\sigma^2 > 0$. If there exists $\beta \in (0, 1/2)$ such that $T_n \Delta_n^{1-2\beta\wedge(1/2)} = \mathrm{o}(1)$ as $n \to \infty$ then $v_n = \Delta_n^\beta$ yields*

$$T_n^{1/2}(\bar{a}_n - a) \xrightarrow{\mathcal{D}} N\big(0, \sigma^2 E_a\big[X_\infty^2\big]^{-1}\big).$$

The estimator is asymptotically efficient.

Example 4.7. Let $L = W + J$, where J is a compound Poisson process

$$J_t = \sum_{i=1}^{N_t} Y_i$$

such that $Y_i \sim F$ are i.i.d. and N_t is a Poisson process with intensity λ. Suppose that F has a bounded Lebesgue density f. Then

$$\int_{-v}^{v} x^2 \mu(\mathrm{d}x) = \lambda \int_{-v}^{v} x^2 f(x)\,\mathrm{d}x \le C v^3$$

for $C > 0$ such that for L Assumption 4.1(i) holds for every $\alpha \in [0, 2)$.

More generally every Lévy process with Blumenthal–Getoor index less than two fulfills Assumption 4.1(i). This includes all Lévy processes commonly used in applications like (tempered) stable, normal inverse Gaussian, variance gamma and also gamma processes.

4.2. Proofs

Asymptotic efficiency of $\bar{a}_n$ follows from the first statement of Theorem 4.6 together with Theorem 2.1 such that it remains to prove the asymptotic normality result. We will divide the proof of Theorem 4.6 into several lemmas. In the proofs in this section, constants may change from line to line or even within one line without further notice.

4.2.1. *A moment bound*

In this section, we derive a moment bound for short time increments of pure jump Lévy processes. Set

$$f(x) = \begin{cases} x^2, & \text{if } |x| \le 1, \\ 0, & \text{if } |x| > 2, \end{cases}$$

and $f(x) \in [0, 2]$ for $|x| \in (1, 2]$ such that $f \in C^\infty(\mathbb{R})$. We scale f to be supported on $[-v, v]$ by

$$f^v(x) = v^2 f(x/v). \tag{19}$$

Proposition 4.8. *Let* $(M_t)_{t\ge 0}$ *be a pure jump Lévy process with Lévy measure* μ *such that* $\operatorname{supp}(\mu) \subset [-1, 1]$ *and Assumption* 4.1(i) *and* (ii) *hold. Then for all* $\beta \in (0, \frac{1}{2})$ *we obtain*

$$E\big[f^{t^\beta}(M_t)\big] = \mathrm{O}\big(t^{1+\beta(2-\alpha)}\big)$$

as $t \downarrow 0$.

Remark 4.9. The estimate in Proposition 4.8 gives actually a bound for the Markov generator of M on the smooth test function f^v.

Proof of Proposition 4.8. Let P^{M_t} denote the distribution of M_t. We apply Plancherel's identity to obtain

$$E\big[f^{t^\beta}(M_t)\big] = \int_{\mathbb{R}} f^{t^\beta}(x) P^{M_t}(\mathrm{d}x) = (2\pi)^{-1} \int_{\mathbb{R}} \mathfrak{F} f^{t^\beta}(u)\overline{\phi_t(u)}\,\mathrm{d}u,$$

where $\mathfrak{F} f = \int_{\mathbb{R}} \mathrm{e}^{\mathrm{i}ux} f(x)\,\mathrm{d}x$ denotes the Fourier transform of f and the characteristic function of M satisfies

$$\phi_t(u) = \exp\bigg(t\int_{-1}^{1} \big(\mathrm{e}^{\mathrm{i}ux} - 1 - \mathrm{i}ux\big)\mu(\mathrm{d}x)\bigg).$$

Let us rewrite ϕ_t as the linearization of the exponential at zero plus a remainder R:

$$\phi_t(u) = 1 + \psi_t(u) + R(t,u)$$

with

$$\psi_t(u) = t\int_{-1}^{1} \big(\mathrm{e}^{\mathrm{i}ux} - 1 - \mathrm{i}ux\big)\mu(\mathrm{d}x).$$

Then,

$$\begin{aligned} E\big[f^{t^\beta}(M_t)\big] &= (2\pi)^{-1} \int_{\mathbb{R}} \mathfrak{F} f^{t^\beta}(u)\big(1 + \overline{\psi_t(u)} + \overline{R(t,u)}\big)\,\mathrm{d}u \\ &= (2\pi)^{-1} \int_{\mathbb{R}} \mathfrak{F} f^{t^\beta}(u)\overline{\psi_t(u)}\,\mathrm{d}u + (2\pi)^{-1} \int_{\mathbb{R}} \mathfrak{F} f^{t^\beta}(u)\overline{R(t,u)}\,\mathrm{d}u. \end{aligned} \tag{20}$$

For the first term on the right-hand side, we obtain

$$\begin{aligned} (2\pi)^{-1} \int_{\mathbb{R}} \mathfrak{F} f^{t^\beta}(u)\overline{\psi_t(u)}\,\mathrm{d}u &= (2\pi)^{-1} t \int_{-1}^{1}\int_{\mathbb{R}} \mathfrak{F} f^{t^\beta}(u)\big(\mathrm{e}^{-\mathrm{i}ux} - 1 + \mathrm{i}ux\big)\,\mathrm{d}u\,\mu(\mathrm{d}x) \\ &= t\int_{-1}^{1}\bigg(f^{t^\beta}(x) + (2\pi)^{-1}\int_{\mathbb{R}} \mathfrak{F}\big(\big(f^{t^\beta}\big)'\big)(u)x\,\mathrm{d}u\bigg)\mu(\mathrm{d}x) \\ &= t\int_{-1}^{1} f^{t^\beta}(x)\mu(\mathrm{d}x) = t\mathrm{O}\big(t^{\beta(2-\alpha)}\big) \end{aligned} \tag{21}$$

by Assumption 4.1(i) and since

$$\int_{\mathbb{R}} \mathfrak{F}\big(\big(f^{t^\beta}\big)'\big)(u)\,\mathrm{d}u = \big(f^{t^\beta}\big)'(0) = 0.$$

It remains to bound the second addend in (20). For $\mathrm{Re}(z) \le 0$ observe that

$$\left|\frac{\mathrm{e}^z - z - 1}{z^2}\right| \le C \tag{22}$$

for constant $C > 0$, since for $|z| \geq 1$

$$\frac{|e^z - z - 1|}{z^2} \leq 2 + \frac{|z|}{z^2} \leq 3.$$

Whereas on the half disk $\{|z| < 1, \mathrm{Re}(z) \leq 0\}$ the continuous function $|e^z - z - 1|$ is bounded and z^2 is bounded except for the singularity at the origin, but at zero we know that $|e^z - z - 1| = \mathrm{O}(z^2)$, that is,

$$\frac{|e^z - z - 1|}{z^2} \leq C < \infty$$

on $\{|z| < 1, \mathrm{Re}(z) \leq 0\}$. Theorem 1.2.5 in [14] implies that $|\psi_t(u)| \leq Ct|u|^\alpha$ such that

$$\left|R(t,u)\right| = \left|e^{\psi_t(u)} - \psi_t(u) - 1\right| \leq \left|\psi_t(u)\right|^2 \leq Ct^2|u|^{2\alpha},$$

where we used (22) and that for every characteristic function $|\exp(\psi_t(u))| = \phi_t(u) \leq 1$ holds. Hence, we obtain

$$\left|\int_{\mathbb{R}} \mathfrak{F}\big(f^{t^\beta}\big)(u)\overline{R(t,u)}\,\mathrm{d}u\right| \leq Ct^2 \int_{\mathbb{R}} \left|\mathfrak{F}\big(f^{t^\beta}\big)(u)\right||u|^{2\alpha}\,\mathrm{d}u. \tag{23}$$

Therefore, it remains to bound $\int_{\mathbb{R}} |\mathfrak{F} f^{t^\beta}(u)||u|^{2\alpha}\,\mathrm{d}u$ in t. From (19) and the scaling property of the Fourier transform it follows that

$$\mathfrak{F}\big(f^v\big)(u) = v^3\mathfrak{F}\big(v^{-1}f(x/v)\big)(u) = v^3\mathfrak{F}(f)(vu).$$

Since $f \in C^\infty(\mathbb{R})$, we obtain $|\mathfrak{F}(f)(u)| \leq C_m|u|^{-m}$ such that

$$\left|\mathfrak{F}\big(f^v\big)(u)\right| \leq C_m v^{3-m}|u|^{-m}$$

for all $u \in \mathbb{R}$ and $m, v > 0$. Then

$$h(v,u) = \left|\mathfrak{F}\big(f^v\big)(u)\right||u|^{2\alpha} \leq C_m v^{3-m}|u|^{2\alpha-m}.$$

If

$$2\alpha + 1 < m \tag{24}$$

holds then $h(v,\cdot) \in L^1(\mathbb{R})$ for all $v \in (0,1)$. Setting $v = t^\beta$ yields

$$t^2\int_{\mathbb{R}} \left|\mathfrak{F}\big(f^{t^\beta}\big)(u)\right||u|^{2\alpha}\,\mathrm{d}u \leq C_m t^{(3-m)\beta+2}$$

for all $m > 0$. Since the first term in (20) is of the order $\mathrm{O}(t^{1+\beta(2-\alpha)})$, we choose m such that

$$(3-m)\beta + 2 \geq 1 + \beta(2-\alpha) \quad \Leftrightarrow \quad m \leq 1 + \beta^{-1} + \alpha.$$

Together with (24) this leads to the condition

$$2\alpha + 1 < 1 + \beta^{-1} + \alpha \quad \Leftrightarrow \quad \alpha < \beta^{-1},$$

which due to $\alpha \in (0,2)$ always holds for $\beta \in (0, 1/2)$. Hence, we obtain

$$\left| \int_{\mathbb{R}} \mathfrak{F}\big(f^{t^\beta}\big)(u)\overline{R(t,u)}\,\mathrm{d}u \right| = \mathrm{O}\big(t^{1+\beta(2-\alpha)}\big).$$

Together with (20) and (21) this yields finally

$$E\big[f^{t^\beta}(M_t)\big] = t\mathrm{O}\big(t^{\beta(2-\alpha)}\big). \qquad \square$$

4.2.2. *Approximating the continuous martingale part*

The main step is to show that the continuous martingale part can be approximated by summing only the increments that are below the threshold v_n. We will use throughout the notation from (17).

Lemma 4.10. *Suppose that the assumptions of Theorem 4.6 hold, then*

$$T_n^{-1/2} \sum_{i=0}^{n-1} X_{t_i} \big(\Delta_i X \mathbf{1}_{\{|\Delta_i X| \le v_n\}} - \Delta_i X^c\big) \xrightarrow{p} 0 \qquad \text{as } n \to \infty.$$

Proof. Let us consider the following decomposition where $\tilde{X} = X_0 + W + D + U$

$$\begin{aligned}
&T_n^{-1/2} \sum_{i=0}^{n-1} X_{t_i} \big(\Delta_i X \mathbf{1}_{\{|\Delta_i X| \le v_n\}} - \Delta_i X^c\big) \\
&\quad = T_n^{-1/2} \sum_{i=0}^{n-1} X_{t_i} \big(\Delta_i \tilde{X} \mathbf{1}_{\{|\Delta_i X| \le v_n\}} - \Delta_i X^c\big) + T_n^{-1/2} \sum_{i=0}^{n-1} X_{t_i} \Delta_i M \mathbf{1}_{\{|\Delta_i X| \le v_n\}} \\
&\quad = T_n^{-1/2} \sum_{i=0}^{n-1} X_{t_i} \big(\Delta_i \tilde{X} \mathbf{1}_{\{|\Delta_i \tilde{X}| \le 2v_n\}} - \Delta_i X^c\big) \\
&\qquad + T_n^{-1/2} \sum_{i=0}^{n-1} X_{t_i} \Delta_i \tilde{X} (\mathbf{1}_{\{|\Delta_i X| \le v_n\}} - \mathbf{1}_{\{|\Delta_i \tilde{X}| \le 2v_n\}}) + T_n^{-1/2} \sum_{i=0}^{n-1} X_{t_i} \Delta_i M \mathbf{1}_{\{|\Delta_i X| \le v_n\}} \\
&\quad = S_1^n + S_2^n + S_3^n.
\end{aligned}$$

Observe that the term S_1^n already appeared in Lemma 3.9 and $\tilde{X} = X_t - M_t$ is a process with finite jump activity. A careful analysis of the proof of Lemma 3.9 reveals that the same estimates apply to S_1^n such that we conclude that S_1^n converges to zero in probability when $n \to \infty$. Let us prove next convergence of

$$S_2^n = T_n^{-1/2} \sum_{i=0}^{n-1} X_{t_i} \Delta_i \tilde{X} \left(-\mathbf{1}_{\{|\Delta_i X| > v_n, |\Delta_i \tilde{X}| \le 2v_n\}} + \mathbf{1}_{\{|\Delta_i X| \le v_n, |\Delta_i \tilde{X}| > 2v_n\}}\right).$$

Let us prove next that the contribution of the second indicator function on the right-hand side tends to zero in probability:

$$\begin{aligned} & P\left(T_n^{-1/2} \sum_{i=0}^{n-1} |X_{t_i} \Delta_i \tilde{X}| \mathbf{1}_{\{|\Delta_i X| \le v_n, |\Delta_i \tilde{X}| > 2v_n\}} > 0\right) \\ & \quad = P\left(\bigcup_{i=0}^{n-1} \{|\Delta_i X| \le v_n, |\Delta_i \tilde{X}| > 2v_n\}\right) \\ & \quad \le \sum_{i=0}^{n-1} P\big(|\Delta_i X| \le v_n, |\Delta_i \tilde{X}| > 2v_n\big). \end{aligned} \tag{25}$$

When $|\Delta_i \tilde{X}| > 2v_n$ then with high probability $|\Delta_i U| > 0$, since by Lemma 3.7 we obtain

$$\begin{aligned} \sum_{i=0}^{n-1} P\big(|\Delta_i \tilde{X}| > 2v_n, |\Delta_i U| = 0\big) & \le \sum_{i=0}^{n-1} P\big(|\Delta_i W + \Delta_i D| > 2v_n\big) \\ & = \mathrm{O}\big(T_n \Delta_n^{1-2\beta}\big). \end{aligned} \tag{26}$$

This together with (25) and that fact that on $\{|\Delta_i X| \le v_n, |\Delta_i \tilde{X}| > 2v_n\}$ necessarily $|\Delta_i M| > v_n$ implies that

$$\begin{aligned} & P\left(T_n^{-1/2} \sum_{i=0}^{n-1} |X_{t_i} \Delta_i \tilde{X}| \mathbf{1}_{\{|\Delta_i X| \le v_n, |\Delta_i \tilde{X}| > 2v_n\}} > 0\right) \\ & \quad \le \sum_{i=0}^{n-1} P\big(|\Delta_i U| \ne 0\big) P\big(|\Delta_i M| > v_n\big) + \mathrm{O}\big(T_n \Delta_n^{1-2\beta}\big) \\ & \quad = \mathrm{O}\big(T_n \Delta_n v_n^{-2}\big) + \mathrm{O}\big(T_n \Delta_n^{1-2\beta}\big) = \mathrm{O}\big(T_n \Delta_n^{1-2\beta}\big), \end{aligned}$$

where we used Markov's inequality and independence of U and M. The remaining term in S_2^n is

$$T_n^{-1/2} \sum_{i=0}^{n-1} X_{t_i} \Delta_i \tilde{X} \mathbf{1}_{\{|\Delta_i X| > v_n, |\Delta_i \tilde{X}| \le 2v_n\}}.$$

Let us prove that on $\{|\Delta_i\tilde{X}| \le 2v_n\}$ the contribution of U is negligible:

$$
\begin{aligned}
&T_n^{-1/2}\sum_{i=0}^{n-1} P\big(|\Delta_i\tilde{X}| \le 2v_n, |\Delta_i U| > 0\big) \\
&\quad = T_n^{-1/2}\sum_{i=0}^{n-1}\big(P\big(|\Delta_i\tilde{X}| \le 2v_n, \Delta_i N = 1\big) + \mathrm{O}\big(\Delta_n^2\big)\big) \qquad (27) \\
&\quad \le T_n^{-1/2}\sum_{i=0}^{n-1}\big(P\big(|\Delta_i W + \Delta_i D| > 1 - 2v_n\big) + \mathrm{O}\big(\Delta_n^2\big)\big) = \mathrm{O}\big(T_n^{1/2}\Delta_n\big)
\end{aligned}
$$

as $n \to \infty$, where N denotes the counting process that counts the jumps of U and the last step follows from Lemma 3.7. Hence, we can assume that $\Delta_i U = 0$ on $\{|\Delta_i\tilde{X}| \le 2v_n\}$ and so $\Delta_i\tilde{X} = \Delta_i W + \Delta_i D$. For $T_n\Delta_n^{1/2-\beta/2} = \mathrm{o}(1)$ it follows from Lemma 4.11 that as $n \to \infty$

$$
T_n^{-1/2}\sum_{i=0}^{n-1} X_{t_i}\Delta_i D\mathbf{1}_{\{|\Delta_i X| > v_n, |\Delta_i\tilde{X}| \le 2v_n\}} \xrightarrow{p} 0.
$$

We have decomposed S_2^n into a term that converges to 0 in probability and a remainder:

$$
S_2^n = T_n^{-1/2}\sum_{i=0}^{n-1} X_{t_i}\Delta_i W\mathbf{1}_{\{|\Delta_i X| > v_n, |\Delta_i\tilde{X}| \le 2v_n\}} + \mathrm{o}_P(1).
$$

For the remainder let us observe that by Lemma 4.12, we obtain

$$
\begin{aligned}
S_2^n &= T_n^{-1/2}\sum_{i=0}^{n-1} X_{t_i}\Delta_i W\mathbf{1}_{\{|\Delta_i X| > v_n, |\Delta_i\tilde{X}| \le 2v_n\}} + \mathrm{o}_P(1) \\
&= T_n^{-1/2}\sum_{i=0}^{n-1} X_{t_i}\Delta_i W\mathbf{1}_{\{|\Delta_i W + \Delta_i D + \Delta_i M| > v_n, |\Delta_i W + \Delta_i D| \le 2v_n\}} + \mathrm{o}_P(1) \\
&= T_n^{-1/2}\sum_{i=0}^{n-1} X_{t_i}\Delta_i W\mathbf{1}_{\{|\Delta_i M| > v_n\}} + \mathrm{o}_P(1)
\end{aligned}
$$

Markov's inequality yields $P(|\Delta_i M| > v_n) \le \Delta_n^{1/2-\beta}$. Independence of X_{t_i}, $\Delta_i W$ and $\Delta_i M$ leads to

$$
E\left[T_n^{-1/2}\sum_{i=0}^{n-1} X_{t_i}\Delta_i W\mathbf{1}_{\{|\Delta_i M| > v_n\}}\right] = 0
$$

and

$$E\left[\left(T_n^{-1/2}\sum_{i=0}^{n-1} X_{t_i}\Delta_i W\mathbf{1}_{\{|\Delta_i M|>v_n\}}\right)^2\right]$$
$$\le T_n^{-1}E\left[\sum_{i=0}^{n-1} X_{t_i}^2(\Delta_i W)^2\mathbf{1}_{\{|\Delta_i M|>v_n\}}\right]$$
$$+ T_n^{-1}E\left[\sum_{i\neq j} X_{t_i}\Delta_i W\mathbf{1}_{\{|\Delta_i M|>v_n\}}X_{t_j}\Delta_j W\mathbf{1}_{\{|\Delta_j M|>v_n\}}\right].$$

Since $X_{t_i}, X_{t_j}, \Delta_i W, \Delta_i M, \Delta_j M \perp \Delta_j W$ for $i<j$ the off-diagonal elements are centered,

$$E\left[\sum_{i\neq j} X_{t_i}\Delta_i W\mathbf{1}_{\{|\Delta_i M|>v_n\}}X_{t_j}\Delta_j W\mathbf{1}_{\{|\Delta_j M|>v_n\}}\right]=0$$

and the diagonal elements can be estimated by

$$T_n^{-1}E\left[\sum_{i=0}^{n-1} X_{t_i}^2(\Delta_i W)^2\mathbf{1}_{\{|\Delta_i M|>v_n\}}\right] \le T_n^{-1}\Delta_n\sum_{i=0}^{n-1}E\big[X_{t_i}^2\big]P\big(|\Delta_i M|>v_n\big)$$
$$\le \sup_i\big\{E\big[X_{t_i}^2\big]\big\}\Delta_n^{1/2-\beta}\to 0$$

as $n\to\infty$. The last step is to show that S_3^n tends to zero in probability as $n\to\infty$. As in (27) it follows that on $|\Delta_i X|\le v_n$ we can assume that $\Delta_i U=0$. Now

$$\sum_{i=0}^{n-1}P\big(|\Delta_i X|\le v_n, \Delta_i U=0\big)\le\sum_{i=0}^{n-1}P\big(|\Delta_i W+\Delta_i D+\Delta_i M|\le v_n, |\Delta_i M|\le 2v_n\big)$$
$$+\sum_{i=0}^{n-1}P\big(|\Delta_i W+\Delta_i D+\Delta_i M|\le v_n, |\Delta_i M|> 2v_n\big).$$

The second addend vanishes, since by Lemma 3.7 we obtain

$$\sum_{i=0}^{n-1}P\big(|\Delta_i W+\Delta_i D+\Delta_i M|\le v_n, |\Delta_i M|> 2v_n\big)$$
$$\le\sum_{i=0}^{n-1}P\big(|\Delta_i W+\Delta_i D|>v_n\big)=\mathrm{O}\big(T_n\Delta_n^{1-2\beta}\big).$$

Thus, S_3^n can be rewritten as

$$S_3^n = T_n^{-1/2} \sum_{i=0}^{n-1} X_{t_i} \Delta_i M \mathbf{1}_{\{|\Delta_i X| \le v_n, |\Delta_i M| \le 2v_n\}} + o_p(1). \tag{28}$$

The convergence of the remaining term in S_3^n is dominated by the behavior of $\Delta_i M$ around the threshold, that is, we prove next that

$$\begin{aligned} &T_n^{-1/2} \sum_{i=0}^{n-1} X_{t_i} \Delta_i M \mathbf{1}_{\{|\Delta_i X| \le v_n, |\Delta_i M| \le 2v_n\}} \\ &\quad = T_n^{-1/2} \sum_{i=0}^{n-1} X_{t_i} \Delta_i M \mathbf{1}_{\{|\Delta_i M| \le 2v_n\}} + o_p(1). \end{aligned}$$

Indeed,

$$\begin{aligned} &T_n^{-1/2} \sum_{i=0}^{n-1} X_{t_i} \Delta_i M (\mathbf{1}_{\{|\Delta_i M| \le 2v_n\}} - \mathbf{1}_{\{|\Delta_i X| \le v_n, |\Delta_i M| \le 2v_n\}}) \\ &\quad = T_n^{-1/2} \sum_{i=0}^{n-1} X_{t_i} \Delta_i M \mathbf{1}_{\{|\Delta_i X| > v_n, |\Delta_i M| \le 2v_n\}}. \end{aligned}$$

That last term tends to zero in probability will be shown in the proof of Lemma 4.12 below following equation (35). To finish the proof, we demonstrate that the first addend on the right-hand side of (28) vanishes asymptotically. Since $X_{t_i}, X_{t_j}, \Delta_i M \perp \Delta_j M$ for $i < j$ the off-diagonal elements vanish by Assumption 4.1(i) such that

$$E\left[\sum_{i \ne j} X_{t_i} \Delta_i M \mathbf{1}_{\{|\Delta_i M| \le 2v_n\}} X_{t_j} \Delta_j M \mathbf{1}_{\{|\Delta_j M| \le 2v_n\}}\right] = 0$$

and the diagonal elements can by Proposition 4.8 be estimated by

$$\begin{aligned} T_n^{-1} E\left[\sum_{i=0}^{n-1} X_{t_i}^2 (\Delta_i M)^2 \mathbf{1}_{\{|\Delta_i M| \le 2v_n\}}\right] &\le T_n^{-1} \sum_{i=0}^{n-1} E[X_{t_i}^2] E[(\Delta_i M)^2 \mathbf{1}_{\{|\Delta_i M| \le 2v_n\}}] \\ &\le \sup_i \{E[X_{t_i}^2]\} v_n^{\alpha} \to 0 \end{aligned}$$

as $n \to \infty$. □

4.2.3. *Approximation of the drift*

The next step is to show that the drift component of X is in the limit not affected by the cut-off.

Lemma 4.11. *If the assumptions of Theorem 4.6 are fulfilled then*

$$T_n^{-1/2}\sum_{i=0}^{n-1} X_{t_i}(\Delta_i D\mathbf{1}_{\{|\Delta_i X|\le v_n\}} - \Delta_i D) \xrightarrow{p} 0 \qquad \text{as } n\to\infty.$$

Proof. We rewrite the sum as follows:

$$T_n^{-1/2}\sum_{i=0}^{n-1} X_{t_i}(\Delta_i D\mathbf{1}_{\{|\Delta_i X|\le v_n\}} - \Delta_i D) = T_n^{-1/2}\sum_{i=0}^{n-1} X_{t_i}\Delta_i D\mathbf{1}_{\{|\Delta_i X|> v_n\}}.$$

Next, we decompose $\Delta_i D$ as follows

$$\Delta_i D = -a\left(\int_{t_i}^{t_{i+1}} (X_s - X_{t_i})\,\mathrm{d}s + \Delta_i X_{t_i}\right)$$

such that by Lemma 4.12 below

$$\begin{aligned} T_n^{-1/2}\sum_{i=0}^{n-1} X_{t_i}\Delta_i D\mathbf{1}_{\{|\Delta_i X|> v_n\}} &= -aT_n^{-1/2}\sum_{i=0}^{n-1} X_{t_i}\int_{t_i}^{t_{i+1}} (X_s - X_{t_i})\,\mathrm{d}s\mathbf{1}_{\{|\Delta_i J|> v_n\}} \\ &\quad - aT_n^{-1/2}\sum_{i=0}^{n-1} X_{t_i}^2\Delta_i\mathbf{1}_{\{|\Delta_i J|> v_n\}} + \mathrm{o}_p(1). \end{aligned} \tag{29}$$

For the second term, we obtain by Markov's inequality and from $v_n = \Delta_n^{\beta}$ that

$$E\left[\left|\sum_{i=0}^{n-1} X_{t_i}^2\Delta_i\mathbf{1}_{\{|\Delta_i J|> v_n\}}\right|\right] \le \sum_{i=0}^{n-1}\Delta_i E\big[X_{t_i}^2\big]P\big(|\Delta_i J| > v_n\big) \le CT_n\Delta_n^{1-2\beta}$$

and so for $T_n^{1/2}\Delta_n^{1-2\beta} = \mathrm{o}(1)$ it follows that

$$T_n^{-1/2}\sum_{i=0}^{n-1} X_{t_i}\Delta_i\mathbf{1}_{\{|\Delta_i J|> v_n\}} = \mathrm{o}_p(1).$$

For the first sum on the right-hand side of (29), we obtain by Hölder's inequality and independence of X_{t_i} and $\Delta_i J$

$$\begin{aligned} &E\left[\left|X_{t_i}\int_{t_i}^{t_{i+1}} (X_s - X_{t_i})\,\mathrm{d}s\mathbf{1}_{\{|\Delta_i J|> v_n\}}\right|\right] \\ &\le E\left[\left(\int_{t_i}^{t_{i+1}} (X_s - X_{t_i})\,\mathrm{d}s\right)^2\right]^{1/2} P\big(|\Delta_i J| > v_n\big)^{1/2} E\big[X_{t_i}^2\big]^{1/2} \\ &= \mathrm{O}\big(\Delta_n^{3/2} v_n^{-1/2}\big) \end{aligned}$$

such that for $T_n^{1/2}\Delta_n^{1/2-\beta/2} = \mathrm{o}(1)$ we can conclude that

$$T_n^{-1/2}\sum_{i=0}^{n-1} X_{t_i}\int_{t_i}^{t_{i+1}} (X_s - X_{t_i})\,\mathrm{d}s\mathbf{1}_{\{|\Delta_i J|>v_n\}} = \mathrm{o}_p(1).$$

□

4.2.4. *Identifying the jumps*

In the following, we will show that the increments of X that are larger than the threshold v_n are dominated by the jump component.

Lemma 4.12.

$$T_n^{-1/2}\sum_{i=0}^{n-1} X_{t_i}\Delta_i X(\mathbf{1}_{\{|\Delta_i X|\le v_n\}} - \mathbf{1}_{\{|\Delta_i M|\le 2v_n\}}) \xrightarrow{p} 0 \qquad \text{as } n\to\infty.$$

Proof. Observe that

$$\begin{aligned} &T_n^{-1/2}\sum_{i=0}^{n-1} X_{t_i}\Delta_i X(\mathbf{1}_{\{|\Delta_i X|\le v_n\}} - \mathbf{1}_{\{|\Delta_i J|\le 2v_n\}}) &(30)\\ &\quad = T_n^{-1/2}\sum_{i=0}^{n-1} X_{t_i}\Delta_i X\mathbf{1}_{\{|\Delta_i X|\le v_n, |\Delta_i J|>2v_n\}} &(31)\\ &\qquad - T_n^{-1/2}\sum_{i=0}^{n-1} X_{t_i}\Delta_i X\mathbf{1}_{\{|\Delta_i X|>v_n, |\Delta_i J|\le 2v_n\}}. \end{aligned}$$

We shall prove in Lemma 4.13 below that

$$T_n^{-1/2}\sum_{i=0}^{n-1} X_{t_i}\Delta_i X\mathbf{1}_{\{|\Delta_i X|\le v_n, |\Delta_i J|>2v_n\}} \xrightarrow{p} 0. \tag{32}$$

In the next step, we show that the contribution of U is negligible, since by independence of $\Delta_i W$, $\Delta_i M$, $\Delta_i U$ and X_{t_i} it follows that

$$\begin{aligned} E\Biggl[\Biggl|\sum_{i=0}^{n-1} X_{t_i}\Delta_i X\mathbf{1}_{\{|\Delta_i J|\le 2v_n, |\Delta_i U|\ne 0\}}\Biggr|\Biggr] &\le \sum_{i=0}^{n-1} E\bigl[|X_{t_i}|\bigr]E\bigl[|\Delta_i W|\bigr]P\bigl(|\Delta_i U|\ne 0\bigr)\\ &\quad + \sum_{i=0}^{n-1} E\bigl[|X_{t_i}|\bigr]E\bigl[|\Delta_i J|\mathbf{1}_{\{|\Delta_i J|\le 2v_n, |\Delta_i U|\ne 0\}}\bigr] &(33)\\ &\quad + \sum_{i=0}^{n-1} E\bigl[|X_{t_i}\Delta_i D|\mathbf{1}_{\{|\Delta_i U|\ne 0\}}\bigr]. \end{aligned}$$

Now U is a compound Poisson process with intensity $\mu(\mathbb{R} \setminus [-1,1]) < \infty$ such that $P(\Delta_i U \neq 0) = \mathrm{O}(\Delta_n)$. We obtain for first addend on the right-hand side

$$T_n^{-1/2} \sum_{i=0}^{n-1} E\big[|\Delta_i W|\big] E\big[|X_{t_i}|\big] P(\Delta_i U \neq 0) = \mathrm{O}\big(T_n^{1/2} \Delta_n^{1/2}\big).$$

We split the second term into the contribution by U and M such that

$$\begin{aligned} T_n^{-1/2} &\sum_{i=0}^{n-1} E\big[|X_{t_i}|\big] E\big[|\Delta_i J| \mathbf{1}_{\{|\Delta_i J| \le 2v_n, \Delta_i U \neq 0\}}\big] \\ &= T_n^{-1/2} \sum_{i=0}^{n-1} E\big[|X_{t_i}|\big] E\big[|\Delta_i M|\big] E[\mathbf{1}_{\{|\Delta_i J| \le 2v_n, \Delta_i U \neq 0\}}] \\ &\quad + T_n^{-1/2} \sum_{i=0}^{n-1} E\big[|X_{t_i}|\big] E\big[|\Delta_i U| \mathbf{1}_{\{|\Delta_i J| \le 2v_n, \Delta_i U \neq 0\}}\big]. \end{aligned}$$

The first sum is of order

$$T_n^{-1/2} \sum_{i=0}^{n-1} E\big[|X_{t_i}|\big] E\big[|\Delta_i M|\big] E[\mathbf{1}_{\{\Delta_i U \neq 0\}}] = \mathrm{O}\big(T_n^{1/2} \Delta_n^{1/2}\big).$$

Höldern's inequality and independence of M and U lead to the following estimate for the second sum:

$$T_n^{-1/2} E \sum_{i=0}^{n-1} |X_{t_i} \Delta_i U \mathbf{1}_{\{|\Delta_i J| \le 2v_n, \Delta_i U \neq 0\}}| = \mathrm{O}\big(T_n^{1/2} \Delta_n^{1/2}\big).$$

To prove convergence of the last addend in (33), we rewrite $\Delta_i D$ as follows

$$\Delta_i D = -a \left(\int_{t_i}^{t_{i+1}} (X_s - X_{t_i}) \, \mathrm{d}s + \Delta_i X_{t_i} \right) \tag{34}$$

and so

$$\begin{aligned} T_n^{-1/2} E\left[\left| \sum_{i=0}^{n-1} X_{t_i} \Delta_i D \mathbf{1}_{\{\Delta_i U \neq 0\}} \right|\right] &\le T_n^{-1/2} a E\left[\left| \sum_{i=0}^{n-1} X_{t_i} \int_{t_i}^{t_{i+1}} (X_s - X_{t_i}) \, \mathrm{d}s \mathbf{1}_{\{\Delta_i U \neq 0\}} \right|\right] \\ &\quad + T_n^{-1/2} E\left[\left| \sum_{i=0}^{n-1} X_{t_i}^2 \Delta_n \mathbf{1}_{\{\Delta_i U \neq 0\}} \right|\right]. \end{aligned}$$

The first term on the right-hand side gives by using Hölder's inequality

$$
\begin{aligned}
&T_n^{-1/2} a E\left[\left|\sum_{i=0}^{n-1} X_{t_i} \int_{t_i}^{t_{i+1}} (X_s - X_{t_i})\,\mathrm{d}s \mathbf{1}_{\{\Delta_i U \neq 0\}}\right|\right] \\
&\quad \le T_n^{-1/2} a \sum_{i=0}^{n-1} E\big[X_{t_i}^2\big]^{1/2} E\left[\left(\int_{t_i}^{t_{i+1}} (X_s - X_{t_i})\,\mathrm{d}s\right)^2\right]^{1/2} \mathbf{1}_{\{\Delta_i U \neq 0\}} \\
&\quad = \mathrm{O}\big(T_n^{1/2}\Delta_n\big).
\end{aligned}
$$

Hence, we obtain

$$
T_n^{-1/2} \sum_{i=0}^{n-1} X_{t_i} \Delta_i D \mathbf{1}_{\{\Delta_i U \neq 0\}} = \mathrm{O}_p\big(T_n^{1/2}\Delta_n^{1/2}\big)
$$

such that it follows that

$$
\sum_{i=0}^{n-1} X_{t_i} \Delta_i X \mathbf{1}_{\{|\Delta_i J| \le 2v_n, |\Delta_i U| \neq 0\}} = \mathrm{o}_p(1).
$$

Since the contribution of U is negligible, we obtain from (30) and (32) that

$$
\begin{aligned}
&T_n^{-1/2} \sum_{i=0}^{n-1} X_{t_i} \Delta_i X (\mathbf{1}_{\{|\Delta_i X| \le v_n\}} - \mathbf{1}_{\{|\Delta_i M| \le v_n\}}) \\
&\quad = T_n^{-1/2} \sum_{i=0}^{n-1} X_{t_i} \Delta_i X \mathbf{1}_{\{|\Delta_i X| > v_n, |\Delta_i J| \le 2v_n, |\Delta_i U| = 0\}} + \mathrm{o}_p(1).
\end{aligned}
$$

Hence, it remains to prove

$$
T_n^{-1/2} \sum_{i=0}^{n-1} X_{t_i} \Delta_i X \mathbf{1}_{\{\Delta_i X| > v_n |\Delta_i M| \le 2v_n, |\Delta_i U| = 0\}} \xrightarrow{p} 0 \qquad \text{as } n \to \infty. \tag{35}
$$

Observe that

$$
\begin{aligned}
&\big\{|\Delta_i M| \le 2v_n, \Delta_i U = 0, |\Delta_i X| > v_n\big\} \\
&\quad \subset \big\{|\Delta_i M| \le 2v_n, \Delta_i U = 0, |\Delta_i W + \Delta_i D| + |\Delta_i M| > v_n\big\} \\
&\quad \subset \big\{|\Delta_i W + \Delta_i D| > v_n/2\big\} \cup \big\{|\Delta_i M| \le 2v_n, |\Delta_i M| > v_n/2\big\}.
\end{aligned}
$$

Therefore, the last two steps will be to show that:

(i) $T_n^{-1/2} \sum_{i=0}^{n-1} X_{t_i} \Delta_i X \mathbf{1}_{\{|\Delta_i W + \Delta_i D| > v_n/2\}} \xrightarrow{p} 0$,
(ii) $T_n^{-1/2} \sum_{i=0}^{n-1} X_{t_i} \Delta_i X \mathbf{1}_{\{|\Delta_i M| > v_n/2, |\Delta_i M| \le 2v_n\}} \xrightarrow{p} 0$.

For the proof of these two convergences, we refer to Lemmas 4.15 and 4.14. □

Lemma 4.13.

$$T_n^{-1/2}\sum_{i=0}^{n-1} X_{t_i}\,\Delta_i X\mathbf{1}_{\{|\Delta_i X|\le v_n,|\Delta_i J|>2v_n\}} \xrightarrow{p} 0 \qquad \textit{as } n\to\infty.$$

Proof. On $\{|\Delta_i X|\le v_n, |\Delta_i J|\ge 2v_n\}$ we have

$$\big||\Delta_i W+\Delta_i D|-|\Delta_i J|\big|\le |\Delta_i X|\le v_n.$$

Hence, we necessarily have $|\Delta_i W+\Delta_i D|>v_n$, that is,

$$\big\{|\Delta_i X|\le v_n, |\Delta_i J|>2v_n\big\}\subset\big\{|\Delta_i W+\Delta_i D|>v_n\big\} \tag{36}$$

such that

$$P\big(|\Delta_i X|\le v_n, |\Delta_i J|>2v_n\big)\le P\big(|\Delta_i W+\Delta_i D|>v_n\big)=\mathrm{O}\big(\Delta_n^{2-\beta}\big). \tag{37}$$

It follows from (36) that

$$\begin{aligned}
&T_n^{-1/2}\left|\sum_{i=0}^{n-1} X_{t_i}\,\Delta_i X\mathbf{1}_{\{|\Delta_i X|\le v_n,|\Delta_i J|>2v_n\}}\right|\\
&\quad\le T_n^{-1/2}\sum_{i=0}^{n-1}|X_{t_i}\,\Delta_i X|\mathbf{1}_{\{|\Delta_i W+\Delta_i D|>v_n\}}\\
&\quad\le T_n^{-1/2}\sum_{i=0}^{n-1}|X_{t_i}\,\Delta_i W|\mathbf{1}_{\{|\Delta_i W+\Delta_i D|>v_n\}}+T_n^{-1/2}\sum_{i=0}^{n-1}|X_{t_i}\,\Delta_i D|\mathbf{1}_{\{|\Delta_i W+\Delta_i D|>v_n\}}\\
&\qquad+T_n^{-1/2}\sum_{i=0}^{n-1}|X_{t_i}\,\Delta_i M|\mathbf{1}_{\{|\Delta_i W+\Delta_i D|>v_n\}}+T_n^{-1/2}\sum_{i=0}^{n-1}|X_{t_i}\,\Delta_i U|\mathbf{1}_{\{|\Delta_i W+\Delta_i D|>v_n\}}\\
&\quad=A_n^1+\cdots+A_n^4.
\end{aligned}$$

For A_n^1 we find by (37), Hölder's inequality and independence of X_{t_i} and $\Delta_i W$ that

$$\begin{aligned}
E\big[|A_n^1|\big]&\le T_n^{-1/2}\Delta_n^{1/2}\sum_{i=0}^{n-1}E\big[X_{t_i}^2\big]^{1/2}P\big(|\Delta_i W+\Delta_i D|>v_n\big)^{1/2}\\
&=\mathrm{O}\big(T_n^{1/2}\Delta_n^{1/2-\beta/2}\big).
\end{aligned}$$

Using (34), we obtain for A_n^2 that

$$E\big[\big|A_n^2\big|\big] \le T_n^{-1/2} a \sum_{i=0}^{n-1} E\bigg[\bigg|X_{t_i}\int_{t_i}^{t_{i+1}}(X_s - X_{t_i})\,\mathrm{d}s\bigg|\mathbf{1}_{\{|\Delta_i W + \Delta_i D| > v_n\}}\bigg]$$

$$+ T_n^{-1/2}\Delta_n a \sum_{i=0}^{n-1} E\big[X_{t_i}^2 \mathbf{1}_{\{|\Delta_i W + \Delta_i D| > v_n\}}\big]$$

Hölder's inequality yields for the first term on the right-hand side

$$\begin{aligned}
&T_n^{-1/2} a \sum_{i=0}^{n-1} E\bigg[\bigg|X_{t_i}\int_{t_i}^{t_{i+1}}(X_s - X_{t_i})\,\mathrm{d}s\bigg|\mathbf{1}_{\{|\Delta_i W + \Delta_i D| > v_n\}}\bigg] \\
&\le T_n^{-1/2} a \sum_{i=0}^{n-1} E\bigg[\bigg(X_{t_i}\int_{t_i}^{t_{i+1}}(X_s - X_{t_i})\,\mathrm{d}s\bigg)^2\bigg]^{1/2} P\big(|\Delta_i W + \Delta_i D| > v_n\big)^{1/2} \\
&= \mathrm{O}\big(T_n^{1/2}\Delta_n^{1-\beta/2}\big)
\end{aligned}$$

for the second addend we find that

$$\begin{aligned}
&T_n^{-1/2}\Delta_n a \sum_{i=0}^{n-1} E\big[X_{t_i}^2 \mathbf{1}_{\{|\Delta_i W + \Delta_i D| > v_n\}}\big] \\
&\le T_n^{-1/2}\Delta_n a \sum_{i=0}^{n-1} E\big[X_{t_i}^4\big]^{1/2} P\big(|\Delta_i W + \Delta_i D| > v_n\big)^{1/2} \\
&= \mathrm{O}\big(T_n^{1/2}\Delta_n^{1/2-\beta/2}\big).
\end{aligned}$$

For A_n^3 we get by a similar estimate as for A_n^1 that

$$E\big[\big|A_n^3\big|\big] = \mathrm{O}\big(T_n^{1/2}\Delta_n^{1/2-\beta/2}\big).$$

The last addend A_n^4 converges to zero in probability, since by independence and Hölder's inequality

$$\begin{aligned}
E\big[\big|A_n^4\big|\big] &\le T_n^{-1/2} \sum_{i=0}^{n-1} E\big[X_{t_i}^2\big]^{1/2} E\big[\Delta_i U^2\big]^{1/2} P\big(|\Delta_i W + \Delta_i D| > v_n\big)^{1/2} \\
&= \mathrm{O}\big(T_n^{1/2}\Delta_n^{1/2-\beta/2}\big).
\end{aligned}$$ □

Now we show that the increments of the continuous part of X are negligible in the limit. This convergence is mainly based in the moment bound that we have derived in Lemma 3.7.

Lemma 4.14.

$$T_n^{-1/2}\sum_{i=0}^{n-1} X_{t_i}\,\Delta_i X\mathbf{1}_{\{|\Delta_i W+\Delta_i D|>v_n\}} \xrightarrow{p} 0 \qquad \textit{as } n\to\infty.$$

Proof. We decompose $\Delta_i X = \Delta_i W + \Delta_i D + \Delta_i M + \Delta_i U$ to obtain

$$\begin{aligned}
&T_n^{-1/2}\sum_{i=0}^{n-1} X_{t_i}\,\Delta_i X\mathbf{1}_{\{|\Delta_i W+\Delta_i D|>v_n\}}\\
&\quad= T_n^{-1/2}\sum_{i=0}^{n-1} X_{t_i}\,\Delta_i W\mathbf{1}_{\{|\Delta_i W+\Delta_i D|>v_n\}} + T_n^{-1/2}\sum_{i=0}^{n-1} X_{t_i}\,\Delta_i D\mathbf{1}_{\{|\Delta_i W+\Delta_i D|>v_n\}}\\
&\qquad+ T_n^{-1/2}\sum_{i=0}^{n-1} X_{t_i}\,\Delta_i M\mathbf{1}_{\{|\Delta_i W+\Delta_i D|>v_n\}} + T_n^{-1/2}\sum_{i=0}^{n-1} X_{t_i}\,\Delta_i U\mathbf{1}_{\{|\Delta_i W+\Delta_i D|>v_n\}}\\
&\quad= V_n^1 + V_n^2 + V_n^3 + V_n^4.
\end{aligned}$$

Lemma 3.7 yields for $\delta = 1/2-\beta$ and $l=2$ that

$$P\big(|\Delta_i W + \Delta_i D| > v_n\big) = \mathrm{O}\big(\Delta_n^{2-2\beta}\big).$$

For V_n^1 we obtain by Hölder's inequality and independence of X_{t_i} and $\Delta_i W$ that

$$\begin{aligned}
E\big[\big|V_n^1\big|\big] &= T_n^{-1/2} E\Bigg[\Bigg|\sum_{i=0}^{n-1} X_{t_i}\,\Delta_i W\mathbf{1}_{\{|\Delta_i W+\Delta_i D|>v_n\}}\Bigg|\Bigg]\\
&\le T_n^{-1/2}\Delta_n^{1/2}\sum_{i=0}^{n-1} E\big[X_{t_i}^2\big]^{1/2} P\big(|\Delta_i W + \Delta_i D| > v_n\big)^{1/2}\\
&= \mathrm{O}\big(T_n^{1/2}\Delta_n^{1/2-\beta}\big).
\end{aligned}$$

To prove convergence of V_n^2 we decompose $\Delta_i D$ as in (34) to obtain

$$\begin{aligned}
E\big[\big|V_n^2\big|\big] &= T_n^{-1/2} E\Bigg[\Bigg|\sum_{i=0}^{n-1} X_{t_i}\,\Delta_i D\mathbf{1}_{\{|\Delta_i W+\Delta_i D|>v_n\}}\Bigg|\Bigg]\\
&\le T_n^{-1/2} a E\Bigg[\Bigg|\sum_{i=0}^{n-1} X_{t_i}\int_{t_i}^{t_{i+1}} (X_s - X_{t_i})\,\mathrm{d}s\mathbf{1}_{\{|\Delta_i W+\Delta_i D|>v_n\}}\Bigg|\Bigg]\\
&\quad+ T_n^{-1/2} a E\Bigg[\Bigg|\sum_{i=0}^{n-1} X_{t_i}^2\,\Delta_i\mathbf{1}_{\{|\Delta_i W+\Delta_i D|>v_n\}}\Bigg|\Bigg].
\end{aligned}$$

Applying Hölder's inequality to the first term on the right-hand side results in

$$T_n^{-1/2} a E\left[\left|\sum_{i=0}^{n-1} X_{t_i} \int_{t_i}^{t_{i+1}} (X_s - X_{t_i})\,\mathrm{d}s \mathbf{1}_{\{|\Delta_i W + \Delta_i D| > v_n\}}\right|\right]$$

$$\le T_n^{-1/2} a \sum_{i=0}^{n-1} E\left[\left(X_{t_i} \int_{t_i}^{t_{i+1}} (X_s - X_{t_i})\,\mathrm{d}s\right)^2\right]^{1/2} P\big(|\Delta_i W + \Delta_i D| > v_n\big)^{1/2}$$

$$= \mathrm{O}\big(T_n^{1/2} \Delta_n^{1-\beta}\big).$$

The remaining term is of the order

$$T_n^{-1/2} a E\left[\left|\sum_{i=0}^{n-1} X_{t_i}^2 \Delta_i \mathbf{1}_{\{|\Delta_i W + \Delta_i D| > v_n\}}\right|\right]$$

$$\le T_n^{-1/2} a \Delta_n \sum_{i=0}^{n-1} E\big[X_{t_i}^4\big]^{1/2} P\big(|\Delta_i W + \Delta_i D| > v_n\big)^{1/2} = \mathrm{O}\big(T_n^{1/2} \Delta_n^{1-\beta}\big).$$

Therefore, we conclude that $V_n^2 \to 0$ as $n \to \infty$. Similar estimates as for V_n^1 can be used for V_n^3 and V_n^4 to show

$$E\big[|V_n^3|\big] = \mathrm{O}\big(T_n^{1/2} \Delta_n^{1-\beta}\big) \quad \text{and} \quad E\big[|V_n^4|\big] = \mathrm{O}\big(T_n^{1/2} \Delta_n^{1-\beta}\big).$$

This concludes the proof. □

The next lemma states that the increments of the jump component that are close to the threshold are negligible in the limit. For the proof, we use the small time moment bound for the jump component from Proposition 4.8. This is the step where Assumption 4.1 on the intensity of small jumps becomes crucial.

Lemma 4.15.

$$T_n^{-1/2} \sum_{i=0}^{n-1} X_{t_i} \Delta_i X \mathbf{1}_{\{v_n/2 < |\Delta_i M| \le 2v_n\}} \xrightarrow{p} 0 \qquad \text{as } n \to \infty.$$

Proof. Let us consider the following decomposition

$$T_n^{-1/2} \sum_{i=0}^{n-1} X_{t_i} \Delta_i X \mathbf{1}_{\{|\Delta_i M| > v_n/2, |\Delta_i M| \le 2v_n\}}$$

$$= T_n^{-1/2} \sum_{i=0}^{n-1} X_{t_i} \Delta_i W \mathbf{1}_{\{|\Delta_i M| > v_n/2, |\Delta_i M| \le 2v_n\}}$$

$$
\begin{aligned}
&+ T_n^{-1/2} \sum_{i=0}^{n-1} X_{t_i} \Delta_i D \mathbf{1}_{\{|\Delta_i M| > v_n/2, |\Delta_i M| \le 2v_n\}} \\
&+ T_n^{-1/2} \sum_{i=0}^{n-1} X_{t_i} \Delta_i U \mathbf{1}_{\{|\Delta_i M| > v_n/2, |\Delta_i M| \le 2v_n\}} \\
&+ T_n^{-1/2} \sum_{i=0}^{n-1} X_{t_i} \Delta_i M \mathbf{1}_{\{|\Delta_i M| > v_n/2, |\Delta_i M| \le 2v_n\}} \\
&= S_n^1 + S_n^2 + S_n^3 + S_n^4.
\end{aligned}
$$

For the probability that $|\Delta_i M|$ lies in $(v_n/2, 2v_n)$, we derive from Proposition 4.8 and Markov's inequality that

$$
\begin{aligned}
P\big(|\Delta_i M| \le 2v_n, |\Delta_i M| > v_n/2\big) &= P\big(|\Delta_i M| \mathbf{1}_{\{|\Delta_i M| \le 2v_n\}} > v_n/2\big) \\
&\le 4 v_n^{-2} E\big[(\Delta_i M)^2 \mathbf{1}_{\{|\Delta_i M| \le 2v_n\}}\big] = \mathrm{O}\big(\Delta_n^{1-\alpha\beta}\big).
\end{aligned} \tag{38}
$$

Hence, by independence of X_{t_i}, $\Delta_i W$, and $\Delta_i M$ we find that $E[S_n^1] = 0$ and the second moment can be estimated as follows.

$$
\begin{aligned}
E\big[\big(S_n^1\big)^2\big] &\le T_n^{-1} E\Bigg[\sum_{i=0}^{n-1} X_{t_i}^2 (\Delta_i W)^2 \mathbf{1}_{\{|\Delta_i M| > v_n/2, |\Delta_i M| \le 2v_n\}}\Bigg] \\
&\quad + T_n^{-1} E\Bigg[\sum_{i \ne j} X_{t_i} \Delta_i W \mathbf{1}_{\{|\Delta_i M| > v_n/2, |\Delta_i M| \le 2v_n\}} X_{t_j} \Delta_j W \mathbf{1}_{\{|\Delta_j M| > v_n/2, |\Delta_j M| \le 2v_n\}}\Bigg].
\end{aligned}
$$

Since $X_{t_i}, X_{t_j}, \Delta_i W, \Delta_j M, \Delta_i M \perp \Delta_j W$ for $i < j$, the off-diagonal elements have zero expectation such that the second addend vanishes. For the diagonal elements, we obtain

$$
\begin{aligned}
T_n^{-1} E\Bigg[\sum_{i=0}^{n-1} X_{t_i}^2 (\Delta_i W)^2 \mathbf{1}_{\{|\Delta_i M| > v_n/2, |\Delta_i M| \le 2v_n\}}\Bigg] &\le T_n^{-1} \Delta_n \sum_{i=0}^{n-1} E\big[X_{t_i}^2\big] \mathrm{O}\big(\Delta_n^{1-\alpha\beta}\big) \\
&= \mathrm{O}\big(\Delta_n^{1-\alpha\beta}\big).
\end{aligned}
$$

This yields the convergence $S_n^1 \xrightarrow{p} 0$ as $n \to \infty$. To prove that $S_n^2 \xrightarrow{p} 0$ as $n \to \infty$, we plug in (34) and obtain

$$
\begin{aligned}
E\big[\big|S_n^2\big|\big] &\le E\Bigg[a T_n^{-1/2} \Bigg|\sum_{i=0}^{n-1} X_{t_i} \int_{t_i}^{t_{i+1}} (X_s - X_{t_i}) \,\mathrm{d}s \mathbf{1}_{\{|\Delta_i M| > v_n/2, |\Delta_i M| \le 2v_n\}}\Bigg|\Bigg] \\
&\quad + E\Bigg[a T_n^{-1/2} \Bigg|\sum_{i=0}^{n-1} X_{t_i}^2 \Delta_i \mathbf{1}_{\{|\Delta_i M| > v_n/2, |\Delta_i M| \le 2v_n\}}\Bigg|\Bigg]
\end{aligned}
$$

and by independence

$$
\begin{aligned}
E\Bigg[aT_n^{-1/2}\Bigg|\sum_{i=0}^{n-1} X_{t_i}^2 \Delta_i \mathbf{1}_{\{|\Delta_i M|>v_n/2, |\Delta_i M|\le 2v_n\}}\Bigg|\Bigg]
&\le aT_n^{-1/2}\sum_{i=0}^{n-1} E\big[X_{t_i}^2\big]\Delta_i P\big(v_n/2 < |\Delta_i M| \le 2v_n\big) \\
&= \mathrm{O}\big(T_n^{1/2}\Delta_n^{1-\alpha\beta}\big).
\end{aligned}
$$

For the second term, Hölder's inequality yields

$$
E\Bigg[aT_n^{-1/2}\Bigg|\sum_{i=0}^{n-1}\int_{t_i}^{t_{i+1}} (X_s - X_{t_i})\,\mathrm{d}s\,\mathbf{1}_{\{|\Delta_i M|>v_n/2, |\Delta_i M|\le 2v_n\}}\Bigg|\Bigg] = \mathrm{O}\big(T_n^{1/2}\Delta_n^{(1-\alpha\beta)/2}\big).
$$

From Assumption 4.1, it follows that S_n^4 is centered for n large enough. Furthermore, from Lemma 4.12 we conclude

$$
\begin{aligned}
E\big[\big(S_n^4\big)^2\big] &= T_n^{-1}\sum_{i=0}^{n-1} E\big[X_{t_i}^2\big]E\big[(\Delta_i M)^2\mathbf{1}_{\{|\Delta_i M|>v_n/2, |\Delta_i M|\le 2v_n\}}\big] \\
&\le T_n^{-1}\sum_{i=0}^{n-1} E\big[X_{t_i}^2\big]E\big[(\Delta_i M)^2\mathbf{1}_{\{|\Delta_i M|\le 2v_n\}}\big] \le \mathrm{O}\big(\Delta_n^{(2-\alpha)\beta}\big) \overset{n\to\infty}{\longrightarrow} 0.
\end{aligned}
$$

Finally, we show that $S_n^3 \overset{n\to\infty}{\longrightarrow} 0$. Independence together with (38) leads to

$$
\begin{aligned}
E\big[\big|S_n^3\big|\big] &= T_n^{-1/2}\sum_{i=0}^{n-1} E\big[|X_{t_i}\Delta_i U \mathbf{1}_{\{|\Delta_i M|>v_n/2, |\Delta_i M|\le 2v_n\}}|\big] \\
&= T_n^{-1/2}\sum_{i=0}^{n-1} E\big[|X_{t_i}|\big]E\big[|\Delta_i U|\big]P\big(|\Delta_i M| > v_n/2, |\Delta_i M| \le 2v_n\big) \\
&= \mathrm{O}\big(T_n^{1/2}\Delta_n^{1-\alpha\beta}\big).
\end{aligned}
$$

□

Proof of Theorem 4.6. Recall that for

$$
\hat{a}_n = -\frac{\sum_{i=1}^n X_{t_i}\Delta_i X^c}{\sum_{i=1}^n X_{t_i}^2 \Delta_i^n}
$$

we already know that $T_n^{1/2}(\hat{a}_n - a) \xrightarrow{\mathcal{D}} N(0, \frac{\sigma^2}{E_a[X_\infty^2]})$ as $n \to \infty$. Therefore, it remains to show

$$
T_n^{1/2}(\bar{a}_n - \hat{a}_n) \xrightarrow{p} 0 \qquad \text{as } n \to \infty. \tag{39}
$$

Observe that

$$T_n^{1/2}(\bar{a}_n - \hat{a}_n) = \frac{T_n^{-1/2}(\sum_{i=1}^n X_{t_i}\Delta_i X \mathbf{1}_{\{|\Delta_i X| \le v_n\}} - \sum_{i=1}^n X_{t_i}\Delta_i X^c)}{T_n^{-1}\sum_{i=1}^n X_{t_i}^2 \Delta_i}.$$

By Lemma 4.10

$$T_n^{-1/2}\left(\sum_{i=1}^n X_{t_i}\Delta_i X \mathbf{1}_{\{|\Delta_i X| \le v_n\}} - \sum_{i=1}^n X_{t_i}\Delta_i X^c\right) \xrightarrow{p} 0 \qquad \text{as } n \to \infty$$

and

$$T_n^{-1}\sum_{i=1}^n X_{t_i}^2 \Delta_i^n \xrightarrow{p} E_a\left[X_\infty^2\right],$$

such that (39) follows. □

5. Simulation results

We investigate the finite sample performance of the estimator from Sections 3 and 4 by means of Monte Carlo simulations. First, we consider Ornstein–Uhlenbeck type processes with finite jump intensity and give mean and standard deviation as well as the number of jumps detected for different parameter values and varying jump intensity. We also take a look at the normalized distribution of the estimation error for finite samples. Then we investigate models with infinite jump activity. In the last part, we compare the performance of the maximum likelihood approach and least squares estimation and find that the jump filtering approach leads to a major improvement of the estimate also for finite samples.

5.1. Finite intensity models

In this section, we perform Monte Carlo simulations for the drift estimator (6) of an Ornstein–Uhlenbeck type process defined by

$$X_t = \mathrm{e}^{-at} X_0 + \int_0^t \mathrm{e}^{-a(t-s)}\,\mathrm{d}L_s, \qquad t \in \mathbb{R}_+. \tag{40}$$

We take a deterministic starting value $X_0 \in \mathbb{R}$ and $a > 0$. The driving Lévy process L is assumed to be of the form

$$L_t = W_t + \sum_{i=1}^{N_t} Y_i,$$

where W is a Wiener process with $E[W_t^2] = \sigma_W^2 t$ and N is a Poisson process with intensity λ and the jump heights Y_i are i.i.d. with $N(0, 2)$-distribution. An advantage of this Ornstein–Uhlenbeck model is that exact simulation algorithms are available both for X and L. We use

Table 1. Mean and standard deviation of $\bar{a}_n$ with $\beta = 0.3$ for an OU process with Gaussian component and compound Poisson jumps

			$a = 2$			$a = 5$		
λ	T	n	Mean	std dev	$\varnothing$ jumps detect	Mean	std dev	$\varnothing$ jumps detect
1	10	1000	2.0	0.3	6.7	5.0	0.5	7.4
		2000	2.0	0.3	7.0	5.0	0.5	7.2
		4000	2.0	0.4	7.0	5.0	0.5	6.8
	20	1000	2.0	0.2	13.1	4.7	0.3	12.5
		2000	2.0	0.2	13.2	4.9	0.4	12.3
		4000	2.0	0.2	13.0	5.0	0.3	13.1
	50	4000	2.0	0.1	31.3	4.8	0.2	31.2
		6000	2.0	0.2	32.2	4.6	0.3	30.1
5	10	1000	1.9	0.2	31.3	4.6	0.3	30.0
		2000	2.0	0.2	31.2	4.8	0.3	30.9
		4000	2.0	0.2	31.6	4.9	0.3	30.9
	20	2000	1.9	0.1	61.4	4.6	0.2	60.2
		4000	2.0	0.1	62.2	4.8	0.2	61.4
	50	4000	1.9	0.1	149	4.6	0.1	145
		6000	1.9	0.1	149	4.7	0.1	148

an exact discretization of the explicit solution (40) to the Langevin equation driven by L on a equidistant time grid $t_i = \Delta_n i$ for $i = 1, \dots, n$. Algorithms for the exact simulation of L can be found in [5], among others.

Table 1 contains means and standard deviations of each 100 realizations of the drift estimator $\bar{a}_n$ from (6). Since the Monte Carlo error is of order $N^{-1/2}$, where N is the number of Monte Carlo iterations, we have chosen a reasonable compromise between precision of the Monte Carlo approximation and computation time. The parameter values are $a = 2$ and 5 and jump intensity λ, time horizon T and number of observations n vary as given in Table 1. We also present the number of increments that were above the threshold $\Delta_n^{0.3}$. This number corresponds to the number of jumps that were detected and we observe that it is relatively stable when T and λ are kept fixed, which suggests that the jump filter works quite reliable for finite intensity models and the threshold exponent $\beta = 0.3$. For the compound Poisson process, the average number of jumps in an interval of length T is $E[N_T] = T\lambda$ and thus is proportional to the jump intensity. This relation is also visible for the simulated data. The average number of filtered jumps is not equal to the expected number of jumps, but lies between 60 and 70% of the latter. This is surprising, since we would expect the average number of detected jumps to approach the expected number as Δ_n tends to zero.

Another interesting finding is that as soon as the step size Δ_n is so small that the discretization error is negligible (cf. Section 4.2.4 in [18] for an analysis of the discretization error), a further increase in the number of observations does not improve the accuracy of the estimator any further. This indicates that the assumption of high-frequency observations is already reasonable when the stochastic error dominates the discretization error at least for finite intensity models.

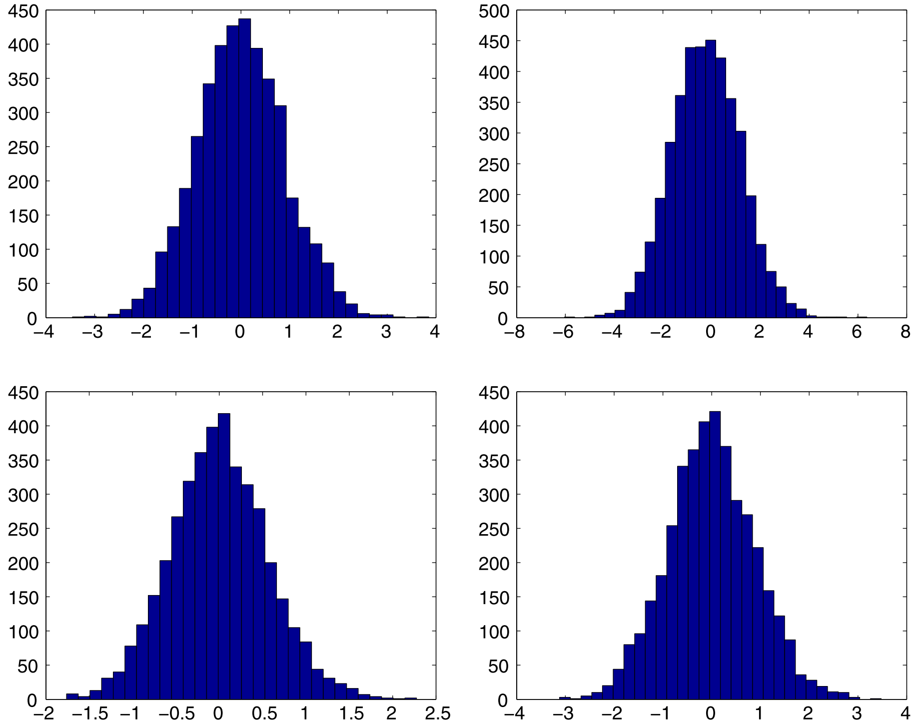

Figure 1. Standardized error distribution of $\bar{a}_n$ for $a = 2$ (left) and $a = 5$ (right) compound Poisson jumps with intensity $\lambda = 1$ (top) and $\lambda = 2$ (bottom).

The distribution of $T^{1/2}(\bar{a}_n - a)$ is depicted in Figure 1 for $T = 70$, $\Delta_n = 0.001$, $\sigma_W = 1$ and $\lambda = 1$ and 2. The histogram on the left corresponds to $a = 2$ whereas on the right we have $a = 5$. From Theorem 3.5 and the Lévy–Itô decomposition, it follows that the asymptotic variance of $\bar{a}_n$ is given by

$$\mathrm{AVAR}(\bar{a}_n) = \bigl(2a\sigma_W^2\bigr)\bigl(\sigma_W^2 + \lambda\sigma_j^2\bigr)^{-1}, \tag{41}$$

where σ_j^2 denotes the variance of the jump heights. Hence, we find that the asymptotic variance is proportional to a, which can also be observed for finite samples in Figure 1. By comparing the results of Figure 1 (top) and (bottom), we find that the variance also scales with the jump intensity as indicated in (41).

Eventually, we find that the estimator performs well even for very short time horizons if the discretization is fine enough. This observation corresponds to the results of Theorem 4.2.12 in [18] that states that the discretization bias is of the order $\mathrm{O}(\Delta_n)$.

5.2. Infinite intensity models

In Section 4, we have proved an asymptotic normality result for the discretized maximum likelihood estimator with jump filter (6) for models that involve a jump component of infinite activity. In this section, we simulate data from an Ornstein–Uhlenbeck model of the form (40) with $L = W + G$, where W is a Wiener process with $E[W_t^2] = \sigma_W^2 t$ and G is a gamma process. Again, we consider an equidistant grid $t_i = i\Delta_n$ for $i = 0, \dots, n$. The gamma process has jumps of infinite activity, paths of finite variation and its Blumenthal–Getoor index is zero. The Lévy measure μ of G has an explicit Lebesgue density given by

$$g(x) = cx^{-1}\mathrm{e}^{-\lambda x}\mathbf{1}_{\{x>0\}} \qquad \text{for } x \in \mathbb{R}.$$

The parameter $c > 0$ controls the jump intensity and $\lambda > 0$ the frequency of large jumps. It follows immediately from this density that G is a subordinator. Exact simulation algorithms are known for increments of gamma processes and we use Johnk's algorithm (cf. [5]).

Table 2 gives mean and standard deviation for different observation lengths and parameter values. The standard deviation scales approximately with $T^{-1/2}$ as expected from Theorem 4.6. In contrast to Table 1, we kept here $\Delta_n = 0.0015$ fixed for all n. As in the finite intensity case we use the threshold exponent $\beta = 0.3$ for the jump filter. We find that the value of a has hardly any impact on the average number of increments that is filtered. When a increases the number of filtered increments also increases slightly, since a greater variability of the drift might push increments with a relatively small jump over the threshold. Histograms of the standardized estimation error of $\bar{a}_n$ are given in Figure 2 for $a = 2$ and $a = 5$ and jumps from a gamma process.

We conclude that the jump filtering approach performs well, also for models with infinite jump activity provided that the maximal observation distance is small.

Table 2. Results of 200 Monte Carlo simulations of $\bar{a}_n$ with $\Delta_n = 0.0015$ and $\beta = 0.3$ for gamma process jumps

		$a = 2$			$a = 5$		
c	T	Mean	std dev	∅ jumps detect	Mean	std dev	∅ jumps detect
0.5	1	2.1	0.8	2.4	5.2	1.2	2.2
	5	2.0	0.4	11.7	5.0	0.6	12.1
	7.5	2.0	0.3	17.7	4.9	0.5	17.8
	10	2.0	0.3	23.7	5.0	0.4	23.9
	20	2.0	0.2	47.2	5.0	0.3	47.6
1	1	2.1	0.8	1.6	5.2	1.4	1.8
	2.5	2.1	0.6	5.2	5.1	1.1	5.9
	5	2.1	0.5	8.4	5.0	0.8	8.6
	7.5	2.0	0.5	12.7	5.0	0.7	13.1
	10	2.0	0.3	17.2	5.0	0.6	17.1

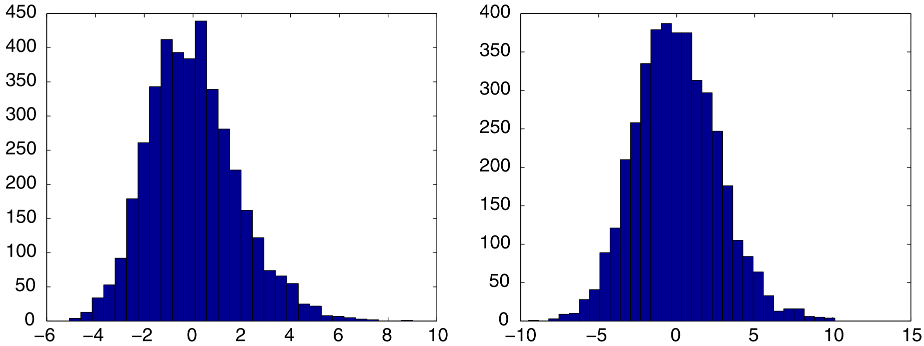

Figure 2. Error distribution of $\bar{a}_n$ for an Ornstein–Uhlenbeck process with $a = 2$ (left) and $a = 5$ (right), $\sigma_W = 1$ and gamma process jumps.

5.3. Maximum likelihood vs. least squares estimation

In this section, we compare maximum likelihood and least squares estimation for the Ornstein–Uhlenbeck type process X defined in (40). For continuously observed X, the least squares estimator for the drift parameter a is given by

$$\hat{a}_T^{\mathrm{LS}} = -\frac{\int_0^T X_s \,\mathrm{d}X_s}{\int_0^T X_s^2 \,\mathrm{d}s}.$$

For Gaussian Ornstein–Uhlenbeck processes, the least squares and the likelihood estimator $\hat{a}_T^{\mathrm{ML}}$ (4) coincide, since the continuous martingale part under P^a equals the process itself. But when the driving process has jumps it follows from Theorem 4.2.10 in [18] that the asymptotic variances of both estimators differ by

$$\mathrm{AVAR}\big(\hat{a}_T^{\mathrm{LS}}\big) - \mathrm{AVAR}\big(\hat{a}_T^{\mathrm{ML}}\big) = E_a\big[X_\infty^2\big]^{-1} \int_{\mathbb{R}} x^2 \mu(\mathrm{d}x) > 0.$$

Hence, the least squares estimator is inefficient when jumps are present.

Figure 3 compares the bias of the MLE and LSE for compound Poisson jumps with varying jump intensity. For each intensity the mean of 500 Monte Carlo simulations is given for an Ornstein–Uhlenbeck process with $\sigma_W = 1$ and jumps with $N(0, 2)$-distribution. The true parameter is $a = 2$ and we find that the MLE has a significantly smaller bias than the LSE.

The standard deviation for both estimators is given in Figure 3 on the right. The jump intensity λ of the compound Poisson part of L varies between one and ten. In this model setup, the difference in asymptotic variance between MLE and LSE is given by

$$\mathrm{AVAR}\big(\hat{a}_T^{\mathrm{LS}}\big) - \mathrm{AVAR}\big(\hat{a}_T^{\mathrm{ML}}\big) = \frac{2a\sigma_j^2\lambda}{\sigma_W^2 + \sigma_j^2\lambda}.$$

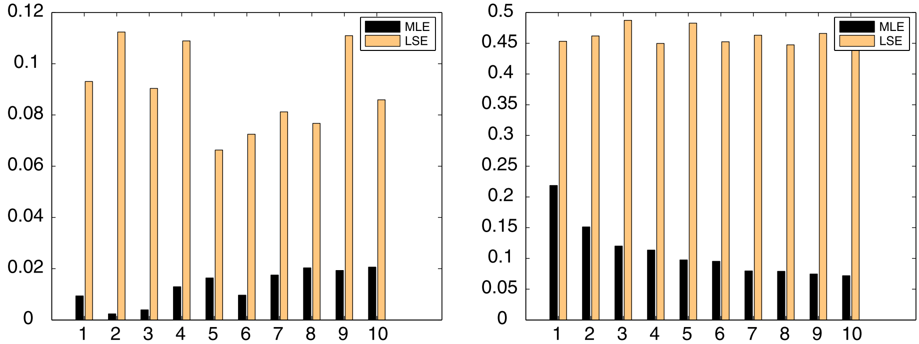

Figure 3. Bias (left) and standard deviation (right) of MLE and LSE for varying jump intensity.

We observe that already for small jump intensities the MLE clearly outperforms the LSE. With growing intensity this efficiency gain becomes even more severe. For $\lambda = 10$, the standard deviation is about five times larger for the least squares estimator.

This simulation example shows the significant gain in efficiency when we use a discretized likelihood estimator with approximation of the continuous part for drift estimation and underlines the importance of jump filtering for jump diffusion models.

Acknowledgments

The author is very grateful to Uwe Küchler and Markus Reiß for stimulating comments and discussions and an anonymous referee for many helpful questions and remarks that have led to considerable improvements. This work was partially supported by Deutsche Forschungsgemeinschaft through IRTG "Stochastic Models of Complex Processes".

References

[1] Aït-Sahalia, Y. and Jacob, J. (2007). Volatility estimators for discretely sampled Lévy processes. *Ann. Statist.* **35** 355–392. MR2332279

[2] Aït-Sahalia, Y. and Jacod, J. (2012). Identifying the successive Blumenthal–Getoor indices of a discretely observed process. *Ann. Statist.* **40** 1430–1464. MR3015031

[3] Barndorff-Nielsen, O.E. and Shephard, N. (2001). Non-Gaussian Ornstein–Uhlenbeck-based models and some of their uses in financial economics. *J. R. Stat. Soc. Ser. B Stat. Methodol.* **63** 167–241. MR1841412

[4] Brockwell, P.J., Davis, R.A. and Yang, Y. (2007). Estimation for nonnegative Lévy-driven Ornstein–Uhlenbeck processes. *J. Appl. Probab.* **44** 977–989. MR2382939

[5] Cont, R. and Tankov, P. (2004). *Financial Modelling with Jump Processes*. *Chapman & Hall/CRC Financial Mathematics Series*. Boca Raton, FL: Chapman & Hall/CRC. MR2042661

[6] Dudley, R.M. (2002). *Real Analysis and Probability*. *Cambridge Studies in Advanced Mathematics* **74**. Cambridge: Cambridge Univ. Press. Revised reprint of the 1989 original. MR1932358

[7] Figueroa-López, J.E. (2008). Small-time moment asymptotics for Lévy processes. *Statist. Probab. Lett.* **78** 3355–3365. MR2479503

[8] Hu, Y. and Long, H. (2009). Least squares estimator for Ornstein–Uhlenbeck processes driven by α-stable motions. *Stochastic Process. Appl.* **119** 2465–2480. MR2532208

[9] Jacod, J. (1985). Grossissements de filtration et processus d'Ornstein–Uhlenbeck généralisé. In *Grossissement de Filtrations: Exemples et Applications. Séminaire de Calcul Stochastique, Paris* 1982/83. *Lecture Notes in Math.* **1118** 36–44. Berlin: Springer.

[10] Jacod, J. (2007). Asymptotic properties of power variations of Lévy processes. *ESAIM Probab. Stat.* **11** 173–196. MR2320815

[11] Jahn, P., Berg, R.W., Hounsgaard, J. and Ditlevsen, S. (2011). Motoneuron membrane potentials follow a time inhomogeneous jump diffusion process. *J. Comput. Neurosci.* **31** 563–579. MR2864731

[12] Jongbloed, G. and van der Meulen, F.H. (2006). Parametric estimation for subordinators and induced OU processes. *Scand. J. Stat.* **33** 825–847. MR2300918

[13] Jongbloed, G., van der Meulen, F.H. and van der Vaart, A.W. (2005). Nonparametric inference for Lévy-driven Ornstein–Uhlenbeck processes. *Bernoulli* **11** 759–791. MR2172840

[14] Kappus, J. (2012). Nonparametric adaptive estimation for discretely observed Lévy processes. Ph.D. thesis, Humboldt-Univ. zu Berlin. Available at http://edoc.hu-berlin.de/docviews/abstract.php?lang=ger&id=39755.

[15] Klenke, A. (2008). *Probability Theory: A Comprehensive Course. Universitext*. London: Springer. MR2372119

[16] Küchler, U. and Sørensen, M. (1997). *Exponential Families of Stochastic Processes. Springer Series in Statistics*. New York: Springer. MR1458891

[17] Lánský, P. and Lánská, V. (1987). Diffusion approximation of the neuronal model with synaptic reversal potentials. *Biol. Cybernet.* **56** 19–26. MR0885360

[18] Mai, H. (2012). Drift estimation for jump diffusions: Time-continuous and high-frequency observations. Ph.D. thesis, Humboldt-Univ. zu Berlin. Available at http://edoc.hu-berlin.de/docviews/abstract.php?id=39623.

[19] Mancini, C. (2009). Non-parametric threshold estimation for models with stochastic diffusion coefficient and jumps. *Scand. J. Stat.* **36** 270–296. MR2528985

[20] Mancini, C. (2011). The speed of convergence of the threshold estimator of integrated variance. *Stochastic Process. Appl.* **121** 845–855. MR2770909

[21] Masuda, H. (2007). Ergodicity and exponential β-mixing bounds for multidimensional diffusions with jumps. *Stochastic Process. Appl.* **117** 35–56. MR2287102

[22] Masuda, H. (2010). Approximate self-weighted LAD estimation of discretely observed ergodic Ornstein–Uhlenbeck processes. *Electron. J. Stat.* **4** 525–565. MR2660532

[23] Sato, K.-i. (1999). *Lévy Processes and Infinitely Divisible Distributions. Cambridge Studies in Advanced Mathematics* **68**. Cambridge: Cambridge Univ. Press. MR1739520

[24] Sato, K.-i. and Yamazato, M. (1984). Operator-self-decomposable distributions as limit distributions of processes of Ornstein–Uhlenbeck type. *Stochastic Process. Appl.* **17** 73–100. MR0738769

[25] Shimizu, Y. (2006). M-estimation for discretely observed ergodic diffusion processes with infinitely many jumps. *Stat. Inference Stoch. Process.* **9** 179–225. MR2249182

[26] Sørensen, M. (1991). Likelihood methods for diffusions with jumps. In *Statistical Inference in Stochastic Processes*. New York: Dekker.

Received July 2012 and revised December 2012