

OTDK-dolgozat

2025

Budapesti Corvinus Egyetem

Pénzügy Intézet



Szemimartingálok ugrási aktivitási indexe – Melyik módszert használjuk?

Jump activity index for semimartingales – Which method should we use?

Székely György József

Konzulens: Badics Milán Csaba

Kézirat lezárásának dátuma: 2023.04.17

Jump activity index for semimartingales: Which method should we use?

Székely György József

1 Introduction

Itô semimartingales are a popular choice for modeling asset prices in financial markets, as they encompass a diverse range of stochastic processes such as diffusion, jump diffusion, and Lévy processes. The tail behavior of return distributions is significantly impacted by extreme movements, commonly known as jumps in asset prices. In recent years, there has been a growing interest in studying semimartingales that incorporate a discontinuous component, with several notable works in the literature, such as those by Aït-Sahalia and Jacod (2007), Fan and Wang (2007), Jacod (2008), Todorov and Tauchen (2011), and Boswijk et al. (2018). Empirical studies and statistical tests on discretely observed asset prices, in high-frequency data, provide strong evidence for the presence of jumps in asset prices. (Barndorff-Nielsen and Shephard (2006), Lee and Mykland (2008), Aït-Sahalia and Jacod (2009a), Mancini and Gobbi (2012), Bibinger and Winkelmann (2015)) In order to study the behavior of the jumps or the jump characteristics in modeling, Aït-Sahalia and Jacod (2009b) defined the jump activity index as a measure of the activity of jumps. I will talk about the applications of the jump activity index in the second section.

In this paper, I examine and compare the accuracy of various high-frequency jump activity estimators, with a focus on their robustness to the presence of microstructure noise, and different sampling frequencies, as well as performance on different data generating processes containing volatility jumps, double price jumps, intensity jump, and self-excitation. After giving a brief overview of the several estimators, I demonstrate the mean squared error and the number of outliers for each scenario.

In Section 2, I summarize the theoretical framework of the high-frequency analysis and present the most important concepts. In Section 3, I give a quick overview of the existing jump activity estimators, and at the end of the section, I propose an alternative method. This estimator is just the first attempt to create a robust estimator, my aim is to construct a well-functioning estimator in future works. In Section 4, I conduct Monte Carlo experiments on the previously presented DGPs.

2 Theoretical Framework

The theoretical framework lies upon the following model. X is a one-dimensional underlying process, usually a log price of an asset, which is an Itô semimartingale. Let's consider as Ait-Sahalia et al. (2009) and Jing et al. (2012), the so-called Grigelionis decomposition of X_t :

$$X_t = X_0 + \int_0^t b_s ds + \int_0^t \sigma_s dW_s + \int_0^t \int_{|x| \leq \epsilon} x(\mu_t - \nu_t)(ds, dx) + \int_0^t \int_{|x| > \epsilon} x\mu(ds, dx),$$

where W_t is a standard Brownian motion, b and σ are locally bounded processes, and μ is a random measure related to the count of jumps with compensator ν given by $\nu(dt, dx) = dtF_t(dx)$, with F_t predictable random measure.

Assumption 1 *There are three constants $\beta \in (0, 2)$, $\beta' \in [0, \beta)$, and $\gamma > 0$, and a locally bounded process $L_t \geq 1$ such that for all (ω, t) , $F_t = F'_t + F''_t$ where:*

1.

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

with some predictable nonnegative processes $a_t^{(+)}$, $a_t^{(-)}$, $z_t^{(+)}$, and $z_t^{(-)}$, and a predictable function $f(\omega, t, x)$ satisfying:

$$\begin{cases} \frac{1}{L_t} \leq z_t^{(+)} \leq 1, \frac{1}{L_t} \leq z_t^{(-)} \leq 1, a_t^{(+)} + a_t^{(-)} \leq L_t \\ 1 + |x|^\gamma f(t, x) \geq 0, |f(t, x)| \leq L_t. \end{cases}$$

2. F''_t is a singular measure with respect to F'_t and satisfies

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

3. The following process is increasing and locally bounded

$$\bar{A}_t = \int_0^t A_s ds, \text{ where } A_t = \frac{a_t^{(+)} + a_t^{(-)}}{\beta}.$$

Remark 1 *We split F_t into two parts, F'_t is close to a Lévy-measure of a beta-stable process in a random interval $(-z_t^{(-)}, z_t^{(+)})$ around 0 with some β , and F''_t is another Lévy-measure with jump activity index lower than $\beta' < \beta$. The measure F''_t determines the large jumps.*

Remark 2 *The processes $a_t^{(\pm)}$ give the intensity, the higher their value, the more the number of small jumps. We have to note that this jump activity can be time-varying and also random. On the set $\{\bar{A}_T > 0\}$ the value of β is constant in the whole interval.*

On a fixed time interval $[0, T]$ we have equidistant observations $X_{i\Delta_n}$ for $i = 0, 1, \dots, [T/\Delta_n]$. We denote the i th increment $|\Delta_i^n X| = X_{i\Delta_n} - X_{(i-1)\Delta_n}$.

Since there is a discontinuous part present, it is natural to investigate the behavior of the jumps or jump characteristics while modeling. Aït-Sahalia and Jacod (2009b) defined the jump activity index to measure the level of jump activity for a generic semimartingale X .

$$\beta_t = \inf \left\{ r \geq 0 ; \sum_{0 \leq s \leq t} |\Delta_s X|^r < \infty \right\},$$

where $|\Delta_s X|$ is the jump size at time s . β is a lower bound for the values of p for which the p th power of the jumps is summable. The jump activity index provides insight into the frequency of small jumps, it characterizes how frequently small jumps occur which is closely linked to the behavior of the Lévy measure for semimartingales near the origin. Notice that for $r = 2$ this expression is always finite, thus $\beta \in [0, 2]$.

For a Lévy-process, the jump activity index coincides with the Blumenthal-Gettoor index defined as

$$\beta = \inf \left\{ r \geq 0 ; \int_{\mathbb{R}} (|x|^r \wedge 1) \nu(dx) < \infty \right\},$$

where $\nu(dx)$ is the density of the Lévy measure. In the particular case when X is a stable process, then β serves as the stable index of the process as well. The value of β corresponds to the rate at which the measure ν diverges near zero. Consequently, a larger value of β implies a greater number of small jumps. This motivated our desire to estimate it.

Many people have already dealt with this topic, these estimators can be categorized into two groups. Let us decompose the semimartingale X into $X_t^c + X_t^d$, where X_t^c and X_t^d represent the continuous and discontinuous parts, respectively. In the first group, the absence of a continuous local martingale part was assumed. ($X_t^c \equiv 0$) Some papers on this topic are Woerner (2011), Todorov and Tauchen (2010) whose work is based on using power variations. Zhao and Wu (2009) proposed a nonparametric estimator for the stable index of a driving stable Lévy process, utilizing two-time scale techniques and a deterministic time-dependent integrand. Cont and Mancini (2007) proposed a test for the finiteness of the variation of the jump part, and Belomestny (2010) provides an estimation when option prices are also available. Theopold and Vetter (2023) developed an estimator considering the irregular time intervals of high-frequency data. The second group assumes that the continuous martingale part is present. ($X_t^c \not\equiv 0$) In the next section, I will go into detail about this category and its related works such as Aït-Sahalia and Jacod (2009b), Jing et al. (2012), and Bull (2016). In recent years, Kolokolov (2022) used different multipower variations to implement a precise estimator of the jump activity index. Aït-Sahalia and Jacod (2014) and Aït-Sahalia and Jacod (2012a) summarize the essence of high-frequency financial econometrics.

The jump activity index has both theoretical and practical significance. In terms of modeling, it can be used to assess whether the jump component of a semimartingale has finite variation or not. Empirical findings suggest that infinite activity jump models that incorporate frequent jumps of all sizes are more suitable for capturing the daily market movements of several financial securities, including stocks, stock indexes, and exchange rates. So the estimation of the jump activity index can give insight into which types of jumps should be incorporated into financial models. In addition, infinite activity jumps can improve option pricing performance in certain cases in options markets. The jump activity affects the estimation of the volatility as well. Higher jump activity makes it more difficult to estimate volatility, requiring stronger assumptions in specialized volatility estimates to overcome this problem. In this paper, I try to summarize the most current estimators and examine their robustness from as many perspectives as possible.

3 Jump activity estimators

3.1 Ait-Sahalia estimators

When the semimartingale X is only observed at certain intervals over $[0, T]$, the estimation problem becomes more difficult due to the presence of a continuous martingale component within X . As we have seen, β characterizes the behavior of ν in the vicinity of zero. Consequently, the small increments of the process provide the most useful information regarding β . However, these small increments contain contributions from both the continuous martingale part of the process and the jumps, which makes distinguishing between them challenging. When estimating β we need to disentangle the continuous part of the semimartingale and even with this gain the best insight into the concentration and quantity of small jumps.

Ait-Sahalia and Jacod (2009b) were one of the first who introduced an estimator for β . In their work, the continuous component of the semimartingale dominates the small increments. They estimate the activity index of the pure jump process where the dominant component is a Brownian motion. They constructed an estimator which is consistent when $\Delta_n \rightarrow 0$, and provide rates of convergence and asymptotic distributions. They assumed that the main part of the Lévy measure near 0 behaves locally like the Lévy measure of a β -stable process.

This estimator lies under the following sum:

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

where $\alpha > 0$ and $0 < \varpi < 1/2$ are two constants. Notice that U counts the number of large increments.

Remark 3 *To better understand why we are doing this, consider the following model: $X = \sigma W + Y$, where Y is a β -stable process. The increments $\Delta_i^n X$ satisfy this equivalence in distribution:*

$$\Delta_i^n X = \Delta_i^n W + \Delta_i^n Y =_d \sigma \Delta_n^{1/2} W_1 + \Delta_n^{1/\beta} Y_1$$

As $\beta < 2$, and $\Delta_n \rightarrow 0$, $\Delta_i^n X$ close $\sigma \Delta_n^{1/2} W_1$ in law with a large probability. Those increments give essentially no information on Y and are of the order of magnitude $\Delta_n^{1/2}$. So if X has a jump exceeding some lower limit, then it is due to dY . Hence we truncate the jumps, throw out the small ones. But β is related to the behavior of ν near 0, hence to the small jumps of Y . This is why we use a cutoff level $\alpha \Delta_n^\varpi$. Asymptotically those increments are big since $\Delta_n^{1/2} \ll \Delta_n^\varpi$.

The key characteristics of $U(\varpi, \alpha, \Delta_n)$ is their convergence in probability:

$$\Delta_n^{\varpi\beta} U(\varpi, \alpha)_t^n \xrightarrow{1} \frac{\bar{A}_t}{\alpha^\beta}.$$

With the help of an $\alpha' > \alpha$, and the previous convergence, they defined:

$$\hat{\beta}_n^{AJ} = \log \left(\frac{U(\varpi, \alpha)_t^n}{U(\varpi, \alpha')_t^n} \right) \bigg/ \log \left(\frac{\alpha'}{\alpha} \right)$$

$\hat{\beta}_n^{AJ}$ is constructed from two U s on the same time scale at two different fixed levels of truncation. They also constructed a second estimator, based on different sampling frequencies:

$$\tilde{\beta}_n^{AJ} = \log \left(\frac{U(\varpi, \alpha)_t^n}{U_2(\varpi, \alpha)_t^n} \right) \bigg/ (\varpi \log(2))$$

where $U_2(\varpi, \alpha)_t^n$ defined analogously to $U(\varpi, \alpha)_t^n$ but now sampling at $2\Delta_n$.

Remark 4 *As we will see below, the choice of the cutoff value α , and the parameter ϖ , are dependent on β . However, asymptotic considerations lead to the selection of $\varpi = 1/5$ as a universal choice that is valid for all possible values of β . When implementing the estimator in practice, the value of ϖ is fixed for any given sample, making α and ϖ , not independent parameters. The level of truncation α can be determined based on the volatility of the continuous part of the semimartingale, which is given by $(t^{-1} \int_0^t \sigma_s^2 ds)^{1/2}$. The aim is to eliminate increments that are mainly due to the continuous part.*

Now we quickly share the asymptotic properties of these estimators. The results presented here are considered model-free as the drift and volatility processes are completely unspecified, except for they're locally bounded. Additionally, the Lévy measures F_t are also not specified, except for the conditions stated in Assumption 1.

Theorem 1 Let $\alpha < \alpha', 0 < \varpi < 1/2$ and $t > 0$. Under Assumption 1, we have $\hat{\beta}_n^{AJ}(t, \varpi, \alpha, \alpha') \xrightarrow{\mathbb{P}} \beta$ on the set $\{\bar{A}_t > 0\}$. Moreover, if

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

then the estimators are $\Delta_n^{\chi-\varepsilon}$ -rate consistent for any $\varepsilon > 0$ on the set $\{\bar{A}_t > 0\}$ in the sense that the sequence of variables $\frac{1}{\Delta_n^{\chi-\varepsilon}} \left(\hat{\beta}_n^{AJ}(t, \varpi, \alpha, \alpha') - \beta \right)_{n \geq 1}$ is bounded in probability in restriction to this set.¹

Theorem 2 Let $\alpha < \alpha'$ and $t > 0$. Assumption 1 holds with $\beta' \in [0, \beta/(2 + \beta))$ and $\gamma > \beta/2$. Then, if $\varpi < 1/(2 + \beta) \wedge 2/(5\beta)$ and in a restriction to the set $\{\bar{A}_t > 0\}$, we have the following stable convergence in law to a centered normal variable independent of X :

$$\frac{1}{\Delta_n^{\varpi\beta/2}} (\hat{\beta}_n^{AJ}(t, \varpi, \alpha, \alpha') - \beta) \rightarrow \mathcal{N} \left(0, \frac{\alpha'^\beta - \alpha^\beta}{\bar{A}_t (\log(\alpha'/\alpha))^2} \right).$$

Theorem 3

$$\frac{\log(\alpha'/\alpha)}{\sqrt{1/U(\varpi, \alpha', \Delta_n)_t - 1/U(\varpi, \alpha, \Delta_n)_t}} \left(\hat{\beta}_n^{AJ}(t, \varpi, \alpha, \alpha') - \beta \right).$$

Remark 5 Similar properties hold for the second estimator, based on the ratio of U s estimated at two different frequencies,

$$\frac{\varpi \log 2}{\sqrt{1/U(\varpi, \alpha, 2\Delta_n)_t - 1/U(\varpi, \alpha, \Delta_n)_t}} \left(\tilde{\beta}_n^{AJ}(t, \varpi, \alpha) - \beta \right)$$

converge stably in law, in a restriction to the set $\{\bar{A}_t > 0\}$, to a standard normal variable independent of X .

3.2 Jing estimator

The Ait-Sahalia estimator however has a significant disadvantage, its major problem is the limited effectiveness of the sample size, even for very large sample sizes. This leads to a reduced number of observations being utilized in estimating β , resulting in a loss of efficiency.

Notice that $U(\varpi, \alpha)$ only counts the number of sufficiently large increments. So if we define an indicator-type weight function $g_0(x) = \mathbb{1}_{|x| > 1}$, we can rewrite U as:

$$U(\varpi, \alpha)_t^n = \sum_{i=1}^{\lfloor t/\Delta_n \rfloor} \mathbb{1} \left(\left| \frac{\Delta_i^n X}{\alpha \Delta_n^\varpi} \right| > 1 \right) = \sum_{i=1}^n g_0 \left(\frac{\Delta_i^n X}{\alpha \Delta_n^\varpi} \right)$$

¹The proof can be found in Ait-Sahalia and Jacod (2009b).

To fully utilize the available data, it may be necessary to take into account not just the large increments but also the small ones. This is because the jump activity index primarily measures small jumps. But as the increment $\Delta_i^n X$ decreases, the contributions from $\Delta_i^n W$ become more significant. As a result, the increment $\Delta_i^n X$ contains less useful information about β . Because of these, Jing et al. (2012) defined the beta-estimators this way:

$$V(\varpi, \alpha, g)_t^n = \sum_{i=1}^{[t/\Delta_n]} g\left(\frac{\Delta_i^n X}{\alpha \Delta_n^\varpi}\right)$$

with a properly chosen g function.

Remark 6 Assume a continuous martingale part is present, then

$$\begin{aligned} \text{for } p > 2 : \sum_{i=1}^n |\Delta_i^n X|^p &\xrightarrow{P} \sum_{0 \leq s \leq T} |\Delta_s X|^p \\ \text{for } p = 2 : \sum_{i=1}^n |\Delta_i^n X|^2 &\xrightarrow{P} \int_0^T \sigma_s^2 ds + \sum_{0 \leq s \leq T} |\Delta_s X|^p \\ \text{for } p < 2 : \Delta_n^{1-p/2} \sum_{i=1}^n |\Delta_i^n X|^p &\xrightarrow{P} \mathbb{E}|\mathcal{N}(0, 1)|^p \int_0^T |\sigma_s|^p ds \end{aligned}$$

Only in the case of $p > 2$ the limits of $\sum_{i=1}^n |\Delta_i^n X|^p$ determined only by jumps, meaning that the influence of the continuous part is entirely eliminated. Thus the suitable choices for weight functions have power form near the origin.

Assumption 2 $g(x) = |x|^p$ if $|x| < a$ for some constant $a > 0$, and even integer $p > 2$, and $g(x)$ is even, non-negative, bounded, and smooth with bounded and Lipschitz continuous first-order derivative.

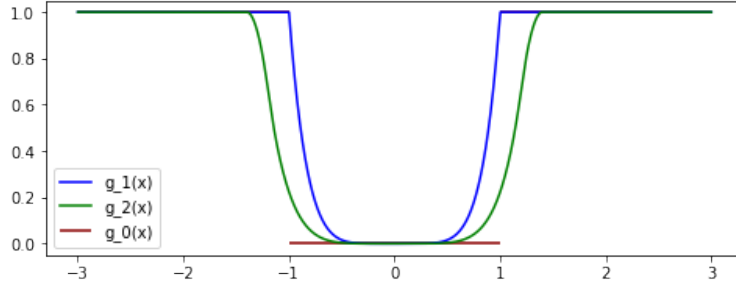
The authors used the following functions:

$$g_1(x) = \begin{cases} |x|^p, & |x| \leq 1 \\ 1, & |x| > 1 \end{cases}$$

$$g_2(x) = \begin{cases} c^{-1}|x|^p, & |x| \leq a \\ c^{-1} \left(a^p + \frac{pa^{p-1}}{2(b-a)} ((b-a)^2 - (|x|-b)^2) \right), & a \leq |x| \leq b \\ 1, & |x| > b \end{cases}$$

where $0 < a < b < \infty$ constants, $c = a^p + pa^{p-1}(b-a)/2$.

Remark 7 Notice, that $g_2(x)$ is a smoother version of $g_1(x)$, if $a = b = 1$ then $g_2(x) = g_1(x)$. Also, $g_1(x) \rightarrow g_0(x)$ as $p \rightarrow \infty$. Finally, the larger value of p means that there is lesser weight on the small increments.



They showed, that under these assumptions,

$$\Delta_n^{\varpi\beta} V(\varpi, \alpha, g)_t^n \xrightarrow{1} \frac{A_t}{\alpha^\beta} \int_0^\infty \frac{g(x)}{x^{1+\beta}} dx.$$

After this theoretical foundation, the following estimate was introduced:

$$\hat{\beta}_n^J = \log \left(\frac{V(\varpi, \alpha, g)_t^n}{V(\varpi, \alpha', g)_t^n} \right) \Bigg/ \log \left(\frac{\alpha'}{\alpha} \right)$$

Remark 8 *The authors derived similar asymptotic theorems like Aït-Sahalia and Jacod (2009b). They found that the convergence rate is $\Delta_n^{\varpi\beta/2}$. Notice, that the choice of ϖ controls the convergence rate, the bigger the value of ϖ the faster the convergence. Aït-Sahalia and Jacod (2009b) required that $\varpi < 1/(\beta+2) \wedge 2/(5\beta)$ so a conservative option is $\varpi < 1/5$. By comparison, Jing et al. (2012) deduce that in their case $\varpi < 1/(\beta+2)$, so $\varpi < 1/4$.*

Remark 9 *Assuming $\gamma > 0$, and when the conditions $p > 2(1 - \varpi\beta)/(1 - 2\varpi)$ and $\beta' < \beta$ are met, the estimator remains consistent. When this is accomplished, so if small jumps dominate the increments of the continuous part, we can extract information on β from small jumps. However, if not, increments of the continuous part will dominate the small jumps.*

Remark 10 *The Jing estimator makes full use of all increments, hence its effective sample size is larger than the Ait-Sahalia estimator.*

3.3 Bull estimator

In the following, I present the Bull estimator. We will assume that β is greater than 0 in the following discussion. The scenario where β equals 0 can be examined independently. Bull's work Bull (2016) builds on Aït-Sahalia and Jacod (2009b), and Jing et al. (2012) papers. Regrettably, the former estimator was unable to yield good convergence rates due to the significant bias in the estimates. This estimator tries to improve on these.

Bull employs three techniques to correct the bias. Firstly, the author symmetrized the data to counteract the effects of high-activity, asymmetrically-distributed jumps.

Secondly, smooth the jump counts to correct for the impact of the roughness in the indicator function, similarly to Jing. Lastly, the most crucial step is to eliminate the remaining bias by canceling between estimates at different time scales.

First, we symmetrize the process X_t :

$$\Delta X_{j,n} = (\Delta X_{j+2/n} - \Delta X_{j+1/n}) - (\Delta X_{j+1/n} - \Delta X_{j/n})$$

Notice, that if X_t is a Lévy process, the random variables $\Delta X_{j,n}$ are symmetric, even if the increments are not. More generally, we can think of $\Delta X_{j,n}$ as symmetrized increments of the process X_t over time intervals with a duration of $2/n$. The next idea is to work with increments of X_t across different time scales simultaneously. Hence Bull defined random variables

$$\Delta X_{j,k,n} = \sum_{l=0}^{k-1} \Delta X_{j+2l,n}.$$

These variables are the symmetrized increments of X_t across time intervals of length $2k/n$. Then the author defined the following smooth function.

$$K(x) = \begin{cases} 1, & |x| < 1 \\ \left(1 + \exp\left(\frac{1}{2-|x|} - \frac{1}{|x|-1}\right)\right)^{-1}, & 1 \leq |x| \leq 2 \\ 0, & |x| > 2 \end{cases}$$

Denote by $m \in \mathbb{N}$ the number of time scales used for bias reduction. Then Bull defined the jump counts for $\tau > 0, \tau = \frac{1}{\alpha \Delta_n^{\frac{1}{\alpha}}}$.

$$\hat{A}_n(\tau) = 0 \vee \hat{A}'_n(\tau), \quad \hat{A}'_n(\tau) = \sum_{j=0}^{n-2m-1} \hat{a}_{j,n}(\tau)$$

where for $j = 0, \dots, n-2m-1$, we set the jump count

$$\hat{a}_{j,n}(\tau) = \sum_{k=1}^m w_k (1 - K(\tau \Delta X_{j,k,n})) , \quad w_k = \frac{(-1)^{k+1}}{2k} \binom{m}{k}.$$

For a constant $\rho = \frac{\alpha'}{\alpha} > 1$, and a sequence $\tau_n > 0$, Bull estimated β by

$$\hat{\beta}_n^B = 0 \vee \log_{\rho} \left(\frac{\hat{A}_n(\rho \tau_n)}{\hat{A}_n(\tau_n)} \right) \wedge 2.$$

Clip the estimate to the $[0, 2]$ interval ensures that the estimator is always reasonable, even when the jump counts are inaccurate. This step reduces its error in finite time.

Remark 11 *Although Ait-Sahalia and Jacod cannot establish results for all semimartingales, they do offer convergence rates assuming that the jumps are controlled by a stochastic integral of a stable-like process. Under these conditions, their estimation*

of the jump activity β attains convergence with a rate of $n^{-\beta/10}$. Bull achieved a convergence rate of $n^{-\beta m/4(m+1)}$ when X_t includes any stable-like jump activity, and naturally, we cannot expect convergence when no stable-like jumps are present. Asymptotically, the rate of convergence will always be improved by choosing m larger, and we should therefore choose m as large as possible. However, in a finite time frame, a higher value of m may require a longer period to attain the asymptotic regime and thus may be less precise in practical applications.

Remark 12 θ tuned to contain a jump larger than a cutoff value $\alpha\Delta_n\varpi$ with some pre-specified tail probability. $\tau_n = Cn^\alpha$, where $\alpha = m/(2m+1)$.

In order to describe the limiting distributions for $\hat{\beta}_n^B$, we have to define the following constants:

$$C_{\beta,\rho} = \frac{K_{\beta,\rho}}{\rho^\beta} \log(\rho)^2 K_\beta^2$$

$$K_\beta = \int_{\mathbb{R}} (1 - K(x)) |x|^{-(1+\beta)} dx$$

$$K_{\beta,\rho} = \int_{\mathbb{R}} (K(x) - K(\rho x))^2 |x|^{-(1+\beta)} dx$$

and for $\hat{\beta}_n^B \in (0, 2]$, the random variables

$$\hat{U}_n(\beta) = \frac{\tau_n^{\hat{\beta}_n^B/2}}{\hat{\sigma}_{\rho,n}} \left(\hat{\beta}_n^B - \beta \right), \quad \hat{\sigma}_{\rho,n}^2 = \frac{C_{\hat{\beta}_n^B,\rho} K_{\hat{\beta}_n^B} \tau_n^{\hat{\beta}_n^B}}{\hat{A}_n(\tau_n)}.$$

when $\hat{\beta}_n^B = 0$, $\hat{U}_n(\beta) = -\infty$. The random variables $\hat{\beta}_n^B$ are always well-defined since $\hat{\beta}_n^B \in [0, 2]$. Under appropriate conditions, the standardised errors $\hat{U}_n(\beta) \xrightarrow{d} \mathcal{N}(0, 1)$.

Remark 13 The author used similar assumptions to Aït-Sahalia and Jacod (2009b), that the log-price process X_t and the volatility process are Itô semimartingales; that the jumps in the log-price process X_t are controlled by a stochastic integral of a stable-like process. The Lévy distribution functions approach those of a beta-stable process; and the drift process, leverage process, and jump integrands exhibit smoothness behavior typical of Itô semimartingales. For precise definitions and notation, we refer to Bull (2016) and respectively Jacod and Protter (2011).

Note, that the presence of two separate one-sided Lévy processes also allows us to describe processes with asymmetric jump activity.

To be precise, we initially define the jump activity processes

$$\bar{A}'_t = \int_0^t \bar{A}'_s, \text{ and } \bar{A}'_t = \frac{1}{2} \left(|a_t^{(+)}|^\beta + |a_t^{(-)}|^\beta \right)$$

It is worth mentioning that the process $\bar{A}'_t$ measures the immediate stable-like jump activity at time t , whereas $\bar{A}'_t$ computes the entire stable-like jump activity during the interval $[0, 1]$.

Theorem 4 *Let $m \in \mathbb{N}$, $\alpha = m/2(m+1)$, $\tau_n = Cn^\alpha$ with some $C > 0$, and $\rho > 0$. Then on the event $\{\bar{A}'_t > 0\}$,*

$$\tau_n^{\beta/2}(\hat{\beta}_n^B - \beta) \xrightarrow{d} \sigma_{\beta,\rho} Z, \quad \hat{\sigma}_{\rho,n}^2 \xrightarrow{p} \sigma_{\beta,\rho}^2, \quad \hat{U}_n(\beta) \xrightarrow{d} Z$$

where the variance

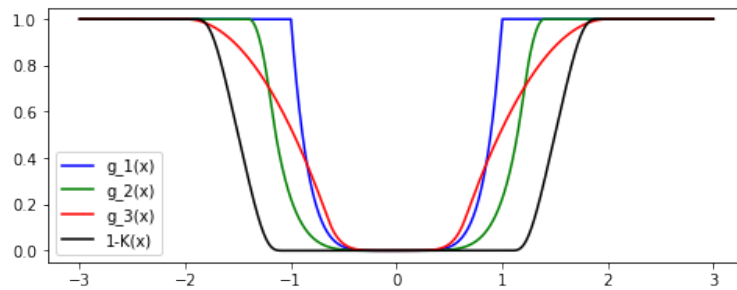
$$\sigma_{\beta,\rho}^2 = C_{\beta,\rho} / \bar{A}'_t,$$

where Z is a random variable defined on a suitable extension of the probability space $(\Omega, \mathcal{F}, \mathbb{B})$, to be standard Gaussian given $\mathcal{F}$.

3.4 Alternative method

Based on the three introduced estimators I construct an alternative method for estimating jump intensity. We know that we can extract the most information about the jumping activity index from small jumps. At the same time the continuous part of the process interferes, if we choose the threshold too low, it will divert the estimator. I used the g_2 function form used by Jing as a base, and I tried to set up an appropriate parameter specification. I tried to choose as low a as possible without ruining the estimation. For b , I wanted to have a smoother transition between 0 and 1. Finally, I choose $a = 0.65$, and $b = 2$. I injected this function into the Bull estimator with the adequate C value based on the data generating process' integrated volatility.

This alternative method worked really well in some cases, as we will see in the next section, however, it has disadvantages. If for some reason small jumps just above the threshold appear, this estimator is particularly susceptible to this, and easily gets biased. I aim to construct a more robust estimator in future works.



4 Monte Carlo experiments

4.1 Base model settings

This section aims to use simulations to compare the performances of the introduced estimators in finite sample sizes. To compare the performances of the estimators, first I generate data from the stochastic volatility model:

$$\begin{cases} dX_t = \sigma_t dW_t + \theta dY_t \\ dv_t = \kappa(\eta - v_t)dt + \gamma v_t^{1/2} dB_t \end{cases}$$

where $v_t = \sigma_t^2$, $E(dW_t dB_t) = \rho$, and Y_t is a centered symmetric β -stable process. I tuned the parameters based on the related literature. (Aït-Sahalia and Jacod (2009b), Jing et al. (2012)) I take $\kappa = 5, \eta = 1/16, \gamma = 0.5, \rho = -0.5, \alpha = 5/16$ and $\alpha' = 2\alpha$. (these parameters adhere to the Feller condition: $2\kappa\eta \geq \gamma$.) I set $T = 1$ (day), which corresponds to 6.5 hours of trading per second, resulting in a total of $N = 23,400$ observations. In the case of the Jing estimator, the parameters of the $g_2(x)$ function are $a = 6/5, b = 7/5$, and $p = 6$.

We calibrate the parameter θ to obtain certain pre-specified values of the tail probability:

$$\mathbb{P}(|\theta \Delta_n Y| > \alpha \Delta_n^\varpi) = \frac{c_\beta \theta^\beta \Delta_n^{1-\varpi\beta}}{\beta \alpha^\beta} + \text{smaller order terms}$$

where

$$c_\beta = \frac{\Gamma(1+\beta)}{\pi} \sin\left(\frac{\beta\pi}{2}\right).$$

I modified these simulation values to fit better with the generated data. These tunings are used to distinguish the continuous increment from the discontinuous increment. These estimators specify truncations ($\alpha \Delta_n^\varpi$) that involve the choice of the truncation root ϖ . In every case, a larger value of omega corresponds to a smaller level of the actual truncation point, implying that a larger number of returns are considered as part of the jump contribution. I fix $\varpi = 0.2$. In the case of the Bull estimator, I calibrate C relative to the size of the diffusion component of X_t , as measured for example by its integrated volatility. In this particular case $C = 0.25$.

First I compare the different estimators in this particular DGP. In the first columns of the tables, we can see the mean squared error (MSE) for each estimator under different values for β in the β -stable process. Specifically $\beta = 1.75, 1.5, 1.25, 1$, and 0.75 . Then the number of outliers, which I define as the number of observations/estimates that are more than 20% away from the original β value. This needs modification when $\beta = 1.75$. I set an upper limit of 2, so any estimate equal to or above 2 is considered an outlier. This is appropriate for estimators that truncate at 2, like the Bull estimator.

The details of these tests can be found in Table 1 Table 2 Table 3.

In high-frequency data, market microstructure noise appears, so the process X is observed with an error: $X_t + \varepsilon_t$. The more frequent the data the bigger the measurement error. In the related literature, we can find a lot of papers (Aït-Sahalia et al. (2012), Jing et al. (2011)), but I did not find such a broad summary analysis. The main difference between those papers is that in our case we have a stochastic volatility process that influences the magnitude of noise. On top of that, the log price process is a mixture of a jump process and a diffusion part, in this case, the estimation of the jump activity index becomes more challenging. I want to measure how disturbing this noise is for these jump activity estimators. I consider three types of additive noise:

$$\varepsilon_t = \begin{cases} 2\sigma_t \Delta_n^{1/2} \varepsilon_t^A & \text{(Gaussian noise)} \\ 2\sigma_t \Delta_n^{1/2} \varepsilon_t^B / \sqrt{\frac{df}{df-2}} & \text{(T-distributed noise)} \\ 2\sigma_t \Delta_n^{1/2} \left(\varepsilon_t^A + \varepsilon_t^B / \sqrt{\frac{df}{df-2}} \right) & \text{(Gaussian-T mixture noise)} \end{cases}$$

where ε_t^A and ε_t^B are independent and identically distributed samples drawn from a standard normal distribution, and a t-distribution with degrees of freedom equal to $df = 2.5$, respectively. The instantaneous standard deviations of the Gaussian and t-distributed noise are twice as much as that of the diffusive increment.

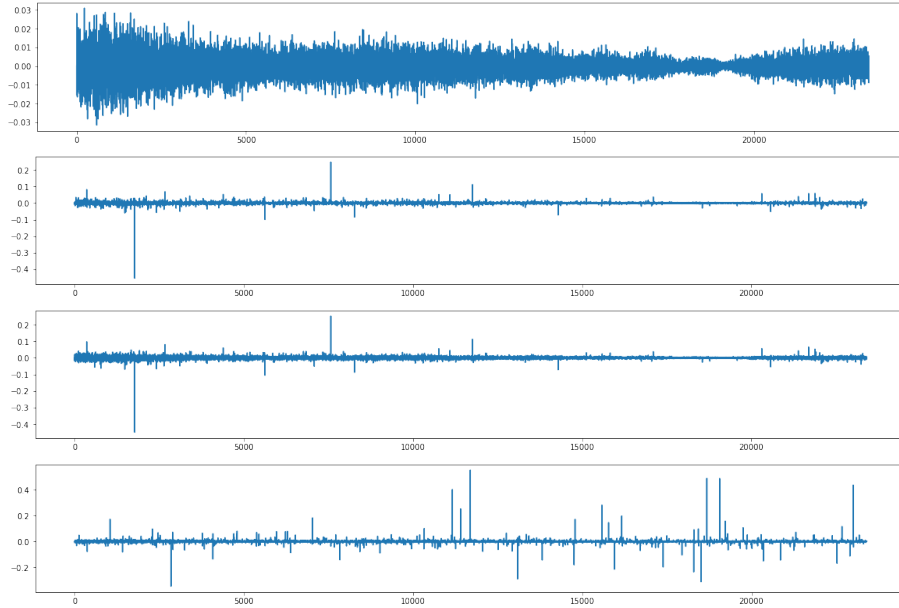


Figure 1: A realization of returns of the noisy process. From top to bottom, the Gaussian noise, T -distributed noise, Gaussian-T mixture noise, and the noiseless process.

As we can see, the Gaussian noise is highly influenced by the starting value of the volatility. The t-distributed noise is added to capture the significant bouncebacks in

the transaction data. If we look at the t-distributed noise we can see, that there are values higher than our threshold. This can lead to spurious results. The results can be seen in Table 5.

After that, I want to test the robustness of these estimators related to the sampling frequency. I consider three different specifications: 1, 3, and 5-second sampling frequency. The results are presented in Table 4.

4.2 Different Data Generating Processes

Now after these thorough examinations, I add one new component to our data generating process and compare the jump activity estimators at these dynamics. First, we have a double price jump process, where the jumps are driven by two independent stable processes $Y^{(1)}$ and $Y^{(2)}$, with different jump activity indices. (β_1, β_2)

$$\begin{cases} dX_t = \sigma_t dW_t + \theta_1 dY_t^{(1)} + \theta_2 dY_t^{(2)} \\ dv_t = \kappa(\eta - v_t)dt + \gamma v_t^{1/2} dB_t \end{cases}$$

I always leave our previous parameters unchanged. I tune θ_1 and θ_2 to deliver different values of tail probability. $P_i = \mathbb{P}(\theta_i \Delta Y_t^{(i)} \geq \alpha \Delta_n^\varpi)$. I consider 3 cases: $P_1 \in \{0.25\%, 0.5\%, 1.00\%\}$, and $P_2 = 0.01P_1$. I examined the effect of β_2 in this DGP. While we still want to estimate β_1 , $\beta_2 = \beta_1 - \mu$, where $\mu \in \{0.1, 0.2, 0.4\}$. I note that estimating β_2 is also possible, but we need an exceptionally large number of observations in order to see through the continuous component and the first index. For related literature, see Aït-Sahalia and Jacod (2012b) and Bull (2016). Table 6 reports the outcomes.

Next, I add a jump component to the volatility process. I am interested in examining the robustness of these estimators in the context of a discontinuous volatility process. To do so, I will test the performance of each estimator on a dataset generated by this type of process.

$$\begin{cases} dX_t = \sigma_t dW_t + \theta dY_t \\ dv_t = \kappa(\eta - v_t)dt + \gamma v_t^{1/2} dB_t + dJ_t^v \end{cases}$$

In the presence of a volatility jump, a single jump occurs on a random arrival time. I assume that the magnitude of the volatility jump increment can be classified as small, with $dJ_t^v = \eta$ or either moderate, with $dJ_t^v = 3\eta$, or large, with $dJ_t^v = 10\eta$. The size of the volatility jump is expressed as a proportion of the long-run variance η . Similar concepts can be found in Maneesoonthorn et al. (2020). Table 7 contains the results.

Based on the work of Boswijk et al. (2018), and Erdemlioglu and Yang (2022) the existing DGPs can be extended with an intensity process, which controls the sizes and frequency of jumps. The theoretical model can be found in the Appendix. (Assumption 3)

$$\begin{cases} dX_t = \sigma_t dW_t + \lambda_t dY_t \\ dv_t = \kappa(\eta - v_t)dt + \gamma v_t^{1/2} dB_t \\ d\lambda_t = \kappa_\lambda(\lambda_\infty - \lambda_t)dt + \gamma_\lambda dB'_t + \xi \mathbb{1}_{\{\tau=t\}} + \zeta \mathbb{1}_{\{\Delta X_t > \varepsilon\}} \end{cases}$$

The main focus of this DGP and the novelty in this setting lies in the third equation where can be found the concept of allowing increments of X that exceed a certain threshold ε to stimulate the jump intensity, provided that ζ is greater than 0. I examine two cases, in the first one $\zeta = 0$, so there is no self-excitation, but an intensity jump ξ in the process, and in the second I will vary the ζ parameter to see its effect on the tests. I set $\kappa_\lambda = 30$, $\gamma_\lambda = 1.5\lambda_\infty$, $\varepsilon = \varepsilon_\lambda \sqrt{\theta} \Delta_n^\varpi$.² The mean reversion level λ_∞ is calibrated to deliver a pre-specified value of the tail probability. B'_t is a standard Brownian motion independent from W, B . I will tune ε_λ later.

The parameter κ_λ governs the rate at which the jump intensity process reverts to its long-run level following a jump in intensity. If the jump intensity reverts faster, the post-jump local average jump intensity will be lower, resulting in a weaker self-excitation effect that is less pronounced and shorter in duration. ζ controls the magnitude of the jump intensity process changes due to the self-excitation effect.

We talk about self-excitation in jumps if a jump in one component raises the probability of future jumps occurring in the same component. The formal definition is the following. An outcome $\omega \in \Omega$ exhibits self-excitation in jumps on the interval $(0, T]$ if λ jumps positively at some (not necessarily all) price jump times, i.e.

$$\sum_{0 \leq t \leq T} \Delta \lambda_t(\omega) \mathbb{1}_{\{\Delta \lambda_t(\omega) \geq 0, \Delta X_t(\omega) \neq 0\}}$$

It is important to note that the condition $\Delta X_t(\omega) \neq 0$ is still required in the indicator function mentioned above to ensure that positive jumps of λ not associated with jumps of X are excluded.

In the case of an intensity jump there is only one jump which occurs at a random time. When we set $\xi = 0$, there is no self-excitation in jumps on any sample path. There may, but need not, be self-excitation in jumps whenever $\xi > 0$.

Remark 14 *Observe that*

$$\frac{\mathbb{P}(|(\lambda_\infty + \xi)\Delta_i^n Y| \geq \alpha \Delta_n^\varpi)}{\mathbb{P}(|\lambda_\infty \Delta_i^n Y| \geq \alpha \Delta_n^\varpi)} \approx \left(\frac{\lambda_\infty + \xi}{\lambda_\infty} \right)^\beta = \left(1 + \frac{\xi}{\lambda_\infty} \right)^\beta.$$

If the ratio of tail probabilities is sufficiently large, it's implying an economically significant change in the tail probability after a large jump. In the case of smaller values, the

²Boswijk et al. (2018) use different parameter settings for the intensity process. The difference is because they have a sampling scheme with $T = 1$ (year), and $T = 5/252$ (1 week).

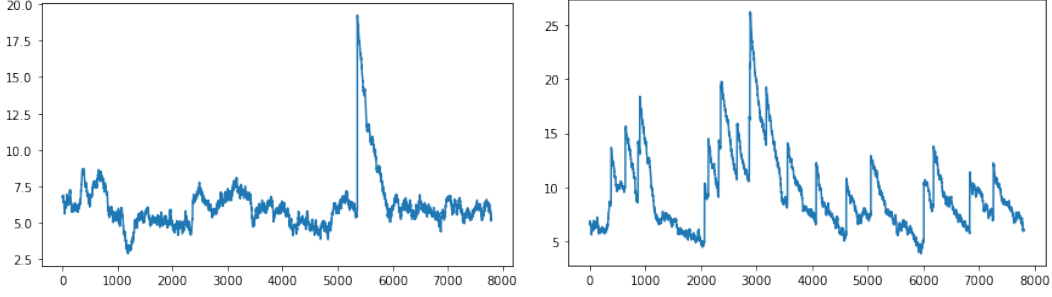


Figure 2: From left to right: jump intensity in case of an intensity jump, and self-excitation with parameters $\beta = 1, p = 1\%, \xi = 2.5\lambda_\infty, \varepsilon_\lambda = 9$.

ratio does not rise significantly, thus the effect is weaker. Based on these observations, I choose the jump values as follows: $\xi \in \{0.5\lambda_\infty, 2.5\lambda_\infty, 5\lambda_\infty\}$. The obtained results can be found in Table 7.

When there is self-excitation, I choose $\zeta = \lambda_\infty$. This ensures that our process behaves normally, and it doesn't explode. If there was no mean-reverting drift term in the intensity process it would be explosive if $\zeta > 0$ which we want to avoid. I vary $\varepsilon_\lambda \in \{9, 12, 18\}$. With these values, the threshold above which the price jumps trigger self-excitation are 0.374, 0.499, and 0.749. However, it does not guarantee with probability one that it will be self-excitation in realization. This parameter set is reasonable given that large log-price jumps, especially the negative ones, are more likely to trigger chain reactions than small ones. For the outcomes, see Table 8.

5 Simulation results

In the following tables, we can see the MSE and the number of outliers for the adequate estimators. In Table 1 I compare the Ait-Sahalia estimators to each other. AS alpha ($\hat{\beta}_n^{AJ}$) is based on two truncation levels, and AS freq ($\tilde{\beta}_n^{AJ}$) is based on two sampling frequencies. $\tilde{\beta}_n^{AJ}$ works well with smaller tail probability, but in some cases, especially in the case of $\beta = 1.75$ completely underestimates. Overall, $\hat{\beta}_n^{AJ}$ works more balanced, so I focus on this estimator from now on.

In general, we can say that higher tail probability generates more increments from the jump process that surpass the cutoff level, resulting in a greater number of observations and leading to a reduction in the standard deviation of the estimates. Also with higher β the error of the estimate is increasing. When β approaches 2, the behavior of the β -stable jump process is similar to a Brownian motion, causing the estimator to degrade in performance.

		MSE					Outliers				
		1.75	1.5	1.25	1	0.75	1.75	1.5	1.25	1	0.75
0.25%	AS alpha	0.287	0.234	0.170	0.072	0.065	44	47	48	43	55
	AS freq	0.183	0.094	0.069	0.073	0.044	47	32	37	50	49
0.50%	AS alpha	0.097	0.086	0.044	0.039	0.023	30	28	19	26	38
	AS freq	0.356	0.065	0.036	0.027	0.019	77	27	22	20	32
1.00%	AS alpha	0.074	0.040	0.019	0.017	0.018	31	11	8	12	25
	AS freq	0.800	0.077	0.024	0.018	0.011	100	33	11	13	15

Table 1: Monte Carlo simulations of the two Ait-Sahalia estimators on 1s sampling frequency on different β and tail probability values. $n = 100, T = 1$.

		MSE					Outliers				
		1.75	1.5	1.25	1	0.75	1.75	1.5	1.25	1	0.75
0.25%	Jing g1	0.086	0.039	0.023	0.026	0.018	33	15	10	21	26
	Jing g2	0.074	0.050	0.035	0.028	0.015	30	16	17	21	23
0.25%	Jing g1	0.053	0.019	0.010	0.008	0.007	23	3	0	1	10
	Jing g2	0.046	0.025	0.012	0.011	0.007	19	9	4	7	8
1.00%	Jing g1	0.099	0.014	0.005	0.004	0.003	57	0	1	0	2
	Jing g2	0.065	0.013	0.006	0.006	0.003	37	0	1	0	1

Table 2: Monte Carlo simulations of

In Table 2, I compare the two Jing estimators, Jing g1 ($\tilde{\beta}_n^J$) based on the g_1 function mentioned above, and Jing g2 ($\hat{\beta}_n^J$) based on g_2 . The two estimators perform similarly, but $\tilde{\beta}_n^J$ works worse in the case $\beta = 1.75$, and g_2 is much smoother than g_1 , so as Jing et al. in their paper I continue to work with $\hat{\beta}_n^J$.

In Table 3, I present the results of the Bull estimator, which offers valuable insights into the impact of various parameters on the estimation performance. Firstly, I analyzed the effect of the choice of m , and we observed that it plays a crucial role. The estimation with $m = 1$ performs much worse than the $m = 3$ case suggested by Bull. I also examined the impact of symmetrization (Bull WS is computed as this step is left out of the process) and found that it has a negligible effect on the results. After this, I looked at the truncation. Bull WT is programmed without truncation at 0 and 2. Naturally, in higher betas, the truncation improves the estimator, and Bull WT performed worse. Finally, I checked how the alternative method works in this case. At smaller beta values it gives really promising results, but it is slightly biased in the case of $\beta = 1.75$.

In Table 4, I stored the results of simulations taken at different frequencies: 1s, 3s, and 5s sampling frequencies, which means 23400, 7800, and 4680 data points. I adapted all parameters to the frequency, including the threshold, and the calibrated values also changed due to this. What immediately stands out is that the Ait-Sahalia and the Jing estimators give very bad results in smaller frequency. This is what we expected based on the previous sections. One of the reasons for the extremely bad values is that we do not truncate at 2, and in many cases, there is a very small number in the denominator (which corresponds to the jump counts with a higher cutoff value), thus significantly overestimating the jump activity index. Sometimes the Ait-Sahalia estimator returned that there is no jump on the trajectory, (the Jing estimator also returned a very small number) in this case, the cutoff level should be reduced. But the number of outliers shows that truncating at 2 would not solve the problem. The Bull estimator is much more robust to changes in frequency based on these simulations. The alternative method also produced very good results.

In Table 5 we can see the results with additive noise. The introduction of Gaussian noise generates a large number of small increments in the process, causing several jumps to cross the cutoff level. This affected the estimates that are most sensitive to changes caused by small jumps. Both Ait-Sahalia and Jing estimators started to get biased, and the alternative method failed to provide accurate estimates for small tail probabilities. The Bull estimator worked better than the others with the T-distributed noise as well, and in the mixture noise, it looks like the alternative method worked the best, but if we look at the number of outliers it tells us that this estimator overestimates especially in the case of $\beta = 1.75$. When I increased the magnitude of the noise, all the estimates quickly got even worse. Notably, the Bull method was the least affected by this deterioration.

			MSE					Outliers				
			1.75	1.5	1.25	1	0.75	1.75	1.5	1.25	1	0.75
Freq 1s	0.25%	AS	0.287	0.234	0.170	0.072	0.065	44	47	48	43	55
		Jing	0.074	0.050	0.035	0.028	0.015	30	16	17	21	23
		Bull	0.094	0.089	0.076	0.048	0.034	43	36	33	33	44
		Bull M	0.068	0.029	0.023	0.022	0.011	16	10	10	16	16
	0.50%	AS	0.097	0.086	0.044	0.039	0.023	30	28	19	26	38
		Jing	0.053	0.019	0.010	0.008	0.007	23	3	0	1	10
		Bull	0.057	0.052	0.037	0.024	0.016	37	23	14	22	25
		Bull M	0.169	0.018	0.008	0.006	0.005	62	1	1	2	6
	1.00%	AS	0.074	0.040	0.019	0.017	0.018	31	11	8	12	25
		Jing	0.065	0.013	0.006	0.006	0.003	37	0	1	0	1
		Bull	0.038	0.042	0.025	0.015	0.007	18	15	14	7	7
		Bull M	0.015	0.013	0.004	0.004	0.002	6	1	0	0	1
Freq 3s	0.25%	AS	14.785	13.144	9.654	8.473	3.292	85	81	77	82	97
		Jing	5.632	2.903	2.303	1.023	0.876	80	75	81	67	77
		Bull	0.356	0.203	0.198	0.166	0.117	72	62			

			MSE					Outliers				
			1.75	1.5	1.25	1	0.75	1.75	1.5	1.25	1	0.75
Gaussian noise	0.25%	AS	0.208	0.159	0.119	0.076	0.056	53	43	36	50	52
		Jing	0.102	0.077	0.067	0.029	0.028	42	31	30	23	32
		Bull	0.095	0.125	0.089	0.042	0.048	35	51	43	35	46
		Bull M	0.921	0.390	0.273	0.323	0.229	88	65	72	85	93
	0.50%	AS	0.127	0.075	0.060	0.033	0.029	41	28	32	29	32
		Jing	0.101	0.034	0.032	0.015	0.008	46	12	13	12	12
		Bull	0.075	0.065	0.047	0.037	0.016	36	24	28	30	29
		Bull M	0.076	0.075	0.073	0.057	0.054	24	30	41	43	61
	1.00%	AS	0.169	0.043	0.031	0.016	0.013	61	15	13	13	18
		Jing	0.133	0.025	0.013	0.009	0.004	74	3	5	3	1
		Bull	0.146	0.038	0.021	0.017	0.009	41	12	8	9	13
		Bull M	0.060	0.015	0.017	0.017	0.011	93	1	6	12	16
T-distributed noise	0.25%	AS	0.268	0.240	0.230	0.167	0.225	57	55	60	66	80
		Jing	0.128	0.150	0.179	0.112	0.161	59	49	71	72	93
		Bull	0.058	0.118	0.099	0.065	0.084	34	49			

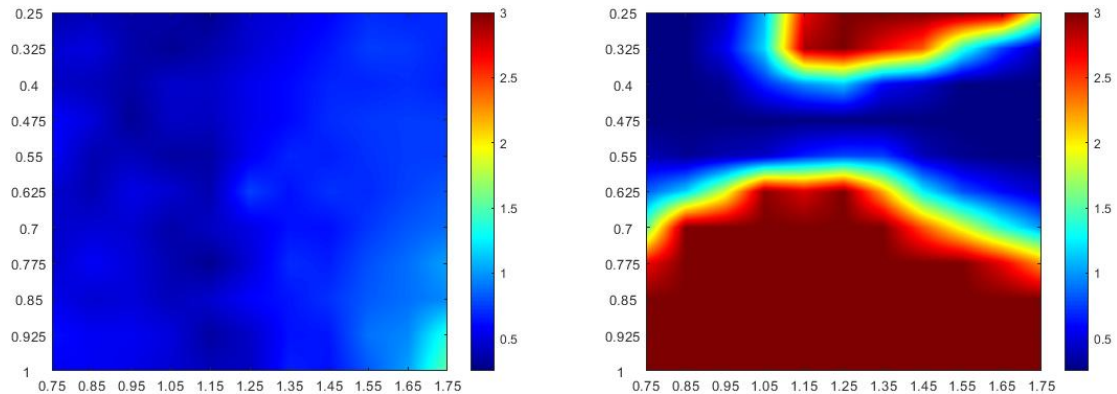
Now I want to check the robustness of the jump activity estimators in the presence of a volatility jump. (Table 7) As a result of volatility jumps, we actually add noise to the process. In general, we can notice that as the noise gets stronger, the Bull estimator works better and better compared to the Jing estimator. I checked higher magnitudes of jumps as well, and these observations only grew stronger. For $\beta = 1.75$, the Ait-Sahalia estimator appeared to be more resilient to a big volatility jump, which is the result of the g_2 function taking into account the smaller increments, which disturb the estimator.

It can be observed from Table 6 that when the DGP is driven by two independent stable processes, estimating in the low tail probability is very challenging. In lower betas, the Jing estimator worked well, but in $\beta = 1.75$, so in the most difficult case to estimate the Bull estimator performed the best. The alternative method was unable to deal with this type of noise.

The presence of an intensity jump made the tail probability random. This case did not make a big difference in the results (Table 8), in fact, the Ait-Sahalia estimator even became a little bit better due to the increase in tail probability followed by the intensity jump. But again, the $\beta = 1.75$ case was the most difficult to handle.

However, when self-excitation is observed in the data, we see something different in Table 9. When $\varepsilon_\lambda = 9$, the intensity process has a lot of jumps and takes on larger values, particularly in the case of higher tail probability where more jumps exceed the cutoff value. As a result, the methods had difficulty estimating the jump activity index accurately. However, as ε_λ increased, the estimators started to return to their natural behavior.

In these heatmaps, we can see the relative MSE of the Jing (left) and the Bull (right) estimators compared to the Ait-Sahalia estimator. Values above 1 mean worse, and below 1 mean better results. The figures are based on the original stochastic volatility model shown in Section 4.1. On the x-axis, the different betas are shown, while the y-axis represents the various tail probabilities.



			MSE					Outliers				
			1.75	1.5	1.25	1	0.75	1.75	1.5	1.25	1	0.75
$\mu = 0.1$	0.25%	AS	0.25	0.136	0.065	0.046	0.051	61	36	26	35	38
		Jing	0.142	0.056	0.032	0.024	0.021	42	18	11	26	28
		Bull	0.081	0.103	0.064	0.043	0.023	58	37	28	31	31
		Bull M	3.024	1.1	0.625	0.575	0.536	100	100	100	100	100
	0.50%	AS	0.138	0.068	0.041	0.033	0.026	44	27	18	21	38
		Jing	0.092	0.027	0.018	0.012	0.009	50	7	4	7	10
		Bull	0.056	0.044	0.031	0.021	0.02	40	16	14	12	33
		Bull M	0.804	0.229	0.131	0.111	0.155	98	84	78	90	100
	1.00%	AS	0.126	0.043	0.019	0.011	0.008	55	15	9	7	11
		Jing	0.122	0.025	0.01	0.005	0.003	71	2	2	0	0
		Bull	0.102	0.025	0.012	0.008	0.01	26	3	2	4	12
		Bull M	0.061	0.012	0.024	0.027	0.031	98	0	4	23	62
$\mu = 0.2$	0.25%	AS	0.182	0.09	0.079	0.055	0.054	50	31	34	32	51
		Jing	0.137	0.045	0.038	0.025	0.017	44	14	16	23	27
		Bull	0.128	0.083	0.077	0.049	0.02	67	33	39	34	

			MSE					Outliers				
			1.75	1.5	1.25	1	0.75	1.75	1.5	1.25	1	0.75
$1*\eta$	0.25%	AS	0.19	0.159	0.134	0.095	0.067	46	51	41	46	55
		Jing	0.072	0.058	0.04	0.039	0.019	27	21	24	28	24
		Bull	0.084	0.099	0.058	0.068	0.031	49	32	29	42	37
		Bull M	0.061	0.036	0.021	0.021	0.012	16	12	10	18	17
	0.50%	AS	0.119	0.071	0.043	0.033	0.038	36	23	17	25	43
		Jing	0.049	0.03	0.022	0.011	0.011	24	11	10	4	17
		Bull	0.059	0.077	0.046	0.023	0.017	37	35	24	14	25
		Bull M	0.151	0.016	0.012	0.007	0.007	56	3	0	1	7
	1.00%	AS	0.085	0.042	0.03	0.016	0.009	37	14	14	13	9
		Jing	0.059	0.016	0.013	0.009	0.003	33	3	1	4	2
		Bull	0.049	0.039	0.031	0.014	0.008	21	12	15	9	11
		Bull M	0.018	0.01	0.007	0.005	0.003	7	0	0	0	1
$3*\eta$	0.25%	AS	0.207	0.117	0.118	0.097	0.052	50	35	40	45	57
		Jing	0.055	0.048	0.053	0.03	0.018	25	18	19	27	21
		Bull	0.084	0.107	0.079	0.056	0.032	45	45	39		

			MSE					Outliers				
			1.75	1.5	1.25	1	0.75	1.75	1.5	1.25	1	0.75
$0.5\lambda_\infty$	0.25%	AS	0.254	0.164	0.107	0.082	0.049	54	42	35	46	50
		Jing	0.091	0.045	0.04	0.03	0.022	32	15	22	22	31
		Bull	0.075	0.1	0.071	0.05	0.034	55	35	36	41	45
		Bull M	0.078	0.027	0.027	0.017	0.013	22	3	11	11	18
	0.50%	AS	0.116	0.063	0.078	0.031	0.028	34	20	26	24	35
		Jing	0.057	0.022	0.02	0.01	0.011	32	4	11	5	13
		Bull	0.048	0.049	0.048	0.025	0.016	34	16	23	17	19
		Bull M	0.08	0.013	0.01	0.006	0.007	19	0	2	2	5
	1.00%	AS	0.105	0.039	0.023	0.02	0.014	42	9	8	15	23
		Jing	0.082	0.019	0.01	0.007	0.005	45	4	1	1	6
		Bull	0.082	0.035	0.022	0.015	0.008	31	10	9	9	11
		Bull M	0.047	0.01	0.006	0.005	0.003	60	0	0	0	1
$2.5\lambda_\infty$	0.25%	AS	0.186	0.135	0.082	0.083	0.051	53	40	36	43	47
		Jing	0.081	0.048	0.032	0.027	0.016	32	14	15	20	23
		Bull	0.084	0.087	0.06	0.06	0.021	41	35	30	39</	

			MSE					Outliers				
			1.75	1.5	1.25	1	0.75	1.75	1.5	1.25	1	0.75
eps = 9	0.25%	AS	0.24	0.169	0.09	0.061	0.055	47	39	38	43	50
		Jing	0.093	0.046	0.035	0.019	0.016	34	15	19	14	22
		Bull	0.086	0.087	0.066	0.039	0.028	56	29	33	31	37
		Bull M	0.06	0.027	0.019	0.011	0.01	18	6	7	7	11
	0.50%	AS	0.14	0.056	0.029	0.02	0.016	48	21	14	17	23
		Jing	0.097	0.023	0.013	0.008	0.005	53	4	4	2	8
		Bull	0.09	0.05	0.04	0.016	0.008	39	17	16	10	8
		Bull M	0.044	0.011	0.006	0.005	0.003	48	1	0	1	1
	1.00%	AS	0.675	1.092	1.01	0.034	0.002	94	68	86	33	0
		Jing	0.634	1.099	1.009	0.031	0.001	99	62	85	26	0
		Bull	0.503	1.131	1.017	0.014	0.002	62	67	86	11	0
		Bull M	0.466	1.09	1.007	0.045	0.001	98	82	85	36	0
eps = 12	0.25%	AS	0.224	0.17	0.077	0.069	0.048	55	42	32	41	43
		Jing	0.079	0.056	0.034	0.03	0.015	27	15	14	22	19
		Bull	0.076	0.087	0.058	0.05	0.026	49	35	28	33	29

6 Conclusions and Future work

The Jing estimator outperforms the Ait-Sahalia estimator in every aspect, as we can see in the heatmap as well. It works fine in nearly every case, but in certain disturbing noises, it also gets slightly biased. This is because the g_2 function is quite sensitive for the small jumps. The Bull estimator works more robustly, especially in smaller sample sizes, but sometimes underperforms the estimator defined by Jing.

If microstructure noise is present in the data, we need a method to smoothen it, because it affects the estimators badly, especially the alternative estimator. Maybe averaging different weighting functions work, and introducing a beta-dependent estimator eliminates the difficulty of estimating $\beta = 0.75$ and 1.75 with the same estimator. In future works, I will try to create a robust estimator and examine the results if additive noise is added to the different DGPs. Additionally, I plan to conduct an empirical analysis using high-frequency cryptocurrency data.

7 Appendix

Assumption 3 *There are three constants $\beta \in (0, 2)$, $\beta' \in [0, \beta)$, and $\gamma > 0$, and a locally bounded process $L_t \geq 1$ such that for all (ω, t) , $F_t = F'_t + F''_t$ where:*

1. $F'_t(\omega, dx) = f'_t(x)\lambda_{t-}(\omega)$, $0 < \lambda_t(\omega) \leq L_t$, and $f'_t(x) = \frac{1+|x|^{\gamma}h_t(x)}{|x|^{1+\beta}}$, with $1 + |x|^{\gamma}h_t(x) \geq 0$, $|h_t(x)| \leq L_t$.
2. F''_t is a singular measure with respect to F'_t and satisfies $\int_{\mathbb{R}}(|x|^{\beta'} \wedge 1)F''_t(dx) \leq L_t$.
3. The process λ_t is an Itô semimartingale given by

$$\lambda_t = \lambda_0 + \int_0^t b'_s ds + \int_0^t \sigma'_s dW_s + \int_0^t \sigma''_s dB_s + \delta' * \mu_t + \delta'' * \mu_t^\perp,$$

where B is a standard Brownian motion independent of W , $\mu_t^\perp$ is orthogonal to μ_t which means that almost surely they share no common Dirac masses lying on the same vertical line, so, in other words, they have no common jump times.

Remark 15 *Take any process of the form*

$$dX_t = b_t dt + \sigma_t dW_t + \delta_{t-} dY_t^{(1)} + \delta'_{t-} dY_t^{(2)}$$

where δ and δ' are càdlàg adapted processes, $Y^{(1)}$ is β -stable process, $Y^{(2)}$ is any other Lévy process whose Lévy measure integrates $|x|^{\beta'}$ near the origin and has an absolutely continuous part whose density is smaller than $K|x|^{\gamma-1-\beta}$ on $[-1, 1]$ for some $\gamma > 0$ (e.g., a stable process with index strictly smaller than β). Then, X will satisfy Assumption 1.

References

- Aït-Sahalia, Y. and Jacod, J. (2007). Volatility estimators for discretely sampled lévy processes.
- Aït-Sahalia, Y. and Jacod, J. (2009a). Estimating the degree of activity of jumps in high frequency data.
- Aït-Sahalia, Y. and Jacod, J. (2009b). Testing for jumps in a discretely observed process. *The Annals of Statistics*, pages 184–222.
- Aït-Sahalia, Y. and Jacod, J. (2012a). Analyzing the spectrum of asset returns: Jump and volatility components in high frequency data. *Journal of Economic Literature*, 50(4):1007–1050.
- Aït-Sahalia, Y. and Jacod, J. (2012b). Identifying the successive blumenthal–gettoor indices of a discretely observed process.
- Aït-Sahalia, Y. and Jacod, J. (2014). High-frequency financial econometrics. In *High-Frequency Financial Econometrics*. Princeton University Press.
- Aït-Sahalia, Y., Jacod, J., and Li, J. (2012). Testing for jumps in noisy high frequency data. *Journal of Econometrics*, 168(2):207–222.
- Barndorff-Nielsen, O. E. and Shephard, N. (2006). Econometrics of testing for jumps in financial economics using bipower variation. *Journal of financial Econometrics*, 4(1):1–30.
- Belomestny, D. (2010). Spectral estimation of the fractional order of a lévy process.
- Bibinger, M. and Winkelmann, L. (2015). Econometrics of co-jumps in high-frequency data with noise. *Journal of Econometrics*, 184(2):361–378.
- Boswijk, H. P., Laeven, R. J., and Yang, X. (2018). Testing for self-excitation in jumps. *Journal of Econometrics*, 203(2):256–266.
- Bull, A. D. (2016). Near-optimal estimation of jump activity in semimartingales.
- Cont, R. and Mancini, C. (2007). Nonparametric tests for analyzing the fine structure of price fluctuations.
- Erdemlioglu, D. and Yang, X. (2022). News arrival, time-varying jump intensity, and realized volatility: Conditional testing approach. *Journal of Financial Econometrics*, *Forthcoming*.

- Fan, J. and Wang, Y. (2007). Multi-scale jump and volatility analysis for high-frequency financial data. *Journal of the American Statistical Association*, 102(480):1349–1362.
- Jacod, J. (2008). Asymptotic properties of realized power variations and related functionals of semimartingales. *Stochastic processes and their applications*, 118(4):517–559.
- Jacod, J. and Protter, P. (2011). *Discretization of processes*, volume 67. Springer Science & Business Media.
- Jing, B.-Y., Kong, X.-B., and Liu, Z. (2011). Estimating the jump activity index under noisy observations using high-frequency data. *Journal of the American Statistical Association*, 106(494):558–568.
- Jing, B.-Y., Kong, X.-B., Liu, Z., and Mykland, P. (2012). On the jump activity index for semimartingales. *Journal of Econometrics*, 166(2):213–223.
- Kolokolov, A. (2022). Estimating jump activity using multipower variation. *Journal of Business & Economic Statistics*, 40(1):128–140.
- Lee, S. S. and Mykland, P. A. (2008). Jumps in financial markets: A new nonparametric test and jump dynamics. *The Review of Financial Studies*, 21(6):2535–2563.
- Mancini, C. and Gobbi, F. (2012). Identifying the brownian covariation from the co-jumps given discrete observations. *Econometric Theory*, 28(2):249–273.
- Maneesoonthorn, W., Martin, G. M., and Forbes, C. S. (2020). High-frequency jump tests: Which test should we use? *Journal of econometrics*, 219(2):478–487.
- Theopold, A. and Vetter, M. (2023). On the estimation of the jump activity index in the case of random observation times. *Japanese Journal of Statistics and Data Science*, pages 1–47.
- Todorov, V. and Tauchen, G. (2010). Activity signature functions for high-frequency data analysis. *Journal of Econometrics*, 154(2):125–138.
- Todorov, V. and Tauchen, G. (2011). Volatility jumps. *Journal of Business & Economic Statistics*, 29(3):356–371.
- Woerner, J. H. (2011). Analyzing the fine structure of continuous time stochastic processes. In *Seminar on Stochastic Analysis, Random Fields and Applications VI: Centro Stefano Franscini, Ascona, May 2008*, pages 473–492. Springer.
- Zhao, Z. and Wu, W. B. (2009). Nonparametric inference of discretely sampled stable lévy processes. *Journal of Econometrics*, 153(1):83–92.