
Volume 166
Issue 2
February 2012
ISSN 0304-4076
ELSEVIER
JOURNAL OF
Econometrics
CONTENTS
K. Akashi and N. Kunitomo, Some properties of the LIML estimator in a dynamic panel structural equation 167
M. Burda, M. Harding and J. Hausman, A Poisson mixture model of discrete choice 184
J.T. Fox, K.i. Kim, S.P. Ryan and P. Bajari, The random coefficients logit model is identified 204
B.-Y. Jing, X.-B. Kong, Z. Liu and P. Mykland, On the jump activity index for semimartingales 213
X. Wei and Y. Yang, Robust forecast combinations 224
CONTENTS CONTINUED ON BACK COVER

Journal of Econometrics 166 (2012) 213–223

Contents lists available at SciVerse ScienceDirect

Journal of Econometrics

journal homepage: www.elsevier.com/locate/jeconom

On the jump activity index for semimartingales

Bing-Yi Jing *, Xin-Bing Kong, Zhi Liu, Per Mykland

Hong Kong University of Science and Technology, Hong Kong
Fudan University, China
Xiamen University, China
University of Chicago, United States

A R T I C L E I N F O

Article history:
Received 16 December 2010
Received in revised form
7 June 2011
Accepted 6 September 2011
Available online 28 September 2011

Keywords:
Semimartingale
Power variation
High frequency
Jump activity index
Stable convergence

A B S T R A C T

Empirical evidence of asset price discontinuities or "jumps" in financial markets has been well documented in the literature. Recently, Aït-Sahalia and Jacod (2009b) defined a general "jump activity index" to describe the degree of jump activities for asset price semimartingales, and provided a consistent estimator when the underlying process contains both a continuous and a jump component. However, only large increments were used in their estimator so that the effective sample size is very small even for large sample sizes. In this paper, we explore ways to improve the Aït-Sahalia and Jacod estimator by making use of all increments, large and small. The improvement is verified through simulations. A real example is also given.

1. Introduction

Ito's semimartingales are widely used in modeling asset prices in financial markets. They are a rich class of stochastic processes including diffusion, jump diffusion, Lévy processes, and so on. Recent years have seen a rapidly increasing interest in semimartingales with the discontinuous part (i.e., jumps) in the literature; see Aït-Sahalia (2004), Aït-Sahalia and Jacod (2007), Fan and Wang (2007), Jacod (2008) and references therein.

Is there indeed a jump part in asset prices? With the availability of high frequency data, many tests have been established in the statistical literature to detect jumps from discretely observed prices, and found evidence of the presence of jumps. See Aït-Sahalia (2002), Jiang and Oomen (2005), Barndorff-Nielsen and Shephard (2006), Lee and Mykland (2007) and Aït-Sahalia and Jacod (2009a), just to name a few. Furthermore, many empirical studies in the literature show strong evidence of the existence of jumps; see, e.g., Carr et al. (2002) and the references therein.

Given that the discontinuous part is present, a natural question for the purpose of modeling is to study the behavior of the jumps or the jump characteristics. As a natural measure of the activity of jumps, Aït-Sahalia and Jacod (2009b) defined a *jump activity index* for a generic semimartingale X as follows:

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

where $\Delta_s X = X_s - X_{s-}$ is the jump size at time s. This jump activity index essentially characterizes how frequently small jumps occur which is closely related to the near origin behavior of the Lévy measure for semimartingales. In particular, when X is a Lévy process, β_t is equivalent to the Blumenthal–Getoor index defined in terms of the Lévy measure as

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

where $\nu(dx)$ is the density of the Lévy measure. For a stable Lévy process, the jump activity index or the Blumenthal–Getoor index is just the stable index.

The jump activity index β_t can be used for different purposes. From a modeling viewpoint, it could be used to judge whether the jump part of a semimartingale has finite variation or not. From a financial viewpoint, jump processes have been introduced since they enable to reproduce various stylized facts of prices such as heavy tails and big jumps. They are also very useful for reproducing empirical features linked to options such as smile or skew under risk neutral measure. For example, Merton's compound Poisson jump specification is suitable to capture large and rare events such as market crashes and corporate defaults. However, as noted by Wu (2008), empirical evidences showed that the inclusion of infinite jumps with all sizes could not only be better suited to capture the movements of many financial securities, but also generate better option pricing performance.

* Correspondence to: Hong Kong University of Science and Technology, Hong Kong.
E-mail addresses: majing@ust.hk (B.-Y. Jing), kongxb@fudan.edu.cn (X.-B. Kong), liuzhi@ust.hk (Z. Liu), per.mykland@gmail.com (P. Mykland).

0304-4076/$ – see front matter © 2011 Elsevier B.V. All rights reserved.
doi:10.1016/j.jeconom.2011.09.036

Various estimators of the jump activity index have been given in the literature for high frequency financial data. We can classify them into two categories. For ease of exposition, let us decompose the semimartingale X into

$$X_t = X_t^c + X_t^d,$$

where X_t^c and X_t^d represent the continuous and discontinuous parts, respectively.

In the first category, it is assumed that the continuous local martingale part is absent, i.e., $X_t^c \equiv 0$. Some literature on estimating the jump activity index is available in this case. Woerner (2006) and Todorov and Tauchen (2010) proposed estimators of the jump activity index for a large class of stable-like semimartingales using power variations; Zhao and Wu (2009) gave a nonparametric estimator of the stable index of the driving stable Lévy process with a deterministic time dependent integrand by two time scale techniques. Incidentally, Todorov and Tauchen (2009) also studied the inference on the activity signature functions for high frequency data.

In the second category, it is assumed that the continuous martingale part is present, i.e., $X_t^c \not\equiv 0$. (We are primarily interested in this category in the present paper as well.) In this case, estimating β_t becomes much more challenging than those in the first category. The reason is that the value of β_t is determined by the small jumps from the discontinuous part X_t^d, which are, unfortunately, "contaminated" with the small increments coming from the continuous component. To get around the problem, Aït-Sahalia and Jacod (2009b) proposed an estimator of β_t by counting the number of "large" increments of a discretely observed semimartingale. It should be noted that, as the threshold goes to zero, their method will capture all the jumps asymptotically. The resulting estimator is consistent and asymptotically normal after being properly standardized.

Despite its initial success, the Aït-Sahalia and Jacod estimator is not trouble free. The major problem is that the effective sample size is very small even for very large sample sizes. As a result, too few observations are used in estimating β_t, resulting in a loss of efficiency. For instance, simulations with 23,400 observations from time 0 to time 1 (corresponding to an intra-day data set with one observation per second) from the Cauchy process retain about 15 usable observations, accounting for a mere 0.0006 proportion of all observations. We will elaborate more on this point in detail later. Of course, these issues were well recognized in Aït-Sahalia and Jacod (2009b), who pointed out, "*this paper represents only a first attempt at measuring the degree of jump activity*".

In this paper, we will explore ways to improve the Aït-Sahalia and Jacod estimator with a view to increase efficiency. The proposed estimator, given in Section 3, makes full use of all increments, both "large" and "small". The "larger" increments are of higher quality, and are used much in the same way as in Aït-Sahalia and Jacod (2009b). However, the "smaller" increments are of lesser quality due to "contaminations" from the continuous component, hence given less weight in the estimator. By doing so, the effective sample size is increased, resulting in more efficiency, which is confirmed both theoretically and in simulations. In fact, our simulations show that the mean squared errors (MSE) of the new estimator have been reduced across the board (by as much as 30% in some cases), compared with those of Aït-Sahalia and Jacod (2009b).

This paper is organized as follows: In Section 2, we specify a semimartingale model and present a review of the Aït-Sahalia and Jacod estimator. Our estimator is presented in Section 3. Asymptotic properties of the new estimator is provided in Section 4. Simulations are carried out in Section 5. A real example is given in Section 6. Proofs are deferred to Section 7.

2. Model settings and a review

Our model setting and assumptions are much the same as in Aït-Sahalia and Jacod (2009b). For completeness, we will briefly list them below.

2.1. Model assumptions

Consider a one-dimensional asset price process X_t on the probability space $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}, P)$, which is an Ito semi-martingale defined by Jacod and Shiryaev (2003) with the form

$$X_t = X_0 + \int_0^t b_s ds + \int_0^t \sigma_s dW_s + \int_0^t \int_{|x|\le 1} x(\mu - \nu)(ds, dx) + \int_0^t \int_{|x|>1} x\mu(ds, dx),$$

where, W_t is a standard Brownian motion, b and σ are optional processes, and μ is a random measure related to the count of jumps with compensator ν given by $\nu(dt, dx) = dtF_t(dx)$. Instead of working with the decomposition involving the jump measure associated to the process, in this paper, we choose to work with the Poisson random measure as done in Aït-Sahalia and Jacod (2009b).

We make the following assumptions.

Assumption 1. b and σ are locally bounded.

Assumption 2. There are three constants $\beta \in (0, 2)$, $\beta' \in [0, \beta)$ and $\gamma > 0$ and a locally bounded process $L_t \ge 1$, such that we have for all (ω, t):

$$F_t = F_t' + F_t'',$$

where

1. F_t' has the form
$$F_t'(dx) = \frac{1 + |x|^\gamma f(t, x)}{|x|^{1+\beta}} (a_t^{(+)} I(0 < x \le z_t^{(+)}) + a_t^{(-)} I(-z_t^{(-)} \le x < 0))dx,$$
for predictable non-negative processes $a_t^{(+)}$, $a_t^{(-)}$, $z_t^{(+)}$ and $z_t^{(-)}$ and some predictable function $f(t, x)$, satisfying for some positive constants $\overline{K}$, $\overline{\eta}$:
$$\begin{cases} a_t^{(+)} + a_t^{(-)} \le L_t, & 1/L_t \le z_t^{(+)} \le 1; \\ 1/L_t \le z_t^{(-)} \le 1; & 1 + |x|^\gamma f(t, x) \ge 0, \quad |f(t, x)| \le L_t, \end{cases}$$
2. F_t'' is singular with respect to F_t', such that $\int (|x|^{\beta'} \wedge 1) F_t''(dx) \le L_t$.

Under Assumption 2, the jumps activity index becomes β. We also need the following notation: $A_t = \int_0^t (a_s^{(+)} + a_s^{(-)}) ds$.

Assume that over a fixed time interval $[0, T]$, we have observations X_{t_i} at equally spaced discrete times $0 = t_0 \le \cdots \le t_i \le t_{i+1} \le \cdots \le t_n = T$ with $\Delta_n = T/n$. Denote the increment in the ith interval by

$$\Delta_i^n X = X_{t_i} - X_{t_{i-1}}.$$

2.2. A review of the Aït-Sahalia and Jacod estimator

The basestone of Aït-Sahalia and Jacod (2009b) is

$$U(\varpi, \alpha)_t^n =: \sum_{i=1}^{[t/\Delta_n]} I(|\Delta_i^n X| > \alpha \Delta_n^\varpi), \tag{2.1}$$

where $\alpha > 0$ and $0 < \varpi < 1/2$ are two constants. Basically, $U(\varpi, \alpha)_t^n$ counts the number of "large" increments which contain information on "large" jumps. Aït-Sahalia and Jacod (2009b) showed that

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

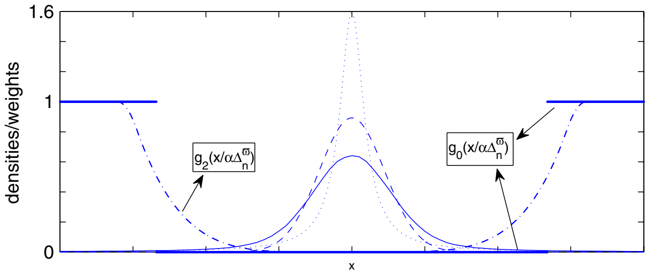

Fig. 1. The pdf's of $\Delta_i^n W$ (broken curve), $\Delta_i^n Y$ (dotted curve) and $\Delta_i^n X$ (solid curve), and the weight functions $g_0(x)$ (piecewise point line) and $g_2(x)$ (dot-dashed curve) which are defined in Section 3. The pdf's of $\Delta_i^n W$ and $\Delta_i^n X$ almost coincide beyond $\alpha\Delta_n^{\varpi}$ (7 standard deviation of $\Delta_i^n W$ suggested by Aït-Sahalia and Jacod (2009b)).

From this and using another α', they proposed to estimate β by:

$$\overline{\beta}_n(t,\varpi,\alpha,\alpha') =: \log\frac{U(\varpi,\alpha)_t^n}{U(\varpi,\alpha')_t^n}\Big/\log\left(\frac{\alpha'}{\alpha}\right), \tag{2.2}$$

and showed that $\overline{\beta}_n$ was consistent and asymptotically normally distributed.

The main drawback of the Aït-Sahalia and Jacod estimator is that the effective sample size utilized by the estimator $\overline{\beta}_n$ is small, even if we sample at a relatively high frequency. To see why, let us consider the special model:

$$X_t = W_t + Y_t, \tag{2.3}$$

where W_t is a standard Brownian motion and Y_t is a β-stable process, so that $\beta_t = \beta$. Any increment $\Delta_i^n X$ satisfies

$$\Delta_i^n X = \Delta_i^n W + \Delta_i^n Y =_d \Delta_n^{1/2} W_1 + \Delta_n^{1/\beta} Y_1, \tag{2.4}$$

where $=_d$ means equivalence in distribution. Given that there is a large increment satisfying $|\Delta_i^n X| \ge \alpha\Delta_n^{\varpi}$, in view of $\varpi < 1/2$, it is almost certain that the major contribution is due to Y. This is because $P(|\Delta_i^n W| \ge \alpha\Delta_n^{\varpi})$ converges to 0 at an exponential rate which is faster than any power of n^{-1}, hence the average number of increments $\ge \alpha\Delta_n^{\varpi}$ is

$$nP(|\Delta_i^n W| \ge \alpha\Delta_n^{\varpi}) \approx 0. \tag{2.5}$$

On the other hand, we have $P(|\Delta_i^n Y| \ge \alpha\Delta_n^{\varpi}) = P(|Y_1| \ge \alpha\,\Delta_n^{\varpi-1/\beta}) \sim 2\alpha^{-\beta}\pi^{-1}/n^{1-\varpi\beta}$ hence the average number of increments $\ge \alpha\Delta_n^{\varpi}$ is

$$nP(|\Delta_i^n Y| \ge \alpha\Delta_n^{\varpi}) \sim 2\alpha^{-\beta}\pi^{-1}n^{\varpi\beta} \longrightarrow \infty. \tag{2.6}$$

Fig. 1 plots the probability densities of $\Delta_i^n W$, $\Delta_i^n Y$ and $\Delta_i^n X$.

For numerical illustration, let us take Y to be a Cauchy process (i.e., $\beta = 1$) with $T = 1$ (day), and $n = 23{,}400$. Then the observations correspond to an intra-day data set with one observation per second. We take $\varpi = 1/5$, $\alpha = 5/16$ and $\alpha' = 2\alpha$, the same values as in Aït-Sahalia and Jacod (2009b) in their simulations. By (2.6), the average number of increments exceeding $\alpha\Delta_n^{\varpi}$ (and $\alpha'\Delta_n^{\varpi}$) are approximately 15 (and 7.5). This accounts for a mere 0.0006 (and 0.0003) proportion of the total observations, which is very small indeed.

3. A new estimator

3.1. Motivation

Note that $U(\varpi,\alpha)_t^n$ in (2.1) in the Aït-Sahalia and Jacod estimator $\overline{\beta}_n$ only counts the number of large increments ($\ge\alpha\Delta_n^{\varpi}$). If we define the weight function as

$$g_0(x) = I\{|x| > 1\},$$

then we can rewrite (2.1) as

$$U(\varpi,\alpha)_t^n =: \sum_{i=1}^{[t/\Delta_n]} I\left(\left|\frac{\Delta_i^n X}{\alpha\Delta_n^{\varpi}}\right| > 1\right) = \sum_{i=1}^{n} g_0\left(\frac{\Delta_i^n X}{\alpha\Delta_n^{\varpi}}\right).$$

We remark that, even though $U(\varpi,\alpha)$ only made use of a fraction of all jumps, it catches more and more jumps as n increases; see (2.6) for example.

In an effort to make complete use of the data, perhaps one should not only consider the "large" increments, but also those "small" ones. After all, the jump activity index is an index of small jumps. From Fig. 1, we note that, if the contribution of the increments of the diffusion term could be controlled properly, there is still room to dig out information on β from relatively small increments. However, the smaller the increment $\Delta_i^n X$ gets, the greater the contributions from $\Delta_i^n W$ become, and consequently, the less usable information about β the increment $\Delta_i^n X$ will contain. This motivates us to define

$$V(\varpi,\alpha,g)_t^n =: \sum_{i=1}^{[t/\Delta_n]} g\left(\frac{\Delta_i^n X}{\alpha\Delta_n^{\varpi}}\right), \tag{3.7}$$

where $g(t)$ decreases to 0 as $|t|$ goes to 0.

How do we choose the weight function $g(x)$ in practice? *Assuming that the continuous martingale is present*, it is known that (c.f., Jacod, 2008),

(i) for $p > 2$, $\sum_{i=1}^n |\Delta_i^n X|^p \to^P \sum_{0\le s\le T} |\Delta_s X|^p$,
(ii) for $p = 2$, $\sum_{i=1}^n |\Delta_i^n X|^2 \to^P \int_0^T \sigma_s^2 ds + \sum_{0\le s\le T} |\Delta_s X|^2$,
(iii) for $p < 2$, $\Delta_n^{1-p/2}\sum_{i=1}^n |\Delta_i^n X|^p \to^P E|\mathcal{N}(0,1)|^p \int_0^T |\sigma_s|^p ds$, where $\mathcal{N}(0,1)$ is a standard normal r.v.

We can see that only for $p > 2$, the limits of $\sum_{i=1}^n |\Delta_i^n X|^p$ (or properly scaled) depend solely on jumps, i.e., the influence from the continuous part is completely eliminated. Therefore, possible weight functions are those having a power form near the origin. In this paper, we will restrict attention to the following simple but flexible class of weight functions:

Assumption 3. $g(x) = |x|^p$ if $|x| \le a$ 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.

Example 1. Let

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

A simple choice of g satisfying Assumption 3 is

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

where $0 < a < b < \infty$ are two constants, and $c = a^p + pa^{p-1}(b-a)/2$. Fig. 2 illustrates the shape of these weight functions. The three weight functions g_i for $i = 0, 1, 2$ are closely related:

(i) $g_2(x)$ is a smoother version of $g_1(x)$; if $a = b = 1$, $g_2(x) = g_1(x)$.
(ii) $g_1(x) \to g_0(x)$ as $p \to \infty$.
(iii) The larger the value of p, the lesser the weight on small increments.

It turns out that, for properly chosen p and under Assumptions 1–3, we have

$$\Delta_n^{\varpi\beta} V(\varpi,\alpha,g)_t^n \longrightarrow^P \frac{A_t}{\alpha^\beta}\int_0^\infty \frac{g(x)}{x^{1+\beta}}dx. \tag{3.8}$$

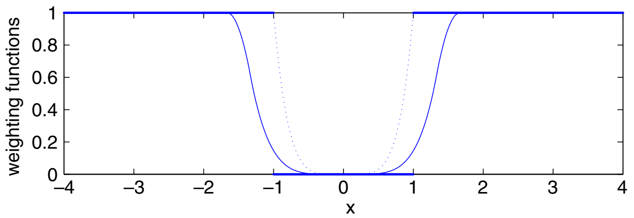

Fig. 2. Plots of $g_0(x)$ (piecewise point line), $g_1(x)$ (dashed curve) and $g_2(x)$ (solid curve) with $a = 6/5$ and $b = 7/5$, where $p = 6$.

From this and for $0 < \alpha < \alpha'$, we can define an estimator of β by

$$\hat{\beta}_n(t, \varpi, \alpha, \alpha') =: \log \frac{V(\varpi, \alpha, g)_t^n}{V(\varpi, \alpha', g)_t^n} \Big/ \log\left(\frac{\alpha'}{\alpha}\right). \tag{3.9}$$

4. Asymptotic results

4.1. General cases

Here, we will establish the consistency and asymptotic normality of the newly proposed estimator $\hat{\beta}_n$. Throughout the paper, we will always assume that

(i) $0 < \alpha < \alpha'$, $0 < \varpi < 1/2$, and $t > 0$;
(ii) Assumptions 1–3 hold;

Before stating our theorems, we introduce some notation. Let $\bar{g}(\alpha, \alpha', x) = g(\alpha x/\alpha')$ and $\tilde{g}(\alpha, \alpha', x) = g(x)\bar{g}(\alpha, \alpha', x)$. Denote $C_\beta(k) = \int_0^\infty g^k(x)/x^{1+\beta}dx$, $k = 1, 2$, and $C'_\beta = \int_0^\infty \tilde{g}(\alpha, \alpha', x)/x^{1+\beta}dx$.

Theorem 1. *Assume that $p > (2 - \varpi\beta)/(1 - 2\varpi)$, $0 \le \beta' < \beta/2$, $\gamma > \beta/2$, and $\varpi < 1/(2 + \beta)$.*

1. *On the set $\{A_t > 0\}$, we have*

$$\frac{1}{\Delta_n^{\varpi\beta/2}}(\hat{\beta}_n(t, \varpi, \alpha, \alpha') - \beta) \longrightarrow v_t \mathcal{N}(0, 1), \quad \text{stably}, \tag{4.10}$$

where $\mathcal{N}(0, 1)$ is a standard normal r.v. independent of X and

$$v_t^2 = \left(\frac{\alpha^\beta}{A_t C_\beta(1)} \frac{1}{\log(\alpha'/\alpha)}\right)^2 \times \left(C_\beta(2) + \left(\frac{\alpha'}{\alpha}\right)^\beta (C_\beta(2) - 2C'_\beta)\right).$$

2. *On the set $\{A_t > 0\}$, we have*

$$\hat{v}_t^{-1}(\hat{\beta}_n(t, \varpi, \alpha, \alpha') - \beta) \longrightarrow \mathcal{N}(0, 1), \quad \text{stably}, \tag{4.11}$$

where $\mathcal{N}(0, 1)$ is a standard normal r.v. independent of X and

$$\hat{v}_t^2 = \left(\frac{1}{\log(\alpha'/\alpha)}\right)^2 \left(\frac{V(\varpi, \alpha, g^2)_t^n}{V^2(\varpi, \alpha, g)_t^n} + \frac{V(\varpi, \alpha, \bar{g}^2)_t^n}{V^2(\varpi, \alpha', g)_t^n} - \frac{2V(\varpi, \alpha, \tilde{g})_t^n}{V(\varpi, \alpha, g)_t^n V(\varpi, \alpha', g)_t^n}\right).$$

Remark 1. The convergence rate in (4.11) is not explicitly stated in the theorem. It is in fact $\Delta_n^{\varpi\beta/2}$, the same as in (4.10).

Remark 2. The choice of ϖ controls the convergence rate in the CLT in the theorem, with the bigger value of ϖ having a faster rate. In Theorem 1, we require that $\varpi < 1/(2 + \beta)$, so a conservative choice is $\varpi < 1/4$. By comparison, Aït-Sahalia and Jacod (2009b) requires that $\varpi < 1/(2 + \beta) \wedge 2/(5\beta)$ and a conservative choice is $\varpi < 1/5$. So potentially, Theorem 1 might offer a faster convergence rate.

Remark 3. The consistency of $\hat{\beta}_n$ in fact holds if $p > 2(1 - \varpi\beta)/(1 - 2\varpi)$, $\beta' < \beta$ and $\gamma > 0$, which is weaker than $p > (2 - \varpi\beta)/(1 - 2\varpi)$ given in Theorem 1. When $p > 2(1 - \varpi\beta)/(1 - 2\varpi)$, small jumps dominate the increments of the continuous part, and so we can extract information on β from small jumps. Otherwise, increments of the continuous part will dominate the small jumps; for example, for $p = 4 < 2(1-\varpi\beta)/(1-2\varpi)$, Cont and Mancini (2007) showed

$$\frac{\Delta_n^{4\varpi-1}}{3} \sum_{i=1}^{[t/n]} \left(\frac{\Delta_i^n X}{\Delta_n^\varpi}\right)^4 I(|\Delta_i^n X| \le \Delta_n^\varpi) \to^P \int_0^T \sigma_t^4 dt. \tag{4.12}$$

Hence, a conservative choice of p is $p \ge 4$.

Remark 4. In Theorem 1, the sample variance $\hat{v}_t^2$ approximates the asymptotic variance v_t^2 very well, as confirmed with our Monte-Carlo simulations.

Remark 5. Our proposed estimator $\hat{\beta}_n$ makes full use of all increments, and hence its effective sample size should be larger than that of $\bar{\beta}_n$. As a result, we expect that $\hat{\beta}_n$ have smaller asymptotic conditional variances than $\bar{\beta}_n$. Denote their asymptotic conditional variances by $v^2(g)$ and $v^2(g_0)$, respectively. From Aït-Sahalia and Jacod (2009b), $v^2(g_0) = (\alpha^\beta - \alpha'^\beta)/\{\bar{A}_t(\log(\alpha'/\alpha))^2\}$. Some simple algebras show that, for a and b sufficiently close to 1,

$$v^2(g) \le v^2(g_0), \quad g = g_1 \text{ or } g_2.$$

We will only prove the case for $g = g_1$ below, since the other case can be done by letting a, b be close to 1. Since $p - \beta < p$, we have

$$v^2(g_1) < \frac{(p-\beta)}{p\bar{A}_t\left(\log\frac{\alpha'}{\alpha}\right)^2}\left((\alpha^\beta + \alpha'^\beta)\frac{2p}{2p-\beta} - 2\alpha^\beta\frac{p}{p-\beta} - 2\alpha'^\beta\left(\frac{\alpha}{\alpha'}\right)^p\left(\frac{\beta}{2p-\beta} - \frac{\beta}{p-\beta}\right)\right).$$

Comparing this with $v^2(g_0)$ and in view of $p > 2$ and $\alpha < \alpha'$, we have

$$v^2(g_0) - v^2(g_1) > \frac{1}{\bar{A}_t(\log(\alpha'/\alpha))^2}\frac{\beta}{2p-\beta} \times \left(\alpha^\beta + \alpha'^\beta - 2\alpha^\beta\left(\frac{\alpha}{\alpha'}\right)^{p-\beta}\right) > 0. \quad \square$$

5. Simulation studies

In this section, we conduct simulations to compare the finite sample performances of our estimator $\hat{\beta}_n$ and the Aït-Sahalia and Jacod estimator, $\bar{\beta}_n$. To do this, we generate data from the stochastic volatility model

$$dX_t = \sigma_t dW_t + \theta dY_t \tag{5.13}$$

and $v_t = \sigma_t^2$ satisfies

$$dv_t = \kappa(\eta - v_t)dt + \gamma v_t^{1/2} dB_t,$$

where $E[dW_t dB_t] = \rho dt$, and Y_t is specified later.

We take $\kappa = 5$, $\eta = 1/16$, $\gamma = 0.5$, $\rho = -0.5$, $\alpha = 5/16$ and $\alpha' = 2\alpha$, which are the same as in Aït-Sahalia and Jacod (2009b) to facilitate comparison. We take $T = 1$ (day), consisting of 6.5 h of trading per second, i.e., $n = 23\,400$. We also choose the weight function $g_2(x)$ in Example 1, and $a = 6/5$ and $b = 7/5$, and set $p = 6$.

We consider two different models for the process Y_t.

Model 1: Y is a β-stable process. In (5.13), we take Y_t to be a symmetric β-stable process with $\beta = 0.25, 0.5, 0.75, 1.00, 1.25, 1.50,$

Table 1
Comparisons of $\hat{\beta}_n$ and $\overline{\beta}_n$, where Y is β-stable.

	Tail prob.	0.10% $\hat{\beta}_n$ ($\overline{\beta}_n$)	0.25% $\hat{\beta}_n$ ($\overline{\beta}_n$)	0.5% $\hat{\beta}_n$ ($\overline{\beta}_n$)	0.75% $\hat{\beta}_n$ ($\overline{\beta}_n$)
$\beta = 1.75$	bias	0.107 (0.113)	0.075 (0.064)	0.105 (0.105)	0.169 (0.149)
	s.e.	0.457 (0.539)	0.263 (0.314)	0.181 (0.211)	0.148 (0.174)
	MSE	0.219 (0.303)	0.075 (0.103)	0.044 (0.056)	0.051 (0.053)
	MSE reduction	27.72%	27.18%	21.43%	3.77%
$\beta = 1.50$	bias	0.062 (0.070)	0.035 (0.031)	0.040 (0.020)	0.038 (0.043)
	s.e.	0.399 (0.471)	0.231 (0.273)	0.163 (0.181)	0.128 (0.150)
	MSE	0.163 (0.226)	0.055 (0.075)	0.028 (0.033)	0.018 (0.024)
	MSE reduction	15.15%	25.00%	27.88%	26.67%
$\beta = 1.25$	bias	0.056 (0.056)	0.009 (0.030)	0.010 (0.019)	0.014 (0.013)
	s.e.	0.346 (0.400)	0.202 (0.226)	0.141 (0.156)	0.114 (0.127)
	MSE	0.123 (0.163)	0.041 (0.052)	0.020 (0.025)	0.013 (0.016)
	MSE reduction	20.00%	18.75%	24.54%	21.15%
$\beta = 1.00$	bias	0.035 (0.051)	0.010 (0.011)	0.005 (0.007)	0.008 (−0.005)
	s.e.	0.295 (0.324)	0.168 (0.196)	0.123 (0.135)	0.100 (0.109)
	MSE	0.088 (0.107)	0.028 (0.039)	0.015 (0.018)	0.010 (0.012)
	MSE reduction	16.67%	16.67%	17.76%	28.21%
$\beta = 0.75$	bias	0.024 (0.025)	0.013 (0.013)	0.002 (0.003)	0.004 (0.005)
	s.e.	0.239 (0.268)	0.143 (0.163)	0.097 (0.114)	0.080 (0.089)
	MSE	0.058 (0.072)	0.020 (0.027)	0.009 (0.013)	0.006 (0.008)
	MSE reduction	30.77%	25.00%	19.44%	22.22%
$\beta = 0.50$	bias	0.005 (0.015)	0.005 (0.007)	0.002 (0.003)	−0.000 (0.003)
	s.e.	0.177 (0.196)	0.109 (0.125)	0.076 (0.089)	0.062 (0.067)
	MSE	0.031 (0.039)	0.012 (0.016)	0.006 (0.008)	0.004 (0.005)
	MSE reduction	25.00%	20.00%	20.51%	25.00%
$\beta = 0.25$	bias	0.006 (0.008)	0.003 (−0.001)	0.002 (−0.001)	0.001 (0.000)
	s.e.	0.117 (0.140)	0.074 (0.083)	0.051 (0.059)	0.043 (0.049)
	MSE	0.014 (0.020)	0.005 (0.007)	0.003 (0.004)	0.002 (0.003)
	MSE reduction	30.00%	28.57%	25.00%	33.33%

and 1.75. We also calibrate θ to deliver some prespecified values of the tail probability

$$P(|\theta\Delta_i^n Y| \geq \alpha\Delta_n^{\varpi}) \sim \frac{c_\beta \theta^\beta \Delta_n^{1-\varpi\beta}}{\beta\alpha^\beta},$$

$$\text{where } c_\beta = \frac{\Gamma(1+\beta)}{\pi}\sin\left(\frac{\beta\pi}{2}\right). \tag{5.14}$$

Model 2: *Y is a CGMY process.* In (5.13), we take $\theta = 1$ and Y_t to be a CGMY process with the Lévy density:

$$f(x) = \frac{c\exp(-gx)}{x^{1+\beta}} I(x > 0) + \frac{c\exp(mx)}{|x|^{1+\beta}} I(x < 0).$$

The trajectories of the CGMY process could be approximately simulated by the time-changed-Brownian-motion algorithm, where the change of time is via the $\beta/2$ stable subordinator which is also a Lévy process with the Lévy density: $f_{ss}(x) = KI(x > 0)/x^{1+\beta/2}$ for some constant K; see Poirot and Tankov (2006). In the simulation, we fix $g = m = 0.1$, and $\beta = 0.5, 0.75, 1.00, 1.25, 1.5$. We calibrate c such that $2\Delta_n \int_{\alpha\Delta_n^{\varpi}}^{\infty} \frac{c}{x^{1+\beta}}(x)dx =$ jump intensity at different levels.

Simulation results for $\hat{\beta}_n$ are summarized in Tables 1–3, with corresponding results for $\overline{\beta}_n$ given in parentheses. In the tables, the biases, standard errors (s.e.'s), and MSEs of $\hat{\beta}_n$ are reported based on 1000 simulations. For illustration, the MSEs given in Table 1 are also presented in Fig. 3.

We make the following observations.

1. *$\hat{\beta}_n$ outperforms $\overline{\beta}_n$ in terms of s.e.'s and MSEs.*
   The biases of both estimators are comparable to each other; $\hat{\beta}_n$ always have smaller s.e.'s (cf. Remark 5) and smaller MSEs. We further note that biases are typically much smaller than their corresponding s.e.'s, so that the MSEs are contributed more from the s.e.'s, and less from the biases. In all cases, $\hat{\beta}_n$ outperforms $\overline{\beta}_n$ in terms of s.e.'s and MSEs.
2. *How much improvement is made by using $\hat{\beta}_n$ instead of $\overline{\beta}_n$?*
   The improvement can be measured by the reduction in MSE, defined by $[\text{MSE}(\hat{\beta}_n) - \text{MSE}(\overline{\beta}_n)]/\text{MSE}(\overline{\beta}_n)$. From both tables, we note that the reductions in MSE can be as high as 30%, and most reductions are between 20% and 30% for both models.
   In a few cases, only modest improvements have been made. A closer inspection in those cases reveals that the bias/s.e. ratios are also relatively high. To further improve performance, one might consider applying some bias reduction techniques, as was done in Aït-Sahalia and Jacod (2009b), although these techniques are somewhat model-based.
3. *The performances of both estimators depend on the value of β.*
   As β increases from 0 to 2, the s.e.'s and MSEs for both estimators all tend to increase. This may not be surprising since, the bigger the β is, the more difficult it is to separate X_t^d from X_t^c, the more contaminations from the continuous part the increments will contain.
4. *The performance is insensitive to the choice of p.*
   We used $p = 6$ in our simulations. Other values of p ranging from 4 to 10 have also been tried, and similar conclusions have been reached, showing that the performance is insensitive to the choice of p. For lack of space, those results are not listed here. We will illustrate this point again with a real example in the next section.

6. A real example

In this section, we apply our procedure to the intra-day data set of Microsoft (MSFT) on December 1, 2000, available from the TAQ database. The log returns are plotted in Fig. 4, which clearly indicates the existence of jumps.

Now we calculate the jump activity index by $\hat{\beta}_n$ and $\overline{\beta}_n$. First, we need to choose the thresholds $\alpha\Delta_n^{\varpi}$ and $\alpha'\Delta_n^{\varpi}$. We take $\Delta_n = 1.46$ s, which is the average duration time between transactions. We choose $\varpi = 1/5$ and $\alpha' = 2\alpha$. In the case of $\hat{\beta}_n$, we take $g(x) = g_2(x)$ with $p = 6$, as given in Example 1.

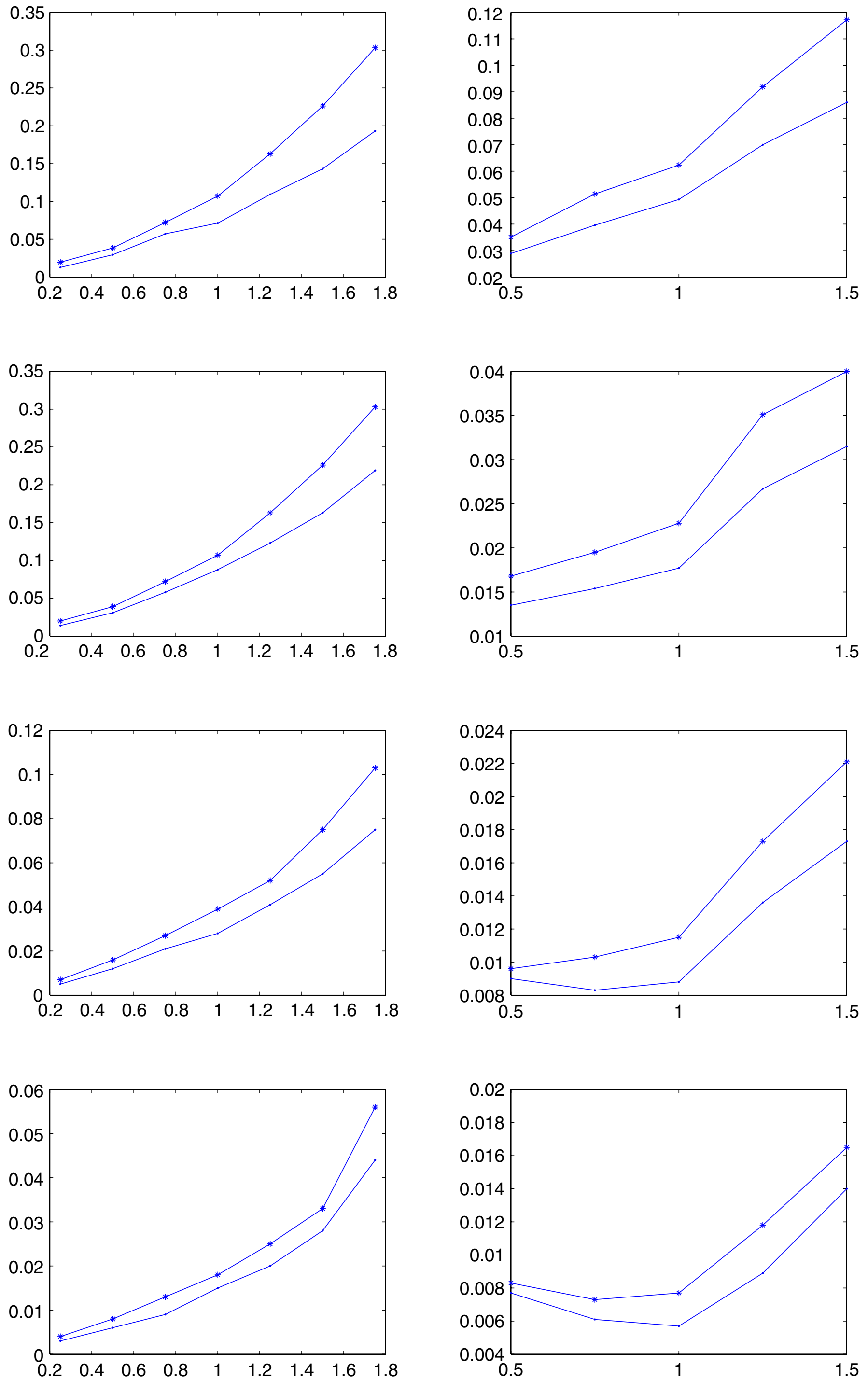

Fig. 3. Plots of MSE vs. β. (1) Left Panels: Y = β-stable; Right Panels: Y = CGMY. (2) Starred lines: MSEs of $\overline{\beta}_n$; Dotted lines: MSEs of $\hat{\beta}_n$.

The purpose of this example is to investigate

(1) how to optimally determine the threshold $\alpha\Delta_n^{\varpi}$ in $\hat{\beta}_n$ and $\overline{\beta}_n$;
(2) how sensitive our estimates $\hat{\beta}_n$ are as p varies.

(1) How to determine the threshold?

We suggest the following procedure to determine the threshold $\alpha\Delta_n^{\varpi}$ (or equivalently α). Note that the MSE of $\hat{\beta}_n$ is $\text{MSE}(\hat{\beta}_n; \alpha) = E(\hat{\beta}_n - \beta)^2 = bias^2(\hat{\beta}_n) + var(\hat{\beta}_n)$, which is a function of α. This can be estimated by

$\widehat{\text{MSE}}(\hat{\beta}_n; \alpha) = \widehat{\text{bias}}^2(\hat{\beta}_n) + \widehat{\text{var}}(\hat{\beta}_n)$.

Here, we take $\widehat{\text{var}}(\hat{\beta}_n) = \hat{\sigma}_T^2$, as in Theorem 1. For $\widehat{\text{bias}}(\hat{\beta}_n)$, we suggest to use a two time-scaled method, which is feasible since, under rather general conditions, we have

$$E\hat{\beta}_n = \beta + C\Delta_n^{1-2\varpi} + o(\Delta_n^{1-2\varpi}). \qquad (6.15)$$

Finally, the optimal α is chosen to minimize the estimated MSE:

$$\alpha^* = \arg\min_{\alpha>0} \widehat{\text{MSE}}(\hat{\beta}_n; \alpha).$$

We now apply this procedure to the real example. We allow α to vary over a fine grid of different values, i.e., $\alpha = 0.0095$, 0.00975, 0.01, 0.011, 0.012, 0.0125, 0.013, 0.014, 0.015, 0.0175,

Table 2
Comparisons of $\hat{\beta}_n$ and $\overline{\beta}_n$, where Y is a CGMY model.

	Jump intensity	0.10% $\hat{\beta}_n$ ($\overline{\beta}_n$)	0.25% $\hat{\beta}_n$ ($\overline{\beta}_n$)	0.5% $\hat{\beta}_n$ ($\overline{\beta}_n$)	0.75% $\hat{\beta}_n$ ($\overline{\beta}_n$)
$\beta = 1.50$	bias	0.060 (0.067)	0.033 (0.034)	0.051 (0.048)	0.065 (0.066)
	s.e.	0.287 (0.336)	0.174 (0.197)	0.121 (0.140)	0.099 (0.110)
	MSE	0.086 (0.0117)	0.032 (0.040)	0.017 (0.022)	0.014 (0.017)
	MSE reduction	26.62%	21.25%	21.72%	15.15%
$\beta = 1.25$	bias	0.043 (0.046)	0.025 (0.029)	0.029 (0.023)	0.032 (0.033)
	s.e.	0.261 (0.300)	0.162 (0.185)	0.115 (0.129)	0.089 (0.104)
	MSE	0.070 (0.092)	0.027 (0.035)	0.014 (0.017)	0.009 (0.012)
	MSE reduction	23.83%	23.93%	21.39%	24.58%
$\beta = 1.00$	bias	0.021 (0.019)	0.008 (0.005)	0.006 (0.007)	0.007 (0.007)
	s.e.	0.221 (0.249)	0.133 (0.151)	0.094 (0.107)	0.075 (0.088)
	MSE	0.049 (0.062)	0.018 (0.023)	0.009 (0.012)	0.006 (0.008)
	MSE reduction	20.87%	22.37%	23.48%	25.97%
$\beta = 0.75$	bias	0.055 (0.059)	0.045 (0.045)	0.043 (0.040)	0.041 (0.038)
	s.e.	0.191 (0.219)	0.116 (0.132)	0.081 (0.093)	0.067 (0.077)
	MSE	0.040 (0.051)	0.015 (0.020)	0.008 (0.010)	0.006 (0.007)
	MSE reduction	22.96%	21.03%	19.42%	16.44%
$\beta = 0.50$	bias	0.070 (0.070)	0.067 (0.067)	0.068 (0.064)	0.069 (0.066)
	s.e.	0.155 (0.174)	0.095 (0.111)	0.066 (0.074)	0.055 (0.063)
	MSE	0.029 (0.035)	0.014 (0.017)	0.009 (0.010)	0.008 (0.008)
	MSE reduction	17.66%	19.64%	6.25%	7.23%

Table 3
Estimates of β when p varies.

p	4	4.5	5	5.5	6	6.5	7	8	10
n	1.532	1.529	1.522	1.512	1.503	1.496	1.487	1.475	1.458
s.e.	0.035	0.038	0.040	0.041	0.042	0.042	0.043	0.044	0.044

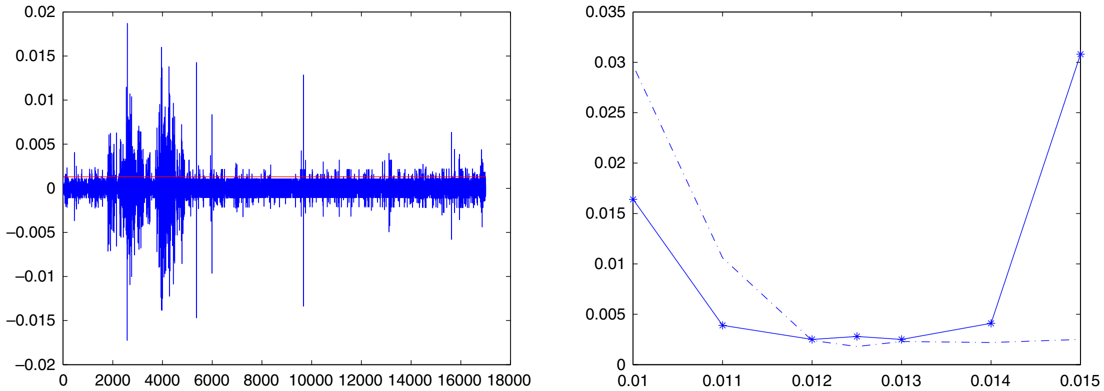

Fig. 4. The left panel: Log returns of the MSFT on 12/01, 2000. The red line corresponds to the threshold $0.012\Delta_n^{1/5}$. The right panel: Estimated MSEs of $\hat{\beta}_n$ (dashed line) and $\overline{\beta}_n$ (starred line) of the MSFT on Dec. 1, 2000. The two estimated MSE curves achieve the minimum at $\alpha^* = 0.012$ (starred) and $\alpha^* = 0.125$ (dashed), respectively. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

0.020, 0.0225, 0.025, 0.0275, 0.030. We restrict α in the range (0.0095, 0.03) since $\overline{\beta}_n$ already shoots above 2 beyond this range. Note that two MSE curves achieve the minimum at $\alpha^* = 0.012$ (for $\overline{\beta}_n$) and $\alpha^* = 0.0125$ (for $\hat{\beta}_n$), respectively. And correspondingly, we obtain

$$\hat{\beta}_n = 1.503, \quad \text{s.e.} = 0.042,$$
$$\overline{\beta}_n = 1.407, \quad \text{s.e.} = 0.046.$$

Fig. 4 plots the $\widehat{\text{MSE}}$s of $\hat{\beta}_n$ (dotted line) and $\widehat{\text{MSE}}$s of $\overline{\beta}_n$ (star-line) for α around α^*'s.

(2) How sensitive our estimate $\hat{\beta}_n$ are as p varies?

Now let us investigate the sensitivity of our estimate $\hat{\beta}_n$ to changes of p. Although Theorem 1 requires that p be an even integer, but here we also include some real-valued p's. The extension of the main results to real-valued p's is still open. Let $p = 4, 4.5, 5, 5.5, 6, 6.5, 7, 8, 10$. For each p, we repeat the above algorithm in finding the optimal α and get the corresponding $\hat{\beta}_n$. They are listed in the Table 3.

From the table, we note that, as p increases from 4 to 10, there is a decreasing trend for the estimate $\hat{\beta}_n$ from 1.532 to 1.458, with a range of 0.074. However, the change is very gradual, indicating that the estimate $\hat{\beta}_n$ is insensitive to the change in p. Furthermore, there is almost no change in the s.e.'s.

Finally, we could treat both α and p as tuning parameters and use the procedure in (1) to find the optimal values α^* and p^*, which minimize the estimated MSEs. Applying such a procedure, we found that the optimal values are around $\alpha^* = 0.013$ and $p^* = 5$.

7. Proofs of main results

We will prove Theorem 1 only, since the proofs for others are simpler and hence omitted. So we assume that all assumptions in Theorem 1 hold in the sequel.

Since the time horizon T is finite, by standard localization method, we can simply prove all the results under the following strengthened assumptions instead of Assumptions 1 and 2.

Assumption 4. The processes b and σ are bounded by some constant L.

Assumption 5. Assumption 2 holds with $L_t(\omega) = L$.

Without loss of generality, we assume that $a = 2$ in Assumption 3, so that $g(z) = z^p$ for $|x| \le 2$.

Throughout the paper, K denotes a constant which may be different at each occurrence. We use E_i^n and E_t respectively as the conditional expectation with respect to $\mathcal{F}_{t_{i-1}}$ and $\mathcal{F}_t$.

Write $X_t = X_t^c + X_t^d$, where $X_t^c = X_0 + \int_0^t b_s ds + \int_0^t \sigma_s dW_s$, and

$$\begin{aligned} X_t^d &= xI(|x| \le 1) \star (\mu - \nu)_t + xI(|x| > 1) \star \mu_t \\ &= xI(|x| > \delta) \star \mu_t + xI(|x| \le \delta) \star (\mu - \nu)_t \\ &\quad + xI(\delta < |x| \le 1) \star \nu_t \\ &=: X_t^d(\delta)' + X_t^d(\delta) + B(\delta). \end{aligned}$$

Finally, we denote $\delta_n = \alpha \Delta_n^{\varpi}$, and $g_n(x) = g(x/\delta_n)$.

An outline of major steps in the proof of Theorem 1

We first prove (3.8) and the asymptotic normality of $\Delta_n^{\varpi\beta} V(\varpi, \alpha, g)_t^n$, which, together with the continuous mapping theorem and the delta method, renders the asymptotic normality of $\hat{\beta}_n(t, \varpi, \alpha, \alpha')$.

To be more precise, write

$$\Delta_n^{\varpi\beta} V(\varpi, \alpha, g)_t^n - \frac{A_t}{\alpha^\beta} C_\beta(1) = I_0 + I_1 + I_2,$$

where

$$I_0 = \Delta_n^{\varpi\beta} V(\varpi, \alpha, g)_t^n - \Delta_n^{\varpi\beta} \sum_{i=1}^{[t/\Delta_n]} E_i^n g_n(\Delta_i^n X),$$

$$I_1 = \Delta_n^{\varpi\beta} \sum_{i=1}^{[t/\Delta_n]} E_i^n \left(g_n(\Delta_i^n X) - \int_{t_i}^{t_{i+1}} \int_R g_n(x) F_t(dx) dt \right),$$

$$I_2 = \Delta_n^{\varpi\beta} \sum_{i=1}^{[t/\Delta_n]} E_i^n \int_{t_i}^{t_{i+1}} \int_R g_n(x) F_t(dx) dt - \frac{A_t}{\alpha^\beta} C_\beta(1).$$

Lemma 1 shows that $I_1 \to^P 0$ if $X = X^d$, i.e., X is a pure jump process. Lemma 2 shows that the presence of a continuous part will not change the convergence of I_1 to 0. Therefore, the combination of Lemmas 1 and 2 proves $I_1 \to^P 0$. $I_2 \to^P 0$ is proved in Lemma 3.

The term I_0 is a martingale w.r.t. $\{\mathcal{F}_{t_i},\ i = 0, \ldots, [t/\Delta_n]\}$, and will be handled in Proposition 1, which gives stable convergence to Gaussian random variable. To prove Theorem 1, we also need the stable convergence mode which is guaranteed by Lemma 4.

7.1. Main lemmas

Before stating Lemma 1, we introduce more notation. Define

$$N_i^n = \sum_{t_{i-1} \le t \le t_i} I(|\Delta_s X| > \delta_n), \quad \text{and}$$

$$C_i^n = \int_{t_i}^{t_{i+1}} \int_R g_n(x) F_t(dx) dt.$$

Let

$$C_{i,1}^n = \int_{t_i}^{t_{i+1}} \int_{|x| \le \delta_n} g_n(x) F_t(dx) dt, \quad \text{and}$$

$$C_{i,2}^n = \int_{t_i}^{t_{i+1}} \int_{|x| > \delta_n} g_n(x) F_t(dx) dt.$$

By Assumption 2, F_t' and F_t'' are mutually singular, hence there exists a predictable subset Φ of $\Omega \times (0, \infty) \times R$ such that F'' and F' are supported on Φ and Φ^c, respectively. Rewrite $X^d(\delta_n) = X_1^d(\delta_n) + X_2^d(\delta_n)$, where

$$X_1^d(\delta_n) = xI(|x| \le \delta_n) I(\Phi^c) \star (\mu - \nu),$$
$$X_2^d(\delta_n) = xI(|x| \le \delta_n) I(\Phi) \star (\mu - \nu).$$

Lemma 1. *Let $\rho = \frac{1}{2}(1 - \varpi\beta) \wedge \varpi(\beta - \beta') \wedge \varpi\gamma$, if $p \ge 2$, we have*

$$|E_i^n(g_n(\Delta_i^n X^d) - C_i^n)| \le K \Delta_n^{1 - \varpi\beta + \rho}.$$

Proof. We have

$$\begin{aligned} E_i^n g_n(\Delta_i^n X^d) &= E_i^n[g_n(\Delta_i^n X^d); N_i^n = 0] + E_i^n[g_n(\Delta_i^n X^d); N_i^n = 1] \\ &\quad + E_i^n[g_n(\Delta_i^n X^d); N_i^n \ge 2]. \end{aligned}$$

By (60) in Aït-Sahalia and Jacod (2009b), the last term $\le K \Delta_n^{2 - 2\varpi\beta}$. To prove Lemma 1, it suffices to prove

$$|E_i^n[g_n(\Delta_i^n X^d); N_i^n = 1] - E_i^n C_{i,2}^n| \le K \Delta_n^{1 - \varpi\beta + \rho}, \tag{7.16}$$

$$|E_i^n[g_n(\Delta_i^n X^d); N_i^n = 0] - E_i^n C_{i,1}^n| \le K \Delta_n^{1 - \varpi\beta + \rho}. \tag{7.17}$$

Proof of (7.16). Note that $X^d = X^d(\delta_n) + X^d(\delta_n)' + B(\delta_n)$, we first show that the contribution of $X^d(\delta_n) + B(\delta_n)$ is negligible. By the mean value theorem,

$$\begin{aligned} g_n(\Delta_i^n X^d) &= g_n(\Delta_i^n X^d(\delta_n)') + g_n'(\Delta_i^n X^d(\delta_n)') \\ &\quad \times [\Delta_i^n X^d(\delta_n) + \Delta_i^n B(\delta_n)] + r_n, \end{aligned} \tag{7.18}$$

where $|r_n| \le K[\Delta_i^n X^d(\delta_n) + \Delta_i^n B(\delta_n)]^2 / \delta_n^2$, since g' is Lipschitz continuous. We first show that the second and the last term of (7.18) are negligible in the mean on $\{N_i^n = 1\}$. Since $g'(0) = 0$,

$$\left[g_n'(\Delta_i^n X^d(\delta_n)')[\Delta_i^n X^d(\delta_n) + \Delta_i^n B(\delta_n)] I(N_i^n = 0)\right] = 0. \tag{7.19}$$

By Assumption 5, note that $X^d(\delta_n)$ is a martingale,

$$\frac{\Delta_i^n B(\delta_n)}{\delta_n} \le \begin{cases} K \Delta_n^{1 - \varpi} & \beta < 1; \\ K \Delta_n^{1 - \varpi - \epsilon} & \beta = 1; \\ K \Delta_n^{1 - \varpi\beta} & \beta > 1; \end{cases} \tag{7.20}$$

and

$$E_i^n \left(\frac{X^d(\delta_n)_u - X^d(\delta_n)_{t_{i-1}}}{\delta_n} \right)^2 \le K \Delta_n^{1 - \varpi\beta}, \tag{7.21}$$

where $t_{i-1} \le u \le t_{t_i}$ and K does not depend on u. By (7.20), (7.21), and (60) in Aït-Sahalia and Jacod (2009b), plus the Cauchy–Schwarz inequality and boundedness of g', we have

$$\begin{aligned} &E_i^n \left(g_n'(\Delta_i^n X^d(\delta_n)')[\Delta_i^n X^d(\delta_n) + \Delta_i^n B(\delta_n)] I(N_i^n \ge 2) \right) \\ &\quad \le K \Delta_n^{3(1 - \varpi\beta)/2}. \end{aligned} \tag{7.22}$$

Rewriting $g_n'(X^d(\delta_n)')$ in the integral form, we have,

$$\begin{aligned} &g_n'(\Delta_i^n X^d(\delta_n)') \\ &\quad = \int_{t_{i-1}}^{t_i} \int_{|x| > \delta_n} [g_n'(X^d(\delta_n)'_{u-} - X^d(\delta_n)'_{t_{i-1}} + x) \\ &\qquad - g_n'(X^d(\delta_n)'_{u-} - X^d(\delta_n)'_{t_{i-1}})] \mu(dx, du). \end{aligned}$$

By the above equation and (7.20), we have

$$|E_i^n[g_n'(\Delta_i^n X^d(\delta_n)') \Delta_i^n B(\delta_n)]| \le K \Delta_n^{1 - \varpi\beta + (1 - \varpi - \epsilon) \wedge (1 - \varpi\beta)}. \tag{7.23}$$

Since $X^d(\delta_n)'$ and $X^d(\delta_n)$ have no common jumps, by the product rule,

$$\begin{aligned} &E_i^n[g_n'(\Delta_i^n X^d(\delta_n)')\Delta_i^n X^d(\delta_n)] \\ &= E_i^n \int_{t_{i-1}}^{t_i}\int_{|x|>\delta_n}[X^d(\delta_n)_u - X^d(\delta_n)_{t_{i-1}}] \\ &\quad\times [g_n'(X^d(\delta_n)'_u - X^d(\delta_n)'_{t_{i-1}} + x) - g_n'(X^d(\delta_n)'_u \\ &\quad - X^d(\delta_n)'_{t_{i-1}})]F_u(dx). \end{aligned}$$

Again using the boundedness of $g'(\cdot)$, (7.21), and the Cauchy–Schwarz inequality, the left side of the above equation is less than $K\Delta_n^{3(1-\varpi\beta)/2}$. This, together with (7.23) results in

$$|E_i^n\left(g_n'(\Delta_i^n X^d(\delta_n)')[\Delta_i^n X^d(\delta_n) + \Delta_i^n B(\delta_n)]\right)| \le K\Delta_n^{3(1-\varpi\beta)/2}, \quad (7.24)$$

which, together with (7.19) and (7.22), results in that

$$\begin{aligned} &|E_i^n(g_n'(\Delta_i^n X^d(\delta_n)')[\Delta_i^n X^d(\delta_n) + \Delta_i^n B(\delta_n)]I(N_i^n = 1))| \\ &\le K\Delta_n^{3(1-\varpi\beta)/2}. \end{aligned}$$

By the estimates on r_n,

$$E_i^n r_n I(N_i^n = 1) \le KE_i^n\left((\Delta_i^n X^d(\delta_n) + \Delta_i^n B(\delta_n))^2 N_i^n\right)/\delta_n^2.$$

Similar to obtaining (7.24), one easily gets $E_i^n r_n I(N_i^n = 1) \le K\Delta_n^{2-2\varpi\beta}$.

On the other hand, note that $X^d(\delta_n)'$ is a random step function,

$$\begin{aligned} &E_i^n g_n(\Delta_i^n X^d(\delta_n)')I(N_i^n = 1) \\ &= E_i^n\left(\sum_{t_{i-1}\le s\le t_i} g_n(\Delta_s X^d(\delta_n)')\right)I(N_i^n = 1) \\ &= E_i^n\left(\sum_{t_{i-1}\le s\le t_i} g_n(\Delta_s X^d(\delta_n)')\right)(1 - I(N_i^n \ne 1)) \\ &= E_i^n\left(\sum_{t_{i-1}\le s\le t_i} g_n(\Delta_s X^d(\delta_n)')\right) \\ &\quad - E_i^n\left(\sum_{t_{i-1}\le s\le t_i} g_n(\Delta_s X^d(\delta_n)')\right)I(N_i^n \ge 2) \\ &= E_i^n C_{i,2}^n - r_n^*, \end{aligned} \quad (7.25)$$

where $r_n^* \le KE_i^n N_i^n I(N_i^n \ge 2)$, so $r_n^* \le K\Delta_n^{2-2\varpi\beta}$ since $N_{t_{i-1}+s}^n - N_{t_{i-1}}^n$, $0 \le s \le \Delta_n$ is a Poisson counting process with cumulative intensity function as $\int_{t_{i-1}}^{t_{i-1}+s}\int_{|x|>\delta_n}F_{t_{i-1}+s}(dx)ds$ which is less than $K\Delta_n^{1-\varpi\beta}$. The combination of (7.22) and (7.25) completes the proof of (7.16). □

Proof of (7.17). Let $A_i^n = \{|\Delta_i^n B(\delta_n)| > \delta_n\}$ and $B_i^n = \{|\Delta_i^n X^d(\delta_n)| > \delta_n\}$. On $\{N_i^n = 0\}$, $X^d = X^d(\delta_n) + B(\delta_n)$. Then by the property of $g(\cdot)$,

$$\begin{aligned} &|E_i^n[g_n(\Delta_i^n X^d); N_i^n = 0] - E_i^n[g_n(\Delta_i^n X^d(\delta_n)); N_i^n = 0]| \\ &\le KE_i^n\left|\frac{\Delta_i^n B(\delta_n)}{\delta_n}\right| I(A_i^n) + \sum_{k=0}^{p-1}\binom{p}{k}E_i^n \\ &\quad\times\left|\frac{\Delta_i^n X^d(\delta_n)}{\delta_n}\right|^k\left|\frac{\Delta_i^n B(\delta_n)}{\delta_n}\right|^{p-k} \\ &\quad + KE_i^n\left|\frac{\Delta_i^n B(\delta_n)}{\delta_n}\right| I(B_i^n). \end{aligned} \quad (7.26)$$

By (7.20), the first term in the last equation is eventually 0. Again using (7.20) and (7.21), both the second and the third terms in the last equation are less than $K\Delta_n^{1-\varpi\beta+(1-\varpi\beta)\wedge(1-\varpi-\epsilon)}$. On the other hand,

$$E_i^n g_n(\Delta_i^n X^d(\delta_n))I(N_i^n = 0) = E_i^n g_n(\Delta_i^n X^d(\delta_n)) + R_n, \quad (7.27)$$

where

$$\begin{aligned} R_n &\le KE_i^n\left(\frac{\Delta_i^n X^d(\delta_n)}{\delta_n}\right)^2 I(N_i^n \ge 1) \\ &\le KE_i^n\left(\frac{\Delta_i^n X^d(\delta_n)}{\delta_n}\right)^2 N_i^n \le K\Delta_n^{2(1-\varpi\beta)} \end{aligned}$$

by product rule and the fact that $N_{t_{i-1}+s}^n - N_{t_{i-1}}^n$ and $X^d(\delta_n)$ have no common jumps. Let $Y_s = X^d(\delta_n)_{t_{i-1}+s} - X^d(\delta_n)_{t_{i-1}}$. By Itô's lemma,

$$\begin{aligned} &E_i^n\left(g_n(Y_{\Delta_n}) - C_{i,1}^n\right) \\ &= E_i^n\int_0^{\Delta_n}\int_{|x|\le\delta_n}\{g_n(Y_s + x) - g_n(Y_s) \\ &\quad - g_n'(Y_s)x - g_n(x)\}F_{t_{i-1}+s}(dx)ds \\ &=: E_i^n\int_0^{\Delta_n}\int_{|x|\le\delta_n}G_n(Y_s, x)F_{t_{i-1}+s}(dx)ds. \end{aligned} \quad (7.28)$$

where $G_n(Y_s, x) = g_n(Y_s + x) - g_n(Y_s) - g_n'(Y_s)x - g_n(x)$ satisfies

$$\begin{aligned} |G_n(Y_s, x)| &\le K(x/\delta_n)^2, \quad |x| \le \delta_n, |Y_s| > \delta_n \\ &= \left|\sum_{k=1}^{p-2}\binom{p}{k}(Y_s/\delta_n)^k(x/\delta_n)^{p-k}\right|, \quad |x| \le \delta_n, |Y_s| \le \delta_n. \end{aligned}$$

It follows from this and Assumption 5 that

$$\begin{aligned} &|E_i^n(g_n(Y_{\Delta_n}) - C_{i,1}^n)| \\ &\le E_i^n\int_0^{\Delta_n}\int_{|x|\le\delta_n}|G_n(Y_s, x)|I(|Y_s| > \delta_n)F_{t_{i-1}+s}(dx)ds \\ &\quad + E_i^n\int_0^{\Delta_n}\int_{|x|\le\delta_n}|G_n(Y_s, x)|I(|Y_s| \le \delta_n)F_{t_{i-1}+s}(dx)ds \\ &\le KE_i^n\int_0^{\Delta_n}I(|Y_s| > \delta_n)\int_{|x|\le\delta_n}(x/\delta_n)^2F_{t_{i-1}+s}(dx)ds \\ &\quad + K\Delta_n^{-\varpi\beta}\sum_{k=1}^{p-2}\int_0^{\Delta_n}\left(E_i^n\left(\frac{Y_s}{\delta_n}\right)^{2k}I(|Y_s| \le \delta_n)\right)^{1/2}ds \\ &\le K\Delta_n^{3(1-\varpi\beta)/2}. \quad \square \end{aligned}$$

Next, we show that Lemma 1 still holds when there is a continuous part.

Lemma 2. *Let $\rho' = \rho \wedge (p(1/2 - \varpi - \epsilon) - (1 - \varpi\beta)) \wedge (1/2 - \varpi - \epsilon)$.*

$$|E_i^n(g_n(\Delta_i^n X) - C_i^n)| \le K\Delta_n^{1-\varpi\beta+\rho'}.$$

Proof. Let $0 < \eta < 1/2 - \varpi$. Now $g_n(\Delta_i^n X) - g_n(\Delta_i^n X^d) = g_n'(\Delta_i^n X^d)\Delta_i^n X^c + \tilde{R}_n$. So we have $E_i^n\left|g_n(\Delta_i^n X) - g_n(\Delta_i^n X^d)\right| \le J_1 + J_2 + J_3$, where

$$\begin{aligned} J_1 &= E_i^n\left|g_n(\Delta_i^n X) - g_n(\Delta_i^n X^d)\right| I(|\Delta_i^n X^c/\delta_n| > \Delta_n^\eta), \\ J_2 &= \left|E_i^n g_n'(\Delta_i^n X^d)(\Delta_i^n X^c)I(|\Delta_i^n X^c/\delta_n| \le \Delta_n^\eta)\right|, \\ J_3 &= \left|E_i^n\tilde{R}_n I(|\Delta_i^n X^c/\delta_n| \le \Delta_n^\eta)\right|. \end{aligned}$$

For J_1, since $|g_n| < K$, for arbitrarily large q, we have $J_1 \le KE_i^n I(|\Delta_i^n X^c/\delta_n| > \Delta_n^\eta) \le K\Delta_n^{q(1/2-\varpi-\eta)}$.

For J_2, by Lemma 1, we have $J_2 \le K\Delta_n^{1-\varpi\beta+\eta}$.

For J_3, by Lemma 1, we have

$$J_3 \le KE_i^n\left(\frac{\Delta_i^nX^c}{\delta_n}\right)^2 I\left(\left|\frac{\Delta_i^nX^c}{\delta_n}\right|\le\Delta_n^\eta\right)I\left(\left|\frac{\Delta_i^nX^d}{\delta_n}\right|>1\right)$$
$$+E_i^n\sum_{k=2}^p\binom{p}{k}\left|\frac{\Delta_i^nX^c}{\delta_n}\right|^k\left|\frac{\Delta_i^nX^d}{\delta_n}\right|^{p-k}$$
$$\times I\left(\left|\frac{\Delta_i^nX^c}{\delta_n}\right|\le\Delta_n^\eta\right)I\left(\left|\frac{\Delta_i^nX^d}{\delta_n}\right|\le 1\right)$$
$$\le K\Delta_n^{2\eta+1-\varpi\beta}+K\Delta_n^{p\left(\frac{1}{2}-\varpi\right)}+K\Delta_n^{(p-1)(1/2-\varpi)+\eta}.$$

Taking $\eta = 1/2-\varpi-\epsilon$ proves the lemma. □

Lemma 3. *Let $\rho''=\frac{1}{2}\wedge\varpi(\beta-\beta')\wedge(\varpi\gamma)$, then*

$$\left|\Delta_n^{\varpi\beta}\sum_{i=1}^{[t/\Delta_n]}E_i^n\int_{t_{i-1}}^{t_i}\int_R g_n(x)F_s(dx)ds-\frac{C_\beta(1)A_t}{\alpha^\beta}\right|\le K\Delta_n^{\rho''}.$$

Proof. By Assumption 2, $E_i^n\int_{t_{i-1}}^{t_i}\int_R g_n(x)F_t(dx)dt =: I_i'+II_i'$, where

$$I_i' = E_i^n\int_{t_{i-1}}^{t_i}\int_R g_n(x)F_t'(dx)dt,$$
$$II_i' = E_i^n\int_{t_{i-1}}^{t_i}\int_R g_n(x)F_t''(dx)dt.$$

First, we look at I_i'. By change of variable,

$$I_i' = E_i^n\int_{t_{i-1}}^{t_i}\int_R g(y)F_t(\delta_n dy)dt$$
$$=\delta_n^{-\beta}C_\beta(1)E_i^n\int_{t_{i-1}}^{t_i}(a_t^{(+)}+a_t^{(-)})dt+D_i^n, \tag{7.29}$$

where $|D_i^n|\le K\Delta_n^{\varpi\gamma}$ by Assumption 5 and boundedness of $g(\cdot)$. Let $A_i^n=\int_{t_{i-1}}^{t_i}(a_t^{(+)}+a_t^{(-)})dt$, then $A_i^n-E_i^nA_i^n$, $1\le i\le n$, are martingale sequences with respect to $\mathcal{F}_{t_i}$, $1\le i\le n$. Since $a_t^{(+)}$ and $a_t^{(-)}$ are bounded, $\sum_{i=1}^{[t/\Delta_n]}E_i^n(A_i^n)^2\le K\Delta_n$. By Doob's inequality, $\left|\sum_{i=1}^{[t/\Delta_n]}(A_i^n-E_i^nA_i^n)\right|\le K\Delta_n^{1/2}$, so by (7.29), $\left|\Delta_n^{\varpi\beta}\sum_{i=1}^{[t/\Delta_n]}I_i'-\alpha^{-\beta}A_tC_\beta(1)\right|\le K\Delta_n^{1/2\wedge\varpi\gamma}$.

Similarly, we have $|\Delta_n^{\varpi\beta}\sum_{i=1}^n II_i'|\le K\Delta_n^{\varpi(\beta-\beta')\wedge 1/2}$. □

To prove the stable convergence, we need the following lemma.

Lemma 4. *Let $h=1-\varpi\beta/2$ and $h'=(-\varpi\beta)\wedge\left(-\varpi\beta-\varpi+\frac{1-\varpi\beta}{2}\right)$. For any bounded martingale M, $0<s<\Delta_n$,*

$$|E_t(M_{t+s}-M_t)g_n(X_{t+s}-X_t)|$$
$$\le K(\Delta_n^{(1-\varpi\beta+\rho')\wedge h}+\epsilon_n\Delta_n^h+\Delta_n^{h'}s(E_t(M_{t+s}-M_t)^2)^{1/2}), \tag{7.30}$$

where $\epsilon_n\downarrow 0$ as $\Delta_n\to 0$.

Proof. By the proof of 2, it suffices to show that (7.30) holds with X replaced by X^d. By Ito's formula and (80) in Aït-Sahalia and Jacod (2009b),

$$(M_{t+s}-M_t)g_n(X_{t+s}^d-X_t^d)=\text{martingale term}+\int_0^s r_u^n du, \tag{7.31}$$

where

$$r_u^n=\int_R F_{t+u}(dx)[(M_{t+u}-M_t)\psi_n(X_{t+u}^d-X_t^d,x)$$
$$+\delta(t+u,x)h_n(X_{t+u}^d-X_t^d,x)], \tag{7.32}$$

where $\delta(\cdot,\cdot)$ is some bounded and predictable function, $h_n(y,x)=g_n(y+x)-g_n(y)$, and $\psi_n(y,x)=g_n(y+x)-g_n(y)-g_n'(y)xI(|x|\le 1)$. It is easy to see that

$$|\psi_n(y,x)|\le K\left(1+\left(\frac{y}{\delta_n}\right)^2\frac{1}{\delta_n}\right)I(|x|>\delta_n)I(|y|\le\delta_n)$$
$$+K\left(1+\frac{1}{\delta_n}\right)I(|x|>\delta_n)I(|y|>\delta_n)$$
$$+K\left(\frac{x}{\delta_n}\right)^2 I(|x|\le\delta_n). \tag{7.33}$$

By the Cauchy–Schwarz inequality, the inequality $|h_n(y,x)|\le K\left(I(|x|>\delta_n)+\frac{|x|}{\delta_n}I(|x|\le\delta_n)\right)$, Assumption 5 and (81) in Aït-Sahalia and Jacod (2009b), we have

$$\int_0^s\int_R F_{t+u}(dx)|\delta(x,t+u)|h_n(y,x)du$$
$$\le K\Delta_n^{1-\varpi\beta/2}\left(\epsilon_n'+\left(E_t(M_{t+s}-M_t)^2\right)^{1/2}\right), \tag{7.34}$$

where $\epsilon_n'^2=\max_{0\le u\le\Delta_n}\int_{|x|\le\delta_n}\delta^2(x,t+u)F_{t+u}(dx)$. Let $\epsilon_n=E_t\epsilon_n'^2$, then by the monotone convergence theorem $\epsilon_n\downarrow 0$. From (7.33), we have

$$\left|\int_R F_{t+u}(dx)\psi_n(y,x)\right|$$
$$\le K\Delta_n^{-\varpi\beta}\left(1+\left(\frac{y}{\delta_n}\right)^2\left(\frac{1}{\delta_n}\right)I(|y|\le\delta_n)+\frac{1}{\delta_n}I(|y|>\delta_n)\right).$$

From this, the Cauchy–Schwarz inequality and Lemma 1, we have

$$E_t\int_0^s\int_R\psi_n(X_{t+u}^d-X_t^d,x)F_{t+u}(dx)|M_{t+u}-M_t|du$$
$$\le K\Delta_n^{h'}s\sqrt{E_t(M_{t+s}-M_t)^2}. \tag{7.35}$$

Combining (7.32) (7.34) and (7.35), we have proved the lemma. □

We now present some results on the bias term I as follows.

Lemma 5. *If M is a bounded martingale, we have*

$$\Delta_n^{-\varpi\beta/2}\left|\Delta_n^{\varpi\beta}\sum_{i=1}^{[t/n]}E_i^n[g_n(\Delta_i^nX)]-\frac{A_t}{\alpha^\beta}C_\beta(1)\right|\to^P 0, \tag{7.36}$$

$$\Delta_n^{\varpi\beta/2}\sum_{i=1}^{[t/n]}|E_i^n[\Delta_i^nM(g_n(\Delta_i^nX))]|\to^P 0. \tag{7.37}$$

Proof. (7.36) is the consequence of Lemmas 1–3. Then it remains to prove (7.37). By (7.30), the left hand side of (7.37) is less than

$$K(\Delta_n^{\rho'-\varpi\beta/2}+\epsilon_n\Delta_n^{h+\varpi\beta/2-1})$$
$$+K\Delta_n^{1/2+h'+\varpi\beta/2}\sum_{i=1}^{[t/n]}\left(\Delta_nE_i^n(\Delta_i^nM)^2\right)^{1/2}. \tag{7.38}$$

By the Cauchy–Schwarz and Jensen inequalities and orthogonality of martingale differences, we have

$$\sum_{i=1}^{[t/n]}\left(\Delta_nE_i^n(\Delta_i^nM)^2\right)^{1/2}$$
$$\le t\left(\sum_{i=1}^{[t/n]}E_i^n(\Delta_i^nM)^2\right)^{1/2}\le t\left(E(M_T-M_0)^2\right)^{1/2}<\infty.$$

From this and choices of ϖ, we see that all the power exponents on Δ_n in (7.38) are positive. This finishes the proof of (7.37). □

Corollary 1. *We have $\Delta_n^{\varpi\beta}\sum_{i=1}^{[t/n]}E_i^ng_n^2(\Delta_i^nX)\to^P\frac{A_t}{\alpha^\beta}C_\beta(2)$.*

7.2. Consistency and asymptotic normality of $\Delta_n^{\varpi\beta}V(\varpi,\alpha)_t^n$

Recall that $\tilde{g}(\alpha,\alpha',x) = g(x)g(\alpha x/\alpha')$, and $\bar{g}(\alpha,\alpha',x) = g(\alpha x/\alpha')$. Further, let $C'_\beta = \int_0^\infty \tilde{g}(\alpha,\alpha',x)/x^{1+\beta}dx$ and $\bar{C}_\beta(k) = \int_0^\infty \bar{g}^k(\alpha,\alpha',x)/x^{1+\beta}dx$.

Proposition 1. *Let $\alpha' > \alpha$. Then, the pair of processes*

$$\Delta_n^{-\varpi\beta/2}\left(\Delta_n^{\varpi\beta}V(\varpi,\alpha)_t^n - \frac{A_t}{\alpha^\beta}C_\beta(1),\ \Delta_n^{\varpi\beta}V(\varpi,\alpha')_t^n - \frac{A_t}{(\alpha')^\beta}C_\beta(1)\right) \tag{7.39}$$

converges stably in law to a continuous Gaussian martingale $(\overline{W},\overline{W}')$ with

$$E_t(\overline{W}^2) = \frac{A_t}{\alpha^\beta}C_\beta(2), \qquad E_t(\overline{W}'^2) = \frac{A_t}{\alpha^\beta}\bar{C}_\beta(2),$$
$$E_t(\overline{W}\,\overline{W}') = \frac{A_t}{\alpha^\beta}C'_\beta. \tag{7.40}$$

Proof. Let

$$\zeta_i^n = \Delta_n^{\varpi\beta/2}(g_n(\Delta_i^nX) - E_i^n g_n(\Delta_i^nX)). \tag{7.41}$$

Define respectively $\zeta_i'^n$ and $g_n(x)'$ by the definition of ζ_i^n and $g_n(x)$ with α replaced by α'. By virtue of (7.36) and the Slutsky theorem, considering the paired triangular arrays $(\zeta_i^n,\zeta_i'^n)_{i=1}^{[t/\Delta_n]}$ is enough. Note that the convergence mode is stable convergence, and that for fixed n, $(\zeta_i^n,\zeta_i'^n)_{i=1}^{[t/\Delta_n]}$ are martingale increments. So by Theorem IX.7.28 of Jacod and Shiryaev (2003), it suffices to show that

$$\begin{cases} \sum_{i=1}^{[t/\Delta_n]} E_i^n(\zeta_i^n)^2 \to^P E_t(\overline{W}^2); \\ \sum_{i=1}^{[t/\Delta_n]} E_i^n(\zeta_i'^n)^2 \to^P E_t(\overline{W}'^2); \\ \sum_{i=1}^{[t/\Delta_n]} E_i^n(\zeta_i^n\zeta_i'^n) \to^P E_t(\overline{W}\,\overline{W}'). \end{cases} \tag{7.42}$$

$$\begin{cases} \sum_{i=1}^{[t/\Delta_n]} E_i^n(\zeta_i^n\Delta_i^nM) \to^P 0; \\ \sum_{i=1}^{[t/\Delta_n]} E_i^n(\zeta_i'^n\Delta_i^nM) \to^P 0. \end{cases} \tag{7.43}$$

By definition, $g_n(x)' = \bar{g}(\alpha,\alpha',x/\delta_n)$. By Lemma 2, $\Delta_n^{\varpi\beta}\sum_{i=1}^{[t/n]}(E_i^n g_n(\Delta_i^nX))^2$, $\Delta_n^{\varpi\beta}\sum_{i=1}^{[t/n]}(E_i^n g_n(\Delta_i^nX)')^2$ and $\Delta_n^{\varpi\beta}\sum_{i=1}^{[t/n]}(E_i^n g_n(\Delta_i^nX))(E_i^n g_n(\Delta_i^nX)')$ are negligible. Therefore, the first two equations of (7.42) are consequences of Corollary 1. From the definition of $g_n(x)$ and $g_n(x)'$, $g_n(x)g_n(x)' = \tilde{g}(\alpha,\alpha',x/\delta_n)$. Replace $g_n(x)$ in (7.36) by $\tilde{g}(\alpha,\alpha',x/\delta_n)$, (7.36) still holds with the right side replaced by $A_tC'_\beta\alpha^{-\beta}$, hence the third equation of (7.42). (7.43) is true due to (7.37). □

7.3. Proof of Theorem 1

Set $\xi_n(\alpha) = \Delta_n^{-\varpi\beta/2}\left(\Delta_n^{\varpi\beta}V(\varpi,\alpha)_t^n - \frac{A_t}{\alpha^\beta}C_\beta(1)\right)$. Then

$$\frac{1}{\Delta_n^{\varpi\beta/2}}(\hat{\beta}_n(t,\varpi,\alpha,\alpha') - \beta) = \frac{-\Delta_n^{\varpi\beta/2}}{\log(\alpha'/\alpha)}\log\frac{1+\Delta_n^{\varpi\beta/2}\frac{\alpha^\beta}{A_t}\frac{1}{C_\beta(1)}\xi_n(\alpha)}{1+\Delta_n^{\varpi\beta/2}\frac{(\alpha')^\beta}{A_t}\frac{1}{C_\beta(1)}\xi_n(\alpha')},$$

which is asymptotically equivalent to

$$\frac{\alpha^\beta}{A_t\log(\alpha'/\alpha)}[\xi_n(\alpha)/C_\beta(1) - \xi_n(\alpha')/\bar{C}_\beta(1)]. \tag{7.44}$$

Then part 1 of Theorem 1 follows from Proposition 1 and (7.44), where we have used the mode of stable convergence.

By the relation among g, $\bar{g}$ and $\tilde{g}$, Proposition 1 and its proof yields the consistency of $\hat{\sigma}_t^2$ to σ_t^2. Using the stable convergence again, Part 2 of Theorem 1 can be obtained readily. □

Acknowledgments

The authors would like to thank the Editor, an Associate Editor and two referees for their constructive suggestions and criticisms which helped to improve this paper considerably.

The first author's research is partially supported by HK RGC grants HKUST6011/07P and HKUST6015/08P, HKUST6019/10P, the National Nature Science Foundation of China (71071155), the Fundamental Research Funds for the Central Universities, and the Research Funds of Renmin University of China (grant 10XNL007), and in part by NSFC (Grant No. 71071155).

References

Aït-Sahalia, Y., 2002. Telling from discrete data whether the underlying continuous-time model is a diffusion. Journal of Finance 57, 2075–2112.
Aït-Sahalia, Y., 2004. Disentangling diffusion from jumps. Journal of Financial Economics 74, 487–528.
Aït-Sahalia, Y., Jacod, J., 2007. Volatility estimators for discretely sampled Levy processes. Annals of Statistics 35, 355–392.
Aït-Sahalia, Y., Jacod, J., 2009a. Testing for jumps in discretely observed process. Annals of Statistics 37, 184–222.
Aït-Sahalia, Y., Jacod, J., 2009b. Estimating the degree of activity of jumps in high frequency data. Annals of Statistics 37, 2202–2244.
Barndorff-Nielsen, O.E., Shephard, N., 2006. Econometrics of testing for jumps in financial economics using bipower variation. Journal of Financial Econometrics 2, 1–48.
Carr, P., Geman, H., Madan, D.B., Yor, M., 2002. The fine structure of asset returns: an empirical investigation. Journal of Business 75, 305–332.
Cont, R., Mancini, C., 2007. Nonparametric tests for probing the nature of asset price processes. Tech. Rep. Universita di Firenze.
Fan, J., Wang, Y., 2007. Multi-scale jump and volatility analysis for high-frequency financial data. Journal of the American Statistical Association 102, 1349–1362.
Jacod, J., 2008. Asymptotic properties of realized powervariations and related functionals of semimartingales. Stochastic Processes and their Applications 118, 517–559.
Jacod, J., Shiryaev, A.N., 2003. Limit Theorems for Stochastic Processes. Springer-Verlag, Berlin.
Jiang, G.J., Oomen, R.C., 2005. A new test for jumps in asset prices. Tech. Rep. The University of Warwick, Warwick Business School.
Lee, S., Mykland, P.A., 2007. Jumps in financial markets: a new nonparametric test and jump dynamics. Review of Financial Studies 21, 2535–2563.
Poirot, J., Tankov, P., 2006. Monte Carlo option pricing for tempered stable (CGMY) processes. Asia-Pacific Financial Markets 4, 327–344.
Todorov, V., Tauchen, G., 2009. Activity signature functions for high-frequency data analysis. Journal of Econometrics 154, 125–138.
Todorov, V., Tauchen, G., 2010. Limit theorems for power variations of pure-jump processes with application to activity estimation. Working Paper.
Woerner, J.H., 2006. Analyzing the fine structure of continuous-time stochastic processes. Tech. Rep. University of Gottingen.
Wu, L., 2008. In: Birge, John, Linetsky, Vadim (Eds.), Modeling Financial Security Returns Using Levy Processes. In: Handbooks in Operations Research and Management Science: Financial Engineering, vol. 15. Elsevier.
Zhao, Z., Wu, W.B., 2009. Nonparametric inference of discretely sampled stable Lévy processes. Journal of Econometrics 153, 83–92.