Journal of Econometrics 153 (2009) 83–92

Contents lists available at ScienceDirect

Journal of Econometrics

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

Nonparametric inference of discretely sampled stable Lévy processes

Zhibiao Zhao *, Wei Biao Wu

Department of Statistics, Penn State University, University Park, PA 16802, United States
Department of Statistics, University of Chicago, Chicago, IL 60637, United States

ARTICLE INFO

Article history:
Received 15 February 2008
Received in revised form
24 April 2009
Accepted 30 April 2009
Available online 13 May 2009

JEL classification:
C13
C14

Keywords:
Bahadur–Kiefer representation
Lévy process
Nonparametric estimation
Quantile regression
Spot volatility
Stable index
Stable process

ABSTRACT

We study nonparametric inference of stochastic models driven by stable Lévy processes. We introduce a nonparametric estimator of the stable index that achieves the parametric $\sqrt{n}$ rate of convergence. For the volatility function, due to the heavy-tailedness, the classical least-squares method is not applicable. We then propose a nonparametric least-absolute-deviation or median-quantile estimator and study its asymptotic behavior, including asymptotic normality and maximal deviations, by establishing a representation of Bahadur–Kiefer type. The result is applied to several major foreign exchange rates.

1. Introduction

Since the celebrated work of Black and Scholes (1973), stochastic models driven by Brownian motions have been an active research area in mathematical finance. See Fan (2005) for an excellent review. Many empirical studies, however, have found that Gaussian processes may not be adequate to capture some important empirical characteristics (for example, leptokurtosis) of financial data. In a pioneering work, Mandelbrot (1963) showed that the tails of the distributions of the cotton price changes are so extraordinarily long that it may be reasonable to assume that the second moment is infinite. He further argued that a good alternative model for cotton price changes is the stable distribution with index 1.7, pioneering the approach of modeling financial data with Lévy processes. See also Fama (1965) for more discussion.

Given discrete observations from a stochastic process $\{X_t\}$ driven by a Lévy process $\{Z_t\}$, we are interested in recovering certain characteristics of the Lévy process $\{Z_t\}$ and the data generating mechanism behind the process $\{X_t\}$. Little is known on statistical inference problems in this context, owing partly to the fact that, except for the special cases of Brownian motion and Cauchy process, the density function of Lévy processes has an intractable form. This problem becomes more complicated in the high-frequency context. Aït-Sahalia and Jacod (2007, 2008) considered this problem in a parametric setting of constant volatility. They studied the behavior of optimal estimators via Fisher's information. In this paper we shall consider this problem under a nonparametric setting.

We suppose that $\{X_t\}$ is from the continuous-time model

$$X_t = \int_0^t \sigma(s)\mathrm{d}Z_s(\alpha), \quad t \geq 0. \tag{1}$$

Here $\sigma(t)$ is a deterministic function representing volatility, and $\{Z_t(\alpha)\}_{t\geq 0}$ is an α-stable Lévy process with index $\alpha \in (0, 2]$ controlling the heaviness of the tail of $Z_t(\alpha)$. In particular, $\{Z_t(2)\}_{t\geq 0}$ is a Brownian motion. For $\alpha < 2$, the process $\{Z_t(\alpha)\}_{t\geq 0}$ has an infinite second moment, and hence model (1) may be more appropriate for datasets with heavy tails or jumps. Mandelbrot (1963) and Aït-Sahalia and Jacod (2007, 2008) considered the special case where $\sigma(t) \equiv \sigma$ is a constant volatility. In Aït-Sahalia and Jacod (2007, 2008), they also added another independent Lévy process dominated by $Z_t(\alpha)$. In comparison, we do not assume any parametric form on σ.

Given observations, our first goal is to estimate the stable index α. The index α measures the heaviness of the underlying process and may serve as an indicator of which model to use: α-stable

* Corresponding address: Department of Statistics, Penn State University, 326 Thomas Building, 16802 University Park, PA, United States. Tel.: +1 814 865 6552; fax: +1 814 863 7114.
E-mail addresses: zuz13@stat.psu.edu, zuz13@psu.edu (Z. Zhao).

0304-4076/$ – see front matter © 2009 Elsevier B.V. All rights reserved.
doi:10.1016/j.jeconom.2009.04.007

process with $\alpha < 2$ or Brownian motion. For example, small α indicates the preference of the former while α close to 2 suggests the latter. There has been an extensive literature on stable index estimation under the parametric setting of independent and identically distributed (iid) stable random variables. See Press (1972) for characteristic function based estimator, McCulloch (1986) for quantile type estimator, Nolan (2001) for maximum likelihood estimate (MLE), and Fan (2006) for U-statistics type methods. For other contributions, see Fama and Roll (1971), De Haan and Resnick (1980) and references therein. Assuming that $\sigma(t) \equiv \sigma$ in (1), Aït-Sahalia and Jacod (2008) showed, by studying Fisher's information, that the optimal rate of index estimators is $\sqrt{n}\log(n)$, but they did not give an explicit estimator. Due to the nonparametric nature of (1), parametric methods, for example, MLE and characteristic function method, are not directly applicable. In this paper, we propose a nonparametric estimator of α based on a two time scales technique. The latter method was applied in Zhang et al. (2005) for estimating integrated volatility with noisy high-frequency data. In contrast to most aforementioned papers that essentially deal with the special case of (1) with $\sigma(t) \equiv 1$, we consider the estimation of α for the nonparametric model (1) without imposing any form on σ. Interestingly, our nonparametric estimator of α achieves the parametric $\sqrt{n}$ rate of convergence.

After estimating α, the next natural question is the nonparametric estimation of the volatility $\sigma(t)$. In this paper we study the problem of spot volatility estimation. For integrated volatility estimation, Woerner (2004) adopted a power variation approach; see also Corcuera et al. (2007). The spot volatility $\sigma(t)$ itself is usually more difficult to deal with than the integrated volatility; see Foster and Nelson (1996). If $\alpha = 2$ and $Z_t(\alpha)$ is a standard Brownian motion, then one can use the classical Nadaraya–Watson type estimator via the expression $\Delta^{-1}\mathbb{E}[(X_{t+\Delta} - X_t)^2] = \sigma^2(t) + O(\Delta)$; see Foster and Nelson (1996) and Stanton (1997) for least-squares (LS) based approaches. For $\alpha < 2$, since X_t does not have finite second moment, the classical LS type methods are not applicable. Foster and Nelson (1996) pointed out that it is quite challenging yet worthwhile to extend their results to heavy-tailed processes. Here we propose a least-absolute-deviation (LAD) or median-quantile regression method. For the proposed LAD estimate, we establish a uniform Bahadur–Kiefer representation, a powerful tool for studying the asymptotic behavior of the estimate. In particular, we obtain the asymptotic normality and maximal deviations of the proposed estimate. Finally, we point out that, since the proposed estimates of α and $\sigma(\cdot)$ depend on the scaling property of stable Lévy processes, it remains unclear how to generalize our methods to deal with more general Lévy processes. We pose this as an open problem.

The rest of the paper is organized as follows. Section 2 briefly introduces stable Lévy processes. In Section 3, we propose a nonparametric estimator of α and establish its asymptotic normality. In Section 4, we study the nonparametric kernel median-quantile estimate of $\sigma(\cdot)$ and its asymptotic properties. In Section 5, we illustrate our method via a simulation study. Section 6 contains an application to several major foreign exchange rates. We defer the proofs to Section 7.

2. Stable Lévy processes

A random variable Z is said to have a symmetric α-stable (SαS) distribution, denoted by $Z \in \mathcal{S}(\alpha, c, \nu)$, if its characteristic function has the form

$$\Psi(u) = \mathbb{E}[\exp(\sqrt{-1}uZ)] = \exp(-|cu|^\alpha + \sqrt{-1}\nu u), \quad u \in \mathbb{R},$$

where $\alpha \in (0, 2]$, $c > 0$ and $\nu \in \mathbb{R}$ are the stable, scale, and location parameters, respectively, and $\sqrt{-1}$ is the imaginary unit. Clearly, $Z \overset{\mathcal{D}}{=} \nu + cZ^*$ for a standard SαS random variable $Z^* \in \mathcal{S}(\alpha, 1, 0)$. Denote by f^*_α the standard SαS density function. By the inversion formula for characteristic functions, $f^*_\alpha(x) = (2\pi)^{-1}\int_{\mathbb{R}} \exp(-\sqrt{-1}ux - |u|^\alpha)\mathrm{d}u$ is symmetric, bounded, and has bounded derivatives of all orders. Let $\Gamma(t) = \int_0^\infty x^{t-1}\mathrm{e}^{-x}\mathrm{d}x$, $t > 0$, be the Gamma function. Then (see Nikias and Shao (1995)), for $x > 0$,

$$f^*_\alpha(x) = \begin{cases} (\pi x)^{-1}\sum_{k=1}^{\infty}(-1)^{k-1}(k!)^{-1}\Gamma(\alpha k + 1)x^{-\alpha k}\sin(k\alpha\pi/2), \\ \quad \alpha \in (0, 1); \\ 1/[\pi(1 + x^2)], \quad \alpha = 1; \\ (\pi\alpha)^{-1}\sum_{k=0}^{\infty}(-1)^k[(2k)!]^{-1}\Gamma[(2k+1)/\alpha]x^{2k}, \\ \quad \alpha \in (1, 2]. \end{cases} \tag{2}$$

A Lévy process has independent and stationary increments. In (1), let $\{Z_t(\alpha)\}_{t\geq 0}$ be a SαS Lévy process such that $Z_1(\alpha) \in \mathcal{S}(\alpha, c, 0)$ has SαS distribution. Then $Z_t(\alpha)$ has characteristic function $\Psi_t(u) = \mathbb{E}\{\exp[\sqrt{-1}uZ_t(\alpha)]\} = \Psi(u)^t = \exp(-t|cu|^\alpha)$. If $\alpha = 2$ and $c = 1/\sqrt{2}$, then $\{Z_t(2)\}_{t\geq 0}$ is a standard Brownian motion. The Cauchy process corresponds to $\alpha = 1$.

Suppose that we observe the process $\{X_t\}_{t\geq 0}$ on the interval $[0, T]$ at discrete time points $0 = t_0 < t_1 < \cdots < t_{n-1} < t_n = T$. For simplicity we shall only consider the case of equi-distant time points $t_i = i\Delta_n$, $0 \leq i \leq n$, and $n\Delta_n = T$. Here $\Delta_n \to 0$ is the time between two consecutive observations. In practice, the time unit is a year and hence $1/\Delta_n$ represents sampling frequency. For example, $\Delta_n = 1/252$ corresponds to daily data and T is the time span in years. Such a framework is appropriate for high-frequency data and may result in quite different asymptotic behavior than the case of $\Delta_n = \Delta$ for a fixed Δ; see Aït-Sahalia and Jacod (2008).

Note that $Z_1(\alpha) \in \mathcal{S}(\alpha, c, 0)$. For notational simplicity, we shall in the following often suppress the dependence on α. For example, we write $Z_t = Z_t(\alpha)$ and $Z_1 = Z_1(\alpha)$. We now list some useful properties of the SαS Lévy process $\{Z_t\}_{t\geq 0}$.

(i) The process $\{Z_t\}_{t\geq 0}$ has independent and stationary increments.

(ii) For real numbers $a < b$ and a piecewise continuous deterministic function g on $[a, b]$, the distributional equivalence holds:

$$\int_a^b g(t)\mathrm{d}Z_t \overset{\mathcal{D}}{=} \left[\int_a^b |g(t)|^\alpha \mathrm{d}t\right]^{1/\alpha} Z_1,$$

$$aZ_1 + bZ_1' \overset{\mathcal{D}}{=} (|a|^\alpha + |b|^\alpha)^{1/\alpha} Z_1, \quad \text{where } Z_1', Z_1 \text{ are iid.}$$

(iii) For all $k \in \mathbb{N}$, $\mathbb{E}(|\log|Z_1||^k) < \infty$.

(iv) For $\alpha < 2$, $\mathbb{E}|Z_1|^p < \infty$ for all $p < \alpha$, and $\mathbb{P}(|Z_1| \geq x) = O(x^{-\alpha})$ as $x \uparrow \infty$.

We now introduce some notation. Let $[\ell_1, \ell_2]$ be an interval. For a function g on $[\ell_1, \ell_2]$, we write $g \in \mathcal{C}^p[\ell_1, \ell_2]$, $p \in \mathbb{N}$, if g has bounded pth order derivatives on $[\ell_1, \ell_2]$, and write $g \in \mathcal{L}^\iota[\ell_1, \ell_2]$ with $\iota \in (0, 1]$ if g is uniformly Hölder continuous with exponent ι in the sense that $\sup_{u,u'\in[\ell_1,\ell_2]} |g(u) - g(u')|/|u - u'|^\iota < \infty$. In particular, $\mathcal{L}^1[\ell_1, \ell_2]$ is the set of Lipschitz continuous functions. Denote by $\mathbf{1}_A$ the indicator function for an event A. For $a \in \mathbb{R}$ let $\lceil a \rceil = \min\{k \in \mathbb{Z} : k \geq a\}$ and $\lfloor a \rfloor = \max\{k \in \mathbb{Z} : k \leq a\}$.

3. Estimate α: A tale of two time scales

To estimate α, we shall apply a two time scales approach. The latter method was applied in Zhang et al. (2005) for estimating integrated volatility.

We now give a heuristic argument. Recall that $n\Delta_n = T$ with fixed T. By (1), $X_{i\Delta_n} - X_{(i-1)\Delta_n} = \gamma_i W_i$, where

$$\gamma_i = \left[\int_{(i-1)\Delta_n}^{i\Delta_n} \sigma^\alpha(s)\mathrm{d}s\right]^{1/\alpha}, \qquad W_i = \gamma_i^{-1}\int_{(i-1)\Delta_n}^{i\Delta_n} \sigma(s)\mathrm{d}Z_s,$$
$$1 \le i \le n. \tag{3}$$

By the properties (i) and (ii) of stable Lévy processes in Section 2, $\{W_i\}_{1\le i\le n}$ are iid random variables distributed as Z_1.

Assume $\sigma \in \mathcal{L}^\iota[0,T]$, $\iota > 0$, and $\inf_{t\in[0,T]}\sigma(t) > 0$. Simple calculations show that

$$\begin{aligned}\log\gamma_i &= \frac{1}{\alpha}\log\{\Delta_n[\sigma^\alpha(i\Delta_n) + O(\Delta_n^\iota)]\}\\ &= \log\sigma(i\Delta_n) + \frac{\log(\Delta_n)}{\alpha} + O(n^{-\iota}).\end{aligned} \tag{4}$$

By (1) and (4),

$$\begin{aligned}S_n &:= \frac{1}{n}\sum_{i=1}^n \log|X_{i\Delta_n} - X_{(i-1)\Delta_n}|\\ &= \frac{1}{n}\sum_{i=1}^n[\log\gamma_i + \log|W_i|]\\ &= \frac{1}{n}\sum_{i=1}^n\log\sigma(i\Delta_n) + \frac{\log(\Delta_n)}{\alpha} + \frac{1}{n}\sum_{i=1}^n\log|W_i| + O(n^{-\iota})\\ &= \frac{1}{T}\int_0^T\log\sigma(t)\mathrm{d}t + \frac{\log(\Delta_n)}{\alpha} + \overline{W}_n + O(n^{-\iota}),\end{aligned} \tag{5}$$

in view of the approximation $\Delta_n\sum_{i=1}^n\log\sigma(i\Delta_n) = \int_0^T\log\sigma(t)\mathrm{d}t + O(n^{-\iota})$, where

$$\overline{W}_n = n^{-1}\sum_{i=1}^n\log|W_i| = \mathbb{E}[\log|Z_1|] + O_p(n^{-1/2}). \tag{6}$$

If $\int_0^T\log\sigma(t)\mathrm{d}t$ were known or can be estimated, then expression (5) can be used to estimate α. However, the estimate of $\sigma(\cdot)$ usually involves the unknown index α itself (see Section 4). To overcome this difficulty, we apply a two time scales technique. Assume that n is an even integer. Write $X_{2i\Delta_n} - X_{2(i-1)\Delta_n} = \gamma_i^* W_i^*$, where

$$\gamma_i^* = \left[\int_{2(i-1)\Delta_n}^{2i\Delta_n} \sigma^\alpha(s)\mathrm{d}s\right]^{1/\alpha}, \qquad W_i^* = \frac{1}{\gamma_i^*}\int_{2(i-1)\Delta_n}^{2i\Delta_n}\sigma(s)\mathrm{d}Z_s,$$
$$1 \le i \le \frac{n}{2}. \tag{7}$$

Then $\{W_i^*\}_{1\le i\le n/2}$ are iid distributed as $|Z_1|$. Similarly as (4) and (5), we obtain

$$\begin{aligned}S_n^* &:= \frac{1}{n/2}\sum_{i=1}^{n/2}\log|X_{2i\Delta_n} - X_{2(i-1)\Delta_n}|\\ &= \frac{1}{T}\int_0^T\log\sigma(t)\mathrm{d}t + \frac{\log(2\Delta_n)}{\alpha} + \overline{W}_n^* + O(n^{-\iota}),\end{aligned} \tag{8}$$

where

$$\overline{W}_n^* = (n/2)^{-1}\sum_{i=1}^{n/2}\log|W_i^*| = \mathbb{E}[\log|Z_1|] + O_p(n^{-1/2}). \tag{9}$$

Clearly, S_n and S_n^* are obtained from samples with different frequencies with S_n^* depending only on the sub-sampling observations with even indices. Taking the difference of (5) and (8), we obtain a natural, easy-to-use and consistent estimate of α:

$$\hat{\alpha} = \frac{\log 2}{S_n^* - S_n} = \alpha + O_p(n^{-1/2} + n^{-\iota}). \tag{10}$$

Theorem 1 provides the asymptotic normality for $\hat{\alpha}$.

Theorem 1. *Let $\alpha \in (0,2]$. Assume that $\sigma \in \mathcal{L}^\iota[0,T]$ with $\iota > 1/2$ and $\inf_{t\in[0,T]}\sigma(t) > 0$. Then, as $n \to \infty$,*

$$\sqrt{n}(\hat{\alpha} - \alpha) \Rightarrow N(0, \varsigma_\alpha^2), \tag{11}$$

where

$$\varsigma_\alpha^2 = \frac{\alpha^4}{2(\log 2)^2}\mathbb{E}\Big[2\log|Z_1 + Z_1'| - \log|Z_1| - \log|Z_1'| - \frac{2\log 2}{\alpha}\Big]^2 < \infty, \tag{12}$$

where $Z_1' \in \mathcal{S}(\alpha, c, 0)$ is independent and identically distributed as Z_1.

An interesting and useful feature of Theorem 1 is that the asymptotic distribution in (11) does not depend on the volatility function $\sigma(t)$. This has two implications: (i) We can construct confidence intervals using (11) without estimating $\sigma(t)$; and (ii) Property (i) further implies that the asymptotic distribution is also independent of the scale parameter c since we can consider the standardized process $Z_t^* = Z_t/c$ with $Z_1^* \in \mathcal{S}(\alpha,1,0)$ and the new volatility $\sigma^*(t) = c\sigma(t)$. The latter property is also easily seen from the fact that ς_α^2 is invariant under transformation $(Z_1, Z_1') \to (Z_1/c, Z_1'/c)$. Therefore, in (12), we can use standard SαS random variables $Z_1^*, Z_1^{*'} \in \mathcal{S}(\alpha,1,0)$, to get the limiting variance ς_α^2. In the context of estimating stable index for iid stable random variables, Fan (2006) adopted a similar framework and established $\sqrt{n}$-normality. In contrast, we deal with discrete samples from the nonparametric model (1) and yet obtain the same convergence rate.

There is no explicit form for ς_α^2. In practice, we may resort to Monte Carlo simulation scheme: generate iid standard SαS random variables $\{Z_i\}_{1\le i\le N+1}$, and estimate ς_α^2 by

$$\hat{\varsigma}_\alpha^2 = \frac{\alpha^4}{2(\log 2)^2}\frac{1}{N}\sum_{i=1}^N\Big[2\log|Z_i + Z_{i+1}| - \log|Z_i| - \log|Z_{i+1}| - \frac{2\log 2}{\alpha}\Big]^2. \tag{13}$$

For example, simulations show that, for $\alpha = 1$ (standard Cauchy random variables) $\varsigma_1^2 \approx 5.1$, and $\varsigma_{1.5}^2 \approx 25.1$ for $\alpha = 1.5$. Based on Theorem 1, a 95% confidence interval for α can be constructed as $\hat{\alpha} \pm 2\hat{\varsigma}_{\hat{\alpha}}/\sqrt{n}$. Given a sample size $n = 1000$, if the estimate $\hat{\alpha} = 1$, then a 95% confidence interval for α is $1 \pm 2\sqrt{5.1}/\sqrt{1000}$, or 1 ± 0.14. In practice, the estimate $\hat{\alpha}$ in (10) might not be in the range $(0,2]$ especially for α close to 0 or 2. Then we may adopt a truncation technique.

Recall $\hat{\alpha}$ in (10), where S_n^* [cf. (8)] only depends on even indices. Alternatively, we can define a similar estimator based on odd indices

$$\hat{\alpha}' = \frac{\log(2)}{S_n' - S_n}, \quad \text{where}$$
$$S_n' := \frac{1}{n/2 - 1}\sum_{i=1}^{n/2-1}\log|X_{(2i+1)\Delta_n} - X_{(2i-1)\Delta_n}|. \tag{14}$$

It can be easily shown that $\hat{\alpha}'$ and $\hat{\alpha}$ have the same asymptotic distribution. We combine the two estimators and propose the following not less efficient estimator

$$\tilde{\alpha} = \frac{\hat{\alpha} + \hat{\alpha}'}{2}. \tag{15}$$

The same argument in the proof of Theorem 1 shows that $\sqrt{n}(\tilde{\alpha} - \alpha) \Rightarrow N(0, \tilde{\varsigma}_\alpha^2)$, where

$$\tilde{\varsigma}_\alpha^2 = \frac{\alpha^4}{(\log 2)^2}\left\{\mathbb{E}\left(\log^2\left|\frac{Z_1 + Z_1'}{2^{1/\alpha}Z_1}\right|\right) + 2\mathbb{E}\left(\log\left|\frac{Z_1 + Z_1'}{2^{1/\alpha}Z_1}\right|\log\left|\frac{Z_1' + Z_1''}{2^{1/\alpha}Z_1'}\right|\right)\right\}.$$

Here $Z_1, Z_1', Z_1'' \in \mathcal{S}(\alpha, c, 0)$ are iid. The limiting variance $\tilde{\varsigma}_\alpha^2$ can be computed using a similar simulation procedure as in (13). By the Cauchy–Schwarz inequality,

$$\mathrm{Var}(\tilde{\alpha}-\alpha) \le \frac{1}{4}\{\mathrm{Var}(\hat{\alpha}-\alpha)+\mathrm{Var}(\hat{\alpha}'-\alpha) + 2[\mathrm{Var}(\hat{\alpha}-\alpha)]^{1/2}[\mathrm{Var}(\hat{\alpha}'-\alpha)]^{1/2}\},$$

we have, by letting $n \to \infty$, that $\tilde{\varsigma}_\alpha^2 = \lim_{n\to\infty} \mathrm{Var}[\sqrt{n}(\tilde{\alpha}-\alpha)] \le \varsigma_\alpha^2$. Our simulation shows that $\tilde{\varsigma}_\alpha^2 \approx 0.65\varsigma_\alpha^2$ for $\alpha \in [1, 2]$ and $\tilde{\varsigma}_\alpha^2/\varsigma_\alpha^2 \in [0.67, 0.78]$ for $\alpha \in [0.1, 1]$. For example, $\tilde{\varsigma}_\alpha^2/\varsigma_\alpha^2 \approx 76\%, 70\%, 67\%, 63\%, 62\%$ for $\alpha = 0.3, 0.7, 1.0, 1.5, 1.8$, respectively.

4. Nonparametric median-quantile estimation of σ

In this section we assume that the index α is known, otherwise we apply the method in Section 3 to estimate it first. We study the nonparametric least-absolute-deviation (LAD) or median-quantile regression estimate of $\sigma(t)$. See Koenker (2005) for an overview of quantile regression.

Recall model (1). Without further constraints the volatility $\sigma(t)$ is not identifiable, because for any $\lambda > 0$, $(\lambda\sigma(t), Z_t(\alpha)/\lambda)$ is also a solution. To ensure identifiability, we impose a quantile assumption on the process $\{Z_t := Z_t(\alpha)\}_{t\ge 0}$ and hence the scale parameter c. For a random variable ε denote its median by $\mathcal{Q}(\varepsilon) = \inf\{t : \mathbb{P}(\varepsilon \le t) \ge 1/2\}$. For the rest of the paper we always assume that $\mathcal{Q}(|Z_1|) = 1$ or equivalently $Z_1 \in \mathcal{G}(\alpha, c, 0)$ with $c = 1/\mathcal{Q}(|Z_1^*|)$, where $Z_1^* \in \mathcal{G}(\alpha, 1, 0)$ is a standard SαS random variable. The requirement $\mathcal{Q}(|Z_1|) = 1$ can always be satisfied after proper scaling.

We now illustrate the intuition of the median-quantile estimator of $\sigma(t)$, $t \in (0, T)$. Assume that $\sigma \in \mathcal{C}^1[0, T]$. By (1) and property (ii) of stable Lévy processes in Section 2,

$$X_{t+\Delta_n} - X_t = \int_t^{t+\Delta_n} \sigma(s)\mathrm{d}Z_s \overset{\mathcal{D}}{=} \left[\int_t^{t+\Delta_n} \sigma^\alpha(s)\mathrm{d}s\right]^{1/\alpha} Z_1. \tag{16}$$

Hence, by the assumption $\mathcal{Q}(|Z_1|) = 1$, we have as $\Delta_n \to 0$ that,

$$\mathcal{Q}(\Delta_n^{-1/\alpha}|X_{t+\Delta_n} - X_t|) = \left[\frac{1}{\Delta_n}\int_t^{t+\Delta_n} \sigma^\alpha(s)\mathrm{d}s\right]^{1/\alpha} = \sigma(t) + O(\Delta_n). \tag{17}$$

Therefore, we can use the local sample median of observations $\Delta_n^{-1/\alpha}|X_{s+\Delta_n} - X_s|$ with $s \approx t$ to estimate $\sigma(t)$. Note that the median $\mathcal{Q}(\varepsilon)$ of ε is a solution to the minimization problem $\mathrm{argmin}_\theta \mathbb{E}(|\varepsilon - \theta| - |\varepsilon|)$. Here, we put $|\varepsilon|$ into the expectation to guarantee that $||\varepsilon - \theta| - |\varepsilon|| \le |\theta|$ always has a finite expectation. Thus, by applying the kernel method, we propose the following estimate of $\sigma(t)$:

$$\hat{\sigma}_{b_n}(t) = \mathrm{argmin}_\theta \sum_{i=1}^n \left|\Delta_n^{-1/\alpha}|X_{i\Delta_n} - X_{(i-1)\Delta_n}| - \theta\right| \times K_{b_n}(i\Delta_n - t), \tag{18}$$

where $K_{b_n}(u) = K(u/b_n)$, K is a kernel function and the bandwidth $b_n \downarrow 0$ and $nb_n \to \infty$.

Unlike least-squares methods, there is no explicit form for $\hat{\sigma}_{b_n}(t)$. In Theorem 2, we establish a Bahadur–Kiefer type representation for $\hat{\sigma}_{b_n}(t)$. The latter result provides a deep insight into the asymptotic properties of $\hat{\sigma}_{b_n}(t)$ by approximating it by a linear form.

4.1. Uniform Bahadur–Kiefer type representation for $\hat{\sigma}_{b_n}$

Under the above setting, denote by F_α and $f_\alpha = F'_\alpha$ the distribution and density functions of $|Z_1| = |Z_1(\alpha)|$, respectively. Since $\mathcal{Q}(|Z_1|) = 1$, $F_\alpha(1) = 1/2$. Recall the standard SαS density f_α^* in (2). Since $Z_1 \overset{\mathcal{D}}{=} cZ_1^*$ with $Z_1^* \in \mathcal{G}(\alpha, 1, 0)$ and $c = 1/\mathcal{Q}(|Z_1^*|)$, we have

$$f_\alpha(x) = 2f_\alpha^*(x/c)/c, \quad x \ge 0, \tag{19}$$

where $c = 1/\mathcal{Q}(|Z_1^*|)$ is uniquely determined by $\int_{1/c}^\infty f_\alpha^*(x)\mathrm{d}x = 1/4$. There are no explicit forms for c, f_α, f'_α and one may use their numerical approximations.

Before we present our results, we need to impose regularity conditions on K and σ. The conditions on K are mild and they are satisfied for rectangle, Epanechnikov, triangle, Parzen and many other popular kernels.

Definition 1. Let $\mathcal{K}_\omega$, $\omega > 0$, be the set of kernels which are bounded, symmetric, and have bounded support $[-\omega, \omega]$ with bounded derivatives. Let $\psi_K = \int_{\mathbb{R}} u^2K(u)\mathrm{d}u/2$ and $\varphi_K = \int_{\mathbb{R}} K^2(u)\mathrm{d}u$, $K \in \mathcal{K}_\omega$.

Condition 1 (*Regularity Conditions*). Assume $\sigma \in \mathcal{C}^4[0, T]$ and $\inf_{t\in[0,T]} \sigma(t) > 0$.

Theorem 2. *Assume $K \in \mathcal{K}_\omega$, $\omega > 0$, and Condition 1. Further assume that the bandwidth b_n satisfies $b_n + (\log n)/(nb_n) \to 0$. Then:*

(i) *We have the following uniform consistency:*

$$\sup_{t\in[\omega b_n, T-\omega b_n]} |\hat{\sigma}_{b_n}(t) - \sigma(t)| = O_{\mathrm{p}}\left\{b_n^2 + \left[\frac{\log n}{nb_n}\right]^{1/2}\right\}. \tag{20}$$

(ii) *The following Bahadur–Kiefer representation holds*

$$\hat{\sigma}_{b_n}(t) - \sigma(t) = -\frac{\sigma(t)}{f_\alpha(1)}Q_{b_n}(t) + b_n^2\psi_K\rho_\sigma(t) + O_{\mathrm{p}}(R_n), \tag{21}$$

uniformly over $t \in [\omega b_n, 1 - \omega b_n]$, where $Y_i = Y_{i,n} = \Delta_n^{-1/\alpha}|X_{i\Delta_n} - X_{(i-1)\Delta_n}|$,

$$\rho_\sigma(t) = \sigma''(t) - [2 + f'_\alpha(1)/f_\alpha(1)][\sigma'(t)]^2/\sigma(t),$$

$$Q_{b_n}(t) = \frac{\Delta_n}{b_n}\sum_{i=1}^n \{\mathbf{1}_{Y_i\le\sigma(t)} - \mathbb{E}[\mathbf{1}_{Y_i\le\sigma(t)}]\}K_{b_n}(i\Delta_n - t),$$

$$R_n = b_n^4 + \left[\frac{\log n}{nb_n}\right]^{3/4} + \left[\frac{b_n\log n}{n}\right]^{1/2}. \tag{22}$$

The statement (20) asserts that the estimate $\hat{\sigma}_{b_n}$ is uniformly consistent provided that the bandwidth satisfies the mild condition $b_n \to 0$ and $\log n = o(nb_n)$, which is only slightly stronger than the natural condition $b_n \to 0$ and $nb_n \to \infty$ in classical nonparametric estimation problems. In (22), if we use the Mean Squared Error (MSE)-optimal bandwidth $b_n \asymp n^{-1/5}$ (see Theorem 3), then the error bound is $O_{\mathrm{p}}[n^{-3/5}(\log n)^{3/4}]$.

For $\alpha < 2$, the second moment is infinite and the classical least-square type methods are not applicable. Our LAD or median-quantile approach provides a useful alternative. The Bahadur–Kiefer representation (21) linearizes $\hat{\sigma}_{b_n}(t) - \sigma(t)$ by approximating it by $b_n^2\psi_K\rho_\sigma(t) - \sigma(t)Q_{b_n}(t)/f_\alpha(1)$. The latter is easier to deal with, due to its additive structure. Such an expansion provides a powerful tool for studying the asymptotic behavior of $\hat{\sigma}_{b_n}$. The following theorem can be used to construct point-wise confidence interval for $\sigma(t)$.

Theorem 3. *Assume that the conditions in Theorem 2 are satisfied and*

$$nb_n^9 + \frac{(\log n)^3}{nb_n} \to 0. \tag{23}$$

Let ρ_σ be as in Theorem 2. Then for each fixed $t \in (0, 1)$, we have the convergence

$$N_t := \frac{2f_\alpha(1)}{\sigma(t)\sqrt{\varphi_K}}\sqrt{b_n/\Delta_n}[\hat{\sigma}_{b_n}(t) - \sigma(t) - b_n^2\psi_K\rho_\sigma(t)] \Rightarrow N(0, 1). \tag{24}$$

Moreover, for distinct values $t_1, t_2, \ldots, t_k \in (0, 1)$, $N_{t_1}, N_{t_2}, \ldots, N_{t_k}$, are asymptotically independent standard normals.

By Theorem 3, the MSE-optimal bandwidth $b_n \asymp n^{-1/5}$. For this optimal bandwidth, condition (23) is satisfied and the speed of convergence in (24) is of order $n^{2/5}$. In general, let $b_n \asymp n^{-\beta}$, then (23) holds for $\beta \in (1/9, 1)$.

Theorem 4 presents a maximal deviation type result for $\hat{\sigma}_{b_n}$. In the context of kernel density estimation, Bickel and Rosenblatt (1973) obtained a very deep maximal deviation result. Since then various results of maximal deviation type have been established under different settings; see Eubank and Speckman (1993) for independent data and Wu and Zhao (2007) and Zhao and Wu (2008) for time series. As demonstrated in these works, such maximal deviation result can be used to construct simultaneous confidence band for $\sigma(\cdot)$. The latter can be used to justify claims on the overall pattern of $\sigma(\cdot)$, such as whether it is indeed time-varying.

Theorem 4. *Assume that the conditions in Theorem 2 are fulfilled and $K \in \mathcal{K}_\omega, \omega > 0$. Let $m_n = \lceil (2\omega b_n)^{-1} - 1\rceil$ and $\mathcal{T}_n = \{\omega b_n + 2j\omega b_n : j = 0, 1, \ldots, m_n - 1\}$. For $n \geq 2$ let*

$$B_n(z) = \sqrt{2\log n} - \frac{1}{\sqrt{2\log n}}\left[\frac{1}{2}\log\log n + \log(2\sqrt{\pi})\right] + \frac{z}{\sqrt{2\log n}}. \tag{25}$$

Further assume that

$$nb_n^9\log n + \frac{(\log n)^3}{nb_n^3} \to 0. \tag{26}$$

Then for all $z \in \mathbb{R}$,

$$\lim_{n\to\infty}\mathbb{P}\left\{\frac{2f_\alpha(1)\sqrt{b_n/\Delta_n}}{\sqrt{\varphi_K}}\sup_{t\in\mathcal{T}_n}\frac{|\hat{\sigma}_{b_n}(t) - \sigma(t) - b_n^2\psi_K\rho_\sigma(t)|}{\sigma(t)} \leq B_{m_n}(z)\right\} = e^{-2e^{-z}}. \tag{27}$$

Condition (26) imposes constraints on the bandwidth b_n. The first part $nb_n^9\log n \to 0$ controls the bias and the second part $(\log n)^3/(nb_n^3) \to 0$ ensures the validity of the moderate deviation principle (see the proofs in Section 7). If $b_n \asymp n^{-\beta}$ for some $\beta \in (1/9, 1/3)$, then (26) holds. In particular, $\beta = 1/5$ corresponds to the optimal bandwidth.

In (21), the bias term $b_n^2\psi_K\rho_\sigma(t)$ contains unknown derivatives of $\sigma(t)$ and is difficult to estimate. To overcome this difficulty, we propose using

$$\tilde{\sigma}_{b_n}(t) = \frac{\hat{\sigma}^2_{b_n}(t)}{\hat{\sigma}_{\sqrt{2}b_n}(t)}. \tag{28}$$

By (21), the asymptotic bias of $\tilde{\sigma}_{b_n}(t) - \sigma(t)$ is of order $O(b_n^4)$. Another bias-correction technique is to use a fourth order kernel K so that $\psi_K = \int_{\mathbb{R}} u^2K(u)\mathrm{d}u = 0$; see Wu and Zhao (2007). An unpleasant problem of the latter approach is that the resulting estimator $\hat{\sigma}_{b_n}$ may be negative. In contrast, our bias-corrected estimator $\tilde{\sigma}_{b_n}$ is always non-negative.

5. A simulation study

In this section we shall conduct a simulation study to access the performance of our stable index estimator $\tilde{\alpha}$ in (15) and the bias-corrected median-quantile estimator $\tilde{\sigma}_{b_n}$ in (28). We use the program STABLE from J. P. Nolan (http://academic2.american.edu/~jpnolan) to generate stable random variables.

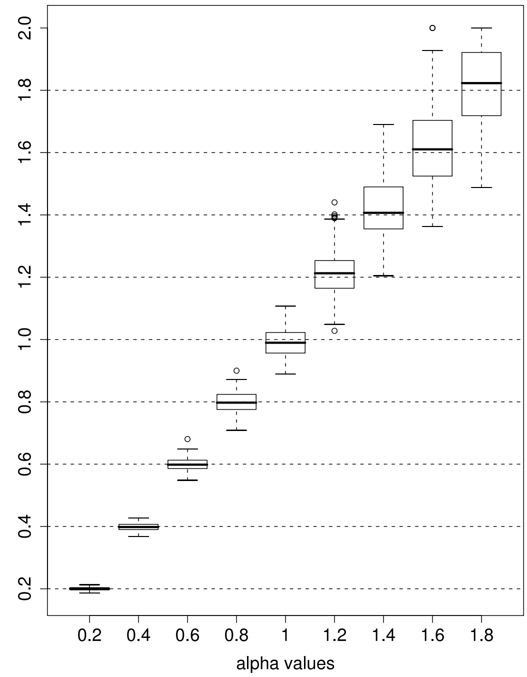

Fig. 1. Boxplots for the estimates of $\alpha = 0.2, 0.4, 0.6, 0.8, 1.0, 1.2, 1.4, 1.6$ and 1.8 with 200 realizations.

In model (1), let $\sigma(t) = 0.5 + 0.3\cos(t)$. To access the performance of $\tilde{\alpha}$ for different values of α, we choose $\alpha = 0.2, 0.4, 0.6, \ldots, 1.6, 1.8$. For each value of α, we generate 200 realizations from model (1) with $\Delta_n = 1/252$ (corresponding to daily observations) and sample size $n = 1500$ ($T \approx 6$ years) each, obtain the estimate $\tilde{\alpha}$ for each realization, and plot the boxplot for these 200 estimates. The boxplots in Fig. 1 show that the estimator performs reasonably well, and the median (the solid line inside the boxes) of those 200 estimates of α is very close to the true value of α. Furthermore, as α increases, the estimator $\tilde{\alpha}$ tends to have larger variations, which is in good agreement with the theoretical result in Theorem 1.

To evaluate the performance of the median-quantile estimator $\tilde{\sigma}_{b_n}$, we fix $\alpha = 1.6$ and use 4 bandwidths $b_n = 0.20, 0.30, 0.50$ and 0.60. Fig. 2 shows that the estimated functions are very close to the true function and they are relatively insensitive to the choice of bandwidth b_n. To further examine the influence of misspecification of α on the estimation of $\sigma(\cdot)$, we perform our estimation procedure with $b_n = 0.4$ and different misspecified values of α: 1.4, 1.5, 1.7, 1.8. Fig. 3 shows that the patterns of the estimated curves of $\sigma(\cdot)$ are reasonably consistent under mild misspecifications of α.

6. An application to foreign exchange rates

The dataset is obtained from http://www.federalreserve.gov/releases/h10/hist/, the website of Federal Reserve Bank of New York. It contains 2013 daily (weekdays) records of Canada Dollar/U.S. Dollar noon buying rates from January 3rd 2000 to December 31st 2007, a period of eight years. Then $n = 2013$, $\Delta_n = 1/252$ and $T = n\Delta_n = 8$ years. Denote by $X_{i\Delta_n}$, $1 \leq i \leq n$, the logarithms of the exchange rates. We assume that $\{X_{i\Delta_n}\}_{1\leq i\leq n}$ are discrete observations from (1). When computing the differences $X_{i\Delta_n} - X_{(i-1)\Delta_n}$ in (5), we subtract the overall median and use $X_{i\Delta_n} - X_{(i-1)\Delta_n} - \text{median}(X_{i\Delta_n} - X_{(i-1)\Delta_n}, 2 \leq i \leq n)$ so that the new differences have median zero.

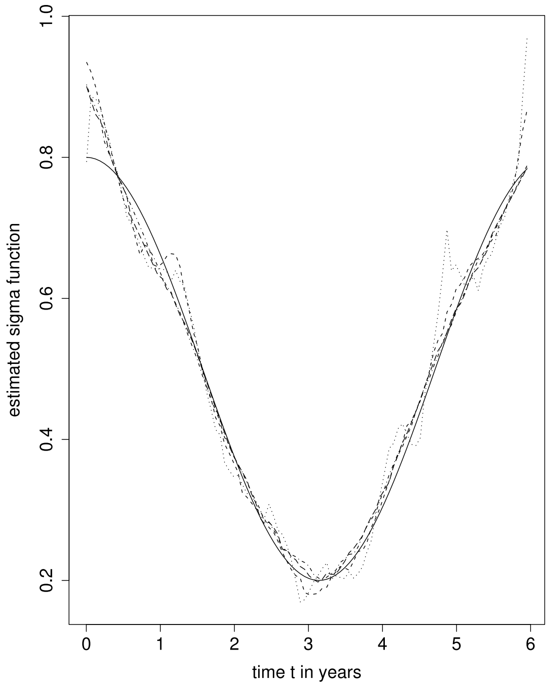

Fig. 2. Median-quantile estimates for $\sigma(t) = 0.5 + 0.3\cos(t)$ in model (1) with $\alpha = 1.6$. The solid curve is the true function $\sigma(t)$, the dotted, dashed, long dashed, and dot-dashed curves correspond to the estimates $\tilde{\sigma}_{b_n}(t)$ with bandwidth $b_n = 0.20, 0.30, 0.50$ and 0.60, respectively.

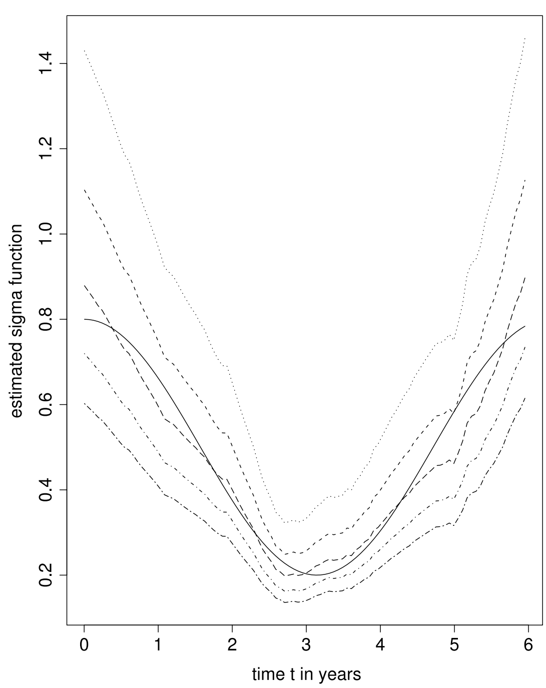

Fig. 3. Median-quantile estimates for $\sigma(t) = 0.5 + 0.3\cos(t)$ in model (1) with misspecified α's. The true $\alpha = 1.6$. The solid curve is the true function $\sigma(t)$, the dotted, dashed, long dashed, dot-dashed, and two-dashed curves (from top to bottom) correspond to the estimates $\tilde{\sigma}_{b_n}(t)$ with $b_n = 0.4$ and misspecified $\alpha = 1.4, 1.5, 1.6, 1.7$ and 1.8, respectively.

Applying the index estimator in (15), we obtain $\tilde{\alpha} = 1.69$ with standard error 0.11. Thus, we strongly reject the null hypothesis of the Brownian motion assumption $\alpha = 2$. Furthermore, we divide the differences $X_{i\Delta_n} - X_{(i-1)\Delta_n} - \text{median}(X_{i\Delta_n} - X_{(i-1)\Delta_n}, 2 \le i \le n)$ into two sub-periods of four years each, and the index estimates for these two sub-periods are both 1.7, which suggests that the index

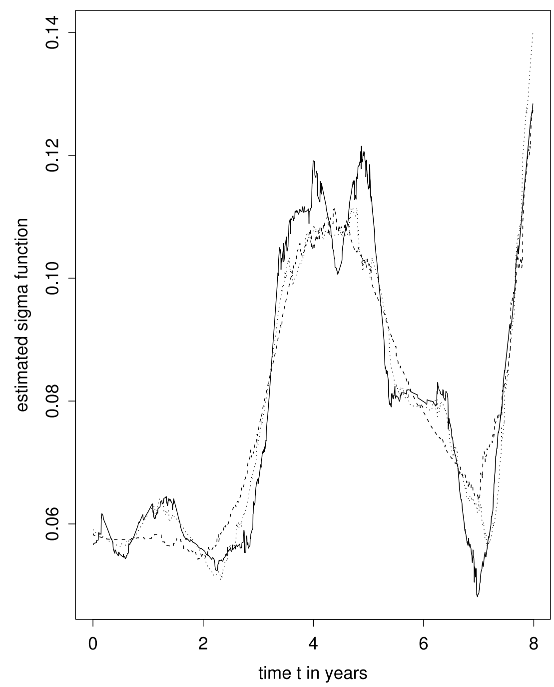

Fig. 4. Median-quantile estimate of the volatility function $\sigma(t)$ for Canadian Dollar/U.S. Dollar exchange rates. The solid, dotted, and dashed curves correspond to $\tilde{\sigma}_{b_n}(t)$ with bandwidth $b_n = 0.25, 0.40$ and 0.60, respectively.

estimator is reasonably consistent. To estimate the spot volatility $\sigma(t)$, we use (28) with different bandwidths $b_n = 0.25, 0.40$ and 0.60. Fig. 4 shows that the volatility changes over time and does not seem to have a simple parametric form. Also, different bandwidths yield quite similar volatility estimates. In what follows, we use $b_n = 0.4$.

Now we check the assumption that the unobservable process $\{Z_t(\alpha)\}_{t\ge 0}$ in (1) is a stable Lévy process with index $\tilde{\alpha} = 1.69$. By (16),

$$\Delta_n^{-1/\alpha}[X_{i\Delta_n} - X_{(i-1)\Delta_n}] \overset{\mathcal{D}}{\approx} \sigma(i\Delta_n) Z_1, \quad Z_1 \in \mathcal{G}(\alpha, c, 0).$$

Hence, under the null hypothesis, the estimated residuals

$$\epsilon_i = \frac{\Delta_n^{-1/\alpha}[X_{i\Delta_n} - X_{(i-1)\Delta_n}]}{\tilde{\sigma}_{b_n}(i\Delta_n)}, \quad 1 \le i \le n, \tag{29}$$

are approximately iid with distribution $\mathcal{G}(\tilde{\alpha}, c, 0)$, where $1/c = \mathcal{Q}(|Z_1(\tilde{\alpha})|) = 0.963$. We perform a Kolmogorov–Smirnov test. Let F_ϵ be the distribution function of $\mathcal{G}(\tilde{\alpha}, c, 0)$, and consider $G_n(u) = n^{-1/2}\sum_{i=1}^n [\mathbf{1}_{F_\epsilon(\epsilon_i)\le u} - u]$. Under the null hypothesis that ϵ_i, $1 \le i \le n$, are iid with distribution $\mathcal{G}(\tilde{\alpha}, c, 0)$, the Kolmogorov–Smirnov test statistic $G_n := \sup_{u\in[0,1]} |G_n(u)| \Rightarrow \sup_{u\in[0,1]} |B_t|$, where $\{B_t\}_{t\ge 0}$ is a Brownian bridge. We find $G_n = 1.086$ with p-value 0.189. Therefore, the null hypothesis is not rejected at 5% significance level.

We also analyze exchange rates between other major currencies and U.S. Dollar during the same time period from January 3rd 2000 to December 31st 2007. Table 1 summarizes our results. We conclude that stable Lévy processes with index between 1.64 and 1.81 can provide a good fit to these currency exchange rates.

Nolan (2001) also used stable distributions to model log returns of several exchange rates during other time periods prior to year 2000. He fitted the log returns directly and hence implicitly assumed a constant volatility. The latter assumption may be problematic in practice. In fact, we find that the estimated volatilities for the above currencies change over time. Marinelli et al. (2001) analyzed tick-by-tick (average time between quotes is about 2 min) exchange rates between U.S. Dollar and Swiss Franc during the period of May 20th 1985–May 20th 1987 and argued that stable

Table 1
Stable index estimates for exchanges rates.

Currencies	Stable index estimate $\tilde{\alpha}$	Standard error
Australian Dollar	1.64	0.106
Canadian Dollar	1.69	0.112
Danish Krone	1.71	0.115
European Euro	1.75	0.120
Japanese Yen	1.77	0.123
British Pound	1.81	0.129
Swiss Franc	1.79	0.126

distributions offer a good fit for the distribution of log returns. Early work by Westerfield (1977) also provided evidence in favor of symmetric stable distributions for exchange rates.

7. Proofs

7.1. *Proof of* Theorem 1

Recall the definition of γ_i, W_i, $\overline{W}_n$, γ_i^*, W_i^* and $\overline{W}_n^*$ in Section 3 [cf. displays (3), (6), (7) and (9)]. For simplicity we only consider the even integer n. Define

$$\overline{W}_n^{**} = \frac{2}{n}\sum_{i=1}^{n/2} \log|W_i^{**}|,$$

$$\text{where } W_i^{**} = \frac{W_{2i-1}+W_{2i}}{2^{1/\alpha}},\ 1 \le i \le \frac{n}{2}. \tag{30}$$

By the properties of SαS process in Section 2, W_i^{**}, $1 \le i \le n/2$, are iid distributed as Z_1. By definition, simple calculations show that $\gamma_{2i-1}/\gamma_i^* = 2^{-1/\alpha} + O(n^{-\iota})$, $\gamma_{2i}/\gamma_i^* = 2^{-1/\alpha} + O(n^{-\iota})$ and $W_i^* = W_{2i-1}\gamma_{2i-1}/\gamma_i^* + W_{2i}\gamma_{2i}/\gamma_i^*$. Thus,

$$|W_i^* - W_i^{**}| = O(n^{-\iota})(|W_{2i-1}| + |W_{2i}|). \tag{31}$$

The following Lemma 1 asserts that $\overline{W}_n^*$ can be well approximated by $\overline{W}_n^{**}$.

Lemma 1. *Let conditions in Theorem 1 hold. Then* $\sqrt{n}(\overline{W}_n^* - \overline{W}_n^{**}) = o_{\mathrm{p}}(1)$.

Proof. It suffices to show that $n\mathbb{E}(\overline{W}_n^* - \overline{W}_n^{**})^2 \to 0$. By definition, for each $1 \le i \le n/2$, (W_i^*, W_i^{**}) only depends on the increments of $\{Z_t\}$ over interval $[2(i-1)\Delta_n, 2i\Delta_n]$ and hence (W_i^*, W_i^{**}), $1 \le i \le n/2$, are independent. We have

$$n\mathbb{E}(\overline{W}_n^* - \overline{W}_n^{**})^2 = \frac{4}{n}\sum_{i=1}^{n/2}\mathbb{E}[(\log|W_i^*| - \log|W_i^{**}|)^2 \mathbf{1}_{|W_i^*|\ge|W_i^{**}|}]$$
$$+ \frac{4}{n}\sum_{i=1}^{n/2}\mathbb{E}[(\log|W_i^*| - \log|W_i^{**}|)^2 \mathbf{1}_{|W_i^*|<|W_i^{**}|}]. \tag{32}$$

We shall only show that the first term converges to zero since the second term can be similarly treated. Let $\epsilon_n = \log n$. Define the events

$$\mathcal{E}_i = \{|W_{2i-1}| + |W_{2i}| \le \epsilon_n\}, \qquad \mathcal{E}_i^* = \{|W_i^{**}| \ge \epsilon_n^{-1}\},$$

and denote by $\mathcal{E}_i^c$ and $\mathcal{E}_i^{*c}$ their corresponding complement events, respectively. Let

$$I_{n,1} = \frac{1}{n}\sum_{i=1}^{n/2}\mathbb{E}[(\log|W_i^*| - \log|W_i^{**}|)^2 \mathbf{1}_{|W_i^*|\ge|W_i^{**}|}\mathbf{1}_{\mathcal{E}_i\cap\mathcal{E}_i^*}],$$

$$I_{n,2} = \frac{1}{n}\sum_{i=1}^{n/2}\mathbb{E}[(\log|W_i^*| - \log|W_i^{**}|)^2 \mathbf{1}_{|W_i^*|\ge|W_i^{**}|}\mathbf{1}_{\mathcal{E}_i^c\cup\mathcal{E}_i^{*c}}].$$

By (31), on the event $\mathcal{E}_i$, $|W_i^* - W_i^{**}| = O(\epsilon_n n^{-\iota})$. Notice that, on the event $\mathcal{E}_i^*$, $|W_i^{**}| \ge \epsilon_n^{-1}$. Therefore, by the inequality $0 \le \log y - \log x = \log[1 + (y-x)/x] \le (y-x)/x$ for $0 < x \le y$, we have $I_{n,1} \to 0$ in view of

$$\mathbb{E}[(\log|W_i^*| - \log|W_i^{**}|)^2 \mathbf{1}_{|W_i^*|\ge|W_i^{**}|}\mathbf{1}_{\mathcal{E}_i\cap\mathcal{E}_i^*}]$$
$$\le \mathbb{E}\Big[\frac{|W_i^* - W_i^{**}|^2}{|W_i^{**}|^2}\mathbf{1}_{\mathcal{E}_i\cap\mathcal{E}_i^*}\Big]$$
$$= O(\epsilon_n^4 n^{-2\iota}). \tag{33}$$

Note that $\mathbb{P}(\mathcal{E}_i^c) \to 0$, $\mathbb{P}(\mathcal{E}_i^{*c}) \to 0$ and the fact that $\mathbb{E}(|\log|Z_1||^k) < \infty$, $k \in \mathbb{N}$, for SαS random variable Z_1. Thus, $I_{n,2} \to 0$ via the following Cauchy–Schwarz inequality:

$$\{\mathbb{E}[(\log|W_i^*| - \log|W_i^{**}|)^2 \mathbf{1}_{|W_i^*|\ge|W_i^{**}|}\mathbf{1}_{\mathcal{E}_i^c\cup\mathcal{E}_i^{*c}}]\}^2$$
$$\le \mathbb{E}[(\log|W_i^*| - \log|W_i^{**}|)^4] \times \mathbb{E}[\mathbf{1}_{\mathcal{E}_i^c\cup\mathcal{E}_i^{*c}}] \to 0,$$
$$\text{as } n \to \infty, \tag{34}$$

entailing the desired result. $\diamond$

Proof of Theorem 1. By the delta-method, it suffices to show the asymptotic normality of $\sqrt{n}(\hat{\alpha}^{-1} - \alpha^{-1})$. By (5), (8) and Lemma 1, we have

$$\begin{aligned}(\log 2)\sqrt{n}(\hat{\alpha}^{-1} - \alpha^{-1}) &= \sqrt{n}(\overline{W}_n^* - \overline{W}_n) + O(n^{1/2-\iota})\\ &= \sqrt{n}(\overline{W}_n^{**} - \overline{W}_n) + o_{\mathrm{p}}(1)\\ &= \frac{1}{\sqrt{n}}\sum_{i=1}^{n/2}[2\log|2^{-1/\alpha}(W_{2i-1} + W_{2i})|\\ &\quad - \log|W_{2i-1}| - \log|W_{2i}|] + o_{\mathrm{p}}(1).\end{aligned}$$

Then the asymptotic normality of $\sqrt{n}(\hat{\alpha}^{-1} - \alpha^{-1})$ easily follows from the fact that W_i, $1 \le i \le n$, are iid random variables distributed as Z_1. $\diamond$

7.2. *Proof of* Theorem 2

Lemmas 2–4 are needed to prove Theorem 2. Recall $\Delta_n = T/n = O(1/n)$.

Lemma 2. *Let $K \in \mathcal{K}_\omega$ and $g(u, v)$, $u, v \in [0, T]$, a bivariate measurable function satisfying*

$$\sup_{u,v\in[0,T]} |\partial^k g(u, v)/\partial u^k| < \infty, \quad k = 0, 1, \ldots, 4.$$

Assume that the bandwidth b_n satisfies $b_n \to 0$ and $nb_n \to \infty$. Then

$$\sup_{t\in[\omega b_n, T-\omega b_n]}\left|\frac{\Delta_n}{b_n}\sum_{i=1}^n K_{b_n}(i\Delta_n - t) - 1\right| = O[(nb_n)^{-1}] \tag{35}$$

and

$$\sup_{t\in[\omega b_n, T-\omega b_n]}\left|\frac{\Delta_n}{b_n}\sum_{i=1}^n g(i\Delta_n, t)K_{b_n}(i\Delta_n - t) - g(t, t) - b_n^2\psi_K \left.\frac{\partial^2 g(u, t)}{\partial u^2}\right|_{u=t}\right|$$
$$= O[(nb_n)^{-1} + b_n^4]. \tag{36}$$

Proof. By the Lipschitz continuity of K, we have, uniformly over $t \in [\omega b_n, T - \omega b_n]$,

$$\left|\frac{\Delta_n}{b_n}\sum_{i=1}^n K_{b_n}(i\Delta_n - t) - \frac{\Delta_n}{b_n}\int_0^n K_{b_n}(x\Delta_n - t)\mathrm{d}x\right|$$
$$\le \frac{\Delta_n}{b_n}\sum_{i=1}^n \int_{i-1}^i \left|K_{b_n}(i\Delta_n - t) - K_{b_n}(x\Delta_n - t)\right| \mathrm{d}x$$

$$\le \frac{\Delta_n}{b_n}\sum_{i\in\mathcal{I}}\int_{i-1}^{i}\left|K_{b_n}(i\Delta_n-t)-K_{b_n}(x\Delta_n-t)\right|\mathrm{d}x$$
$$=O[(nb_n)^{-1}], \tag{37}$$

where $\mathcal{I}=\{i\in\mathbb{N}: t-\omega b_n\le i\Delta_n\le t+\omega b_n+\Delta_n\}$ since the summand vanishes for i outside $\mathcal{I}$. The expression (35) then follows from the observation $\Delta_n b_n^{-1}\int_0^n K_{b_n}(x\Delta_n-t)\mathrm{d}x=\int_{-t/b_n}^{(T-t)/b_n}K(u)\mathrm{d}u=\int_{-\omega}^{\omega}K(u)\mathrm{d}u=1$ for $t\in[\omega b_n, T-\omega b_n]$. The expression (36) follows from the same argument as in (35) via Taylor's expansion $g(i\Delta_n,t)=\sum_{k=0}^{3}(i\Delta_n-t)^k/k!\partial^k g(u,t)/\partial u^k|_{u=t}+O(b_n^4)$ for $|i\Delta_n-t|\le\omega b_n$ in conjunction with the symmetry of K. We omit the details. ◇

Let $Y_i=\Delta_n^{-1/\alpha}|X_{i\Delta_n}-X_{(i-1)\Delta_n}|$ be as in Theorem 2. Recall that F_α, f_α are the distribution and density functions of $|Z_1(\alpha)|$, respectively, and that $F_\alpha(1)=1/2$; see Section 4. For $t\in[0,T]$ define

$$L_n(s,t)=\sum_{i=1}^{n}K_{b_n}(i\Delta_n-t)\mathbf{1}_{Y_i\le\sigma(t)+s}, \tag{38}$$

$$J_n(s,t)=\mathbb{E}[L_n(s,t)],\quad s\in\mathbb{R},\quad\text{and}$$
$$J_n(t)=\sum_{i=1}^{n}K_{b_n}(i\Delta_n-t). \tag{39}$$

Lemma 3. *Let $K\in\mathcal{K}_\omega$ and Condition 1 hold. Then the following expressions hold uniformly over $t\in[\omega b_n, T-\omega b_n]$:*

$$J_n(0,t)=J_n(t)/2-\Delta_n^{-1}b_n^3\psi_K\rho_\sigma(t)f_\alpha(1)/\sigma(t)+O(b_n+nb_n^5), \tag{40}$$

$$J_n(s,t)-J_n(0,t)=\Delta_n^{-1}b_n s[f_\alpha(1)/\sigma(t)+O(s+b_n^2+n^{-1})]. \tag{41}$$

Proof. Let γ_i be as in (3). By Taylor's expansion, we have $\Delta_n^{-1/\alpha}\gamma_i=\sigma(i\Delta_n)+O(1/n)$. By (16), simple calculations show that

$$\begin{aligned}J_n(0,t)&=\sum_{i=1}^{n}\mathbb{P}\{Y_i\le\sigma(t)\}K_{b_n}(i\Delta_n-t)\\&=\sum_{i=1}^{n}F_\alpha\left\{\frac{\sigma(t)}{\Delta_n^{-1/\alpha}\gamma_i}\right\}K_{b_n}(i\Delta_n-t)\\&=\sum_{i=1}^{n}[F_\alpha\{\sigma(t)/\sigma(i\Delta_n)\}+O(n^{-1})]K_{b_n}(i\Delta_n-t),\end{aligned} \tag{42}$$

which entails (41) in view of Lemma 2 with $g(u,v)=F_\alpha\{\sigma(v)/\sigma(u)\}$ and the assumption $F_\alpha(1)=1/2$. Similarly, we can derive an expansion for $J_n'(0,t):=\partial J_n(s,t)/\partial s|_{s=0}$, and the second expression then follows from $J_n(s,t)-J_n(0,t)=sJ_n'(0,t)+O(nb_ns^2)$. Details are omitted. ◇

Lemma 4. *Let $K\in\mathcal{K}_\omega$ and $\sigma\in\mathcal{C}^1[0,T]$. Assume that $b_n\to 0$ and $\delta_n\to 0$ are positive sequences satisfying $\sup_n\log n/(nb_n\delta_n)<\infty$. Then*

$$\sup_{t\in[0,T],|s|\le\delta_n}|[L_n(s,t)-J_n(s,t)]-[L_n(0,t)-J_n(0,t)]|$$
$$=O_{\mathrm{p}}(\sqrt{nb_n\delta_n\log n}) \tag{43}$$

and

$$\sup_{t\in[0,T]}|L_n(0,t)-J_n(0,t)|=O_{\mathrm{p}}(\sqrt{nb_n\log n}). \tag{44}$$

Proof. We only prove (43) since (44) follows similarly. Without loss of generality we consider positive $s\in[0,\delta_n]$. Let $N=\lfloor n/b_n\rfloor$ and $s_i=i\delta_n/N$, $t_i=iT/N$, $i=0,1,\ldots,N$. Then $\{s_i, 0\le i\le N\}$ and $\{t_i, 0\le i\le N\}$ divide the intervals $[0,\delta_n]$ and $[0,T]$, respectively, into N subintervals. For any $s\in[0,\delta_n]$ and $t\in[0,T]$, there exists unique j and k ($j=\lfloor Ns/\delta_n\rfloor$ and $k=\lfloor Nt/T\rfloor$) such that $s\in[s_j,s_{j+1})$ and $t\in[t_k,t_{k+1})$. Notice that $|t-t_k|\le|t_{k+1}-t_k|=T/N$. By the Lipschitz continuity of K and σ, there exists a constant c_1 such that $\sup_{t\in[0,T]}|\sigma(t)-\sigma(t_k)|\le c_1/N$ and $\sup_{t,u\in[0,T]}|K_{b_n}(u-t)-K_{b_n}(u-t_k)|\le c_1/(Nb_n)$. Write $[L_n(s,t)-J_n(s,t)]-[L_n(0,t)-J_n(0,t)]=\sum_{i=1}^{n}\{D_i(s,t)-\mathbb{E}[D_i(s,t)]\}$, where

$$D_i(s,t)=K_{b_n}(i\Delta_n-t)\mathbf{1}_{\sigma(t)<Y_i\le\sigma(t)+s},$$
$$\bar{A}_{ijk}=K_{b_n}(i\Delta_n-t_k)\mathbf{1}_{\sigma(t_k)-c_1/N<Y_i\le\sigma(t_k)+c_1/N+s_{j+1}},$$
$$\underline{A}_{ijk}=K_{b_n}(i\Delta_n-t_k)\mathbf{1}_{\sigma(t_k)+c_1/N<Y_i\le\sigma(t_k)-c_1/N+s_j}.$$

Then $\underline{A}_{ijk}-c_1/(Nb_n)\le D_i(s,t)\le\bar{A}_{ijk}+c_1/(Nb_n)$. Consequently,

$$\begin{aligned}\underline{A}_{ijk}-\mathbb{E}(\bar{A}_{ijk})-\frac{2c_1}{Nb_n}&\le D_i(s,t)-\mathbb{E}[D_i(s,t)]\\&\le\bar{A}_{ijk}-\mathbb{E}(\underline{A}_{ijk})+\frac{2c_1}{Nb_n}.\end{aligned} \tag{45}$$

Clearly, there exists a constant c_2 such that $0\le\mathbb{E}(\bar{A}_{ijk})-\mathbb{E}(\underline{A}_{ijk})\le c_2/N\le c_1/(Nb_n)$ for large n. Thus, by (45), we have

$$\begin{aligned}\underline{A}_{ijk}-\mathbb{E}(\underline{A}_{ijk})-\frac{3c_1}{Nb_n}&\le D_i(s,t)-\mathbb{E}[D_i(s,t)]\\&\le\bar{A}_{ijk}-\mathbb{E}(\bar{A}_{ijk})+\frac{3c_1}{Nb_n},\end{aligned}$$

which implies

$$\begin{aligned}\underline{U}_n(j,k)-\frac{3c_1n}{Nb_n}&\le\sum_{i=1}^{n}\{D_i(s,t)-\mathbb{E}[D_i(s,t)]\}\\&\le\bar{U}_n(j,k)+\frac{3c_1n}{Nb_n},\end{aligned} \tag{46}$$

where

$$\bar{U}_n(j,k)=\sum_{i=1}^{n}[\bar{A}_{ijk}-\mathbb{E}(\bar{A}_{ijk})]\quad\text{and}\quad\underline{U}_n(j,k)=\sum_{i=1}^{n}[\underline{A}_{ijk}-\mathbb{E}(\underline{A}_{ijk})].$$

Write $d_n=nb_n\delta_n$. Since $n/(Nb_n)=O(1)=o[(d_n\log n)^{1/2}]$, by (46), it remains to show that $\max_{0\le j,k\le N}|\bar{U}_n(j,k)|=O_{\mathrm{p}}[(d_n\log n)^{1/2}]$ and $\max_{0\le j,k\le N}|\underline{U}_n(j,k)|=O_{\mathrm{p}}[(d_n\log n)^{1/2}]$. We shall only prove the first expression since the second one follows similarly. Because Y_i has bounded density, $\mathbb{P}\{\sigma(t_k)-c_1/N<Y_i\le\sigma(t_k)+c_1/N+s_{j+1}\}=O(1/N+\delta_n)=O(\delta_n)$, where the second equality follows from $1/N=O(b_n/n)=o(\delta_n)$ in view of the assumption $\sup_n\log n/(nb_n\delta_n)<\infty$. Hence, by Lemma 2, there exists a constant c_3 such that

$$\begin{aligned}\sum_{i=1}^{n}\mathbb{E}[\bar{A}_{ijk}-\mathbb{E}(\bar{A}_{ijk})]^2&\le\sum_{i=1}^{n}\mathbb{E}(\bar{A}_{ijk}^2)=O(\delta_n)\sum_{i=1}^{n}K_{b_n}^2(i\Delta_n-t_k)\\&\le c_3d_n.\end{aligned} \tag{47}$$

Here the constant c_3 can be taken large enough so that $|\bar{A}_{ijk}-\mathbb{E}(\bar{A}_{ijk})|\le c_3$. The independent increments property of Lévy processes entails the independence of $\bar{A}_{ijk}$, $0\le i\le n$. Let $c>0$. By Bernstein's inequality (Bennett, 1962), we have

$$\begin{aligned}&\mathbb{P}\left\{\max_{0\le j,k\le N}|\bar{U}_n(j,k)|\ge c\sqrt{d_n\log n}\right\}\\&\le\sum_{0\le j,k\le N}\mathbb{P}\left\{|\bar{U}_n(j,k)|\ge c\sqrt{d_n\log n}\right\}\\&\le 2\sum_{0\le j,k\le N}\exp\left\{-\frac{c^2d_n\log n}{2c_3d_n+3cc_3(d_n\log n)^{1/2}}\right\}\\&\le 2(N+1)^2n^{-\lambda},\end{aligned} \tag{48}$$

where $\lambda = c^2/(2c_3 + 3cc_3c_4)$ and $c_4 = \sup_n(\log n/d_n)^{1/2} < \infty$. The desired result then follows by choosing sufficiently large c. $\diamond$

Proof of Theorem 2. Let $\delta_n = b_n^2 + (nb_n/\log n)^{-1/2}$ and k_n be a sequence of positive numbers such that $k_n \to \infty$ and $k_n\delta_n \to 0$. Let $\omega_n(t) := \hat{\sigma}_{b_n}(t) - \sigma(t)$. We first show that $\sup_{t\in[\omega b_n, T-\omega b_n]} |\omega_n(t)| = O_p(\delta_n)$. Recall $L_n(s,t)$ and $J_n(s,t)$ in (38) and (39). Because $\hat{\sigma}_{b_n}(t)$ is a solution to the minimization problem (18), by Koenker (2005, pp. 32–33), $\sup_{t\in[0,T]}|L_n(\omega_n(t),t) - J_n(t)/2| \le \sup_{t\in[0,T]}\sum_{i=1}^n K_{b_n}(i\Delta_n - t)\mathbf{1}_{Y_i=\hat{\sigma}_{b_n}(t)} \le \sup_{t\in[0,T]}K(t)\times$ (maximum number of identical Y_i's) $= O_p(1)$ by the continuity of Y_i's. Since $k_n \to \infty$, $k_n\delta_n \to 0$ and $\Delta_n = T/n$, by Lemmas 3 and 4, we have, uniformly over $t \in [\omega b_n, T - \omega b_n]$,

$$\begin{aligned}
&L_n(k_n\delta_n, t) - J_n(t)/2\\
&\quad = J_n(k_n\delta_n, t) - J_n(t)/2 + L_n(0,t) - J_n(0,t) + O_p(\sqrt{nb_n\log n})\\
&\quad = [J_n(0,t) - J_n(t)/2] + [J_n(k_n\delta_n, t) - J_n(0,t)]\\
&\qquad + O_p(\sqrt{nb_n\log n})\\
&\quad = O(nb_n^3 + b_n) + \Delta_n^{-1}b_nk_n\delta_n[f_\alpha(1)/\sigma(t) + o(1)]\\
&\qquad + O_p(\sqrt{nb_n\log n})\\
&\quad = \Delta_n^{-1}b_nk_n\delta_n[f_\alpha(1)/\sigma(t) + o(1)] + o_p(\Delta_n^{-1}b_nk_n\delta_n) \to \infty \qquad (49)
\end{aligned}$$

in view of $k_n \to \infty$, $nb_n\delta_n \to \infty$, and $\inf_{t\in[0,T]}\sigma(t) > 0$. Similarly, we have $L_n(-k_n\delta_n, t) - J_n(t)/2 \to -\infty$, uniformly over $t \in [\omega b_n, 1 - \omega b_n]$. Since $L_n(s,t)$ is non-decreasing in s and $\sup_{t\in[0,T]}|L_n(\omega_n(t),t) - J_n(t)/2| = O_p(1)$, we conclude that $\sup_{t\in[\omega b_n, T-\omega b_n]}|\omega_n(t)| = O_p(k_n\delta_n)$. Since $k_n \to \infty$ can be arbitrarily slow, we have $\sup_{t\in[\omega b_n,T-\omega b_n]}|\omega_n(t)| = O_p(\delta_n)$, completing the proof of (i).

To prove (ii), by (i), we have $\omega_n(t) = O_p(\delta_n)$. Apply Lemmas 3 and 4 again,

$$\begin{aligned}
&L_n(0,t) - J_n(0,t)\\
&\quad = L_n(\omega_n(t),t) - J_n(\omega_n(t),t) + O_p(\sqrt{nb_n\delta_n\log n})\\
&\quad = [L_n(\omega_n(t),t) - J_n(t)/2] - [J_n(0,t) - J_n(t)/2]\\
&\qquad - [J_n(\omega_n(t),t) - J_n(0,t)] + O_p(\sqrt{nb_n\delta_n\log n})\\
&\quad = O_p(1) + \Delta_n^{-1}b_n^3\psi_K\rho_\sigma(t)f_\alpha(1)/\sigma(t) - \Delta_n^{-1}b_n\omega_n(t)f(1)/\sigma(t)\\
&\qquad + O_p[\sqrt{nb_n\delta_n\log n} + nb_n(\delta_n^2 + n^{-1})]. \qquad (50)
\end{aligned}$$

The Bahadur–Kiefer representation (21) follows by solving $\omega_n(t)$ in the above equation. $\diamond$

7.3. Proofs of Theorems 3 and 4

Proof of Theorem 3. Let $Q_{b_n}(t)$ be as in Theorem 2. Note that $\{Y_i\}_{1\le i\le n}$ are independent. Under condition (23), the error term R_n in (22) satisfies $R_n\sqrt{nb_n} \to 0$. By the Bahadur–Kiefer representation (21), it suffices to show the asymptotic normality for $\sqrt{b_n/\Delta_n}Q_{b_n}(t)$. The boundedness of K entails the Lindeberg condition. By (16) and Lemma 2, elementary calculations show that $(b_n/\Delta_n)\mathrm{Var}[Q_{b_n}(t)] \to \varphi_K/4$. Then $\sqrt{b_n/\Delta_n}Q_{b_n}(t) \Rightarrow N(0, \varphi_K/4)$, and the desired result follows. $\diamond$

Proof of Theorem 4. Under condition (26), we have $R_n(nb_n\log n)^{1/2} \to 0$. So, by Slutsky's theorem and the Bahadur–Kiefer representation (21), it suffices to show that

$$\lim_{n\to\infty}\mathbb{P}\left\{\sup_{t\in\mathcal{T}_n}|V_n(t)| \le B_{m_n}(z)\right\} = \mathrm{e}^{-2\mathrm{e}^{-z}},$$

$$\text{where } V_n(t) = \frac{2}{\sqrt{\varphi_K}}\sqrt{b_n/\Delta_n}Q_{b_n}(t). \qquad (51)$$

Write $V_n(t) = \sum_{i=1}^n \eta_i(t)$, where

$$\eta_i = \frac{2}{\sqrt{\varphi_K b_n/\Delta_n}}\{\mathbf{1}_{Y_i\le\sigma(t)} - \mathbb{E}[\mathbf{1}_{Y_i\le\sigma(t)}]\}K_{b_n}(i\Delta_n - t). \qquad (52)$$

For each fixed $t \in [\omega b_n, T - \omega b_n]$, notice that $\{\eta_i(t)\}_{1\le i\le n}$ are independent random variables, and in particular they form martingale differences with respect to $\mathcal{F}_i$, the sigma-field generated by $\{Y_j\}_{j\le i}$. Since $F_\alpha(1) = 1/2$, by (16) and Lemma 2, we have as in (42),

$$\begin{aligned}
\sum_{i=1}^n\mathbb{E}[\eta_i^2(t)|\mathcal{F}_{i-1}] &= \frac{4\Delta_n}{\varphi_K b_n}\sum_{i=1}^n\{\mathbb{P}[Y_i \le \sigma(t)] - \mathbb{P}^2[Y_i \le \sigma(t)]\}\\
&\quad \times K_{b_n}^2(i\Delta_n - t)\\
&= \frac{4\Delta_n}{\varphi_K b_n}\sum_{i=1}^n\left[F_\alpha\left\{\frac{\sigma(t)}{\sigma(i\Delta_n)}\right\} - F_\alpha^2\left\{\frac{\sigma(t)}{\sigma(i\Delta_n)}\right\}\right]\\
&\quad \times K_{b_n}^2(i\Delta_n - t) + O(n^{-1})\\
&= 1 + O(r_n), \qquad (53)
\end{aligned}$$

uniformly over $t \in [\omega b_n, T - \omega b_n]$, where $r_n = b_n^2 + (nb_n)^{-1}$. Elementary manipulations show that $\sum_{i=1}^n\mathbb{E}|\eta_i^3(t)| = O[(nb_n)^{-1/2}]$. Let z be fixed and write $\Xi_n = [1 + B_{m_n}(z)]^4\exp[B_{m_n}^2(z)/2][r_n^{3/2} + (nb_n)^{-1/2}]$. Under condition (26), $\Xi_n \to 0$. Let $t_j = \omega b_n + 2j\omega b_n$, $0 \le j \le m_n - 1$. Denote by E_j the event $\{|V_n(t_j)| \le B_{m_n}(z)\}$ and by E_j^c its complement event. Let W be a standard normal random variable. By Theorem 1 in Grama and Haeusler (2006),

$$\mathbb{P}[E_j^c] = [1 + O(\Xi_n^{1/4})]\mathbb{P}[|W| \ge B_{m_n}(z)] = \frac{2\mathrm{e}^{-z}}{m_n}[1 + o(1)], \qquad (54)$$

uniformly in j, in view of $\mathbb{P}(|W| \ge x) = [1 + o(1)]\phi(x)/x$ as $x \to \infty$, where ϕ is the standard normal density function.

Denote by $\mathcal{I}_j = \{i : \eta_i(t_j) \ne 0\}$ the set of indices of non-vanishing summands in $V_n(t_j)$. For $j \ne j'$, Since K has support $[-\omega, \omega]$ and $|t_j - t_{j'}| \ge 2\omega b_n$, $\mathcal{I}_j$ and $\mathcal{I}_{j'}$ have no common elements. Therefore, the independence of $\{Y_i\}_{1\le i\le n}$ implies the independence of $V_n(t_j)$ and $V_n(t_{j'})$. So, (51) follows from

$$\begin{aligned}
&\mathbb{P}\{\sup_{t\in\mathcal{T}_n}|V_n(t)| \le B_{m_n}(z)\} = \prod_{j=0}^{m_n-1}\{1 - \mathbb{P}[E_j^c]\}\\
&\quad = \left\{1 - \frac{2\mathrm{e}^{-z}}{m_n}[1 + o(1)]\right\}^{m_n} \to \mathrm{e}^{-2\mathrm{e}^{-z}}. \quad \diamond \qquad (55)
\end{aligned}$$

Acknowledgements

We are grateful to Professor Jean Jacod for valuable suggestions and for providing us with his manuscripts. We also thank an Associate Editor and an anonymous referee for their very helpful comments on an earlier version of this paper. The work is supported in part by NSF grant DMS-0478704.

References

Aït-Sahalia, Y., Jacod, J., 2007. Volatility estimators for discretely sampled Lévy processes. The Annals of Statistics 35, 355–392.

Aït-Sahalia, Y., Jacod, J., 2008. Fisher's information for discretely sampled Lévy processes. Econometrica 76, 727–761.

Bennett, G., 1962. Probability inequalities for the sum of independent random variables. Journal of the American Statistical Association 57, 33–45.

Bickel, P.J., Rosenblatt, M., 1973. On some global measures of the deviations of density function estimates. The Annals of Statistics 1, 1071–1095.

Black, F., Scholes, M., 1973. The pricing of options and corporate liabilities. Journal of Political Economy 81, 637–654.

Corcuera, J.M., Nualart, D., Woerner, J.H., 2007. A functional central limit theorem for the realized power variation of integrated stable processes. Stochastic Analysis and Applications 25, 169–186.

De Haan, L., Resnick, S., 1980. A simple asymptotic estimate for the index of a stable distribution. Journal of the Royal Statistical Society, B 42, 83–88.
Eubank, R.L., Speckman, P.L., 1993. Confidence bands in nonparametric regression. Journal of the American Statistical Association 88, 1287–1301.
Fama, E.F., 1965. The behavior of stock-market prices. Journal of Business 38, 34–105.
Fama, E.F., Roll, R., 1971. Parameter estimates for symmetric stable distributions. Journal of the American Statistical Association 66, 331–338.
Fan, J., 2005. A selective overview of nonparametric methods in financial econometrics. Statistical Science 20, 317–337.
Fan, Z., 2006. Parameter estimation of stable distributions. Communications in Statistics: Theory and Methods 35, 245–255.
Foster, D.P., Nelson, D.B., 1996. Continuous record asymptotics for rolling sample variance estimators. Econometrica 64, 139–174.
Grama, I.G., Haeusler, E., 2006. An asymptotic expansion for probabilities of moderate deviations for multivariate martingales. Journal of Theoretical Probability 19, 1–44.
Koenker, R., 2005. Quantile Regression. Cambridge University Press, New York.
Mandelbrot, B., 1963. The variation of certain speculative prices. Journal of Business 36, 394–419.
Marinelli, C., Rachev, S.T., Roll, R., 2001. Subordinated exchange rate models: Evidence for heavy tailed distributions and long-range dependence. Mathematical and Computer Modelling 34, 955–1001.
McCulloch, J.H., 1986. Simple consistent estimators of stable distribution parameters. Communication in Statistics: Simulation and Computation 15, 1109–1136.
Nikias, C.L., Shao, M., 1995. Signal Processing with Alpha-stable Distributions and Applications. Wiley, New York.
Nolan, J.P., 2001. Maximum likelihood estimation of stable parameters. In: Barndorff-Nielsen, O.E., Mikosch, T., Resnick, I. (Eds.), Lévy Processes: Theory and Applications. Birkhäuser, Boston, pp. 379–400.
Press, J.S., 1972. Estimation in univariate and multivariate stable distributions. Journal of The American Statistical Association 67, 842–846.
Stanton, R., 1997. A nonparametric model of term structure dynamics and the market price of interest rate risk. Journal of Finance 52, 1973–2002.
Westerfield, J.M., 1977. An examination of foreign exchange risk under fixed and floating rate regimes. Journal of International Economics 7, 181–200.
Woerner, J.H., 2004. Purely discontinuous Lévy processes and power variation: Inference for integrated volatility and the scale parameter. Manuscript.
Wu, W.B., Zhao, Z., 2007. Inference of trends in time series. Journal of the Royal Statistical Society, B 69, 391–410.
Zhang, L., Mykland, P., Aït-Sahalia, Y., 2005. A tale of two time scales: Determining integrated volatility with noisy high-frequency data. Journal of the American Statistical Association 472, 1394–1411.
Zhao, Z., Wu, W.B., 2008. Confidence bands in nonparametric time series regression. The Annals of Statistics 36, 1854–1878.