ECONOMETRIC REVIEWS
2018, VOL. 37, NO. 10, 1051–1066
http://dx.doi.org/10.1080/07474938.2016.1188876

The estimation for Lévy processes in high frequency data

Jing Zheng[a], Wentao Gu[b], Baolin Xu[c], and Zongwu Cai[d]

[a]Institute of Applied Mathematics, Hangzhou Dianzi University, Hangzhou, China; [b]Department of Statistics, Zhejiang Gongshang University, Hangzhou, China; [c]Department of Finance, School of Economics, Xiamen University, Xiamen, China; [d]Department of Economics, University of Kansas, Lawrence, Kansas, USA

ABSTRACT
In this article, a generalized Lévy model is proposed and its parameters are estimated in high-frequency data settings. An infinitesimal generator of Lévy processes is used to study the asymptotic properties of the drift and volatility estimators. They are consistent asymptotically and are independent of other parameters making them better than those in Chen et al. (2010). The estimators proposed here also have fast convergence rates and are simple to implement.

KEYWORDS
Financial data; high frequency; jump; Lévy measure; Lévy process

JEL CLASSIFICATION
C13

1. Introduction

The Lévy process is currently of considerable interest to researchers in econometric and financial fields due to its ability to explain abrupt jumps in financial markets and due to its mathematical tractability. For example, it is used in the analysis of discontinuous changes in price, believed to be an essential component of financial asset-price dynamics. See Cont and Tankov (2003) for more detail.

Let $\{X_t,\ t \in R^+\}$ be a Lévy process taking values in R, whose characteristic function is given as follows:

$$E(e^{i\theta X_t}) = \exp\left\{t\left[ib\theta - \frac{\sigma^2\theta^2}{2} + \int_R (e^{i\theta u} - 1 - i\theta u I_{|u|<1})\rho(du)\right]\right\}. \tag{1.1}$$

Here, $b \in R$, $\sigma \geq 0$, and ρ is a measure on R satisfying $\rho(\{0\}) = 0$ and the integrability condition

$$\int_R (1 \wedge |u|^2)\rho(du) < \infty.$$

This formula is known as the Lévy-Khintchine formula. The process $\{X_t, t \in R^+\}$ is characterized by the generating triplet $(b, \sigma, \rho(\cdot))$, where b is the drift of X_t, and ρ is the Lévy measure. Please refer to Bertoin (1996) and Sato (1999) for precise definitions and more details concerning properties of Lévy processes.

Aït-Sahalia and Jacod (2007, 2009), studied many properties of Lévy processes using high frequency financial data: the degree of activity, the test of jump, and the estimator of σ. However, it is difficult to identify the estimators of the drift and the Lévy measure from observations of finite time interval, as indicated by Aït-Sahalia and Jacod (2009).

Because of the explicit expression of the Lévy-Khintchine formula, some researchers have estimated the parameters of the Lévy process based on the characteristic function expression of (1.1). Chen et al. (2010) proposed a nonparametric approach to estimating the characteristics of Lévy processes using regression of the empirical characteristic function. However, in their article, the estimator of the drift was dependent on other unknown parameters. According to the authors, "The regression models used here to define estimators seem perhaps rather naive, and in particular, the polynomial models considered there may seem to be crude approximations" (Chen et al., 2010, p. 259).

CONTACT Wentao Gu zjgsu-guwentao@hotmail.com Department of Statistics, Zhejiang Gongshang University, Hangzhou, China.

In this article, the following model with the characteristic function is considered:

$$E(e^{i\theta X_t}) = \exp\left\{t\left[ib\theta - \frac{\sigma^2}{2}\theta^2 + \int_{\{|x|\le\epsilon\}}(e^{i\theta u} - 1 - i\theta u I_{|u|<1})cu^{-\alpha-1}(du)\right.\right.$$
$$\left.\left. + \int_{\{|x|>\epsilon\}}(e^{i\theta u} - 1 - i\theta u I_{|u|<1})\rho_2(du)\right]\right\}.$$

Here, $0 < \epsilon \le 1$ and $0 < \alpha < 2$. The current model can approximate the general Lévy process, and it is more general than that described by Chen et al. (2010). Another advantage of this model is that it can provide one possible way of estimating the Lévy measure. It is well known that Lévy measure can be a infinity measure and explode near zero; here, we approximate it with a stable measure near zero, similar to that in Aït-Sahalia and Jacod (2009).

Using the infinitesimal generator of Lévy processes, we propose estimators of the drift and the volatility, which are consistent under high frequency data (a Lévy process is observed at equidistant time points $\Delta_n, \ldots, n\Delta_n$ with $\lim_{n\to\infty}\Delta_n = 0$, $\lim_{n\to\infty} n\Delta_n = \infty$), and the rates of convergence are faster than those of Chen et al. (2010). We will discuss this method in detail in Sections 2 and 3.

For the estimation of c, the estimator was identified using the small-time ergodic property of the Lévy process:

$$\lim_{t\to 0}\frac{1}{t}EX_t^2 I_{\{|X_t|<\epsilon\}} = \sigma^2 + \int_{\{|x|<\epsilon\}} x^2\rho(dx).$$

This estimator is asymptotically normal.

Recently, related problems have been treated by Comte et al. (2009), Ueltzhöfer (2013), and Schmisser (2014). Comte et al. (2009) proposed estimators of b, σ^2, and ρ by Fourier inversion and kernel smoothing. Their models involved finite jump activity and had absolutely continuous Lévy measures. They considered a superposition of a compound Poisson process and an independent Brownian motion, which is treated as a special case in the current model. Comte and Genon-Catalot (2011) studied nonparametric estimation of the Lévy measure at high frequencies using an adaptive deconvolution method. In their model, the strong moment condition ($E|X_1|^4 < \infty$) must be satisfied. Chen et al. (2010) used a nonparametric method to estimate the parameters of a class of Lévy processes.

The article is organized as follows. The proposed model is given in Section 2. In Section 3, the construction of consistent estimators of parameters b, c, and measure ρ is described. In Section 4, the theoretical properties of the estimators are discussed. Section 5 reports a simulation experiment. A real data example is given in Section 6. Technical details are deferred to the Appendix.

Throughout this article, $X = \{X_t,\ t \in R^+\}$ is a Lévy process and C always stands for a positive constant, whose value may be different at different contexts. The expectation operator under P and P_x are denoted by E and E_x. $\varphi_\eta(\cdot)$ stands for the characteristic function of a random variable (a distribution) η.

2. Model specification

Let $X = \{X_t,\ t \in R^+\}$ be a Lévy process taking values in R on the naturally filtered space generated by itself, whose characteristic function is given by (1.1).

For any fixed $0 < \epsilon \le 1$, X can be expressed as

$$X_t = bt + \sigma W_t + Z_t^1 + Z_t^2. \tag{2.1}$$

Here, b and σ are the drift and the volatility of X, respectively. W is a Brownian motion. Z^1 is a Lévy process that has fewer than ϵ jumps and is called the compensated sum of jumps. Z^2 is another Lévy process which has jumps more than ϵ. It is well known that W, Z^1, and Z^2 are independent. In fact, Z^1 is a square-integrable martingale with expectation zero, and Z^2 is a compound Poisson process. When $\epsilon = 1$, (2.1) is the Lévy-Itô decomposition. It is written as (2.1) in order to avoid the need for complex symbols. In this article, we assume that Z^1 is a symmetrical and stable-like process whose Lévy measure is stable and supported in $[-\epsilon, \epsilon]$. Hence, the Lévy measure and the characteristic function of X can be written as follows:

$$\rho(dx) = cx^{-\alpha-1} I_{\{|x|<\epsilon\}} dx + \rho_2 I_{\{|x|\geq\epsilon\}} dx, \tag{2.2}$$

and

$$E(e^{i\theta X_t}) = \exp\left\{ t \left[ib\theta - \frac{\sigma^2}{2}\theta^2 + \int_{\{|x|\leq\epsilon\}} (e^{i\theta u} - 1 - i\theta u I_{|u|<1}) cu^{-\alpha-1}(du) + \int_{\{|x|>\epsilon\}} (e^{i\theta u} - 1 - i\theta u I_{|u|<1}) \rho_2(du) \right] \right\}. \tag{2.3}$$

Here, ρ_2 is the Lévy measure of Z_2.

One advantage of this model is its ability to estimate the Lévy measure. When the Lévy measure explodes near 0, it is not possible to estimate it with finite data. However, a great deal of financial data tell us that infinite activity exists and the model above could be used to approximate general Lévy processes. This idea is similar to the conclusions drawn by Aït-Sahalia and Jacod (2009), who use an activity index to describe small jumps. The activity index defined by Aït-Sahalia and Jacod (2009) was equal to α in (2.2), but the current model is more general than that proposed in that article. When $b = 0$ and either σ or Z^1 is 0, it is appopriate to use an earlier model also proposed by Aït-Sahalia and Jacod (2007), who assesses the volatility estimators using Fisher information. Chen et al. (2010) considered the following model, whose characteristic function is given as follows:

$$E(e^{i\theta X_t}) = \exp\{ib\theta - c|\theta|^\alpha + \psi(\theta)\},$$

where $\psi(\theta)$ is the characteristic function of a compound Poisson process. This model is a special case of the currently proposed model when σ or c is zero.

The following examples are provided to show how a general Lévy process can be approximated by our model with a choice of small enough ϵ.

2.1. Example 1

Drift + Brownian motion + Stable process

Consider the Lévy measure $\rho(dx) = c|x|^{-\alpha-1}dx$ for $x \neq 0$. X can be written as follows:

$$X_t = bt + \sigma W_t + Z_t^1 + Z_t^2,$$

where Z^1 is a pure jump martingale with Lévy measure $c|x|^{-\alpha-1} I_{\{|x|\leq\epsilon\}} dx$, and Z^2 is a compound Poisson process with finite measure $cx^{-\alpha-1} I_{\{|x|>\epsilon\}} dx$, which is another way of describing the currently proposed model.

2.2. Example 2

Drift + Brownian motion + compound Poisson process

Let the following be true:

$$X_t = bt + \sigma W_t + \sum_{i=1}^{N_t} Y_i,$$

where $N = \{N_t, t \in R^+\}$ is a Poisson process with constant intensity λ, and Y_i is a sequence of independent and identically distributed (*i.i.d.*) random variables with density f, which is independent of the Poisson process N. Then $\sum_{i=1}^{N_t} Y_i$ is a compound Poisson process and X is a Lévy process with Lévy measure $\lambda f(x)dx$. Let $Z_t^1 = 0$ and $Z_t^2 = \sum_{i=1}^{N_t} Y_i$. This model can be approximated by (2.3), and the error of ρ in (2.2) is $O(\epsilon)$.

2.3. Example 3

Drift + Brownian motion + Gamma process

Consider the Lévy measure $\rho(dx) = a\exp(-bx)x^{-1}I_{x>0}dx$. X can be written as

$$X_t = bt + \sigma W_t + Z_t^1 + Z_t^2,$$

where Z^1 is a pure jump martingale with Lévy measure $a\exp(-bx)x^{-1}I_{\{0<x\leq\epsilon\}}dx$, and Z^2 is a compound Poisson process with finite measure $a\exp(-bx)x^{-1}I_{\{x>\epsilon\}}dx$. The density function of this model has a semi-heavy (right) tail and has an infinite number of small jumps because the Lévy measure explodes near 0. In this model, the index α is one. It can also be approximated by (2.3) and the error of ρ in (2.2) is $o(\epsilon)$.

2.4. Example 4

Drift + Brownian motion + Variance gamma process

Let the Lévy measure $\rho(dx) = a\exp(Ax - B|x|)|x|^{-1}dx$ with $A = \frac{\theta}{\sigma^2}$ and $B = \frac{(\theta^2+2\sigma^2/a)^{1/2}}{\sigma^2}$. X can be written as follows:

$$X_t = bt + \sigma W_t + Z_t^1 + Z_t^2,$$

where Z^1 is a pure jump martingale with Lévy measure $a\exp(Ax - B|x|)|x|^{-1}I_{\{|x|\leq\epsilon\}}dx$, and Z^2 is a compound Poisson process with finite measure $a\exp(Ax-B|x|)|x|^{-1}I_{\{|x|>\epsilon\}}dx$. The index α here is one. It can also be approximated using (2.3), and the error of ρ in (2.2) is $O(\epsilon)$.

3. Methodology

In this section, the parameters of the proposed model are estimated using (2.3). First, we define some notations. Denote the observed n independent increments $Y_i = X_{i\Delta_n} - X_{(i-1)\Delta_n}$, $i = 1, 2 \ldots n$. When $\Delta = \Delta_n$ tends to 0, and $n\Delta$ tends to infinity, the sequence Y_i is a high frequency datum, such as can be found in financial data. Hence Δ and Y_i depend on n. However, to simplify notation, we omit the dependence on n and simply write Δ, Y_i. Because the Lévy process has independent and stationary increments, $Y_i, i = 1, 2, \ldots n$, is *i.i.d.* with the characteristic function of (1.1), where t is replace with Δ.

3.1. Estimation of b and σ^2

In most phenomena that can be modeled using a Lévy process, the way in which the process moves from point to point can be grasped intuitively. The infinitesimal generator may be one means of describing how the process moves from point to point in an infinitesimally small time interval.

For Lévy processes, the infinitesimal generator

$$(Lf)(x) := \lim_{t\to 0}\frac{1}{t}\{Ef(X_t + x) - f(x)\} \tag{3.1}$$

exists, and the following is true:

$$\begin{aligned}(Lf)(x) &= bf'(x) + \frac{\sigma^2}{2}f''(x)\\ &\quad + \int_R [f(x+y) - f(x) - f'(x)yI_{\{|y|<1\}}]\rho(dy),\end{aligned} \tag{3.2}$$

where $f(x) \in \mathfrak{C}_0^2(\mathfrak{b})$, which is twice continuously differentiable, and it is bounded together with its derivatives and tends to 0 at infinity. For example, $f(x)$ could be the density function of normal distribution. The following is true:

$$\left\|(Lf)(x) - \frac{1}{t}[Ef(X_t + x) - f(x)]\right\| \to 0$$

as $t \to 0$ (see Revuz and Yor, 2005, p. 281),

$$\left\|(Lf)(x) - \frac{1}{t}[Ef(X_t + x) - f(x)]\right\| < \epsilon(t).$$

When $f \in \mathfrak{C}_0^\infty(K)$, the convergence rate for (3.1) can be established (see Theorem A.1). This suggests that regression can be used to estimate the parameters of (3.2). Let

$$(\hat{L}f)(x) = \frac{1}{\Delta}\left\{\frac{1}{n}\sum_{i=1}^{n} f(Y_i + x) - f(x)\right\}.$$

Then the following equation is true:

$$(\hat{L}f)(x) = bf'(x) + \frac{\sigma^2}{2}f''(x) + \int_R [f(x+y) - f(x) - f'(x)yI_{\{|y|<1\}}]\rho(dy) + \epsilon(x),$$

where

$$\epsilon(x) = (\hat{L}f)(x) - (Lf)(x). \tag{3.3}$$

Choosing a set Γ of values x and minimizing the following equation:

$$\int_\Gamma |(\hat{L}f)(x) - bf'(x) - \frac{\sigma^2}{2}f''(x) - \int_R [f(x+y) - f(x) - f'(x)yI_{\{|y|<1\}}]\rho(dy)|^2 dx, \tag{3.4}$$

this produces the estimator of b. Also (3.4) could be rewritten as follows:

$$\int_\Gamma |(\hat{L}f)(x) - bf'(x)|^2 dx. \tag{3.5}$$

The estimator of b is provided as follows:

$$\hat{b} = \frac{\int_\Gamma f'(x)(\hat{L}f)_(x)dx}{\int_\Gamma [f'(x)]^2 dx}.$$

Remark 3.1. In practice, each regression needs to be implemented over a set Γ of values x. It is important to choose the fitting function $f(x)$ and set Γ. In the current simulation, Γ is symmetrical and far enough from zero. The length of Γ much greater than that of $(n\Delta)$, which makes the value of $\epsilon(x)$ in (3.3) small. $f(x)$ must be in the domain of the infinitesimal generator L. Choosing suitable $f(x)$ and Γ can improve the accuracy of the estimation. This is discussed in Section 4.

Estimators of σ^2 based on power variations of X have been proposed and mostly studied in high frequency data for which t is fixed. Aït-Sahalia and Jacod (2007) identified the estimator of σ^2 under parametric and semiparametric settings with complex constructions and proofs. σ^2 was estimated from (3.2) by extension of this methodology. Similar to the estimation of b, let

$$(\hat{L}f)(x) = \frac{1}{\Delta}\left\{\frac{1}{n}\sum_{i=1}^{n} f(Y_i + x) - f(x)\right\}.$$

The regression can be used to determine the estimator of σ^2 as follows:

$$\hat{\sigma^2} = \frac{\int_\Gamma f''(x)(\hat{L}f)(x)dx}{\int_\Gamma [f''(x)]^2 dx}. \tag{3.6}$$

Like the estimation of b, it is important to choose suitable $f(x)$ and Γ in (3.6) in order to render the bias small. Before offering other estimators, a few remarks should be made.

Remark 3.2. When $\int_{|x|<1} |x|\rho(dx) < \infty$, the jump part of X has bounded variation on every compact time interval *a.s.*, and the characteristic function of X can be written as follows:

$$E(e^{i\theta X_t}) = \exp\left\{t\left[ib'\theta - \frac{\sigma^2\theta^2}{2} - \int_R (e^{i\theta u} - 1)\rho(du)\right]\right\},$$

$b' = b + \int_{|x|<1} x\rho(dx)$, some authors call it as the generous drift. As described in Section 5, data can be simulated with Poison jumps and Cauchy jumps, and b can be estimated in this paper.

Remark 3.3. Comte and Genon-Catalot (2011) identified the estimators of b and σ^2 using empirical moments under the assumption of finite moments. In the current article, finite moment restrictions are not needed.

Remark 3.4. By (Bertoin, 1996, p. 24), the Fourier transform of $(Lf)f(x)$ is

$$\mathfrak{F}(Lf)(\theta) = -\psi(-\theta)\mathfrak{F}f(\theta),$$

where $\psi(\theta) = t^{-1}\log E(e^{i\theta X_t})$. Hence, the equivalent estimation of Chen et al. (2010) can be determined based on empirical characteristic function by a weighted least-squares method. The regression method in this article does not require using the logarithm.

3.2. Estimation of α and c

Two important parameters in a Lévy process are α and c. In financial time series, α is used to explain the long-memory property of financial data. When $\alpha = 2$, the model is a Brownian motion which implies the efficient market (see Fama, 1970). The activity exponent is another name for α which describes the small jump.

There are many methods of estimating α. For example, Hill (1975) provided a way to estimate it; Csörgő et al. (1985) and Groeneboom et al. (2003) used the kernel method to estimate α. Because the model in Aït-Sahalia and Jacod (2009) is moderately close to our model (2.3), it is here suggested that the method proposed by Aït-Sahalia and Jacod (2009) be used to estimate α. They define this method as an instantaneous jump activity index. Using high-frequency data, they proposed the following method of estimating α:

$$\hat{\alpha}_n(t, \varpi, \epsilon) = \frac{\log(U(\varpi,\epsilon)_t^n / U_2(\varpi,\epsilon)_t^n)}{\varpi \log 2},$$

where

$$U(\varpi,\epsilon)_t^n = \sum_{i=1}^{t/\Delta_n} I_{\{|Y_i|>\epsilon\Delta_n^{\varpi}\}}$$

and $U_2(\varpi,\epsilon)_t^n$ is defined analogously to $U(\varpi,\epsilon)_t^n$, except that sampling at Δ_n is replaced by sampling at $2\Delta_n$. In this article, α is assumed to be known.

The estimator of c can be determined using the small-time ergodic property of Lévy processes:

$$\lim_{t\to 0} \frac{1}{t} E|X_t|^k I_{\{|X_t|<\epsilon\}} = \int_{|x|<\epsilon} |x|^k \rho(dx) \qquad \text{for } k \geq 3. \tag{3.7}$$

This can also be shown using (2.2)

$$\rho(dx)I_{\{|x|<\epsilon\}} = c|x|^{-\alpha-1},$$

This produces the estimator of c:

$$c^* = \frac{k-\alpha}{2\epsilon^{k-\alpha}n\Delta}\sum_{i=1}^{n}|Y_i|^3 I_{\{|Y_i|<\epsilon\}}.$$

In practice, c^* is not a good estimator of c because it depends heavily on Δ, and has a steady bias. In case of the bias, the following property is considered:

$$\lim_{t\to 0}\frac{1}{t}EX_t^2 I_{\{|X_t|<\epsilon\}} = \sigma^2 + \int_{\{|x|<\epsilon\}} x^2\rho(dx).$$

Hence, the estimator of c can be given as follows:

$$\hat{c} = \frac{2-\alpha}{(2-2^{\alpha-1})\epsilon^{2-\alpha}n\Delta}\sum_{i=1}^{n}(|Y_{2i}|^2 I_{\{|Y_{2i}|<\epsilon\}} - |Y_{2i-1}|^2 I_{\{|Y_{2i-1}|<\frac{\epsilon}{2}\}}). \tag{3.8}$$

Here the sample size is assumed to be $2n$.

3.3. Estimation of Lévy measure

For every fixed $a > 0$, the measure $\epsilon^{-1}\mathbb{P}_0(X_\epsilon \in dx)$ converges vaguely to $\rho(dx)$ on $\{|x| > a\}$ as $\epsilon \to 0+$ (Bertoin, 1996). Using high-frequency data and letting $n\Delta \to \infty$, facilitates approximation of $\mathbb{P}_0(X_\epsilon \in dx)$ using the empirical distribution function $\frac{1}{n}\sum_{i=1}^{n} I_{\{|Y_i|>a\}}$ and finding the estimator of Lévy tail measure:

$$\hat{\rho}[a,\infty) = \Delta^{-1}\frac{1}{n}\sum_{i=1}^{n} I_{\{|Y_i|>a\}}.$$

However, this estimator is not very good in practice as indicated by our simulations. Hence, we suggest the method of deconvolution used in Chen et al. (2010) for the estimation of Lévy measure ρ.

4. Properties of estimators

This section focuses on the properties of these proposed estimators, and $\epsilon = 1$ in model (2.3). In practice, each regression needs to be implemented over a set Γ which usually depends on $f(x)$. For $f \in \mathfrak{C}_0^2(\mathfrak{b})$, we choose $f(x)$ such that $f(x) > 0, f'(x)$ is an odd (or even) function, and $f''(x)$ is an even (or odd) function, which can reduce the error when Γ is a symmetric set. Let the following be true:

$$\delta_1 = \int_\Gamma |f'(x)|dx \int_{|y|<1}[f(x+y)-f(x)-yf'(x)]|y|^{-1-\alpha}dy \Big/ \int_\Gamma [f'(x)]^2 dx,$$

$$\delta_2 = \int_\Gamma |f'(x)|dx \int_{|y|\geq 1}[f(x+y)-f(x)]\rho(dy) \Big/ \int_\Gamma [f'(x)]^2 dx,$$

and

$$\delta_1' = \int_\Gamma |f''(x)|dx \int_{|y|<1}[f(x+y)-f(x)-yf'(x)]|y|^{-1-\alpha}dy \Big/ \int_\Gamma [f''(x)]^2 dx,$$

$$\delta_2' = \int_\Gamma |f''(x)|dx \int_{|y|\geq 1}[f(x+y)-f(x)]\rho(dy) \Big/ \int_\Gamma [f''(x)]^2 dx.$$

For the estimators of b and σ^2, as indicated by the regression of the infinitesimal generator, the following properties can be stated as a theorem.

Theorem 4.1. *If Δ tends to 0, $n\Delta$ tends to infinity. $f \in \mathfrak{C}_0^2(\mathfrak{b})$ and Γ is a symmetric subset of $\mathbb{R}$. Then the following is true:*

$$\hat{b} - b = O_p\left\{\delta_1 + \delta_2 + (\Delta n)^{-\frac{1}{2}} \frac{\int_\Gamma |f'(x)|dx}{\int_\Gamma [f'(x)]^2 dx}\right\}, \tag{4.1}$$

$$\hat{\sigma^2} - \sigma^2 = O_p\left\{\delta_1' + \delta_2' + (\Delta n)^{-\frac{1}{2}} \frac{\int_\Gamma |f''(x)|dx}{\int_\Gamma [f''(x)]^2 dx}\right\}.$$

For $f(x)$, the following two types are considered:

Type A. $f''(x)/f'(x) \to 0$ as $x \to \infty$, $f'(x)$ and $f''(x)$ are integrable. For example, $f(x) = \frac{1}{1+x^2}$.

Type B. $f'(x)/f''(x) \to 0$ as $x \to \infty$, $f'(x)$ and $f''(x)$ are integrable. For example, $f(x) = e^{-x^2/2}$.

In general, $f(x)$ was chosen in Type A to estimate b and $f(x)$ in Type B for σ^2. If something is known about the Lévy measure ρ, the estimators can be improved upon. The implications of Theorem 4.1 are discussed next.

Corollary 4.1. *If $\rho(|x| \geq 1) = 0$ in model (2.3), let $f(x) = \frac{1}{1+x^2}$ and $\Gamma = [x, +\infty) \cup (-\infty, -x]$. Then the following is true:*

$$\hat{b} - b = O_p\{x^{-1} + (\Delta n)^{-\frac{1}{2}}[f'(x)]^{-1}\}.$$

Corollary 4.2. *If $c = 0$ and $\rho(-\infty, -1) = 0$ in model (2.2), let $f(x) = e^{-x^2/2}$ and $\Gamma = [x, +\infty)$. Then*

$$\hat{\sigma^2} - \sigma^2 = O_p\{x^{-2} + (\Delta n)^{-\frac{1}{2}}[f''(x)]^{-1}\}.$$

In Corollary 4.1, the process X has only small jumps; and X in Corollary 4.2 is a compound Poisson process with positive jumps. For stable processes, the following theorem is used.

Theorem 4.2. *If X is an α-stable process, let $f(x) = \frac{\sin x}{1+x^2}$ and $\Gamma = [2k\pi, k \in \mathbb{Z}^+]$, where k is far enough from zero. Then for $x \in \Gamma$,*

$$\hat{b} - b = O_p\{x^{-(\alpha\wedge 1)} + (\Delta n)^{-\frac{1}{2}}[\min_{x\in\Gamma} f'(x)]^{-1}\}.$$

The last theorem concerns the property of estimator of c using the small-time ergodic property of the Lévy process.

Theorem 4.3. *If Δ tends to 0, $n\Delta$ tends to infinity, and $n\Delta^3$ tends to 0, the sample Y_i, $i = 1, 2, \cdots 2n$ are defined as above. Let the following be true:*

$$Z_t^n = \sum_{i=1}^{[nt]} \left(n\Delta)^{-1/2} \left[\frac{2-\alpha}{(2-2^{\alpha-1})\epsilon^{2-\alpha}} (|Y_{2i}|^2 I_{\{|Y_{2i}|<\epsilon\}} - |Y_{2i-1}|^2 I_{\{|Y_{2i-1}|<\frac{\epsilon}{2}\}} \right) - c \right].$$

Then the process $\{Z_t^n, t \in R^+\}$ converges in distribution to a Wiener process with variance $t\sigma_3^2$, where $\sigma_3^2 = \frac{(2-\alpha)^2}{(2-2^{\alpha-1})^2\epsilon^{4-2\alpha}}[\int_{\{|x|<\epsilon\}} x^4\rho(dx) + \int_{\{|x|<\epsilon/2\}} x^4\rho(dx)]$.

5. A simulation study

To demonstrate the estimation methods proposed in the previous sections, the following simulations were performed.

The first simulated example involved estimating the drift coefficient b based on (3.5). The data generation process involves drift + Brownian motion + compound Poisson process,

$$X_t = bt + \sigma W_t + \sum_{i=1}^{N_t} Z_i,$$

where $b = 3, \sigma = 1$, and $N = \{N_t, t \in R^+\}$ are a Poisson process with constant intensity $\lambda = 0.5$. $\{Z_i, i = 1, 2 \ldots\}$ is a sequence of *i.i.d.* normal random variables with mean = 4 and standard deviation = 10. The function $f(x)$ was selected:

$$f(x) = \frac{sin(x)}{1 + x^2}.$$

A sequence $\{x_i\}_{i=1}^{T}$ with $T = 50$ chosen to be $\{-3 + \frac{i}{25}, i = 1, \ldots, 25\} \bigcup \{2 + \frac{i}{25}, i = 1, \ldots, 25\}$. For each x_i, $n = 500$ observations of Lévy process $\{X_t\}$ were generated by setting $\triangle t = 5/3{,}600$. Let $Y_t = X_t - X_{t-1}$. Then the left side of (3.1) could be estimated as follows:

$$(\hat{Lf})(x_i) \approx \frac{1}{\triangle t}\left(\frac{1}{n-1}\sum_{t=2}^{n} f(Y_t + x_i) - f(x_i)\right).$$

This yields 50 pairs of the values $\{f'(x_i), (\hat{Lf})(x_i)\}$, where $f'(x_i)$ is the derivative of f at x_i. By running a simple linear regression without intercept for $(\hat{Lf})(x_i)$ on $f'(x_i)$ the estimator of the drift coefficient b is found.

In order to confirm performance, the simulation was run with two different sample sizes, $n = 500$ and 1,000. Different $\triangle t$ values were also used: 10/3,600, 20/3,600, 40/3,600 and 60/3,600, respectively. The results are shown in Fig. 1, where each of the two groups are for $n = 500$ and 1,000, respectively, under different $\triangle t$ from small to large.

The results given above show that the estimated medians of the groups are all quite close to the true parameter $b = 3$. Under fixed $\triangle t$, the variance decreases as the sample size increases. The performance seems to be very good for all the chosen values of $\triangle t$.

Next, simulation was performed to estimate b using a Cauchy process instead of compound Poisson process as in the data generation process:

$$X_t = bt + \sigma W_t + C_t(A, B).$$

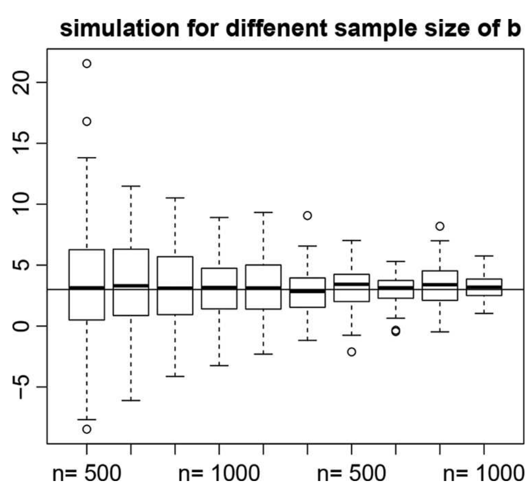

Figure 1. Regression for b with compound poison process, $\triangle t = 5/3{,}600$, 10/3,600, 20/3,600, 40/3,600, and 60/3,600 from left to right. With each $\triangle t$, the results of sample size $n = 500$ and $n = 1{,}000$ are shown.

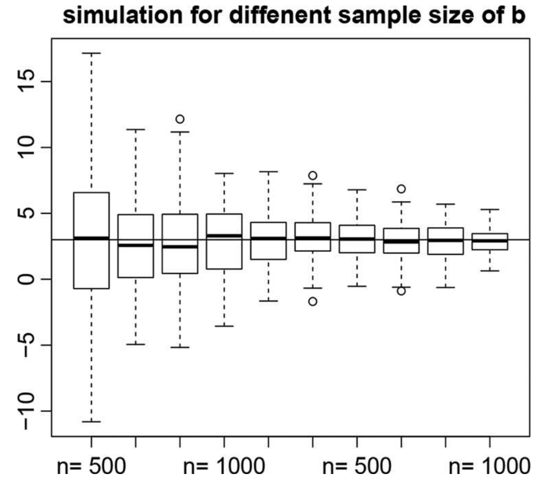

Figure 2. Regression for b with Cauchy process, $\triangle t = 5/3{,}600$, 10/3,600, 20/3,600, 40/3,600 and 60/3,600 from left to right. With each $\triangle t$ the results of sample size $n = 500$ and $n = 1{,}000$ are shown.

Here $b = 3$, $\sigma = 4$, and $C_t(A, B)$ are a Cauchy process with the location parameter $A = 0$ and scale parameter $B = \triangle t$. The same sample size and $\triangle t$ as above were used, and the results in Fig. 2 were produced.

As shown in Fig. 2, the medians of the groups are all close to the true parameter $b = 3$. And under fixed $\triangle t$, the variance is decreasing as the sample size increases. The performance seems to be very good for all values of $\triangle t$.

It is very similar to estimating σ^2 by the regression setup as in (3.6). With the same models and same choices of sample size and $\triangle t$, we have the results shown in Fig. 3.

Results show that the bias for the estimation of σ^2 is better for the Lévy process with Poisson jump than with the Cauchy jump. Again the bias increases when $\triangle t$ increases. For the Lévy process with Cauchy jump, the bias is still obvious, even when the $\triangle t$ is small. This might be because of the greater amount of noise produced by small $\triangle t$ in the random simulation.

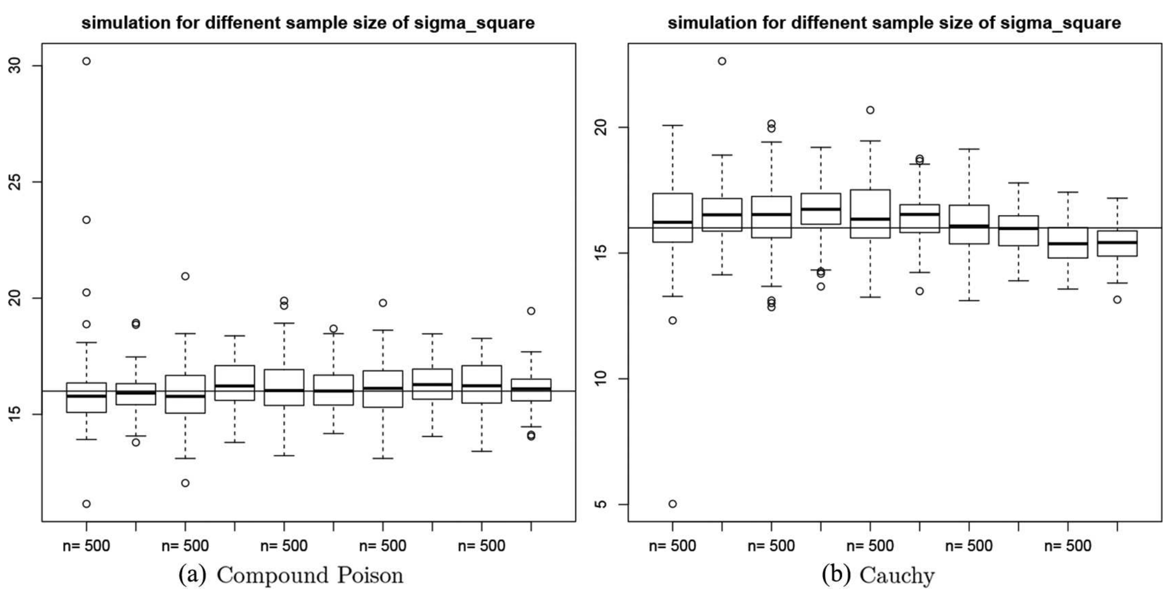

Figure 3. Regression for σ^2, $\triangle t = 5/3{,}600$, 10/3,600, 20/3,600, 40/3,600, and 60/3,600 from left to right. With each $\triangle t$ the results of sample size $n = 500$ and $n = 1{,}000$ are shown.

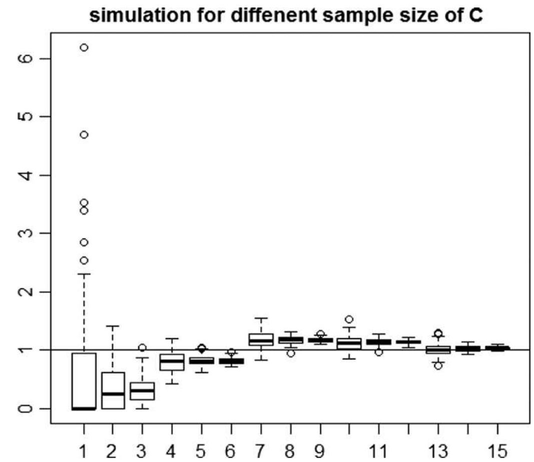

Figure 4. Estimation of C, $\triangle t = 5/3{,}600$, 200/3,600, 500/3,600, 600/3,600, and 8,000/3,600 from left to right. With each $\triangle t$ the results of sample size $n = 200$, $n = 1{,}000$ and $n = 2{,}000$ are shown.

The final simulation involved estimating the parameter c in a Lévy process with the Cauchy process serving as the jump part. The procedure is based on (3.8). Here, the Cauchy process in the data generating process also involved location 0 and scale $\triangle t$ and the parameter c was theoretically 1. The $\triangle t$ was set to 5/3,600, 200/3,600, 500/3,600, 600/3,600, and 800/3,600 with the total number of observations N to be 200, 1,000, and 2,000, respectively.

As shown in Fig. 4, groups 1, 2, and 3 were associated with $n = 200$, $n = 1,000$ and $n = 2,000$, respectively, when $\triangle t = 5/3,600$. Groups 4, 5 and 6 were for those values of n under $\triangle t = 200/3,600$. Groups 7, 8 and 9 were under $\triangle t = 500/3,600$. Groups 10, 11 and 12 were under $\triangle t = 600/3,600$, and groups 13, 14 and 15 were under $\triangle t = 800/3,600$. As shown, the results improved when $\triangle t$ increases.

In simulations involving estimating b and σ^2, x_i must be selected carefully. As shown in the proof, values of x_i can be determined from intervals symmetrical to zero.

Results also showed the estimations to be sensitive to the choice of $\triangle t$, especially for c and σ^2, even when the sample size n was already large. In practical applications, relatively large $\triangle t$, such as 5-minute data, should be used for the estimation of C. The identification of the optimal $\triangle t$ to estimate the parameters is out of the scope for this article. This should be addressed in further research.

6. A real data example

The current method was used on daily tick-by-tick data collected between May 6 and May 8, 2009, for the Shanghai Stock Exchange (SSE) Constituent Index and the Financial Index, which were constructed in June 2002 by the Shanghai Stock Exchange to promote the long-term development of infrastructure and the standardization of the security market.

A 30-second interval served as data frequency, and total sample size was set to $n = 1,446$. Figure 5 shows the plots of original price and log return.

The parameter α of the asset price data is usually between 1.6 and 2. Here the regression procedure was run with $\alpha = 1.8$. The estimation method was applied to the function

$$f(x) = \frac{sin(x)}{1 + x^2},$$

This produced the estimated values $\hat{b} = 1.90 \times 10^{-4}$, and $\hat{\sigma}^2 = 3.96 \times 10^{-5}$. By taking $\epsilon = 0.004$ in (3.8), the estimator of parameter c was found to be $\hat{c} = 2.078793 \times 10^{-5}$.

A different $\triangle t$ was used to estimate c because, in the current simulation, the estimation of c is better when $\triangle t$ is relatively large. Figure 6 shows the two plots of the Y_t, which is used to estimate c. The left

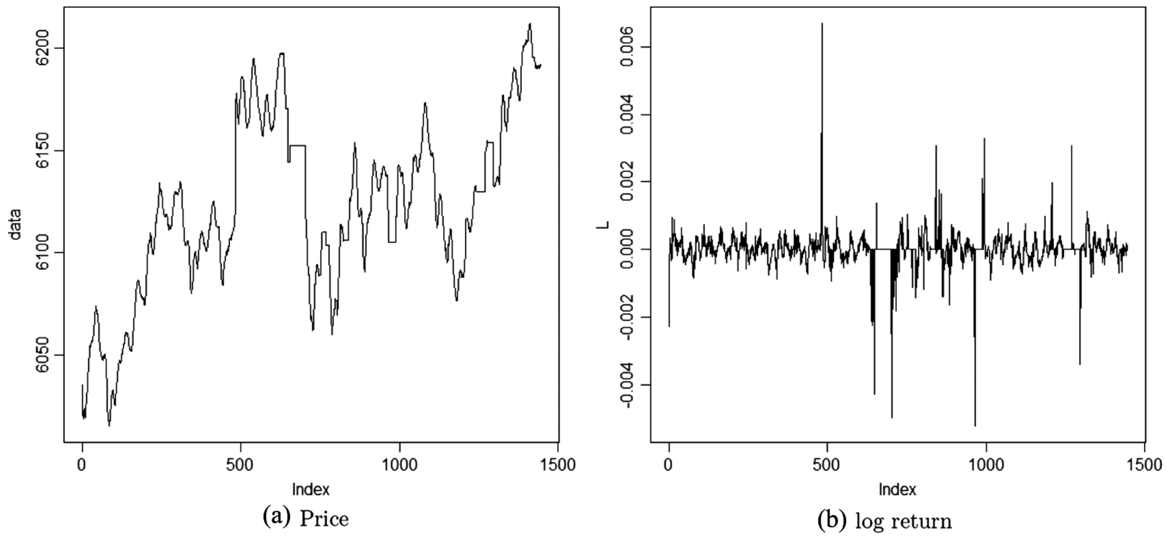

Figure 5. Thirty-second high frequency data from May 6 to May 8, 2009.

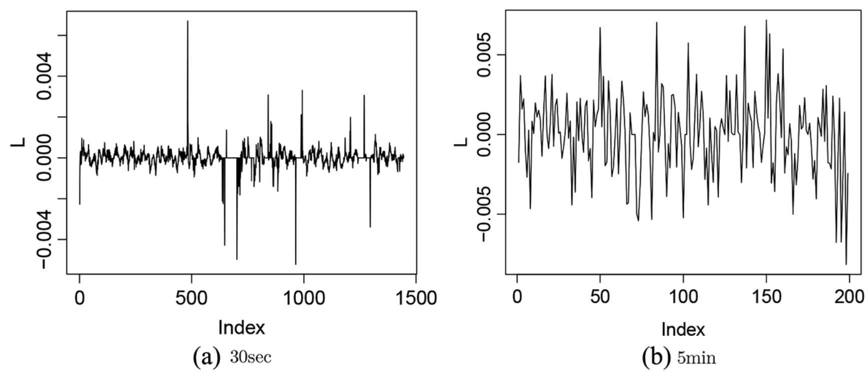

Figure 6. The Y_t data for estimating C.

panel is at $\triangle t = 30$ seconds, and the right one is at $\triangle t = 5$ minutes. The estimator $\hat{c}$ of the 5-minute data was 1.005×10^{-4}. As expected, the lower frequency estimates were much more variable, since they relied on a smaller number of large increments in total.

A. Appendix: Proofs

A lemma reported by Figueroa-López and by Jacod is used here (Figueroa-López, 2008; Jacod, 2007, p. 185). It concerns the small-time ergodic property of Lévy process, and will be used in the proofs of Theorem 4.1 and Theorem 4.3.

Lemma A.1. *Let $\{X_t, t \in R^+\}$ be a Lévy process, and with Lévy triplet (b, σ^2, ρ). Then*

$$\lim_{t \to 0} \frac{1}{t} E X_t^2 I_{\{|X_t| < \epsilon\}} = \sigma^2 + \int_{\{|x| < \epsilon\}} x^2 \rho(dx), \tag{A.1}$$

and

$$\lim_{t \to 0} \frac{1}{t} E X_t^k I_{\{|X_t| < \epsilon\}} = \int_{\{|x| < \epsilon\}} x^k \rho(dx) \qquad \text{for } k > 2. \tag{A.2}$$

The following lemma comes from Figueroa-López (2008).

Lemma A.2. *Let $\{X_t, t \in R^+\}$ be a Lévy process, and with Lévy triplet (b, σ^2, ρ), let f be twice continuous differentiable such that the following statements hold:*
(i) *f vanishes in a neighborhood of the origin;*
(ii) *$|f^{(i)}|$ is bounded for $i = 0, 1, 2$.*
Then

$$\lim_{t\to 0} \frac{1}{t}\left\{\frac{1}{t}Ef(X_t) - \int f(x)\rho(dx)\right\} = C(f, \rho). \tag{A.3}$$

Now, let $g(x) = x^2 I_{\{\epsilon/2<x<\epsilon\}}$. We can find f_h and f_h' with $f_h \le g \le f_h'$, which meet the conditions of Lemma A.2, and

$$\lim_{h\to 0} f_h = \lim_{h\to 0} f_h' = g.$$

By the dominated convergence theorem and (A.3), for $g(x) = x^2 I_{\{\epsilon/2<x<\epsilon\}}$, we have

$$\lim_{t\to 0} \frac{1}{t}\left\{\frac{1}{t}Eg(X_t) - \int_{\epsilon/2}^{\epsilon} x^2\rho(dx)\right\} = C(g, \rho). \tag{A.4}$$

A.1. Proof of Theorem 4.3

Proof. Consider the process $\{Z_t^n, t \in R^+\}$, where

$$Z_t = \sum_{i=1}^{[nt]} (n\Delta)^{-1/2}\left[\frac{2-\alpha}{(2-2^{\alpha-1})\epsilon^{2-\alpha}}(|Y_{2i}|^2 I_{\{|Y_{2i}|<\epsilon/2\}} - |Y_{2i-1}|^2 I_{\{|Y_{2i-1}|<\epsilon\}}) - c\right].$$

Z_t^n is the sum of $[nt]$ *i.i.d.* variables. We have $|(n\Delta)^{-1/2}[\frac{2-\alpha}{(2-2^{\alpha-1})\epsilon^{2-\alpha}}(|Y_{2i}|^2 I_{\{|Y_{2i}|<\epsilon/2\}} - |Y_{2i-1}|^2 I_{\{|Y_{2i-1}|<\epsilon\}}) - c]| \le C(n\Delta)^{-1/2} \to 0$ for all $i = 1, 2, \cdots [nt]$, which implies the uniformly tightness of Z_t. According to Theorem 14.1 of Billingsley (1968), what we need to prove looks like a Central Limit Theorem (CLT), but this is a bit difficult since the centering is not explicit.

Suppose $\zeta_i = [\frac{2-\alpha}{(2-2^{\alpha-1})\epsilon^{2-\alpha}}(|Y_{2i}|^2 I_{\{|Y_{2i}|<\epsilon/2\}} - |Y_{2i-1}|^2 I_{\{|Y_{2i-1}|<\epsilon\}})]$, By a standard CLT,

$$(n\Delta)^{-1/2}\sum_{i=1}^{[nt]}(\zeta_i - E\zeta_i) \sim N(0, \sigma_3^2), \tag{A.5}$$

where $\sigma_3^2 = \frac{(2-\alpha)^2}{(2-2^{\alpha-1})^2\epsilon^{4-2\alpha}}[\int_{\{|x|<\epsilon\}} x^4\rho(dx) + \int_{\{|x|<\epsilon/2\}} x^4\rho(dx)]$.

Notice (A.4), and by the condition $n\Delta^3 \to 0$, we can get

$$(n\Delta)^{-1/2} nt(E\zeta_i - \Delta c) = (n\Delta)^{1/2}\left(\frac{1}{\Delta}E\zeta_i - c\right)t = Cn^{1/2}\Delta^{3/2} \to 0. \tag{A.6}$$

Hence, $\{Z_t^n, t \in R^+\}$ converges in law to a Wiener process with variance σ_3^2. □

A.2. Proof of Theorem 4.1

Before the proof of Theorem 4.1, we first show the consistency of $(\hat{L}f)(x)$ with $(Lf)(x)$.

Theorem A.1. *Let $X = \{X_t, t \in R^+\}$ be a Lévy process, and with Lévy triplet (b, σ^2, ρ), $f \in \mathfrak{C}_0^\infty(K)$, L is the infinitesimal generator of X. Then*

$$(Lf)(x) - \frac{1}{t}[Ef(X_t + x) - f(x)] = O(t) \text{ as } t \to 0.$$

Proof. By a simple calculation (or see Bertoin, 1996, p. 24), we can get the Fourier transform of $Ef(X_t+x)$,

$$\mathfrak{F}Ef(X_t+x)(\theta)=\exp\{-t\psi(-\theta)\}\mathfrak{F}f(\theta),$$

and

$$\mathfrak{F}(Lf)(x)(\theta)=-\psi(-\theta)\}\mathfrak{F}f(\theta),$$

where $\psi(\theta)=t^{-1}\log E(e^{i\theta X_t})$. Using the Fourier inversion,

$$\begin{aligned}Ef(X_t+x)-f(x)&=\frac{1}{2\pi}\int_R e^{-i\theta x}\exp\{-t\psi(-\theta)\}\mathfrak{F}f(\theta)d\theta-f(x)\\&=\frac{1}{2\pi}\int_R e^{-i\theta x}[\exp\{-t\psi(-\theta)\}-1]\mathfrak{F}f(\theta)d\theta\\&=\frac{1}{2\pi}\int_R e^{-i\theta x}[-t\psi(-\theta)\}+t^2\psi^2(-\theta)+o(t^2)]\mathfrak{F}f(\theta)d\theta\\&=t(Lf)(x)+Ct^2+o(t^2).\end{aligned}$$

□

Now, we begin the proof of Theorem 4.1, we just prove (i), and the proof of (ii) is similar.

Proof. Let

$$(\hat{L}f)(x)=bf'(x)+\epsilon_1(x),$$

where

$$\epsilon_1(x)=\frac{\sigma^2}{2}f''(x)+\int_R[f(x+y)-f(x)-f'(x)yI_{\{|y|<1\}}]\rho(dx)+\epsilon(x), \tag{A.7}$$

and $\epsilon(x)$ is defined in (3.3). We write $\epsilon(x)$ as

$$\epsilon(x)=(\hat{L}f)(x)-\frac{1}{\Delta}[Ef(X_t+x)-f(x)]+\frac{1}{\Delta}[Ef(X_t+x)-f(x)]-(Lf)(x),$$

and $(\hat{L}f)(x)$ is defined as (3.6). By the definition of $\hat{b}$ and (A.7),

$$\hat{b}^*-b=\frac{\int_\Gamma f'(x)\epsilon_1(x)dx}{\int_\Gamma[f'(x)]^2dx},$$

and

$$\begin{aligned}\epsilon_1(x)=&\frac{\sigma^2}{2}f''(x)+\int_R[f(x+y)-f(x)-f'(x)yI_{\{|y|<1\}}]\rho(dx)\\&+\frac{1}{\Delta}\left\{\frac{1}{n}\sum_{i=1}^{n}[f(Y_i+x)-Ef(Y_i+x)]\right\}+\epsilon(\Delta).\end{aligned} \tag{A.8}$$

Hence

$$\int_\Gamma f'(x)\epsilon_1(t)dx=\int_\Gamma\frac{\sigma^2}{2}f'(x)f''(x)dx+\int_\Gamma\int_R[f(x+y)-f(x)-f'(x)yI_{\{|y|<1\}}]\rho(dy)f'(x)dx+R,$$

where

$$R=\frac{1}{\Delta}\int_\Gamma\left\{\frac{1}{n}\sum_{i=1}^{n}[f(Y_i+x)-Ef(Y_i+x)]\right\}f'(x)dx+\epsilon(\Delta).$$

Using the definition of the infinitesimal generator and the property of $f(x)$, we can get that $Var f(Y_i + x) = O(\Delta)$. Similar to the proof of Theorem 4.2, since $f(Y_i)$, $i = 1, 2, \cdots n$ are *i.i.d.*, we have

$$\frac{1}{n\Delta}\sum_{i=1}^{n}[f(Y_i+x)-Ef(Y_i+x)] = O[(n\Delta)^{-\frac{1}{2}}]. \tag{A.9}$$

Hence

$$R\left[\int_\Gamma [f'(x)]^2 dx\right]^{-1} \le \frac{(n\Delta)^{-1/2}\int_\Gamma |f'(x)|dx}{\int_\Gamma [f'(x)]^2 dx} = O\left\{(n\Delta)^{-1/2}\frac{\int_\Gamma |f'(x)|dx}{\int_\Gamma [f'(x)]^2 dx}\right\}. \tag{A.10}$$

Remeber $f'(x)$ and $f''(x)$ are bounded, and $f'(x)f''(x)$ is an odd function,

$$\int_\Gamma \frac{\sigma^2}{2} f'(x) f''(x) dx = 0 \tag{A.11}$$

and

$$\begin{aligned}
&\left|\int_\Gamma \int_R [f(x+y) - f(x) - f'(x) y I_{\{|y|<1\}}]\rho(dy) f'(x) dx\right| \\
&\le \int_\Gamma \int_{|y|<1} |[f(x+y) - f(x) - f'(x)y]||y|^{-1-\alpha}|f'(x)|dx \\
&+ \int_\Gamma \int_{|y|\ge 1} [|f(x+y) - f(x)|]\rho(dy)|f'(x)|dx.
\end{aligned}$$

This implies (4.1).

A.3. Proof of Theorem 4.2

Proof. $f(x) = \frac{\sin x}{1+x^2}$, so $f'(x) = \frac{\cos x}{1+x^2} - \frac{2x\sin x}{(1+x^2)^2}$, and $f''(x) = \frac{-\sin x}{1+x^2} - \frac{4x\cos x + 2\sin x}{(1+x^2)^2} + \frac{8x^2 \sin x}{(1+x^2)^3}$. For $x \in [2k\pi, k \in \mathbb{Z}]$, we have

$$|f''(x)| \le Cx^{-3} \qquad \text{as } k \to \infty; \tag{A.12}$$

because X is an α-stable process, the Lévy measure $\rho(dx) = c|x|^{-1-\alpha}dx$,

$$\begin{aligned}
&\int_{\{|y|<1\}} [f(x+y) - f(x) - f'(x)y]|y|^{-1-\alpha} dy \\
&= \int_{\{|y|<1\}} \left[\frac{\sin y}{1+(x+y)^2} - \frac{y}{1+x^2}\right]|y|^{-1-\alpha} dy \\
&= \int_0^1 \left[\frac{\sin y}{1+(x+y)^2} - \frac{\sin y}{1+(x-y)^2}\right] y^{-1-\alpha} dy \\
&= \int_0^1 \left[\frac{4xy\sin y}{[1+(x+y)^2][1+(x-y)^2]}\right] y^{-1-\alpha} dy \\
&= O(x^{-3}),
\end{aligned} \tag{A.13}$$

and

$$\begin{aligned}
&\int_{\{|y|\ge 1\}} [f(x+y) - f(x)]|y|^{-1-\alpha} dy \\
&= \int_{\{|y|\ge 1\}} \left[\frac{\sin y}{1+(x+y)^2}\right]|y|^{-1-\alpha} dy \\
&= \int_1^\infty \left[\frac{4xy\sin y}{[1+(x+y)^2][1+(x-y)^2]}\right] y^{-1-\alpha} dy \\
&= O(x^{-[3\wedge(2+\alpha)]}),
\end{aligned} \tag{A.14}$$

together with (A.8), (A.10) we finish the proof.

Funding

Supported by National Social Science Fund of China (11CTJ009).

References

Aït-Sahalia, Y., Jacod, J. (2009). Estimating the degree of activity of jumps in high frequency data. *Annals of Statistics* 37:2202–2244.

Aït-Sahalia, Y., Jacod, J. (2007). Volatility Estimator for discretely sampled Lévy processes. *Annals of Statistics* 35:355–392.

Bertoin, J. (1996). *Lévy Processes*. Cambridge: Cambridge University Press.

Billingsley, P. (1968). *Convergence of Probability Measures*. New York: Wiley.

Chen, S. X., Delaigle, A., Hall, P. (2010). Nonparametric estimation for a class of Lévy processes. *Journal of Econometrics* 157:257–271.

Comte, F., Genon-Catalot, V. (2009). Nonparametric estimation for pure jump Lévy processes based on high frequency data. *Stochastic Processes & Their Applications* 119(12):4088–4123.

Comte, F., Genon-Catalot, V. (2011). Estimation for Lévy processes from high frequency data within a long time interval. *Annals of Statistics* 39(2):803–837; 8:1–67.

Cont, R., Tankov, P. (2003). Financial Modelling With Jump Processes. In: Chapman and Hall/CRC Financial Mathematics Series.

Csörgő, S., Deheuvels, P., Mason, D. (1985). Kernel estimates of the tail index of a distribution. *Annals of Statistics* 13: 1050–1077.

Fama, E., 1970. Efficient capital markets: A review of theory and empirical work. *Journal of Finance* 25:383–417.

Figueroa-López, J. E. (2008). Small-time moment asymptotics for Lévy processes. *Statistics and Probability Letters* 78: 3355–3365.

Florian A. J., Ueltzhöfer, F. A. J. (2013). On non-parametric estimation of the Lévy kernel of Markov processes. *Stochastic Processes and their Applications* 123:3663–3709.

Groeneboom, P., Lopuhaa, H. P., De Wolf, P. P. (2003). Kernel-type estimators for the extreme value index. *Annals of Statistics* 31:1956–1995.

Jacod, J. (2007). Asymptotic properties of power variations of Lévy processes. *ESAIM: Probability and Statistics* 11:173–196.

Hill, B. M. (1975). A simple general approach to inference about the tail of a distribution. *Annals of Statistics* 3:1163–1174.

Revuz, D., Yor, M. (2005). *Continuous Martingales and Brownian Motion*. Springer.

Sato, K.-I. (1999). *Lévy Processes and Infinitely Divisible Distributions*. Cambridge: Cambridge University Press.

Schmisser, E. (2014). Non-parametric adaptive estimation of the drift for a jump diffusion process. *Stochastic Processes and Their Applications* 124:883–914.