Nonparametric Estimation of Second-Order Jump-Diffusion Model[*]

Lin Zheng-Yan, Song Yu-Ping,[†] Wang Han-Chao

Department of Mathematics, Zhejiang University,

Hangzhou, China, 310027

June 29, 2011

Abstract: We study the nonparametric estimators of the infinitesimal coefficients of the second-order jump-diffusion models. Under the mild conditions, we obtain the weak consistency and the asymptotic normalities of the estimators.

Keyword: Second-order jump-diffusion; N-W estimator; Weak consistency; Asymptotic normality.

Mathematics Subject Classification: 60J60; 62G15

1 Introduction

Nonstationary time series play an important role in economics and finance. Many statisticians have studied the asymptotic properties on nonstationary time series model such as the integrated time series. The integrated stochastic processes would explain some modern econometric phenomena better since the current observation usually behaves as the cumulation of all past perturbation. Especially, when testing unit roots, if null hypothesis holds, we have

$$Y_t = Y_0 + \sum_{k=1}^{t} X_k, \tag{1.1}$$

where $X_k = \alpha + \varepsilon_k$, $\{\varepsilon_k : k = 0, 1, 2, \cdots\}$ is a sequence of i.i.d random variables. Obviously, the sequence $\{Y_t; t = 0, 1, 2, \cdots\}$ is not stationary. One common method to deal with the nonstationary time series is to use a differenced model (e.g. $\Delta Y_t = X_t$). The difference transformation is practical in research. This model can be utilized in modern econometric analysis such as logarithmic stock price random walk and annual capital GDP. For further motivation and study on asymptotic properties, see Park and Phillips (2001).

Recently, some authors have studied the continuous case of model (1.1) since many financial models are continuous-time. Naturally, model (1.1) can be extended to

$$Y_t = Y_0 + \int_0^t X_s ds. \tag{1.2}$$

The differenced-data model of (1.2) is $dY_t = X_t dt$. Considering stock price index at some time, X_s represents stock price index; Y_t indicates the cumulation of stock price index. In model (1.2), X_s is usually assumed as a stationary continuous-time process such as diffusion process. The model accommodates nonstationary integrated stochastic processes that can be made stationary by differencing.

[†]Corresponding author, songyuping2011@gmail.com

Many popular models in financial time series involve continuous-time diffusion processes as solutions to stochastic differential equations below, such as option prices, interest rates, exchange rates and inter alia, see Baxter and Rennie (1996). A diffusion process X_t is represented by the following stochastic differential equation:

$$dX_t = \mu(X_t)dt + \sigma(X_t)dW_t, \tag{1.3}$$

where W_t is a standard Brownian motion, $\mu(\cdot)$ and $\sigma(\cdot)$ are the infinitesimal conditional drift and variation respectively. Based on the discret-time observations $\{X_{i\Delta_n}; i = 1, 2, \cdots\}$, some authors have studied the statistical inference for the coefficients, e.g., Bandi and Phillips (2003), Fan and Zhang (2003) and Xu (2009, 2010).

However, all sample functions of a diffusion process driven by a Brownian motion are of unbounded variation and nowhere differentiable. So it is difficult to model integrated and differentiated diffusion processes which are useful in empirical finance. According to (1.2) and (1.3), Nicolau (2007) discussed a second-order diffusion process satisfying the following second-order stochastic differential equation:

$$\begin{cases} dY_t = X_t dt, \\ dX_t = \mu(X_t)dt + \sigma(X_t)dW_t, \end{cases} \tag{1.4}$$

where W_t is a standard Brownian motion and X_t is (by hypothesis) a stationary process. We observe that the model (1.4) can be written as a second-order stochastic differential equation: $d(dY_t/dt) = \mu(X_t)dt + \sigma(X_t)dW_t$. The second-order diffusion process overcomes the difficulties associated with the nondifferentiability of a Brownian motion.

For model (1.4), the estimators for $\mu(x)$ and $\sigma^2(x)$ have been considered based on discrete-time observations. Nicolau (2007) considered the Nadaraya-Watson estimators for $\mu(x)$ and $\sigma^2(x)$; Wang and Lin (2011) presented local linear estimators for these unknown quantities, Wang; Zhang and Wang (2011) studied the empirical likelihood inference for them.

In fact, in many economic models the path of X_t is not continuous, so besides $\mu(x), \sigma^2(x)$ we need some coefficients to characterize the jump, for example, jump measure, intensity measure and conditional impact of a jump. Compared with estimators for model (1.4), we should estimate various coefficients such as $\mu(x), \sigma(x)$ and $c(x, z)$ for the following second-order jump-diffusion model based on discrete high-frequent observations, that is, considering the model

$$\begin{cases} dY_t = X_t dt, \\ dX_t = \mu(X_{t-})dt + \sigma(X_{t-})dW_t + \int_{\mathscr{E}} c(X_{t-}, z)r(\omega, dt, dz), \end{cases} \tag{1.5}$$

where $\mathscr{E} = \mathbb{R} \setminus \{0\}, W_t$ is a Wiener process, $r(\omega, dt, dz) = (p - q)(dt, dz),\ p(dt, dz)$ is a time-homogeneous Poisson random measure on $\mathbb{R}_+ \times \mathbb{R}$ independent of W_t, and $q(dt, dz)$ is its intensity measure, that is, $E[p(dt, dz)] = q(dt, dz) = f(z)dzdt,\ f(z)$ is a Lévy density.

Recently, many literature have been involved in the research on the following popular jump-diffusion model:

$$dX_t = \mu(X_{t-})dt + \sigma(X_{t-})dW_t + \int_{\mathscr{E}} c(X_{t-}, z)r(\omega, dt, dz), \tag{1.6}$$

in the sense that this model can accommodate impact of sudden and large shocks to financial markets, such as macroeconomic announcements and a dramatic interest rate cut by the Federal Reserve, see Bakshi (1997), Duffie (2000), Eraker, Johannes and Polson (2003) Johannes (2004).

The statistical inferences for the coefficients of the model are based on the following relations:

$$\lim_{\Delta \to 0} E[\frac{X_{t+\Delta} - X_t}{\Delta}|X_t = x] = \mu(x), \tag{1.7}$$

$$\lim_{\Delta \to 0} E[\frac{(X_{t+\Delta} - X_t)^2}{\Delta}|X_t = x] = \sigma^2(x) + \int_{\mathscr{E}} c(x,z)f(z)dz. \tag{1.8}$$

Based on these infinitesimal moments of the jump-diffusion model and discrete-time observations $\{X_{i\Delta_n}; i = 1, 2, \cdots\}$, many literature have been involved in this field to estimate the coefficients from different aspects, see Bandi and Nguyen (2003), Lin and Wang (2010). Mancini (2004, 2009) with Renò (2010) introduced threshold estimation for the coefficients of the model (1.6). Parametric estimations of the coefficients for the model (1.6) have been presented by Shimizu (2006).

Similarly, we should consider model (1.5) to overcome the nondifferentiability of a Brownian motion and to model integrated and differentiated jump-diffusion processes for stock prices, compounded return or log return of an asset. Estimation of the coefficients in model (1.5) gives rise to new challenges for two main reasons. On the one hand, different from model (1.3) or model (1.6), although we usually get observations $\{Y_{i\Delta_n}; i = 1, 2, \cdots\}$, which are the cumulation of all the past perturbations, we cannot obtain the value of X at time t_i from $Y_{t_i} = Y_0 + \int_0^{t_i} X_s ds$ in a fixed sample intervals. In addition, estimation of the coefficients cannot in principle be based on the observations $\{Y_{i\Delta_n}; i = 1, 2, \cdots\}$ because conditional distribution of Y is generally unknown even if that of X is known. On the other hand, for model (1.5) whether we have the similar relations as (1.7) and (1.8) which are the basic ones to establish Nadaraya-Watson estimators? Fortunately these are settled satisfactorily.

We usually obtain discrete-time observations $\{Y_{i\Delta_n}; i = 1, 2, \cdots\}$ rather than $\{X_{i\Delta_n}; i = 1, 2, \cdots\}$ for model (1.5). As Nicolau (2007) showed, with discrete-time observations $\{Y_{i\Delta_n}; i = 1, 2, \cdots\}$ and given that

$$Y_{i\Delta_n} - Y_{(i-1)\Delta_n} = \int_0^{i\Delta_n} X_u du - \int_0^{(i-1)\Delta_n} X_u du = \int_{(i-1)\Delta_n}^{i\Delta_n} X_u du,$$

we can get an approximation value of X at instant $t_i = i\Delta_n$ by

$$\widetilde{X}_{i\Delta_n} = \frac{Y_{i\Delta_n} - Y_{(i-1)\Delta_n}}{\Delta_n}, \tag{1.9}$$

which is close to the value of X at $t = i\Delta_n$ or $t = (i-1)\Delta_n$ with Δ_n tends to zero. Our estimation procedures should be based on the samples $\{\widetilde{X}_{i\Delta_n}; i = 0, 1, 2, \cdots\}$.

We can also build the similar relations for model (1.5) as (1.7) and (1.8) seen in Remark 3.1, that is

$$E[\frac{\widetilde{X}_{(i+1)\Delta_n} - \widetilde{X}_{i\Delta_n}}{\Delta_n}|\mathscr{F}_{(i-1)\Delta_n}] = \mu(X_{(i-1)\Delta_n}) + O_p(\Delta_n), \tag{1.10}$$

$$E[\frac{(\widetilde{X}_{(i+1)\Delta_n} - \widetilde{X}_{i\Delta_n})^2}{\Delta_n}|\mathscr{F}_{(i-1)\Delta_n}] = \frac{2}{3}\sigma^2(X_{(i-1)\Delta_n}) + \frac{2}{3}\int_{\mathscr{E}} c^2(X_{(i-1)\Delta_n}, z)f(z)dz + O_p(\Delta_n). \tag{1.11}$$

where $\mathscr{F}_t = \sigma\{X_s, s \leq t\}$. Based on these relations and (1.9), we can establish Nadaraya-Watson estimators for $\mu(x)$ and $\sigma^2(x) + \int_{\mathscr{E}} c^2(x,z)f(z)dz$, and further, under appropriate conditions we verify the consistency and asymptotic normality for these proposed estimators.

The remainder of this paper is organized as follows. In Section 2, ordinary assumptions and N-W estimators for model (1.5) are introduced. In Section 3, we present some useful preliminary results. In this section, consistency and asymptotic normality of the estimators in our paper are given. The proofs will be collected in Section 4. We give the brief calculation for the main equations (3.3) and (3.4) in Appendix.

2 Nadaraya-Watson Estimators and Assumptions

We briefly discuss the Nadaraya-Watson estimators for the coefficients in model (1.5) based on $\{\widetilde{X}_{i\Delta_n}; i = 0, 1, 2, \cdots\}$. We firstly construct Nadaraya-Watson estimators for them based on equations (1.10) and (1.11). The Nadaraya-Watson estimators for $\mu(x)$ and $M(x) = \sigma^2(x) + \int_{\mathscr{E}} c(x, z)f(z)dz$ are solutions to the following optimal problem:

$$\arg\min_{\mu(x)} \Sigma_{i=1}^n \Big(\frac{\widetilde{X}_{(i+1)\Delta_n} - \widetilde{X}_{i\Delta n}}{\Delta_n} - \mu(x) \Big)^2 K\Big(\frac{x - \widetilde{X}_{(i-1)\Delta_n}}{h_n} \Big), \tag{2.1}$$

$$\arg\min_{M(x)} \Sigma_{i=1}^n \Big(\frac{\frac{3}{2}(\widetilde{X}_{(i+1)\Delta_n} - \widetilde{X}_{i\Delta_n})^2}{\Delta_n} - M(x) \Big)^2 K\Big(\frac{x - \widetilde{X}_{(i-1)\Delta_n}}{h_n} \Big). \tag{2.2}$$

The solutions to (2.1) and (2.2) are

$$\hat{a}_n(x) = \frac{A_n(x)}{\hat{p}_n(x)}, \qquad \hat{b}_n(x) = \frac{B_n(x)}{\hat{p}_n(x)} \tag{2.3}$$

where $A_n(x) = \frac{1}{nh_n} \sum_{i=1}^n K\Big(\frac{x-\widetilde{X}_{(i-1)\Delta_n}}{h_n} \Big) \frac{(\widetilde{X}_{(i+1)\Delta_n} - \widetilde{X}_{i\Delta_n})}{\Delta_n}, B_n(x) = \frac{1}{nh_n} \sum_{i=1}^n K\Big(\frac{x-\widetilde{X}_{(i-1)\Delta_n}}{h_n} \Big) \frac{\frac{3}{2}(\widetilde{X}_{(i+1)\Delta_n} - \widetilde{X}_{i\Delta_n})^2}{\Delta_n}$,
$\hat{p}_n(x) = \frac{1}{nh_n} \sum_{i=1}^n K\Big(\frac{x-\widetilde{X}_{(i-1)\Delta_n}}{h_n} \Big)$.

The paper is devoted to study the following estimators $\hat{p}_n(x), \hat{a}_n(x), \hat{b}_n(x)$ for $p(x)$, $\mu(x)$ and $\sigma^2(x) + \int_{\mathscr{E}} c^2(x, z)f(z)dz$ respectively, where $p(x)$ be the stationary probability measure for X_t. There are some differences between the estimators given in this paper for model (1.6) and the ones in Bandi and Nguyen (2003). Take the estimator for $\mu(x)$ as an example. In Bandi and Nguyen (2003), if the observations are based on $\{\widetilde{X}_{i\Delta_n}; i = 0, 1, 2, \cdots\}$, the Nadaraya-Watson estimator for $\mu(x)$ should be $\hat{a}'_n(x) = \frac{A'_n(x)}{\hat{p}_n(x)}$ where $A'_n(x) = \frac{1}{nh_n} \sum_{i=1}^n K\Big(\frac{x-\widetilde{X}_{(i-1)\Delta_n}}{h_n} \Big) \frac{(\widetilde{X}_{i\Delta_n} - \widetilde{X}_{(i-1)\Delta_n})}{\Delta_n}$. We easily see that $\frac{(\widetilde{X}_{(i+1)\Delta_n} - \widetilde{X}_{i\Delta_n})}{\Delta_n}$ and $\frac{(\widetilde{X}_{i\Delta_n} - \widetilde{X}_{(i-1)\Delta_n})}{\Delta_n}$ in $A_n(x)$ and $A'_n(x)$ are different. The model considered in this article is more complex than the one in Bandi and Nguyen (2003). The consistence and asymptotic normality of the estimators may be not obtained by the method exploited for stationary case in Bandi and Nguyen (2003), but we can get them by means of the method introduced by Nicolau (2007). This is due to the sense that we need to compute some meaningful conditional expect values of the estimator in the proof which may be not obtained if $A_n(x)$ is the same as $A'_n(x)$. In addition, there are some contacts between the estimators here and in Bandi and Nguyen (2003). $\widetilde{X}_{i\Delta_n}$ and $\widetilde{X}_{(i-1)\Delta_n}$ are close to the value of $X_{(i-1)\Delta_n}$, which intuitively to a certain extent guarantees the desired result in the article reasonably from the result in Bandi and Nguyen (2003) (rigorous proof theoretically seen in the following Lemmas and Theorems).

We now present some assumptions used in the paper. In what follows, let $\mathscr{D} = (l, u)$ with $l \geq -\infty$ and $u \leq \infty$ denote the admissible range of the process X_t.

Assumption 1.

(i) (Local Lipschitz continuity) For each $n \in \mathbb{N}$, there exist a constant L_n and a function $\zeta_n : \mathscr{E} \to \mathbb{R}_+$ with $\int_{\mathscr{E}} \zeta_n^2(z)f(z)dz < \infty$ such that, for any $|x| \leq n, |y| \leq n$,

$$|\mu(x) - \mu(y)| + |\sigma(x) - \sigma(y)| \leq L_n|x - y|, \quad |c(x, z) - c(y, z)| \leq \zeta_n(z)|x - y|.$$

(ii) (Linear growthness) For each $n \in \mathbb{N}$, there is ζ_n as above, such that for all $x \in \mathbb{R}$,

$$|c(x, z)| \leq \zeta_n(z)(1 + |x|).$$

Remark 2.1. Assumption 1 guarantees the existence and uniqueness of a solution to X_t in Eq.(1.5) on the probability space $(\Omega, \mathscr{F}, P)$, see Jacod and Shiryayev (1987).

Assumption 2. The process X_t is ergodic and stationary with a finite invariant measure $\phi(x)$.

Remark 2.2. The finite invariant measure implies that the process X_t is positive Harris recurrent with the stationary probability measure $p(x) = \frac{\phi(x)}{\phi(\mathscr{D})}$, $\forall x \in \mathscr{D}$. The hypothesis that X_t is a stationary process is obviously a plausible assumption because for major integrated time series data, a simple differentiation generally assures stationarity.

Assumption 3. The process X_t is $\rho-$mixing with $\sum_{i \geq 1} \rho(i\Delta_n) < \infty$.

Remark 2.3. We notice that the process $\{\widetilde{X}_{i\Delta n}; i = 1, 2, \cdots\}$ is stationary and $\rho-$mixing with the same mixing size. Similarly as Ditlevsen and Sørensen (2004) pointed out, by stationarity, the law of the process X_t is invariant under time translations, which easily implies that $\{\widetilde{X}_{i\Delta n}; i = 1, 2, \cdots\}$ is a stationary process. Since the σ-algebra generated by $\{\widetilde{X}_{i\Delta_n}, i = 1, \cdots, n\}$ is contained in the σ-algebra generated by $\{X_u, 0 \leq u \leq n\Delta_n\}$, and the σ-algebra generated by $\{\widetilde{X}_{i\Delta_n}, i = n+1, \cdots\}$ is contained in the σ-algebra generated by $\{X_u, u \geq n\Delta_n\}$, the process $\widetilde{X}_{i\Delta_n}$ is $\rho-$mixing with the the same mixing size. For instance, the $\rho-$mixing process X_t with exponentially decreasing mixing coefficients satisfies the condition, see Hansen and Scheinkman (1995), Chen, Hansen and Carrasco (2010).

Assumption 4. The kernel $K(\cdot) : \mathbb{R} \to \mathbb{R}^+$ is a positive, symmetric and continuously differentiable function satisfying $\int K(u)du = 1, \int uK(u) = 0$ and $K_2 := \int K^2(u)du < \infty$.

Assumption 5. For any $2 \leq i \leq n$, $\lim_{h \to 0} E\left[\frac{1}{h} |K'(\xi_{n,i})|^\alpha\right] < \infty$ where $\alpha = 2, 4$ or 8 and $\xi_{n,i} = \theta(\frac{x-X_{(i-1)\Delta_n}}{h}) + (1-\theta)(\frac{x-\widetilde{X}_{(i-1)\Delta_n}}{h})$, $0 \leq \theta \leq 1$.

Remark 2.4. As Nicolau (2007) pointed out this assumption is generally satisfied under very weak conditions. For instance, with a Gaussian kernel and a Cauchy stationary density (which has heavy tails) we still have $\lim_{h \to 0} E\left[(\frac{1}{h})|K'(\frac{X}{h})|^8\right] < \infty$. Notice that the expectation with respect to the distribution $\xi_{n,i}$ depends on the stationary densities of X and $\widetilde{X}$ because $\xi_{n,i}$ is a convex linear combination of X and $\widetilde{X}$.

Assumption 6. For all $p \geq 1$, $\sup_{t \geq 0} E[|X_t|^p] < \infty$, and $\int_{\mathscr{E}} |z|^p f(z)dz < \infty$.

Remark 2.5. This assumption guarantees that Lemma 3.1 can be used properly throughout the article. If X_t is a Lévy process with bounded jumps (i.e., $\sup_t |\Delta X_t| \leq C < \infty$ almost surely, where C is a nonrandom constant), then $E\{|X_t^n|\} < \infty \; \forall n$, that is, X_t has bounded moments of all orders. This condition is widely used in the estimation of an ergodic diffusion or jump-diffusion from discrete observations, see Florens-Zmirou (1989), Kessler (1997), Shimizu and Yoshida (2006).

Assumption 7. $\Delta_n \to 0$, $h_n = \Delta_n^{\frac{2}{11}}$, $(\frac{n\Delta_n}{h_n})(\Delta_n \log(\frac{1}{\Delta_n}))^{\frac{1}{2}} \to 0$, $h_n n\Delta_n \to \infty$, as $n \to \infty$.

3 Some Technical Lemmas and Asymptotic Results

We lay out some notations. For $x = (x_1, \cdots, x_d)$, $\partial_{x_j} := \frac{\partial}{\partial x_j}$, $\partial_{x_j}^2 := \frac{\partial^2}{\partial x_j^2}$, $\partial_{x_i x_j}^2 := \frac{\partial^2}{\partial x_i \partial x_j}$, $\partial x := (\partial_{x_1}, \cdots, \partial_{x_d})^*$, and $\partial_x^2 f(x) = (\partial_{x_i x_j}^2)_{1 \leq i,j \leq d}$, where $*$ stands for the transpose.

Lemma 3.1. (Yasutaka Shimizu, 2006) Let Z be a d-dimensional solution-process to the stochastic differential equation

$$Z_t = Z_0 + \int_0^t \mu(Z_{s-})ds + \int_0^t \sigma(Z_{s-})dW_s + \int_0^t \int_{\mathscr{E}} c(Z_{s-}, z)r(\omega, dt, dz),$$

where Z_0 is a random variable, $\mathscr{E} = \mathbb{R}^d \setminus \{0\}$, $\mu(x), c(x,z)$ are d-dimensional vectors defined on $\mathbb{R}^d, \mathbb{R}^d \times \mathscr{E}$ respectively, $\sigma(x)$ is a $d \times d$ diagnonal matrix defined on $\mathbb{R}^d$, and W_t is a d-dimensional vector of independent

Brownian motions. Let g be a $C^{2(l+1)}$-class function whose derivatives up to $2(l+1)$th are of polynomial growth. Assume that the coefficient $\mu(x), \sigma(x)$, and $c(x,z)$ are C^{2l}-class function whose derivatives with respective to x up to $2l$th are of polynomial growth. Under Assumption 6, the following expansion holds

$$E[g(Z_t)|\mathscr{F}_s] = \sum_{j=0}^{l} L^j g(Z_s) \frac{\Delta_n^j}{j!} + R, \tag{3.1}$$

for $t > s$ and $\Delta_n = t-s$, where $R = \int_0^{\Delta_n} \int_0^{u_1} \ldots \int_0^{u_l} E[L^{l+1}g(Z_{s+u_{l+1}})|\mathscr{F}_s]du_1 \ldots du_{l+1}$ is a stochastic function of order Δ_n^{l+1}, $Lg(x) = \partial_x^* g(x)\mu(x) + \frac{1}{2}tr[\partial_x^2 g(x)\sigma(x)\sigma^*(x)] + \int_{\mathscr{E}}\{g(x+c(x,z)) - g(x) - \partial_x^* g(x)c(x,z)\}f(z)dz$.

Remark 3.1. Consider a particularly important model:

$$\begin{cases} dY_t = X_t dt, \\ dX_t = \mu(X_{t-})dt + \sigma(X_{t-})dW_t + \int_{\mathscr{E}} c(X_{t-}, z)r(w, dt, dz). \end{cases}$$

As $d = 2$, we have

$$\begin{aligned} Lg(x,y) &= x(\partial g/\partial y) + \mu(x)(\partial g/\partial x) + \frac{1}{2}\sigma^2(x)(\partial^2 g/\partial x^2) \\ &\quad + \int_{\mathscr{E}}\{g(x+c(x,z),y) - g(x,y) - \frac{\partial g}{\partial x} \cdot c(x,z)\}f(z)dz. \end{aligned} \tag{3.2}$$

Based on the second-order differential operator (3.2), we can calculate many mathematical expectations involving $\widetilde{X}_{i\Delta_n}$, for instance (see Appendix for details):

$$E[\frac{\widetilde{X}_{(i+1)\Delta_n} - \widetilde{X}_{i\Delta_n}}{\Delta_n}|\mathscr{F}_{(i-1)\Delta_n}] = \mu(X_{(i-1)\Delta_n}) + O_p(\Delta_n), \tag{3.3}$$

$$E[\frac{(\widetilde{X}_{(i+1)\Delta_n} - \widetilde{X}_{i\Delta_n})^2}{\Delta_n}|\mathscr{F}_{(i-1)\Delta_n}] = \frac{2}{3}\sigma^2(X_{(i-1)\Delta_n}) + \frac{2}{3}\int_{\mathscr{E}} c^2(X_{(i-1)\Delta_n}, z)f(z)dz + O_p(\Delta_n). \tag{3.4}$$

In fact, the above equations provide the basis for estimators (2.2) and (2.3).

Lemma 3.2. (Bandi and Nguyen, 2003) Assume that Assumptions 1, 2, 4, 6, 7 hold, let

$$\hat{p}_n^0(x) = \frac{1}{nh_n}\sum_{i=1}^{n} K\left(\frac{x - X_{(i-1)\Delta_n}}{h_n}\right), \quad \hat{a}_n^0(x) = \frac{A_n^0(x)}{\hat{p}_n^0(x)}, \quad \hat{b}_n^0(x) = \frac{B_n^0(x)}{\hat{p}_n^0(x)} \tag{3.5}$$

where

$$A_n^0(x) = \frac{1}{nh_n}\sum_{i=1}^{n} K\left(\frac{x - X_{(i-1)\Delta_n}}{h_n}\right)\frac{(X_{i\Delta_n} - X_{(i-1)\Delta_n})}{\Delta_n},$$

$$B_n^0(x) = \frac{1}{nh_n}\sum_{i=1}^{n} K\left(\frac{x - X_{(i-1)\Delta_n}}{h_n}\right)\frac{(X_{i\Delta_n} - X_{(i-1)\Delta_n})^2}{\Delta_n}.$$

We have

$$\hat{p}_n^0(x) \xrightarrow{p} p(x), \quad \hat{a}_n^0(x) \xrightarrow{p} \mu(x), \quad \hat{b}_n^0(x) \xrightarrow{p} \sigma^2(x) + \int_{\mathscr{E}} c^2(x,z)f(z)dz,$$

$$\sqrt{h_n n \Delta_n}(\hat{a}_n^0(x) - \mu(x)) \xrightarrow{d} N\left(0, K_2\frac{\sigma^2(x) + \int_{\mathscr{E}} c^2(x,z)f(z)dz}{p(x)}\right),$$

$$\sqrt{h_n n \Delta_n}(\hat{b}_n^0(x) - (\sigma^2(x) + \int_{\mathscr{E}} c^2(x,z)f(z)dz)) \xrightarrow{d} N\left(0, K_2\frac{\int_{\mathscr{E}} c^4(x,z)f(z)dz}{p(x)}\right).$$

Lemma 3.3. Assumptions 1-3 and 5-7 lead to the following results,

(i) $\frac{1}{nh_n}\sum_{i=1}^{n} K\left(\frac{x-\widetilde{X}_{(i-1)\Delta_n}}{h_n}\right) - \frac{1}{nh_n}\sum_{i=1}^{n} K\left(\frac{x-X_{(i-1)\Delta_n}}{h_n}\right) \xrightarrow{p} 0,$

(ii) $\frac{1}{nh_n}\sum_{i=1}^{n} K\left(\frac{x-\widetilde{X}_{(i-1)\Delta_n}}{h_n}\right)\frac{(\widetilde{X}_{(i+1)\Delta_n}-\widetilde{X}_{i\Delta_n})}{\Delta_n} - \frac{1}{nh_n}\sum_{i=1}^{n} K\left(\frac{x-X_{(i-1)\Delta_n}}{h_n}\right)\frac{(\widetilde{X}_{(i+1)\Delta_n}-\widetilde{X}_{i\Delta_n})}{\Delta_n} \xrightarrow{p} 0,$

(iii) $\frac{1}{nh_n}\sum_{i=1}^{n} K\left(\frac{x-\widetilde{X}_{(i-1)\Delta_n}}{h_n}\right)\frac{\frac{3}{2}(\widetilde{X}_{(i+1)\Delta_n}-\widetilde{X}_{i\Delta_n})^2}{\Delta_n} - \frac{1}{nh_n}\sum_{i=1}^{n} K\left(\frac{x-X_{(i-1)\Delta_n}}{h_n}\right)\frac{\frac{3}{2}(\widetilde{X}_{(i+1)\Delta_n}-\widetilde{X}_{i\Delta_n})^2}{\Delta_n} \xrightarrow{p} 0.$

Theorem 3.1. If Assumptions 1 - 7 hold, then

$$\hat{p}_n(x) := \frac{1}{nh_n}\sum_{i=1}^{n} K\left(\frac{x-\widetilde{X}_{(i-1)\Delta_n}}{h_n}\right) \xrightarrow{p} p(x).$$

Theorem 3.2.

(i) Under the Assumptions 1 - 7, we have

$$\hat{a}_n(x) = \frac{A_n(x)}{\hat{p}_n(x)} \xrightarrow{p} \mu(x).$$

(ii) Under the Assumptions 1 - 7 and $h_n n \Delta_n^3 \to 0$, we have

$$\sqrt{h_n n \Delta_n}\left(\hat{a}_n(x) - \mu(x)\right) \xrightarrow{d} N\left(0, K_2 \frac{\sigma^2(x) + \int_{\mathcal{E}} c^2(x,z)f(z)dz}{p(x)}\right).$$

Theorem 3.3.

(i) Under the Assumptions 1 - 7, we have

$$\hat{b}_n(x) = \frac{B_n(x)}{\hat{p}_n(x)} \xrightarrow{p} \sigma^2(x) + \int_{\mathcal{E}} c^2(x,z)f(z)dz.$$

(ii) Under the assumptions 1 - 7 and $h_n n \Delta_n^3 \to 0$, we have

$$\sqrt{h_n n \Delta_n}\left(\hat{b}_n(x) - (\sigma^2(x) + \int_{\mathcal{E}} c^2(x,z)f(z)dz)\right) \xrightarrow{d} N\left(0, K_2 \frac{\int_{\mathcal{E}} c^4(x,z)f(z)dz}{p(x)}\right).$$

Remark 3.2. Contrary to the second-order diffusion model without jumps (João Nicolau, 2007), the second infinitesimal moment estimator has a rate of convergence that is the same as the rate of convergence of the first infinitesimal moment estimator. Apparently, this is due to the presence of discontinuous breaks that have an equal impact on all the functional estimates.

4 Proofs.

Proof of Lemma 3.3.

Let $\varepsilon_{1,n} = \hat{p}_n(x) - \hat{p}_n^0(x) = \frac{1}{nh_n}\sum_{i=1}^{n} K\left(\frac{x-\widetilde{X}_{(i-1)\Delta_n}}{h_n}\right) - \frac{1}{nh_n}\sum_{i=1}^{n} K\left(\frac{x-X_{(i-1)\Delta_n}}{h_n}\right).$

By the mean-value theorem, stationarity and Hölder's inequality, we obtain:

$$
\begin{aligned}
E[|\varepsilon_{1,n}|] &\leq E[\frac{1}{nh_n}\Sigma_{i=1}^{n}|K^{'}(\xi_{n,i})\frac{\widetilde{X}_{(i-1)\Delta_n} - X_{(i-1)\Delta_n}}{h_n}|] \\
&= E[\frac{1}{h_n}|K^{'}(\xi_{n,2})\frac{\widetilde{X}_{\Delta_n} - X_{\Delta_n}}{h_n}|]
\end{aligned}
$$

$$\leq \left(E\left[\frac{1}{h_n}K'^2(\xi_{n,2})\right]\right)^{\frac{1}{2}}\left(E\left[\frac{(\widetilde{X}_{\Delta_n}-X_{\Delta_n})^2}{h_n^3}\right]\right)^{\frac{1}{2}},$$

where $\xi_{n,2}=\theta(\frac{x-X_{\Delta_n}}{h})+(1-\theta)(\frac{x-\widetilde{X}_{\Delta_n}}{h})\ 0\leq\theta\leq 1$.

Using Lemma 3.1 with $d=2$, we get

$$E\left[\frac{(\widetilde{X}_{\Delta_n}-X_{\Delta_n})^2}{h_n^3}\right]=\frac{\Delta_n}{3h_n^3}\left(E[\sigma^2(X_0)]+E\left[\int_{\mathscr{E}}c^2(X_0,z)f(z)dz\right]+O(\Delta_n)\right)\to 0$$

from Assumptions 5 - 7, thus $E[|\varepsilon_{1,n}|]\to 0$. Lemma 3.3 (i) follows from Chebyshev's inequality.

Now we prove (ii), write

$$\varepsilon_{2,n}=A_n(x)-A_n^0(x) = \frac{1}{nh_n}\sum_{i=1}^{n}K\left(\frac{x-\widetilde{X}_{(i-1)\Delta_n}}{h_n}\right)\frac{(\widetilde{X}_{(i+1)\Delta_n}-\widetilde{X}_{i\Delta_n})}{\Delta_n}$$
$$-\frac{1}{nh_n}\sum_{i=1}^{n}K\left(\frac{x-X_{(i-1)\Delta_n}}{h_n}\right)\frac{(\widetilde{X}_{(i+1)\Delta_n}-\widetilde{X}_{i\Delta_n})}{\Delta_n}.$$

We have

$$E[\varepsilon_{2,n}] = E\left[\frac{1}{nh_n}\sum_{i=1}^{n}\left(K\left(\frac{x-\widetilde{X}_{(i-1)\Delta_n}}{h_n}\right)-K\left(\frac{x-X_{(i-1)\Delta_n}}{h_n}\right)\right)E\left[\frac{\widetilde{X}_{(i+1)\Delta_n}-\widetilde{X}_{i\Delta_n}}{\Delta_n}\Big|\mathscr{F}_{(i-1)\Delta_n}\right]\right]$$
$$= E\left[\frac{1}{nh_n}\sum_{i=1}^{n}\left(K\left(\frac{x-\widetilde{X}_{(i-1)\Delta_n}}{h_n}\right)-K\left(\frac{x-X_{(i-1)\Delta_n}}{h_n}\right)\right)\left(\mu(X_{(i-1)\Delta_n})+O_p(\Delta_n)\right)\right]$$
$$= E\left[\frac{1}{h_n}K'(\xi_{n,2})\frac{\widetilde{X}_{\Delta_n}-X_{\Delta_n}}{h_n}\left(\mu(X_{\Delta_n})+O_p(\Delta_n)\right)\right]$$

by (3.3), the mean-value theorem and stationarity. Hence

$$\left|E[\varepsilon_{2,n}]\right| \leq E\left[\frac{1}{h_n}|K'(\xi_{n,2})|\frac{|\widetilde{X}_{\Delta_n}-X_{\Delta_n}|}{h_n}|\mu(X_{\Delta_n})+O_p(\Delta_n)|\right]$$
$$\leq \frac{1}{h_n^2}\left(E[K'^4(\xi_{n,2})]\right)^{\frac{1}{4}}\left(E[|\mu(X_{\Delta_n})+O_p(\Delta_n)|^4]\right)^{\frac{1}{4}}\left(E[(\widetilde{X}_{\Delta_n}-X_{\Delta_n})^2]\right)^{\frac{1}{2}}$$
$$= \left(\frac{E[\sigma^2(X_0)]+E[\int_{\mathscr{E}}c^2(X_0,z)f(z)dz]}{3}+O(\Delta_n)\right)^{\frac{1}{2}}\left(\frac{\Delta_n^{\frac{1}{2}}}{h_n^{\frac{7}{4}}}\right)$$
$$\cdot\left(E\left[\frac{1}{h_n}K'^4(\xi_{n,2})\right]\right)^{\frac{1}{4}}\left(E[|\mu(X_0)|^4]+O(\Delta_n)\right)^{\frac{1}{4}}\to 0,$$

by Hölder inequality, Lemma 3.1 and Assumption 5. So $E[\varepsilon_{2,n}]\to 0$.

Moreover

$$Var[\varepsilon_{2,n}] = \frac{1}{n\Delta_n h_n}Var\left[\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\frac{1}{\sqrt{h_n}}K'(\xi_{n,i})\frac{\widetilde{X}_{(i-1)\Delta_n}-X_{(i-1)\Delta_n}}{h_n}\sqrt{\Delta_n}\frac{\widetilde{X}_{(i+1)\Delta_n}-\widetilde{X}_{i\Delta_n}}{\Delta_n}\right]$$
$$=: \frac{1}{n\Delta_n h_n}Var\left[\frac{1}{\sqrt{n}}\sum_{i=1}^{n}f_i\right].$$

We find $Var[\frac{1}{\sqrt{n}}\sum_{i=1}^{n}f_i] = \frac{1}{n}\sum_{i=1}^{n}Var[f_i] + \varepsilon_n$ where ε_n represents the sum of $\frac{2}{n}\sum_{j=1}^{n-1}\sum_{i=j+1}^{n}$ terms involving the autovariances. Under Assumptions 2 and 3, the series ε_∞ is absolutely convergent, and one easily obtains $Var[\frac{1}{\sqrt{n}}\sum_{i=1}^{n}f_i] < \infty$ if $E[f_i^2] < \infty$.

Using Lemma 3.1 with $d = 2$, simple but tedious calculations enable us to get

$$E[(\widetilde{X}_{(i-1)\Delta_n} - X_{(i-1)\Delta_n})^4|\mathscr{F}_{(i-2)\Delta_n}] = \frac{\Delta_n \int_{\mathscr{E}} c^4(X_{(i-2)\Delta_n}, z)f(z)dz}{3} + O(\Delta_n^2).$$

Now we calculate $E[f_i^2]$:

$$
\begin{aligned}
E[f_i^2] &= E\Big[\frac{1}{h_n}K'^2(\xi_{n,i})\frac{(\widetilde{X}_{(i-1)\Delta_n} - X_{(i-1)\Delta_n})^2}{h_n^2}\frac{(\widetilde{X}_{(i+1)\Delta_n} - \widetilde{X}_{i\Delta_n})^2}{\Delta_n}\Big] \\
&= E\Big[\frac{1}{h_n}K'^2(\xi_{n,i})\frac{(\widetilde{X}_{(i-1)\Delta_n} - X_{(i-1)\Delta_n})^2}{h_n^2}\Big(\frac{2}{3}\sigma^2(X_{(i-1)\Delta_n}) + O_p(\Delta_n)\Big)\Big] \\
&\leq \frac{C}{h_n^3}\Big(E[K'^8(\xi_{n,i})]\Big)^{\frac{1}{4}}\Big(E[(\widetilde{X}_{(i-1)\Delta_n} - X_{(i-1)\Delta_n})^4]\Big)^{\frac{1}{2}}\Big(E[(\sigma^2(X_{(i-1)\Delta_n}) + O_p(\Delta_n))^4]\Big)^{\frac{1}{4}} \\
&= C\frac{\Delta_n^{\frac{1}{2}}}{h_n^{\frac{11}{4}}}\Big(E[\frac{1}{h_n}K'^8(\xi_{n,i})]\Big)^{\frac{1}{4}}\Big(E[\sigma^8(X_{(i-1)\Delta_n})] + O(\Delta_n)\Big)^{\frac{1}{4}} < \infty
\end{aligned}
$$

by (3.4), Hölder inequality and Assumptions 5 and 7.

In conclusion, $Var[\varepsilon_{2,n}] = \frac{1}{n\Delta_n h_n}Var[\frac{1}{\sqrt{n}}\sum_{i=1}^{n}f_i] \to 0$ as $n \to \infty$ by Assumption 7.

So Lemma 3.3 (ii) can be deduced from the above considerations.

Finally, write

$$
\begin{aligned}
\varepsilon_{3,n} = B_n(x) - B_n^0(x) &= \frac{1}{nh_n}\sum_{i=1}^{n}K\Big(\frac{x - \widetilde{X}_{(i-1)\Delta_n}}{h_n}\Big)\frac{(\widetilde{X}_{(i+1)\Delta_n} - \widetilde{X}_{i\Delta_n})^2}{\Delta_n} \\
&\quad - \frac{1}{nh_n}\sum_{i=1}^{n}K\Big(\frac{x - X_{(i-1)\Delta_n}}{h_n}\Big)\frac{(\widetilde{X}_{(i+1)\Delta_n} - \widetilde{X}_{i\Delta_n})^2}{\Delta_n}.
\end{aligned}
$$

Lemma 3.3 (iii) now follows:

$$
\begin{aligned}
E[|\varepsilon_{3,n}|] &\leq E\Big[\frac{1}{nh_n}\sum_{i=1}^{n}\Big|K\Big(\frac{x - \widetilde{X}_{(i-1)\Delta_n}}{h_n}\Big) - K\Big(\frac{x - X_{(i-1)\Delta_n}}{h_n}\Big)\Big|\frac{(\widetilde{X}_{(i+1)\Delta_n} - \widetilde{X}_{i\Delta_n})^2}{\Delta_n}\Big] \\
&\leq CE\Big[\frac{1}{h_n}K'(\xi_{n,2})\frac{|\widetilde{X}_{\Delta_n} - X_{\Delta_n}|}{h_n}\Big(\sigma^2(X_{\Delta_n}) + O_p(\Delta_n)\Big)\Big] \\
&\leq \frac{C}{h_n^2}\Big(E[K'^4(\xi_{n,2})]\Big)^{\frac{1}{4}}\Big(E[(\widetilde{X}_{\Delta_n} - X_{\Delta_n})^2]\Big)^{\frac{1}{2}}\Big(E[\sigma^8(X_{\Delta_n})] + O(\Delta_n)\Big)^{\frac{1}{4}} \\
&= C\frac{\Delta_n^{\frac{1}{2}}}{h_n^{\frac{7}{4}}}\Big(E[\frac{1}{h_n}K'^4(\xi_{n,2})]\Big)^{\frac{1}{4}}\Big(E[\sigma^8(X_{\Delta_n})] + O(\Delta_n)\Big)^{\frac{1}{4}} \to 0,
\end{aligned}
$$

which implies Lemma 3.3 (iii).

Proof of Theorem 3.1. One can easily prove Theorem 3.1 by Lemma 3.2 and Lemma 3.3 (i).

Proof of Theorem 3.2.
(i) From Theorem 3.1 we get $\hat{p}_n(x) \xrightarrow{p} p(x)$, hence to prove

$$\hat{a}_n(x) = \frac{A_n(x)}{\hat{p}_n(x)} \xrightarrow{p} \mu(x)$$

it is sufficient to verify that
$$A_n(x) \xrightarrow{p} \mu(x)p(x).$$

From Lemma 3.2. we obtain
$$A_n^0(x) \xrightarrow{p} \mu(x)p(x).$$

Now we prove $A_n(x) - A_n^0(x) \xrightarrow{p} 0$. By Lemma 3.3 (ii) $A_n(x) - A_n^0(x)$ has the same limit in probability as

$$
\begin{aligned}
\delta_{1,n}(x) \quad &:= \quad \frac{1}{nh_n} \sum_{i=1}^{n} K\Big(\frac{x - X_{(i-1)\Delta_n}}{h_n}\Big) \frac{(\widetilde{X}_{(i+1)\Delta_n} - \widetilde{X}_{i\Delta_n})}{\Delta_n} \\
&\quad - \frac{1}{nh_n} \sum_{i=1}^{n} K\Big(\frac{x - X_{(i-1)\Delta_n}}{h_n}\Big) \frac{(X_{i\Delta_n} - X_{(i-1)\Delta_n})}{\Delta_n} \\
&= \quad \frac{1}{nh_n} \sum_{i=1}^{n} K\Big(\frac{x - X_{(i-1)\Delta_n}}{h_n}\Big) \Big(\frac{(\widetilde{X}_{(i+1)\Delta_n} - \widetilde{X}_{i\Delta_n})}{\Delta_n} - \frac{(X_{i\Delta_n} - X_{(i-1)\Delta_n})}{\Delta_n}\Big).
\end{aligned}
$$

We now prove that $\delta_{1,n}(x) \xrightarrow{p} 0$. Using Assumptions 1, 2, 4, 6 and Lemma 3.1, we find

$$
\begin{aligned}
E[\delta_{1,n}(x)] \;=\; & E\Big[\frac{1}{h_n} K\Big(\frac{x - X_{(i-1)\Delta_n}}{h_n}\Big) \Big(\frac{(\widetilde{X}_{(i+1)\Delta_n} - \widetilde{X}_{i\Delta_n})}{\Delta_n} - \frac{(X_{i\Delta_n} - X_{(i-1)\Delta_n})}{\Delta_n}\Big)\Big] \\
=\; & E\Big[\frac{1}{h_n} K\Big(\frac{x - X_{(i-1)\Delta_n}}{h_n}\Big) E\Big[E\Big[\Big(\frac{(\widetilde{X}_{(i+1)\Delta_n} - \widetilde{X}_{i\Delta_n})}{\Delta_n} - \frac{(X_{i\Delta_n} - X_{(i-1)\Delta_n})}{\Delta_n}\Big)\Big|\mathscr{F}_{i\Delta_n}\Big]\Big|\mathscr{F}_{(i-1)\Delta_n}\Big]\Big] \\
=\; & \frac{\Delta_n}{2} E\Big[\frac{1}{h_n} K\Big(\frac{x - X_{(i-1)\Delta_n}}{h_n}\Big) \Big(\mu(X_{(i-1)\Delta_n})\mu'(X_{(i-1)\Delta_n}) + \frac{1}{2}\sigma^2(X_{(i-1)\Delta_n})\mu''(X_{(i-1)\Delta_n}) \\
& + \int_{\mathscr{E}} \{\mu(X_{(i-1)\Delta_n} + c(X_{(i-1)\Delta_n}, z)) - \mu(X_{(i-1)\Delta_n}) - \mu'(X_{(i-1)\Delta_n}) \cdot c(X_{(i-1)\Delta_n}, z)\} f(z)dz\Big)\Big] \\
=\; & \frac{\Delta_n}{2} \int \frac{1}{h_n} K\Big(\frac{x - u}{h_n}\Big) \Big(\mu(u)\mu'(u) + \frac{1}{2}\sigma^2(u)\mu''(u) \\
& + \int_{\mathscr{E}} \{\mu(u + c(u, z)) - \mu(u) - \mu'(u) \cdot c(u, z)\} f(z)dz\Big) p(u)du \\
=\; & O(\Delta_n)
\end{aligned}
$$

and

$$
\begin{aligned}
& Var[\delta_{1,n}(x)] \\
=\; & \frac{1}{n\Delta_n h_n} Var\Big[\frac{1}{\sqrt{n}} \sum_{i=1}^{n} \frac{1}{\sqrt{h_n}} K\Big(\frac{x - X_{(i-1)\Delta_n}}{h_n}\Big) \sqrt{\Delta_n}\Big(\frac{(\widetilde{X}_{(i+1)\Delta_n} - \widetilde{X}_{i\Delta_n})}{\Delta_n} - \frac{(X_{i\Delta_n} - X_{(i-1)\Delta_n})}{\Delta_n}\Big)\Big] \\
=:\; & \frac{1}{n\Delta_n h_n} Var\Big[\frac{1}{\sqrt{n}} \sum_{i=1}^{n} g_i\Big].
\end{aligned}
$$

We have $Var[\frac{1}{\sqrt{n}}\sum_{i=1}^{n} g_i] = \frac{1}{n}\sum_{i=1}^{n} Var[g_i] + \varepsilon_n$ where ε_n represents the sum of $\frac{2}{n}\sum_{j=1}^{n-1}\sum_{i=j+1}^{n}$ terms involving the autovariances. Under Assumptions 2 and 3, the series ε_∞ is absolutely convergent, and one easily obtains $Var[\frac{1}{\sqrt{n}}\sum_{i=1}^{n} g_i] < \infty$ if $E[g_i^2] < \infty$. In fact

$$E[g_i^2] = E\Big[\frac{1}{h_n} K^2\Big(\frac{x - X_{(i-1)\Delta_n}}{h_n}\Big) \Delta_n \Big(\frac{(\widetilde{X}_{(i+1)\Delta_n} - \widetilde{X}_{i\Delta_n})}{\Delta_n} - \frac{(X_{i\Delta_n} - X_{(i-1)\Delta_n})}{\Delta_n}\Big)^2\Big]$$

is finite because

$$E\Big[\Delta_n\Big(\frac{(\widetilde{X}_{(i+1)\Delta_n} - \widetilde{X}_{i\Delta_n})}{\Delta_n} - \frac{(X_{i\Delta_n} - X_{(i-1)\Delta_n})}{\Delta_n}\Big)^2\Big] = \frac{2}{3}\Big(E[\sigma^2(X_0)] + E[\int_{\mathscr{E}} c^2(X_0, z)f(z)dz]\Big) + O(\Delta_n)$$

by Lemma 3.1. In conclusion, $Var[\delta_{1,n}] = \frac{1}{n\Delta_n h_n} Var[\frac{1}{\sqrt{n}} \sum_{i=1}^{n} g_i] \to 0$ as $n \to \infty$ by Assumption 7.

(ii) By Lemma 3.2,

$$U_n^0(x) := \sqrt{h_n n \Delta_n}(\hat{a}_n^0(x) - \mu(x)) \xrightarrow{d} N\left(0, K_2 \frac{\sigma^2(x) + \int_{\mathscr{E}} c^2(x,z)f(z)dz}{p(x)}\right)$$

and by the asymptotic equivalence theorem, it suffices to prove that

$$U_n(x) - U_n^0(x) \xrightarrow{p} 0$$

where $U_n(x) = \sqrt{h_n n \Delta_n}(\hat{a}_n(x) - \mu(x))$. From part (i) we know that $U_n(x) - U_n^0(x) = \sqrt{h_n n \Delta_n}\left(\frac{\delta_{1,n}(x)}{\hat{p}_n(x)}\right)$ and $\frac{\delta_{1,n}(x)}{\hat{p}_n(x)} = O_p(\Delta_n)$. Thus the assumption $h_n n \Delta_n^3 \to 0$ leads to the result.

Proof of Theorem 3.3.

(i) From Theorem 3.1 we know $\hat{p}_n(x) \xrightarrow{p} p(x)$, hence to prove

$$\hat{b}_n(x) = \frac{B_n(x)}{\hat{p}_n(x)} \xrightarrow{p} \sigma^2(x) + \int_{\mathscr{E}} c^2(x,z)f(z)dz$$

it is sufficient to verify that

$$B_n(x) \xrightarrow{p} (\sigma^2(x) + \int_{\mathscr{E}} c^2(x,z)f(z)dz)p(x).$$

Following Lemma 3.2 under the conditions of the theorem we obtain

$$B_n^0(x) \xrightarrow{p} (\sigma^2(x) + \int_{\mathscr{E}} c^2(x,z)f(z)dz)p(x).$$

So, we only need to prove that

$$B_n(x) - B_n^0(x) \xrightarrow{p} 0.$$

By Lemma 3.3 (iii) $B_n(x) - B_n^0(x)$ has the same limit in probability as

$$
\begin{aligned}
\delta_{2,n}(x) \quad &:= \quad \frac{1}{nh_n} \sum_{i=1}^{n} K\left(\frac{x - X_{(i-1)\Delta_n}}{h_n}\right) \frac{\frac{3}{2}(\widetilde{X}_{(i+1)\Delta_n} - \widetilde{X}_{i\Delta_n})^2}{\Delta_n} \\
&\quad - \frac{1}{nh_n} \sum_{i=1}^{n} K\left(\frac{x - X_{(i-1)\Delta_n}}{h_n}\right) \frac{(X_{i\Delta_n} - X_{(i-1)\Delta_n})^2}{\Delta_n} \\
&= \quad \frac{1}{nh_n} \sum_{i=1}^{n} K\left(\frac{x - X_{(i-1)\Delta_n}}{h_n}\right) \left(\frac{\frac{3}{2}(\widetilde{X}_{(i+1)\Delta_n} - \widetilde{X}_{i\Delta_n})^2}{\Delta_n} - \frac{(X_{i\Delta_n} - X_{(i-1)\Delta_n})^2}{\Delta_n}\right).
\end{aligned}
$$

We now prove that $\delta_{2,n}(x) \xrightarrow{p} 0$. Let $q_i = \frac{\frac{3}{2}(\widetilde{X}_{(i+1)\Delta_n} - \widetilde{X}_{i\Delta_n})^2}{\Delta_n} - \frac{(X_{i\Delta_n} - X_{(i-1)\Delta_n})^2}{\Delta_n}$. By Lemma 3.1

$$E\left[E\left[\frac{\frac{3}{2}(\widetilde{X}_{(i+1)\Delta_n} - \widetilde{X}_{i\Delta_n})^2}{\Delta_n}\Big|\mathscr{F}_{i\Delta_n}\right]\Big|\mathscr{F}_{(i-1)\Delta_n}\right] = \sigma^2(X_{(i-1)\Delta_n}) + \int_{\mathscr{E}} c^2(X_{(i-1)\Delta_n})f(z)dz + O(\Delta_n),$$

$$E\left[\frac{(X_{i\Delta_n} - X_{(i-1)\Delta_n})^2}{\Delta_n}\Big|\mathscr{F}_{(i-1)\Delta_n}\right] = \sigma^2(X_{(i-1)\Delta_n}) + \int_{\mathscr{E}} c^2(X_{(i-1)\Delta_n})f(z)dz + O(\Delta_n).$$

One has

$$E[q_i] = E\left[\left(\frac{\frac{3}{2}(\widetilde{X}_{(i+1)\Delta_n} - \widetilde{X}_{i\Delta_n})^2}{\Delta_n} - \frac{(X_{i\Delta_n} - X_{(i-1)\Delta_n})^2}{\Delta_n}\right)\right] = O(\Delta_n).$$

We easily verifies that $E[\delta_{2,n}(x)] \to 0$ by stationarity and the above equations.

On the other hand,

$$Var[\delta_{2,n}(x)]$$

$$= \frac{1}{n\Delta_n h_n}Var\left[\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\frac{1}{\sqrt{h_n}}K\left(\frac{x-X_{(i-1)\Delta_n}}{h_n}\right)\sqrt{\Delta_n}\left(\frac{\frac{3}{2}(\widetilde{X}_{(i+1)\Delta_n}-\widetilde{X}_{i\Delta_n})^2}{\Delta_n}-\frac{(X_{i\Delta_n}-X_{(i-1)\Delta_n})^2}{\Delta_n}\right)\right]$$

$$=: \frac{1}{n\Delta_n h_n}Var\left[\frac{1}{\sqrt{n}}\sum_{i=1}^{n}s_i\right].$$

Using the same arguments as in the proof of Theorem 3.2 (i), it is easy to conclude that the condition:

$$E\left[\Delta_n\left(\frac{\frac{3}{2}(\widetilde{X}_{(i+1)\Delta_n}-\widetilde{X}_{i\Delta_n})^2}{\Delta_n}-\frac{(X_{i\Delta_n}-X_{(i-1)\Delta_n})^2}{\Delta_n}\right)^2\right] = O(1)$$

assures $Var[\delta_{2,n}(x)] \to 0$.

(ii) By Lemma 3.2,

$$U_n^0(x) := \sqrt{h_n n\Delta_n}\left(\hat{b}_n^0(x) - (\sigma^2(x) + \int_{\mathcal{E}}c^2(x,z)f(z)dz)\right) \xrightarrow{d} N\left(0, K_2\frac{\int_{\mathcal{E}}c^4(x,z)f(z)dz}{p(x)}\right)$$

and by the asymptotic equivalence theorem, it suffices to prove that

$$U_n(x) - U_n^0(x) \xrightarrow{p} 0$$

where $U_n(x) = \sqrt{h_n n\Delta_n}(\hat{b}_n(x) - (\sigma^2(x) + \int_{\mathcal{E}}c^2(x,z)f(z)dz))$.

From part (i) we know that

$$U_n(x) - U_n^0(x) = \sqrt{h_n n\Delta_n}\left(\frac{\delta_{2,n}(x)}{\hat{p}_n(x)}\right)$$

and

$$\frac{\delta_{2,n}(x)}{\hat{p}_n(x)} = O_p(\Delta_n).$$

Thus the assumption $h_n n\Delta_n^3 \to 0$ leads to the result.

5 Appendix.

In this part, we give some details for equations (3.3) and (3.4).

$$E[\frac{\widetilde{X}_{(i+1)\Delta_n}-\widetilde{X}_{i\Delta_n}}{\Delta_n}|\mathscr{F}_{(i-1)\Delta_n}] = E[E[\frac{(Y_{(i+1)\Delta_n}-2Y_{i\Delta_n}+Y_{(i-1)\Delta_n})}{\Delta_n^2}|\mathscr{F}_{i\Delta_n}]|\mathscr{F}_{(i-1)\Delta_n}]. \qquad (5.1)$$

We firstly deal with the following part by equations (3.1) and (3.2):

$$E[\frac{(Y_{(i+1)\Delta_n}-2Y_{i\Delta_n}+Y_{(i-1)\Delta_n})}{\Delta_n^2}|\mathscr{F}_{i\Delta_n}] = \frac{1}{\Delta_n^2}(Y_{(i-1)\Delta_n}-Y_{i\Delta_n})+\frac{1}{\Delta_n}X_{i\Delta_n}+\frac{1}{2}\mu(X_{i\Delta_n})+O_p(\Delta_n). \qquad (5.2)$$

By (3.1) and (5.2), (5.1) can be written as:

$$E[\frac{1}{\Delta_n^2}(Y_{(i-1)\Delta_n}-Y_{i\Delta_n})|\mathscr{F}_{(i-1)\Delta_n}]+E[\frac{1}{\Delta_n}X_{i\Delta_n}|\mathscr{F}_{(i-1)\Delta_n}]+\frac{1}{2}\mu(X_{(i-1)\Delta_n})+O_p(\Delta_n). \qquad (5.3)$$

We have: $E[\frac{1}{\Delta_n^2}(Y_{(i-1)\Delta_n}-Y_{i\Delta_n})|\mathscr{F}_{(i-1)\Delta_n}] = -\frac{1}{\Delta_n}X_{(i-1)\Delta_n}-\frac{1}{2}\mu(X_{(i-1)\Delta_n})+O_p(\Delta_n),$ \qquad (5.4)

$$E[\frac{1}{\Delta_n}X_{i\Delta_n}|\mathscr{F}_{(i-1)\Delta_n}] = \frac{1}{\Delta_n}X_{(i-1)\Delta_n}+\mu(X_{(i-1)\Delta_n})+O_p(\Delta_n). \qquad (5.5)$$

Finally we can deduce (3.3) by the above equations (5.1) - (5.5).

For (3.4), we know the following equations by the tower property of conditional expectation:

$$E[\frac{(\widetilde{X}_{(i+1)\Delta_n} - \widetilde{X}_{i\Delta_n})^2}{\Delta_n}|\mathscr{F}_{(i-1)\Delta_n}] = E[E[\frac{(Y_{(i+1)\Delta_n} - 2Y_{i\Delta_n} + Y_{(i-1)\Delta_n})^2}{\Delta_n^3}|\mathscr{F}_{i\Delta_n}]|\mathscr{F}_{(i-1)\Delta_n}]. \qquad (5.6)$$

We deal with the inner part in (5.6) by equations (3.1) and (3.2):

$$
\begin{aligned}
E[\frac{(Y_{(i+1)\Delta_n} - 2Y_{i\Delta_n} + Y_{(i-1)\Delta_n})^2}{\Delta_n^3}|\mathscr{F}_{i\Delta_n}] &= \frac{1}{\Delta_n^3}(Y_{i\Delta_n} - Y_{(i-1)\Delta_n})^2 + \frac{2}{\Delta_n^2}(Y_{(i-1)\Delta_n} - Y_{i\Delta_n})X_{i\Delta_n} \\
&\quad + \frac{1}{\Delta_n}[X_{i\Delta_n}^2 + \mu(X_{i\Delta_n})(Y_{(i-1)\Delta_n} - Y_{i\Delta_n})] \\
&\quad + \frac{1}{6}[6X_{i\Delta_n}\mu(X_{i\Delta_n}) + 2\sigma^2(X_{i\Delta_n}) + 4\int_{\mathscr{E}} c^2(X_{i\Delta_n}, z)f(z)dz] + O_p(\Delta_n) \\
&:= A_1 + A_2 + A_3 + A_4
\end{aligned}
$$

For $E[A_1|\mathscr{F}_{(i-1)\Delta_n}]$, denote $g(x,y) = (y - Y_{(i-1)\Delta_n})^2$, we have:

$Lg(x,y) = 2(y - Y_{(i-1)\Delta_n})x$,

$L^2g(x,y) = 2x^2 + 2\mu(x)(y - Y_{(i-1)\Delta_n})$,

$L^3g(x,y) = 2\mu(x)x + \mu(x)[4x + 2\mu'(x)(y - Y_{(i-1)\Delta_n})] + \frac{1}{2}\sigma^2(x)[4 + 2\mu''(x)(y - Y_{(i-1)\Delta_n})]$

$\quad + \int_{\mathscr{E}}\{2(x+c(x,z))^2 - 2x^2 + [2\mu(x+c(x,z)) - 2\mu(x)](y - Y_{(i-1)\Delta_n}) - [4x + 2\mu'(x)(y - Y_{(i-1)\Delta_n})]c(x,z)\}f(z)dz$.

By Lemma 3.1, we can obtain

$E[A_1|\mathscr{F}_{(i-1)\Delta_n}] = \frac{X_{(i-1)\Delta_n}^2}{\Delta_n} + \mu(X_{(i-1)\Delta_n})X_{(i-1)\Delta_n} + \frac{\sigma^2(X_{(i-1)\Delta_n})}{3} + \frac{1}{3}\int_{\mathscr{E}} c^2(X_{(i-1)\Delta_n}, z)f(z)dz + O_p(\Delta_n)$,

By constructing functions $g_2(x,y) = (Y_{(i-1)\Delta_n} - y)x$, $g_3(x,y) = x^2 + \mu(x)(Y_{(i-1)\Delta_n} - y)$, we can get:

$E[A_2|\mathscr{F}_{(i-1)\Delta_n}] = -\frac{2X_{(i-1)\Delta_n}^2}{\Delta_n} - 3\mu(X_{(i-1)\Delta_n})X_{(i-1)\Delta_n} - \sigma^2(X_{(i-1)\Delta_n}) - \int_{\mathscr{E}} c^2(X_{(i-1)\Delta_n}, z)f(z)dz + O_p(\Delta_n)$,

$E[A_3|\mathscr{F}_{(i-1)\Delta_n}] = \frac{X_{(i-1)\Delta_n}^2}{\Delta_n} + X_{(i-1)\Delta_n}\mu(X_{(i-1)\Delta_n}) + \sigma^2(X_{(i-1)\Delta_n}) + \int_{\mathscr{E}} c^2(X_{(i-1)\Delta_n}, z)f(z)dz + O_p(\Delta_n)$.

We easily have:

$E[A_4|\mathscr{F}_{(i-1)\Delta_n}] = \mu(X_{(i-1)\Delta_n})X_{(i-1)\Delta_n} + \frac{\sigma^2(X_{(i-1)\Delta_n})}{3} + \frac{1}{3}\int_{\mathscr{E}} c^2(X_{(i-1)\Delta_n}, z)f(z)dz + O_p(\Delta_n)$.

Following the above calculation, we known (3.4) holds, that is:

$E[\frac{(\widetilde{X}_{(i+1)\Delta_n} - \widetilde{X}_{i\Delta_n})^2}{\Delta_n}|\mathscr{F}_{(i-1)\Delta_n}] = \frac{2}{3}\sigma^2(X_{(i-1)\Delta_n}) + \frac{2}{3}\int_{\mathscr{E}} c^2(X_{(i-1)\Delta_n}, z)f(z)dz + O_p(\Delta_n)$.

Other equations calculated analogue with the above, we omit them here.

6 References

Baskshi, G., Cao, Z.,Chen, Z. (1997) Empirical performance of alternative option pricing models. *Journal of Finance* 52, 2003-2049

Bandi, F., Nguyen, T.(2003) On the functional estimation of jump diffusion models. *Journal of Econometrics* 116, 293-328.

Bandi, F., Phillips, P. (2003) Fully nonparametric estimation of scalar diffusion models. *Econometrica* 71, 241-283.

Baxter, M., Rennie, A. (1996) Financial Calculus. An Introduction to Derivative Pricing. Cambridge, U.K.: Cambridge University Press.

Chen, X., Hansen, L., Carrasco, M. (2010) Nonlinearity and temporal dependence. *Journal of Econometrics* 155, 155-169.

Ditlevsen, S., Sørensen, M. (2004) Inference for observations of integrated diffusion processes.*Scandinavian Journal of Statistics* 31, 417-429.

Duffie, D., Pan, J. and Singleton, K. (2000) Transform analysis and asset pricing for affine jump-diffusions. *Econometrica* 68, 1343-1376.

Eraker, B., Johannes, M. and Polson, N. (2003) The impact of jumps in volatility and returns. *Journal of Finance* 58, 1269-1300.

Fan, J. and Zhang, C. (2003) A re-examination of diffusion estimators with applications to financial model validation. *Journal of the American Statistical Association*, 98, 118- 134.

Florens-Zmirou, D. (1989) Approximate discrete time schemes for statistics of diffusion processes. *Statistics* 20, 547-557.

Hansen, L., Scheinkman, J. (1995) Back to the future: Generating moment implications for continuous-time Markov processes. *Econometrica* 63, 767-804

Jacod, J., Shiryaev, A. (1987) Limit Theory for Stochastic Processes. Springer, Berlin-Heidelberg-New York-Hong Kong-London-Milan-Paris-Tokyo.

Johannes, M.S. (2004) The economic and statistical role of jumps to interest rates. *Journal of Finance* 59, 227-260.

Kessler, M. (1997) Estimation of an ergodic diffusion from discrete observations. *Scandinavian Journal of Statistics* 24, 211-229.

Lin, Z., Wang, H. (2010) Empirical Likelihood Inference for Diffusion Processes with Jumps. *Science China Mathematics* 53, 1805-1816.

Mancini, C., (2004) Estimation of the parameters of jump of a general Poisson-diffusion model. *Scandinavian Actuarial Journal* 1, 42-52

Mancini, C., (2009) Non-parametric threshold for models with stochastic diffusion coefficient and jumps. *Scandinavian Journal of Statistics* 36(2), 270-296

Mancini, C., Renò, R., (2010) Threshold estimation of Markov models with jumps and interest rate modeling. *Journal of Econometrics* 10, 3-19

Nicolau, J. (2003) Bias reduction in nonparametric diffusion coefficient estimation. *Econometric Theory* 19, 754-777.

Nicolau, J. (2007) Nonparametric estimation of scend-order stochastic differential equations. *Econometric Theory* 23, 880-898.

Park, J., Phillips, P. (2001) Nonlinear regressions with integrated time series. *Econometrica* 69, 117-161.

Protter, P. (2005) Stochastic Integration and Differential Equations. Springer, New York.

Rogers, L., Williams, D. (2000) Diffusions, Markov processes, and Martingales, Vol 2. Cambridge University Press.

Shimizu, Y. and Yoshida, N. (2006) Estimation of parameters for diffusion processes with jumps from discrete observation. *Statistical Inference of Stochastic Processes* 9, 227-277.

Xu, K., (2009) Empirical likelihood based inference for nonparametric recurrent diffusions. *Journal of Econometrics* 153, 65-82.

Xu, K., (2010) Re-weighted Functional Estimation of Diffusion Models. *Econometric Theory* 26, 541-563.

Wang, H., Lin, Z. (2011) Local linear estimation of second-order diffusion models. *Communications in Statistics-Theory and Methods* 40, 394-407.

Wang, Y., Zhang, L., Wang, H. (submitted) Empirical likelihood based inference for second-order diffusion models.