QUADERNI

Università degli Studi di Siena
DIPARTIMENTO DI ECONOMIA POLITICA

CECILIA MANCINI
ROBERTO RENO'

Nonparametric estimation in models with Lévy type jumps and stochastic volatility

n. 451 - Maggio 2005

Abstract - We introduce a nonparametric estimator of the volatility function in univariate processes with L´evy type jumps and stochastic volatility when we observe the state variable at discrete times. Our results rely on the fact that it is possible to recognize the discontinuous part of the state variable from those squared increments between observations exceeding a suitable threshold. We discuss the implementation of the estimator with high-frequency data.

JEL Classification: C14,C6

Cecilia Mancini, Dipartimento di Matematica per le Decisioni, Università di Firenze
Roberto Renò, Dipartimento di Economia Politica, Università di Siena

1 Introduction

In this paper, we focus on nonparametric estimation of a univariate model with stochastic volatility and jumps such as

$$dX_t = a(X_t)dt + \sigma(X_t)dW_t + dJ_t, \qquad t \in [0, T],$$

defined on a fixed time span $[0, T]$. X describes the evolution of an economic variable as an interest rate or a logarithmic asset price. The semimartingale X is composed by a drift component which is a deterministic function $a(\cdot)$ of X_t, by a continuous, diffusive Brownian motion with stochastic volatility in the form of a deterministic function $\sigma(\cdot)$ of X_t, and by a discontinuous part in form of jumps driven by a Lévy process J_t. These kind of models turn out to be very useful when the state variable varies in form of small, continuous changes (modelled by the Brownian motion) as well as with abrupt, discontinuous variations (modelled by the jump component).

Stochastic volatility models with jumps are used in a variety of financial applications. For interest rate modelling, Das (2002); Piazzesi (2005) show that the role of jumps is relevant in incorporating newly released information in interest rate levels. The statistical and economic role of jumps in interest rate modelling is further discussed in Johannes (2004). Bond pricing for jump-diffusions is discussed in Eberlein and Raible (1999). Jumps are also very important for derivative pricing, since option writers and buyers are aware of the possibility of sudden changes of the underlying, so that they demand an higher risk premium which affects the term structure of implied volatility, see Bakshi et al. (1997); Bates (2000); Eraker et al. (2002); Andersen et al. (2002); Pan (2002). Also pure-jump processes received a lot of attention, see (Madan, 1999; Carr et al., 2002). An important problem for the risk management and for the construction of hedging strategies is to identify the contribution given to X separately by each component.

In the recent literature, many intriguing methodologies have been proposed to separate the variations in the state variable X_t due to the diffusive part from those due to jumps with a feasible econometric technique.
In a parametric model parameterized functional forms are imposed for the drift and diffusion functions and for the jump component. In that framework Jiang and Oomen (2004) develop estimators based on a weighted sum of squared increments in an affine assets price model where the jump part has finite activity. Ait-Sahalia (2004) shows that it is possible to disentangle jumps from the continuous variations using a maximum likelihood approach. Aït-Sahalia and Jacod (2005) construct a threshold based estimator of the volatility when each source of randomness is a stable Levy process.
In a nonparametric framework, Barndorff-Nielsen and Shephard (2004a,b); Woerner (2003); Mancini (2004a) estimate the *integrated* volatility. Barndorff-Nielsen and Shephard (2004b) show that, when the volatility is independent of the leading Brownian motion, the power variation of the state variable is a consistent estimator of the integral of the corresponding power of the volatility, even in the presence of a finite activity jump process. Woerner (2003) shows that the power variation is consistent even when the power is strictly less than 2 and the jump part belongs to a specified class of infinite activity processes. Barndorff-Nielsen and Shephard (2004a) develop the original theory of the bi-power variation, which also allows them to construct a test for the presence of jumps.
In this paper, we want to estimate the *local* volatility. Nonparametric estimation of the drift $a(\cdot)$ and the diffusion coefficient $\sigma(\cdot)$ has been studied, in absence of the jump component, by Florens-Zmirou (1993); Jiang and Knight (1997); Stanton

(1997); Bandi and Phillips (2003); Renò (2004). In presence of jumps, the only nonparametric estimators we are aware of has been proposed by Bandi and Nguyen (2003) and studied by Johannes (2004). Their model contains a finite activity jump part, not necessarily of Lévy type, since it has stochastic jump intensity. They base themselves on nonparametric kernel estimation of unconditional moments of the state variable, and the presence of jumps is identified by an estimate of the excess kurtosis.
Our estimator is basically different. We build on the work of Mancini (2004a,b), who shows that when the interval between two observations shrinks, since the diffusive part tends to zero at known rate, it is possible to establish when there were some jumps and to identify asymptotically both jump times and sizes. Therefore our idea is to get an estimate of the continuous path of the state variable, and to perform nonparametric estimation on that. We also discuss an extension which can be implemented with high-frequency data.

2 Assumptions and preliminary results

In this Section, we set up the model, the assumptions and we recall some preliminary results which will be useful in constructing our estimator. We model the evolution of an (observable) state variable by a stochastic process X_t in the time interval $[0, T]$. While we leave the possibility to X_t to be any variable, in financial applications it can be thought as the short rate, or the logarithm of an asset price, of a stock index, or of a foreign exchange rate.

We work in a filtered probability space (Ω, $(\mathcal{F}_t)_{t\in[0,T]}$, $\mathcal{F}$, P), satisfying the usual conditions (Protter, 1990), where W is a standard Brownian motion and J is a pure jump Lévy process. We then assume that $(X_t)_{t\in[0,T]}$ is a real process such that $X_0 \in \mathbb{R}$ and

$$dX_t = a(X_t)dt + \sigma(X_t)dW_t + dJ_t, \; t \in [0, T]. \tag{2.1}$$

The Lévy process J can be decomposed as the sum of the jumps bigger than one and the sum of the compensated jumps smaller than one. Following this decomposition, we write $J \equiv J_1 + \tilde{J}_2$, where

$$J_{1s} \doteq \int_0^s \int_{|x|>1} x\mu(dt, dx) \text{ and } \tilde{J}_{2s} \doteq \int_0^s \int_{|x|\le 1} x[\mu(dt, dx) - \nu(dx)dt],$$

μ being the jumps random measure of X, and ν being the Lévy measure of J. The process J_1 of the jumps with size bigger than one is a finite activity compound Poisson process, $J_{1s} = \sum_{\ell=1}^{N_s^1} \gamma_\ell^1$ (see Cont and Tankov (2004)). For simplicity we will write N in place of N^1, and γ in place of γ^1.

Typically, we do not observe X_t continuously, but in form of n discrete time observations $\{X_0, X_{t_1}, ..., X_{t_{n-1}}, X_{t_n}\}$. We develop our theory for the case of equally spaced observations, that is $t_i = i\delta$, where $\delta = T/n$. However our results still hold when the data are not equally spaced (Mancini (2004a), Florens-Zmirou (1993), Jiang and Knight (1997)).

For a given process Z, we use the following notations:

- $\Delta_i Z = Z_{t_i} - Z_{t_{i-1}}$, the increment of Z between t_{i-1} and t_i
- $\Delta Z_t = Z_t - Z_{t-}$ the size of the jump, if any, at time t

- $\Delta_{i,j}Z = Z_{t_i+s^i_j} - Z_{t_i+s^i_{j-1}}$, the increment of Z between $t_i + s^i_{j-1}$ and $t_i + s^i_j$ where $\{s^i_j, j = 1..m\}$ is a partition of $]t_{i-1}, t_i]$. In the case of equally spaced observations we will have $\Delta_{i,j}Z = Z_{\frac{1}{n}(i+\frac{j}{m})} - Z_{\frac{1}{n}(i+\frac{j-1}{m})}$
- Z^c is the continuous martingale part of Z
- We denote by $(\tau_j)_{j\in\mathbb{N}}$ the jump instants of J_1 and by $\tau^{(i)}$ the instant of the first jump in $]t_{i-1}, t_i]$, if $\Delta_i N \geq 1$
- $H.W$ is the process given by the stochastic integral $\left(\int_0^t H_s dW_s\right)_{t\in[0,T]}$.

We require the following assumption throughout all the paper.

Assumption 2.1 *$a_t \doteq a(X_t)$, $\sigma_t \doteq \sigma(X_t)$ are progressively measurable processes with cadlag paths such to guarantee that the SDE (2.1) has a unique strong solution which is adapted and right continuous with left limits on $[0,T]$ (see Ikeda and Watanabe (1981)).*

Definition 2.2 *A bandwidth parameter is a sequence of real numbers h such that as $n \to \infty$ we have $h \to 0$ and $nh \to \infty$.*

An example of bandwidth parameter which is very popular in applications (Scott, 1992) is the following:

$$h = h_s \hat{\sigma} n^{-\frac{1}{5}} \tag{2.2}$$

where h_s is a real constant to be tuned, and $\hat{\sigma}$ is the sample standard deviation.

When $J \equiv 0$ we denote our process by Y. In the case of no jumps, Florens-Zmirou (1993) proves the following proposition

Proposition 2.3 *(Florens-Zmirou, 1993) Define*

$$S^n_t(x) \doteq \frac{n\sum_{i=1}^n I_{\{|Y_{t_i}-x|<h_n\}}(\Delta_i Y)^2}{T\sum_{i=1}^n I_{\{|Y_{t_i}-x|<h_n\}}}. \tag{2.3}$$

If
• the coefficient function $a(x)$ is bounded and has two continuous bounded derivatives;
• the local volatility function $\sigma(x)$ is uniformly bounded and bounded away from zero and has three continuous bounded derivatives;
• if the bandwidth parameter satisfies $nh^4 \to 0$ as $n \to \infty$,
then $S^n_t(x)$ converges to $\sigma^2(x)$ in probability for any x visited by Y.

In fact she shows that

$$\frac{1}{2h}\sum_{i=1}^n I_{\{|Y_{t_i}-x|<h\}}\frac{T}{n} \tag{2.4}$$

is a consistent approximation of the local time $L_T(x)$, the time the process Y stands near x, and

$$\frac{1}{2h}\sum_{i=1}^n I_{\{|Y_{t_i}-x|<h\}}(\Delta_i Y)^2 \tag{2.5}$$

is a consistent approximation of $\sigma^2(x)L_T(x)$. Loosely speaking, the local time of Y measures how many observations Y_{t_i} are near x. It is defined, for every t, as the following almost sure limit:

$$L_t(x) = \lim_{\varepsilon \to 0} \frac{1}{2\varepsilon} \int_0^t I_{]x-\varepsilon, x+\varepsilon[}(Y_\tau) d\tau, \tag{2.6}$$

see e.g. Revuz and Yor (1998). On the other hand it is well known that $\sum_{i=1}^n (\Delta_i Y)^2 \to_P \int_0^T \sigma_u^2 du$: if we take only those increments for which Y_{t_i} was near x, we get pointwise information about $\sigma(x)$ instead of about the integrated volatility over the full path. This is justified by the fact that (see Revuz and Yor (1998))

$$\lim_{\varepsilon \to 0} \frac{1}{2\varepsilon} \int_0^t I_{\{x-\varepsilon, x+\varepsilon\}}(Y_s) d[Y]_s = \lim_{\varepsilon \to 0} \frac{1}{2\varepsilon} \int_0^t I_{\{x-\varepsilon, x+\varepsilon\}}(Y_s) \sigma^2(Y_s) ds = \sigma^2(x) L_T(x).$$

Indeed, $S_t^n(x)$ is a weighted average of the squared increments $(\Delta_i Y)^2$ of the diffusion Y where the weights are higher when Y_{t_i} happened to be near x.

Using the indicator function is not necessary. For example, Jiang and Knight (1997) showed a result analogous to proposition 2.3 with a continuous kernel replacing the indicator function.

Definition 2.4 *A kernel $K(\cdot)$ is a bounded non negative real function such that $\int_{-\infty}^{+\infty} K(s)$ $ds = 1$.*

An example of smooth kernel is the Gaussian function. The indicator function used by Florens-Zmirou (1993), namely $K(u) = I_{\{|u|<1\}}$, is non smooth. The choice of the kernel function is usually found to be irrelevant in applications. Much more important is the choice of the bandwidth parameter.

Jiang and Knight (1997) show that for proper kernels K

$$\frac{T}{nh} \sum_{i=1}^{n} K\left(\frac{Y_{t_i} - x}{h}\right) \to_{a.s.} L_T(x) \tag{2.7}$$

and

$$\frac{1}{h} \sum_{i=1}^{n} K\left(\frac{Y_{t_i} - x}{h}\right) (\Delta_i Y)^2 \to_{L^2} \sigma^2(x) L_T(x) \tag{2.8}$$

and therefore they validate the following result

Proposition 2.5 *(Jiang and Knight, 1997) Define ${}^K S_t^n(x)$ as:*

$${}^K S_t^n(x) = \frac{n \sum_{i=0}^{n-1} K\left(\frac{Y_{t_i} - x}{h}\right) (\Delta_i Y)^2}{T \sum_{i=1}^{n} K\left(\frac{Y_{t_i} - x}{h}\right)} \tag{2.9}$$

If
- *the assumptions of 2.3 on a, σ and the bandwidth h are satisfied;*
- *the kernel K is positive, continuously differentiable with $\lim_{x\to+\infty} K(x) = \lim_{x\to-\infty} K(x)$ $= 0$;*

then $S_t^n(x)$ converges to $\sigma^2(x)$ in probability for any x visited by Y.

Under the further assumption that $nh^3 \to 0$ both Florens-Zmirou (1993) and Jiang and Knight (1997) prove also the asymptotic normality for $S_t^n(x)$ and ${}^K S_t^n(x)$.

3 Nonparametric estimation of the diffusion coefficient

3.1 The case of finite activity

We now focus on defining a nonparametric estimator of $\sigma^2(\cdot)$ when our model contains also a jump part. Let us begin considering the case in which J is a finite activity jump process, that is it is a compound Poisson process (see Cont and Tankov (2004)). Our model is now

$$X_t = Y_t + J_{1,t}, \tag{3.1}$$

where $Y_t = \int_0^t a_u du + \int_0^t \sigma_u dW_u$ is the diffusion part of X and $J_{1,t} = \sum_{k=1}^{N_t} \gamma_k$, with $J_{1,0^-} = 0$.
A fundamental tool for our aim is to disentangle, from the discrete observations of X, the contributions given by the jumps and those given by the diffusion part. For that we borrow some results from Mancini (2004a).

Theorem 3.1 *(Mancini, 2004a) If $r(u)$ is a real deterministic function such that*

$$\lim_{n\to\infty} r\left(\frac{T}{n}\right) = 0 \quad \text{and} \quad \lim_{n\to\infty} \frac{\frac{T}{n}\log\frac{n}{T}}{r\left(\frac{T}{n}\right)} = 0$$

then for P*-almost all* ω $\exists \bar{n}(\omega)$ *s.t.* $\forall n \geq \bar{n}(\omega)$ *we have*

$$\forall i = 1, ..., n, \quad I_{\{\Delta_i N = 0\}}(\omega) = I_{\{(\Delta_i X)^2 \leq r(\frac{T}{n})\}}(\omega). \tag{3.2}$$

•

This enables us to say that some jumps occurred within $]t_{i-1}, t_i]$ if and only if the squared increment $(\Delta_i X)^2$ is larger than $r(\frac{T}{n})$. As a first consequence, the cumulative sum of the properly small squared increments $(\Delta_i X)^2$ will be equal, for large n, to the sum of the squared increments of the continuous part Y and this intuition has been pursued in Mancini (2004a) to provide an approximation of the quadratic variation of Y, which is in fact the *integrated* volatility of X. Moreover we reach an approximation of the whole jump process J using

$$\hat{\gamma}_{\tau^{(i)}} \doteq \Delta_i X I_{\{(\Delta_i X)^2 > r(\frac{1}{n})\}}.$$

In fact the following results hold.

Corollary 3.2 *(Mancini, 2004a) If we choose r as in Theorem 3.1 then*

$$\sum_{i=1}^{n} (\Delta_i X)^2 I_{\{(\Delta_i X)^2 \leq r(\frac{T}{n})\}} \to_P \int_0^T \sigma^2(X_s)ds$$

Proposition 3.3 *(Mancini (2004b): we remark that part a) of theorem 3.1 of Mancini (2004b) in the present framework with fixed time horizon T still holds. We can reformulate the result as follows)*
Assume that J is a compound Poisson process, let $\delta = \frac{T}{n} \to 0$ and the threshold function r be such that $\frac{\delta \ln\frac{\delta}{\delta}}{r(\delta)} \to_\delta 0$. Then, for any $\varepsilon > 0$,

$$P\left(\bigcup_{i=1}^{n} \{n^k |\hat{\gamma}_{\tau^{(i)}} - \gamma_{\tau^{(i)}} I_{\{\Delta_i N \geq 1\}}| > \varepsilon\}\right) \to_P 0 \quad \forall k \in [0, \frac{1}{2}[.$$

•

Another consequence of theorem 3.1 here is that now we can reach an estimation $\hat{Y}$ of the continuous part Y of (2.1) and we can apply to it, for instance, the estimators proposed in Florens-Zmirou (1993) or in Jiang and Knight (1997) for the *local* volatility $\sigma^2(x)$.
In fact, to get $\sigma^2(x)L_T(x)$ in the spirit of proposition 2.3 of Florens-Zmirou (1993), we will select, among the "small" squared increments $(\Delta_i X)^2$ (having index i such that no jumps occurred within the interval $]t_{i-1}, t_i]$), those for which the continuous part Y of X was close to x at time t_i.

From a practical point of view $\hat{J}_{1,t} \doteq \sum_{i=1}^{t_i \wedge t} \Delta_i X \; I_{\{(\Delta_i X)^2 > r(\frac{1}{n})\}}$ is a consistent approximation of the jump part $J_{1,t}$ of X_t, so that $\hat{Y} = X - \hat{J}_1$ is a consistent estimator of Y which allows us to detect whether Y_t is near to x or not. Note that a continuous semimartingale Y a.s. reaches any x belonging to $[\min_{t\in[0,T]} Y_t, \max_{t\in[0,T]} Y_t]$ during the period $[0,T]$ (Karatzas and Shreve (1988)). Moreover this happens infinitely many times within $[0,T]$. When X contains a finite activity jump process J_1, there are only finite many jumps within $[0,T]$, and thus X a.s. reaches any x belonging to $]\inf_{t\in[0,T]} X_t, \sup x_{t\in[0,T]} X_t[$ anyway infinitely often.
The following theorem is our first result and validates our criterion.

Theorem 3.4 *Let X be a jump-diffusion process as in (3.1).*

- *Let the assumptions of 2.3 for a and σ and for the bandwidth parameter hold*
- *let further the bandwidth parameter be such that $\exists \beta > 1 : nh^\beta \to \infty$*
- *let the threshold function r satisfy the assumptions of theorem 3.1.*

Then for any x visited by X

$$\hat{\sigma}_n^2(x) = \frac{n\sum_{i=1}^n K\left(\frac{X_{t_i} - \hat{J}_{1,t_i} - x}{h}\right)(\Delta_i X)^2 I_{\{(\Delta_i X)^2 \le r(\frac{T}{n})\}}}{T\sum_{i=1}^n K\left(\frac{X_{t_i} - \hat{J}_{1,t_i} - x}{h}\right)} \to_P \sigma^2(x) \qquad (3.3)$$

as soon as the kernel K is continuous and satisfies the conditions in proposition 2.5 or alternatively $K(u) = I_{\{|u|<1\}}$.

Lemma 3.5 *Mancini (2004a) Take $T = 1$.*

$$\sup_i |\Delta_i \sigma.W| \le \sqrt{2}(\bar{\sigma}+1)\sqrt{\frac{1}{n}\ln n} \doteq M(\omega)\sqrt{\frac{1}{n}\ln n},$$

where $\bar{\sigma}(\omega) = \sup_{s\in[0,T]} |\sigma_s(\omega)|$ is a.s. finite since σ is cadlag. •

Proof of Theorem 3.4 We set, without loss of generality, $T = 1$ and $t_i = \frac{T}{n} = \frac{1}{n}$. Set $\hat{Y} = X - \hat{J}_1$ and $L \doteq L_1$.
Step 1. We see that a.s. for n big enough, uniformly in i,

$$\frac{\hat{Y}_{t_i} - Y_{t_i}}{h} \to 0.$$

In fact

$$\frac{X_{t_j} - \hat{J}_{1,t_j} - Y_{t_j}}{h} = \frac{J_{1,t_j} - \hat{J}_{1,t_j}}{h} =$$

$$\frac{1}{h}\left[\sum_{k=1}^{N_{t_j}}\gamma_k-\sum_{i=1}^{j}\left(\int_{t_{i-1}}^{t_i}a_udu+\int_{t_{i-1}}^{t_i}\sigma_udW_u+\sum_{\ell=1}^{\Delta_iN}\gamma_\ell\right)I_{\{(\Delta_iX)^2>r(\frac{1}{n})\}}\right]. \quad (3.4)$$

By theorem 3.1 we have that a.s. for n big enough, uniformly in i, $I_{\{(\Delta_iX)^2>r(\frac{1}{n})\}}=I_{\{\Delta_iN\neq 0\}}$; however, since J_1 has finite activity, for big n, uniformly in i, $I_{\{\Delta_iN\neq 0\}}=I_{\{\Delta_iN=1\}}$. Therefore for big n (3.4) coincides with

$$\frac{1}{h}\left[\sum_{k=1}^{N_{t_j}}\gamma_k-\sum_{i=1}^{j}\left(\int_{t_{i-1}}^{t_i}a_udu+\int_{t_{i-1}}^{t_i}\sigma_udW_u+\gamma_{\tau^{(i)}}\right)I_{\{\Delta_iN=1\}}\right]=$$

$$\sum_{i=1}^{j}\frac{1}{h}\left(\int_{t_{i-1}}^{t_i}a_udu+\int_{t_{i-1}}^{t_i}\sigma_udW_u\right)I_{\{\Delta_iN=1\}}.$$

Set $\bar{a}(\omega)=\sup_{u\in[0,T]}|a_u(\omega)|$: $\bar{a}$ is a.s. finite, since a is a cadlag function of t. Then the last sum tends to zero uniformly in j, since, for big n,

$$\frac{1}{h}\sum_{i=1}^{j}\int_{t_{i-1}}^{t_i}a_uduI_{\{\Delta_iN=1\}}=\sum_{i=1}^{j}\int_{t_{i-1}}^{t_i}a_udu\Delta_iNI_{\{\Delta_iN=1\}}\leq\frac{\bar{a}}{nh}N_1\to 0.$$

Moreover

$$\frac{1}{h}\sum_{i=1}^{j}\int_{t_{i-1}}^{t_i}\sigma_udW_uI_{\{\Delta_iN=1\}}\leq\sup_i\frac{|\Delta_i\sigma.W|}{h}N_T\to 0$$

uniformly on j, since by lemma 3.5 $\sup_i\frac{|\Delta_i\sigma.W|}{h}\leq M(\omega)\sqrt{\frac{1}{nh^2}\ln n}\to_{a.s.}0$.

Consequence of step 1 for the indicator functions kernel. Remember that L is the local time of Y. We have in particular that a.s. for each $\varepsilon>0$ for n big enough $\sup_i|\hat{Y}_{t_i}-Y_{t_i}|\leq\varepsilon$. We now show that

$$Plim\,\frac{n\sum_{i=1}^n I_{\{|\hat{Y}_{t_i}-x|<h\}}(\Delta_iY)^2}{\sum_{i=1}^n I_{\{|\hat{Y}_{t_i}-x|<h\}}}=Plim\,\frac{n\sum_{i=1}^n I_{\{|Y_{t_i}-x|<h\}}(\Delta_iY)^2}{\sum_{i=1}^n I_{\{|Y_{t_i}-x|<h\}}}$$

Let us preliminarily remark that for every $\varepsilon>0$ then $\forall\ i$ if $|\hat{Y}_{t_i}-x|<h$, since $|Y_{t_i}-x|\leq|\hat{Y}_{t_i}-Y_{t_i}|+|\hat{Y}_{t_i}-x|$, then $|Y_{t_i}-x|<h+\varepsilon$. On the other hand, if $|Y_{t_i}-x|<h-\varepsilon$ then $|\hat{Y}_{t_i}-x|\leq|Y_{t_i}-\hat{Y}_{t_i}|+|Y_{t_i}-x|<h$.
That means: asymptotically Y_{t_i} is distant from x less than h if and only if $\hat{Y}_{t_i}$ does. In particular, choosing for instance $\varepsilon=\varepsilon(h)=h^4$, we have $I_{\{|Y_{t_i}-x|<h-h^4\}}\leq I_{\{|\hat{Y}_{t_i}-x|<h\}}$ and $I_{\{|\hat{Y}_{t_i}-x|<h\}}\leq I_{\{|Y_{t_i}-x|<h+h^4\}}$, and thus

$$\frac{n\sum_{i=1}^n I_{\{|\hat{Y}_{t_i}-x|<h\}}(\Delta_iY)^2}{\sum_{i=1}^n I_{\{|\hat{Y}_{t_i}-x|<h\}}}\leq$$

$$\frac{n\sum_{i=1}^n I_{\{|Y_{t_i}-x|<h+h^4\}}(\Delta_iY)^2}{\sum_{i=1}^n I_{\{|Y_{t_i}-x|<h-h^4\}}}=$$

$$\frac{n\sum_{i=1}^n I_{\{|Y_{t_i}-x|<h+h^4\}}(\Delta_iY)^2}{\sum_{i=1}^n I_{\{|Y_{t_i}-x|<h+h^4\}}}\frac{\sum_{i=1}^n I_{\{|Y_{t_i}-x|<h+h^4\}}}{\sum_{i=1}^n I_{\{|Y_{t_i}-x|<h-h^4\}}}:$$

the first factor tends in probability to $\sigma^2(x)$ by proposition 2.3, since $n(h+h^4)\to\infty$. The second factor, by (2.4), has Plim equal to

$$Plim\,\frac{\frac{L(x)}{2(h+h^4)n}}{\frac{L(x)}{2(h-h^4)n}}=1.$$

On the other hand we have

$$\frac{n\sum_{i=1}^n I_{\{|\hat{Y}_{t_i}-x|<h\}}(\Delta_i Y)^2}{\sum_{i=1}^n I_{\{|\hat{Y}_{t_i}-x|<h\}}} \geq$$

$$\frac{n\sum_{i=1}^n I_{\{|Y_{t_i}-x|<h-h^4\}}(\Delta_i Y)^2}{\sum_{i=1}^n I_{\{|Y_{t_i}-x|<h-h^4\}}} \frac{\sum_{i=1}^n I_{\{|Y_{t_i}-x|<h-h^4\}}}{\sum_{i=1}^n I_{\{|Y_{t_i}-x|<h+h^4\}}},$$

where again the first factor tends to $\sigma^2(x)$ and the second one to 1.

Consequence of step 1 for continuous kernels. We show that

$$Plim \frac{\sum_{i=1}^n K\left(\frac{\hat{Y}_{t_i}-x}{h}\right)(\Delta_i Y)^2}{\sum_{i=1}^n K\left(\frac{\hat{Y}_{t_i}-x}{h}\right)\frac{1}{n}} = Plim \frac{\sum_{i=1}^n K\left(\frac{Y_{t_i}-x}{h}\right)(\Delta_i Y)^2}{\sum_{i=1}^n K\left(\frac{Y_{t_i}-x}{h}\right)\frac{1}{n}}$$

In fact, from the results about the approximation of stochastic integrals with respect to a continuous semimartingale (see Metivier (1982)),

$$Plim \sum_{i=1}^n \left| K\left(\frac{\hat{Y}_{t_i}-x}{h}\right) - K\left(\frac{Y_{t_i}-x}{h}\right)\right| (\Delta_i Y)^2 =$$

$$\int_0^1 Plim \left| K\left(\frac{\hat{Y}_{t_i}-x}{h}\right) - K\left(\frac{Y_{t_i}-x}{h}\right)\right| d[Y]_s$$

which is zero, since $Plim \left| K\left(\frac{\hat{Y}_{t_i}-x}{h}\right) - K\left(\frac{Y_{t_i}-x}{h}\right)\right| = 0$ a.s. and, uniformly in i, $\left| K\left(\frac{\hat{Y}_{t_i}-x}{h}\right) - K\left(\frac{Y_{t_i}-x}{h}\right)\right| \leq \bar{K} \doteq \sup_{[0,T]} K(u) < \infty$.
Moreover, analogously,

$$Plim \sum_{i=1}^n \left| K\left(\frac{\hat{Y}_{t_i}-x}{h}\right) - K\left(\frac{Y_{t_i}-x}{h}\right)\right| \frac{1}{n} =$$

$$Plim \sum_{i=1}^n \left| K\left(\frac{\hat{Y}_{t_i}-x}{h}\right) - K\left(\frac{Y_{t_i}-x}{h}\right)\right| \Delta[W]_{t_i} =$$

$$\int_0^1 Plim \left| K\left(\frac{\hat{Y}_{t_i}-x}{h}\right) - K\left(\frac{Y_{t_i}-x}{h}\right)\right| ds = 0.$$

Step 2. For any kernel function K, we have that $\hat{\sigma}_n^2(x)$ has the same limit in probability as

$$\frac{n\sum_{i=1}^n K\left(\frac{Y_{t_i}-x}{h}\right)(\Delta_i Y)^2 I_{\{\Delta_i N=0\}}}{\sum_{i=1}^n K\left(\frac{Y_{t_i}-x}{h}\right)} =$$

$$\frac{\frac{1}{h}\sum_{i=1}^n K\left(\frac{Y_{t_i}-x}{h}\right)(\Delta_i Y)^2}{\frac{1}{h}\sum_{i=1}^n K\left(\frac{Y_{t_i}-x}{h}\right)\frac{1}{n}} - \frac{\sum_{i=1}^n K\left(\frac{Y_{t_i}-x}{h}\right)(\Delta_i Y)^2 I_{\{\Delta_i N\neq 0\}}}{\sum_{i=1}^n K\left(\frac{Y_{t_i}-x}{h}\right)\frac{1}{n}}. \quad (3.5)$$

The Plim of the first term coincides with $\sigma^2(x)$ by Florens-Zmirou (1993) in the case of indicator functions and by Jiang and Knight (1997) in the case of continuous kernel, while the second term is negligible, since

$$\frac{\sup_i (\Delta_i Y)^2}{\frac{1}{n}\ln n} \leq \sup_i \frac{2(\int_{t_{i-1}}^{t_i} a_u du)^2 + 2(\int_{t_{i-1}}^{t_i} \sigma_u dW_u)^2}{\frac{1}{n}\ln n} \leq$$

$$\sup_i \frac{2\frac{\bar{a}^2}{n^2} + 2M^2 \frac{1}{n} \ln n}{\frac{1}{n} \ln n} \leq C(\omega),$$

and therefore

$$\left| \frac{n \sum_{i=1}^n K\left(\frac{Y_{t_i}-x}{h}\right) (\Delta_i Y)^2 I_{\{\Delta_i N \neq 0\}}}{\sum_{i=1}^n K\left(\frac{Y_{t_i}-x}{h}\right)} \right| \leq$$

$$nC(\omega)\, \frac{1}{n} \ln n \, \frac{\sum_{i=1}^n K\left(\frac{Y_{t_i}-x}{h}\right) I_{\{\Delta_i N \neq 0\}}}{\sum_{i=1}^n K\left(\frac{Y_{t_i}-x}{h}\right)} \leq$$

$$C(\omega)\; \ln n\; \bar{K} \frac{\sum_{i=1}^n I_{\{\Delta_i N \neq 0\}}}{\sum_{i=1}^n K\left(\frac{Y_{t_i}-x}{h}\right)} \leq$$

$$C(\omega) \bar{K} \ln n \frac{N_T}{\sum_{i=1}^n K\left(\frac{Y_{t_i}-x}{h}\right)},$$

and the ratio above has the same Plim as

$$Plim\; \ln n \frac{N_T}{2nhL_T(x)} = Plim \frac{n^\alpha}{nh} \frac{\ln n}{n^\alpha},$$

which is zero as soon as $\alpha = 1 - \frac{1}{\beta}$. •

While the estimator (3.3) is consistent, we define another estimator which asymptotically behaves in the same way as the former, but potentially has a much larger relevance in practical application. The idea is just to replace $(\Delta_i X)^2 I_{\{(\Delta_i X)^2 \leq r(\frac{1}{n})\}}$ with m jump-free squared increments in the interval $]t_{i-1}, t_i]$ as follows:

$$\hat{\sigma}^2_{n,m}(x) = \frac{n \sum_{i=1}^n \left[K\left(\frac{X_i - \hat{J}_{1,t_i} - x}{h}\right) \sum_{j=1}^m (\Delta_{i,j} X)^2 I_{\{(\Delta_i \tilde{J}_2)^2 \leq r(\frac{T}{mn})\}} \right]}{T \sum_{i=1}^n K\left(\frac{X_i - \hat{J}_{1,t_i} - x}{h}\right)} \tag{3.6}$$

In fact since the quantity $\sum_{j=1}^m (\Delta_{i,j} X)^2 I_{\{(\Delta_i \tilde{J}_2)^2 \leq r(\frac{T}{mn})\}}$ tends to the integrated volatility in the interval $]t_{i-1}, t_i]$, it is the analogous of the very popular realized volatility within $]t_{i-1}, t_i]$ (see for instance (Andersen et al., 2003)) when extended for the presence of jumps in the driving equation, and theorem 3.4 allows to directly use realized volatility measures for estimating non-parametrically the diffusion coefficient in (2.1). The following proposition validates our $\hat{\sigma}^2_{n,m}(x)$.

Proposition 3.6 *Let $m \to \infty$ be such that $mnh^4 \to 0$. Then, under the same assumptions of theorem 3.4, we have*

$$\underset{|(n,m)| \to \infty}{Plim} \; \hat{\sigma}^2_{n,m}(x) = \sigma^2(x).$$

Proof. Set again $T = 1$. Let us start considering the indicator functions kernel:

$$\hat{\sigma}^2_{n,m}(x) = \frac{n \sum_{i=1}^n I_{\{|X_{t_i} - \hat{J}_{1,t_i} - x| < h\}} \sum_{j=1}^m (\Delta_{i,j} X)^2 I_{\{(\Delta_{i,j} X)^2 \leq r(\frac{1}{mn})\}}}{\sum_{i=1}^n I_{\{|X_{t_i} - \hat{J}_{1,t_i} - x| < h\}}},$$

has Plim coinciding with

$$Plim \frac{n \sum_{i,j} I_{\{|Y_{t_i}-x|<h\}} (\Delta_{i,j}Y)^2 I_{\{\Delta_{i,j}N=0\}}}{\sum_{i=1}^n I_{\{|Y_{t_i}-x|<h\}}} =$$

$$= Plim \frac{n \sum_{i,j} I_{\{|Y_{t_i}-x|<h\}} (\Delta_{i,j}Y)^2}{\sum_{i=1}^n I_{\{|Y_{t_i}-x|<h\}}}, \tag{3.7}$$

since the sum of the terms for which some jumps occurred is negligible.
Note that as soon as $|Y_{t_i} - x| < h$ we have that for any $\varepsilon > 0$, for large n, $|Y_{t_{i,j}} - x| \le |Y_{t_i} - x| + |Y_{t_i} - Y_{t_{i,j}}| < h + \varepsilon$, uniformly in i and j, since Y is uniformly continuous on $[0, T]$. Therefore we can show that

$Plim\ \hat{\sigma}^2_{n,m}(x) \le \sigma^2(x)$. In fact (3.7) is dominated by

$$Plim\ \frac{n \sum_{i,j} I_{\{|Y_{t_{i,j}}-x|<h+\varepsilon\}} (\Delta_{i,j}Y)^2}{\sum_{i=1}^n I_{\{|Y_{t_i}-x|<h\}}} =$$

$$Plim\ \frac{h+\varepsilon}{h} \frac{\frac{\sum_{i,j} I_{\{|Y_{t_{i,j}}-x|<h+\varepsilon\}} (\Delta_{i,j}Y)^2}{2(h+\varepsilon)}}{\frac{\sum_{i=1}^n I_{\{|Y_{t_i}-x|<h\}}}{2nh}} =$$

$$Plim\ (1 + \frac{\varepsilon}{h}) \frac{L_1(x)\sigma^2(x)}{L_1(x)},$$

by Florens-Zmirou (1993), as soon as we choose a sequence $\varepsilon = \varepsilon(h)$ such that $\varepsilon/h \to 0$, so that $nm(h+\varepsilon)^4 \to 0$.
Therefore, for any such sequence of εs, (3.7) is dominated by $Plim\ (1 + \frac{\varepsilon}{h})\sigma^2(x) = \sigma^2(x)$.

On the other hand for any $\varepsilon \in]0, h[$ as soon as $|Y_{t_{i,j}} - x| < h - \varepsilon$ then, for large n, $|Y_{t_i} - x| \le |Y_{t_{i,j}} - x| + |Y_{t_i} - Y_{t_{i,j}}| < h$, uniformly in i and j, and thus

$$Plim\ \frac{n \sum_{i,j} I_{\{|Y_{t_i}-x|<h\}} (\Delta_{i,j}Y)^2}{\sum_{i=1}^n I_{\{|Y_{t_i}-x|<h\}}}$$

$$\ge Plim\ \frac{\frac{\sum_{i,j} I_{\{|Y_{t_{i,j}}-x|<h-\varepsilon\}} (\Delta_{i,j}Y)^2}{2(h-\varepsilon)}}{\frac{\sum_{i=1}^n I_{\{|Y_{t_i}-x|<h\}}}{2nh}} \frac{h-\varepsilon}{h};$$

choosing a sequence $\varepsilon(h) > 0$ such that $\varepsilon(h)/h \to 0$ such a Plim coincides with

$$= Plim\ \frac{L_1(x)\sigma^2(x)}{L_1(x)} \left(1 - \frac{\varepsilon}{h}\right) = \sigma^2(x).$$

We analogously deal with a continuous kernel in place of the indicators. •

3.2 The case of infinite activity

Now we have

$$X = Y + J_1 + \tilde{J}_2, \tag{3.8}$$

with $\tilde{J}_{2s} \doteq \int_0^s \int_{|x|\le 1} x[\mu(dt,dx) - \nu(dx)dt]$ and $J_{1s} = \int_0^s \int_{|x|>1} x\mu(dt,dx) = \sum_{k=1}^{N_t} \gamma_k$. Even in this case we can estimate the continuous part from our discrete observations and thus we can apply to $\hat{Y}$ the nonparametric estimator of Florence-Zmirou or of Jiang and Knight.
Set

$$\hat{J}_t^{(>r)} \doteq \sum_{i=1}^{t_i \wedge t} \Delta_i X I_{\{(\Delta_i X)^2 > r(\frac{T}{n})\}}. \tag{3.9}$$

Now $\hat{J}_t^{(>r)}$ is a consistent approximation not only of $J_{1,t}$ but also of part of $\tilde{J}_2$. To give an idea, for small $\delta = \frac{T}{n}$ we have substantially that, for each i, $I_{\{(\Delta_i X)^2 \le r(\delta)\}} = I_{\{(\Delta_i \tilde{J}_2)^2 \le 4r(\delta), \Delta_i J_1 = 0\}}$, however within $\hat{J}_t^{(>r)}$ the contribution of the terms $I_{\{\Delta_i J_1 = 0\}}$ is negligible. So

$$\sum_{i=1}^{t_i \wedge t} \Delta_i X I_{\{(\Delta_i X)^2 \le r(\frac{T}{n})\}} \approx \sum_{i=1}^{t_i \wedge t} \Delta_i X I_{\{(\Delta_i \tilde{J}_2)^2 \le 4r(\delta), \Delta_i J_1 = 0\}}$$

$$= \sum_{i=1}^{t_i \wedge t} (\Delta_i Y + \Delta_i \tilde{J}_2) I_{\{(\Delta_i \tilde{J}_2)^2 \le 4r(\delta), \Delta_i J_1 = 0\}} \approx \sum_{i=1}^{t_i \wedge t} (\Delta_i Y + \Delta_i \tilde{J}_2) I_{\{(\Delta_i \tilde{J}_2)^2 \le 4r(\delta)\}}$$

$$\approx \sum_{i=1}^{t_i \wedge t} \Delta_i Y + \sum_{i=1}^{t_i \wedge t} \Delta_i \tilde{J}_2 I_{\{(\Delta_i \tilde{J}_2)^2 \le 4r(\delta)\}}$$

and

$$\hat{J}_t^{(>r)} = \sum_{i=1}^{t_i \wedge t} \Delta_i X - \sum_{i=1}^{t_i \wedge t} \Delta_i X I_{\{(\Delta_i X)^2 \le r(\frac{T}{n})\}}$$

$$\approx \sum_{i=1}^{t_i \wedge t} (\Delta_i J_1 + \Delta_i \tilde{J}_2 I_{\{(\Delta_i \tilde{J}_2)^2 > 4r(\frac{T}{n})\}}).$$

Now from (Mancini, 2004a) we have

$$\Delta_i \tilde{J}_2 I_{\{(\Delta_i \tilde{J}_2)^2 \le 4r(\frac{1}{n})\}} = \Delta_i \tilde{J}_{2m} - \Delta_i \tilde{J}_{2c}$$

where

$$\Delta_i \tilde{J}_{2m} \doteq \int_{t_{i-1}}^{t_i} \int_{|x| \le 2\sqrt{r(\delta)}} x \tilde{\mu}(dx, dt), \quad \Delta_i \tilde{J}_{2c} \doteq \int_{t_{i-1}}^{t_i} \int_{2\sqrt{r(\delta)} < |x| \le 1} x \nu(dx) dt,$$

and consequently $\Delta_i \tilde{J}_2 I_{\{(\Delta_i \tilde{J}_2)^2 > 4r(\delta)\}} = \int_{t_{i-1}}^{t_i} \int_{2\sqrt{r(\delta)} < |x| \le 1} x\mu(dx, dt)$. So $\hat{J}_t^{(>r)}$ is given by all the jumps of J_1 and all the jumps of $\tilde{J}_2$ bigger in absolute value than $2\sqrt{r(\delta)}$.
Since for $r(\delta) \to 0$ the jumps bigger in absolute value than $2\sqrt{r(\delta)}$ are all jumps of $\Delta_i \tilde{J}_2$, the tool is to approximate Y with $X - \hat{J}_t^{(>r)} = \sum_{i=1}^{t_i \wedge t} \Delta_i X I_{\{(\Delta_i X)^2 \le r(\frac{T}{n})\}}$. We have the following result.

Theorem 3.7 *Let X be as in (3.8). Under the same assumptions of theorem 3.4. Specify $r(\delta) = \delta^\eta$, $\eta \in]0, 1[$. We then have again that*

$$\hat{\sigma}_n^2(x) = \frac{n \sum_{i=1}^n K\left(\frac{X_{t_{i-1}} - \hat{J}_{t_{i-1}}^{(>r)} - x}{h}\right) (\Delta_i X)^2 I_{\{(\Delta_i X)^2 \le r(\frac{T}{n})\}}}{T \sum_{i=1}^n K\left(\frac{X_{t_{i-1}} - \hat{J}_{t_{i-1}}^{(>r)} - x}{h}\right)} \to_P \sigma^2(x) \tag{3.10}$$

as soon as

$$\frac{\sqrt{r(\delta)}}{h} \to 0$$

the kernel K is continuous and satisfies the conditions in proposition 2.5 and

$$\frac{K\left(\frac{z-x}{h}\right)}{h} \le C_x \text{ for each } x\text{, uniformly on } z$$

or alternatively $K(u) = I_{\{|u|<1\}}$.

Remarks. Note that the kernel $K(u) = e^{-u^2}$ satisfies the assumptions of the previous theorem.
$r(\delta)$ and h satisfy the requested assumptions for $\eta \in]2/3, 1[$.

Proof. Set $T = 1$. Let α be the Blumenthal-Gatoor index of J (see e.g. Cont and Tankov (2004)).
First of all note that on $\{(\Delta_i \tilde{J}_2)^2 \le 4r(\delta), \Delta_i N = 0\}$ we have

$$(\Delta_i X)^2 K\left(\frac{X_{t_{i-1}} - \hat{J}^{(>r)}_{t_{i-1}} - x}{h}\right) \le (\Delta_i \tilde{J}_2 + \Delta_i Y)^2 \bar{K} = O(r(\delta)),$$

so that by lemma 3.5 in Cont and Mancini (2005) (see the appendix) the limit in probability of $\hat{\sigma}^2_n(x)$ coincides with

$$Plim \frac{n\sum_{i=1}^n K\left(\frac{X_{t_{i-1}} - \hat{J}^{(>r)}_{t_{i-1}} - x}{h}\right)(\Delta_i X)^2 I_{\{(\Delta_i \tilde{J}_2)^2 \le 4r(\delta), \Delta_i N=0\}}}{\sum_{i=1}^n K\left(\frac{X_{t_{i-1}} - \hat{J}^{(>r)}_{t_{i-1}} - x}{h}\right)}. \tag{3.11}$$

We now show that we can replace $K\left(\frac{X_{t_{i-1}} - \hat{J}^{(>r)}_{t_{i-1}} - x}{h}\right)$ with $K\left(\frac{Y_{t_{i-1}} - x}{h}\right)$ for each i. By lemma 5.2 and the continuity of K it is sufficient to show that for each t_k

$$\frac{X_{t_{k-1}} - \hat{J}^{(>r)}_{t_{k-1}} - Y_{t_{k-1}}}{h} = \frac{\sum_{i=1}^k (\Delta_i X I_{\{(\Delta_i X)^2 \le r(\delta)\}} - \Delta_i Y)}{h} \to 0.$$

However we can proceed through the following steps.

First of all

$$\frac{\sum_{i=1}^k \Delta_i X I_{\{(\Delta_i X)^2 \le r(\delta)\}}}{h}$$

has the same Plim as

$$\frac{\sum_{i=1}^k (\Delta_i \tilde{J}_2 + \Delta_i Y) I_{\{(\Delta_i X)^2 \le r(\delta), |\Delta_i J| \le 2\sqrt{r(\delta)}, \Delta_i N=0\}}}{h}. \tag{3.12}$$

In fact on $\{(\Delta_i X)^2 \le r(\delta)\}$ we have $|\Delta_i J| - |\Delta_i Y| \le |\Delta_i X| \le \sqrt{r(\delta)}$ and thus, for small δ, $|\Delta_i J| \le 2\sqrt{r(\delta)}$. Therefore for small δ

$$\sum_{i=1}^k \Delta_i X I_{\{(\Delta_i X)^2 \le r(\delta)\}} = \sum_{i=1}^k \Delta_i X I_{\{(\Delta_i X)^2 \le r(\delta), |\Delta_i J| \le 2\sqrt{r(\delta)}\}}.$$

Moreover

$$Plim\frac{\sum_{i=1}^{k}|\Delta_i X| I_{\{(\Delta_i X)^2\le r(\delta),|\Delta_i J|\le 2\sqrt{r(\delta)},\Delta_i N\ne 0\}}}{h}\le\frac{\sqrt{r(\delta)}}{h}N_T\to 0.$$

Secondly let us show that in fact

$$Plim\frac{\sum_{i=1}^{k}(\Delta_i\tilde{J}_2+\Delta_i Y) I_{\{(\Delta_i X)^2> r(\delta),|\Delta_i J|\le 2\sqrt{r(\delta)},\Delta_i N=0\}}}{h}=0,$$

which will imply that (3.12) coincides with

$$Plim\frac{\sum_{i=1}^{k}(\Delta_i\tilde{J}_2+\Delta_i Y) I_{\{|\Delta_i J|\le 2\sqrt{r(\delta)},\Delta_i N=0\}}}{h}. \tag{3.13}$$

In fact we have

$$\{(\Delta_i\tilde{J}_2)^2\le 4r(\delta),\Delta_i N=0,(\Delta_i X)^2>r(\delta)\}\subset$$

$$\{|\Delta_i\tilde{J}_2|\le 2\sqrt{r(\delta)},|\Delta_i Y|+|\Delta_i\tilde{J}_2|>\sqrt{r(\delta)}\}\subset$$

$$\{|\Delta_i\tilde{J}_2|\le 2\sqrt{r(\delta)},|\Delta_i Y|>\sqrt{r(\delta)}/2\}\cup\{|\Delta_i\tilde{J}_2|\le 2\sqrt{r(\delta)},|\Delta_i\tilde{J}_2|>\sqrt{r(\delta)}/2\}: \tag{3.14}$$

and, passing to a subsequence, both sets are empty, a.s., uniformly in i, for small δ since on one hand $\{|\Delta_i Y|>\sqrt{r(\delta)}/2\}=\{\frac{|\Delta_i Y|}{\sqrt{\delta\ln\frac{1}{\delta}}}>\frac{1}{2}\frac{\sqrt{r(\delta)}}{\sqrt{\delta\ln\frac{1}{\delta}}}\}$ and $\frac{\sqrt{r(\delta)}}{\sqrt{\delta\ln\frac{1}{\delta}}}\to\infty$ by assumption, while $\frac{|\Delta_i Y|}{\sqrt{\delta\ln\frac{1}{\delta}}}\le\bar{a}+M<\infty$.

On the other hand $\{|\Delta_i\tilde{J}_2|\le 2\sqrt{r(\delta)},|\Delta_i\tilde{J}_2|>\sqrt{r(\delta)}/2\}\subset\{(\Delta_i\tilde{J}_2)^2>4r(\delta)\}$. However on $\{(\Delta_i\tilde{J}_2)^2>4r(\delta)\}$ we have that $\Delta_i\tilde{J}_2=\int_{t_{i-1}}^{t_i}\int_{|x|>2\sqrt{r(\delta)}}x\mu(dt,dx)=\Delta_i M+\Delta_i C$, where

$$\Delta_i M\doteq\int_{t_{i-1}}^{t_i}\int_{|x|>2\sqrt{r(\delta)}}x\tilde{\mu}(dt,dx),\quad \Delta_i C\doteq\int_{t_{i-1}}^{t_i}\int_{|x|>2\sqrt{r(\delta)}}x\nu(dt,dx),$$

so

$$\{(\Delta_i\tilde{J}_2)^2>4r(\delta)\}\subset\{2(\Delta_i M)^2+2(\Delta_i C)^2>4r(\delta)\}\subset$$
$$\{(\Delta_i M)^2>r(\delta)\}\cup\{(\Delta_i C)^2>r(\delta)\}.$$

Now, $\{(\Delta_i C)^2>r(\delta)\}=\{\delta r(\delta)^{-\frac{\alpha}{2}}>1\}=\emptyset$ for sufficiently small δ, uniformly in i, since $\frac{d}{r(\delta)^{\frac{\alpha}{2}}}=\left(\frac{\delta}{r(\delta)}\right)^{\frac{\alpha}{2}}\delta^{1-\frac{\alpha}{2}}\to 0$.

Moreover $\{(\Delta_i M)^2>r(\delta)\}$ in probability coincides with $\{(M_\delta)^2>r(\delta)\}$ which is a subset of

$$\{\sup_{s\le\delta}M_s^2>r(\delta)\}=\{\sup_{s\le\delta}\frac{|M_s|}{\delta^{1/2-\gamma}(1-r(\delta)^{1-\alpha/2})}\delta^{1/2-\gamma}(1-r(\delta)^{1-\alpha/2})>\sqrt{r(\delta)}\}, \tag{3.15}$$

where we take $\gamma>0$ such that $1-\eta-2\gamma>0$. Since by the Doob inequality $\sup_{s\le\delta}\frac{|M_s|}{\delta^{1/2-\gamma}(1-r(\delta)^{1-\alpha/2})}\to_P 0$, passing to a subsequence we have the a.s. convergence to zero, and in particular $\sup_{s\le\delta}\frac{|M_s|}{\delta^{1/2-\gamma}(1-r(\delta)^{1-\alpha/2})}<1$ a.s. for small δ, so (3.15) is subset of

$$\{\delta^{1/2-\gamma}(1-r(\delta)^{1-\alpha/2})>\sqrt{r(\delta)}\}=\{\delta^{1/2-\gamma}r(\delta)^{-1/2}(1-r(\delta)^{1-\alpha/2})>1\}$$

which is empty for small δ.

Now we deal with (3.13), which like as before coincides with

$$Plim\frac{\sum_{i=1}^{k}(\Delta_i\tilde{J}_2+\Delta_i Y)I_{\{(\Delta_i\tilde{J}_2)^2\le 4r(\delta)\}}}{h},$$

and also with

$$Plim\frac{\sum_{i=1}^{k}\Delta_i Y+\sum_{i=1}^{k}\Delta_i\tilde{J}_2 I_{\{(\Delta_i\tilde{J}_2)^2\le 4r(\delta)\}}}{h} \tag{3.16}$$

Now (3.16) can be written as

$$Plim\frac{\sum_{i=1}^{k}\Delta_i Y\left(1+\frac{\Delta_i\tilde{J}_{2c}}{\Delta_i Y}+\frac{\Delta_i\tilde{J}_{2m}}{\Delta_i Y}\right)}{h}=Plim\frac{\sum_{i=1}^{k}\Delta_i Y}{h},$$

where last equality holds since uniformly with respect to i we have on one hand, for any $\varepsilon>0$, $E\left[\left|\frac{\Delta_i\tilde{J}_{2m}}{\Delta_i Y}\right|\right]\le E\left[\frac{|\Delta_i\tilde{J}_{2m}|}{(1-\varepsilon)\sqrt{\delta\ln|\ln\delta|}}\right]$ which, by the Doob inequality is dominated by

$$\frac{\sqrt{\delta}\sqrt{\int_{|x|\le\sqrt{r(\delta)}}x^2\nu(dx)}}{(1-\varepsilon)\sqrt{\delta\ln|\ln\delta|}}=O(\sqrt{\frac{r(\delta)^{1-\alpha/2}}{\ln|\ln\delta|}}\to 0).$$

On the other hand $E\left[\left|\frac{\Delta_i\tilde{J}_{2c}}{\Delta_i Y}\right|\right]\le\frac{\delta\int_{|x|\in]2\sqrt{r(\delta)},1]}x\nu(dx)}{(1-\varepsilon)\sqrt{\delta\ln|\ln\delta|}}\to 0.$

Now (3.11) is reduced to

$$Plim\frac{\sum_{i=1}^{n}K\left(\frac{Y_{t_i}-x}{h}\right)(\Delta_i Y+\Delta_i\tilde{J}_2)^2 I_{\{(\Delta_i\tilde{J}_2)^2\le 4r(\delta)\}}}{h\frac{\sum_{i=1}^{n}K\left(\frac{Y_{t_i}-x}{h}\right)}{nh}}=$$

$$=Plim\frac{\frac{1}{h}\sum_{i=1}^{n}K\left(\frac{Y_{t_i}-x}{h}\right)(\Delta_i Y)^2}{L_1^Y}-Plim\frac{\sum_{i=1}^{n}\frac{K\left(\frac{Y_{t_i}-x}{h}\right)}{h}(\Delta_i Y)^2 I_{\{(\Delta_i\tilde{J}_2)^2>4r(\delta)\}}}{L_1^Y}$$

$$+Plim\frac{\sum_{i=1}^{n}\frac{K\left(\frac{Y_{t_i}-x}{h}\right)}{h}(\Delta_i\tilde{J}_2)^2 I_{\{(\Delta_i\tilde{J}_2)^2\le 4r(\delta)\}}}{L_1^Y}$$

$$+2Plim\frac{\sum_{i=1}^{n}\frac{K\left(\frac{Y_{t_i}-x}{h}\right)}{h}\Delta_i Y\Delta_i\tilde{J}_2 I_{\{(\Delta_i\tilde{J}_2)^2\le 4r(\delta)\}}}{L_1^Y}:$$

the first term tends to $\sigma^2(x)$ by Florens-Zmirou (1993), while each one of the other terms tends to zero in probability. The second one vanishes like as before. Define

$$K_s=K\left(\frac{Y_s-x}{h}\right).$$

Then the numerator of the third term has Plim coinciding with

$$Plim\int_0^1\frac{K_s}{h}\,d[\tilde{J}_{2,m}]_s=Plim\int_0^1\int_{|z|\le 2\sqrt{r(\delta)}}\frac{K_s}{h}\,z^2\mu(dz),$$

whose expectation is $Plim \int_0^1 \int_{|z|\le 2\sqrt{r(\delta)}} \frac{K_s}{h}\, z^2\nu(dz) = 0$.
Finally the numerator of the last term has Plim coinciding with

$$2Plim \int_0^1 \frac{K_s}{h}\, d[Y, \tilde{J}_{2,m}]_s = 0,$$

since $[Y, \tilde{J}_{2,m}]_t =< Y^c, \tilde{J}^c_{2m} >_t + \sum_{s\le t} \Delta Y_s \Delta \tilde{J}_{2m,s} = 0$. •

4 Conclusions

In this preliminary version paper, we introduce nonparametric estimation of the local volatility in univariate processes with stochastic volatility and jumps in the case of Lévy jump part, extending the preceding literature on nonparametric estimation of continuous diffusions and of diffusions plus finite activity jump processes. We build on the results of Mancini (2004a) who shows how to identify jump times and sizes of a discretely observed process. Our results can be useful in the framework of interest rate modelling and option pricing, and further research on this topic is under development.

5 Appendix

Lemma 5.1 (Lemma 3.5 in Cont and Mancini (2005)) *If* $r(\delta) \to_\delta 0 \sup_{i=1..n} |a_{ni}| 0(r(\delta))$ *then*

$$Plim \sum_i |a_{ni}| I_{\{(\Delta_i X)^2 \le r(\frac{1}{n})\}} = Plim \sum_i |a_{ni}| I_{\{(\Delta_i \tilde{J}_2)^2 \le 4r(\delta), \Delta_i N = 0\}}.$$

Lemma 5.2 *Let $K^{(\delta)}$ be a uniformly bounded sequence of semimartingales, that is $|K_t^{(\delta)}| \le \bar{K}$ a.s. for each δ and t. Let $K_t^{(\delta)}$ converge, as $\delta \to 0$, to a semimartingale K, a.s. for every t.*
a) If Z is a semimartingale then

$$\sum_{i=1}^n K^{(\delta)}_{t_{i-1}} (\Delta_i Z)^2 \to_P \int_0^T K_s d[Z]_s.$$

b) If $Z^{(\delta)}$ is a sequence of semimartingales for which $E[(\Delta_i Z^{(\delta)})^2] = \delta u_\delta$ with $u_\delta \to 0$ then

$$\sum_{i=1}^n K^{(\delta)}_{t_{i-1}} (\Delta_i Z^{(\delta)})^2 \to_{L^1} 0.$$

Proof. a) Note that $\sum_{i=1}^n K_{t_{i-1}} (\Delta_i Z)^2 \to_P \int_0^T K_s d[Z]_s$.
Following now the lines in Metivier (1982), p.177, we have

$$\sum_{i=1}^n (K^{(\delta)}_{t_{i-1}} - K_{t_{i-1}})(\Delta_i Z)^2 = \sum_{i=1}^n (K^{(\delta)}_{t_{i-1}} - K_{t_{i-1}}) \Delta_i (Z^2) - 2 \sum_{i=1}^n (K^{(\delta)}_{t_{i-1}} - K_{t_{i-1}}) Z_{t_{i-1}} \Delta_i Z =$$

$$= \int_0^T \phi_s^{(\delta)} d(Z_s^2) - 2\int_0^T \psi_s^{(\delta)} dZ_s$$

where $\phi_s^{(\delta)} = \sum_{i=1}^n (K_{t_{i-1}}^{(\delta)} - K_{t_{i-1}}) I_{]t_{i-1},t_i]}(s)$ is adapted and left continuous, bounded by $2\bar{K}$ and tends to 0 a.s. for each t; analogously $\psi_s^{(\delta)} = \sum_{i=1}^n (K_{t_{i-1}}^{(\delta)} - K_{t_{i-1}}) Z_{t_{i-1}} I_{]t_{i-1},t_i]}(s)$ is adapted and left continuous bounded a.s. by $2\bar{K} \sup_{t\in[0,T]} |Z_t| < \infty$ and tends to 0 a.s. for each t. Therefore, by the "dominated convergence" theorem (theorem 24.2 in Metivier (1982)) the integrals above converge in probability, as $\delta \to 0$, to 0.

b)

$$E[\sum_{i=1}^n K_{t_{i-1}}^{(\delta)} (\Delta_i Z^{(\delta)})^2] \le \bar{K} n E[(\Delta_i Z^{(\delta)})^2] = 2\bar{K} n \delta u_\delta \to 0.$$

•

References

Ait-Sahalia, Y. (2004). Disentangling diffusion from jumps. *Journal of Financial Economics 74*, 487–528.

Aït-Sahalia, Y. and J. Jacod (2005). Discretely sampled levy processes: estimation. Working Paper.

Andersen, T., L. Benzoni, and J. Lund (2002). Towards an empirical foundation for continous-time equity return models. *Journal of Finance 57*, 1239–1284.

Andersen, T., T. Bollerslev, F. Diebold, and P. Labys (2003). Modeling and forecasting realized volatility. *Econometrica 71*, 579–625.

Bakshi, G., C. Cao, and Z. Chen (1997). Empirical performance of alternative option pricing models. *Journal of Finance 52*, 2003–2049.

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

Bandi, F. and P. Phillips (2003). Fully nonparametric estimation of scalar diffusion models. *Econometrica 71*(1), 241–283.

Barndorff-Nielsen, O. E. and N. Shephard (2004a). Econometrics of testing for jumps in financial economics using bipower variation. *Journal of Financial Econometrics*. Forthcoming.

Barndorff-Nielsen, O. E. and N. Shephard (2004b). Power and bipower variation with stochastic volatility and jumps. *Journal of Financial Econometrics 2*, 1–48.

Bates, D. (2000). Post-'87 crash fears in the S&P 500 futures option market. *Journal of Econometrics 94*, 181–238.

Carr, P., H. Geman, D. Madan, and M. Yor (2002). The fine structure of asset returns: an empirical investigation. *Journal of Business 75*(2).

Cont, R. and C. Mancini (2005). Detecting the presence of a diffusion and the nature of jumps in asset prices. Working paper.

Cont, R. and P. Tankov (2004). *Financial Modelling with Jump Processes*. Chapman & Hall - CRC.

Das, S. (2002). The surprise element: Jumps in interest rates. *Journal of Econometrics 106*, 27–65.

Eberlein, E. and S. Raible (1999). Term structure models driven by general Lévy processes. *Mathematical Finance 9*, 31–53.

Eraker, B., M. Johannes, and N. Polson (2002). The impact of jumps in volatility and returns. *Journal of Finance*. Forthcoming.

Florens-Zmirou, D. (1993). On estimating the diffusion coefficient from discrete observations. *Journal of Applied Probability 30*, 790–804.

Ikeda, N. and S. Watanabe (1981). *Stochastic differential equations and diffusion processes*. North Holland.

Jiang, G. and J. Knight (1997). A nonparametric approach to the estimation of diffusion processes, with an application to a short-term interest rate model. *Econometric Theory 13*, 615–645.

Jiang, G. and R. Oomen (2004). Estimating latent variables and jump diffusion models using high frequency data. Working Paper.

Johannes, M. (2004). The statistical and economic role of jumps in continuous-time interest rate models. *Journal of Finance 59*, 227–260.

Karatzas, I. and E. Shreve (1988). *Brownian Motion and Stochastic Calculus*. Springer-Verlag.

Madan, D. (1999). Purely discontinuous asset price processes. Working Paper.

Mancini, C. (2004a). Estimating the integrated volatility in stochastic volatility models with Lévy type jumps. Working Paper, Università di Firenze.

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

Metivier, M. (1982). *Semimartingales: a course on stochastic processes*. De Gruyter.

Pan, J. (2002). The jump-risk premia implicit in options: Evidence from an integrated time series study. *Journal of Financial Economics 63*, 3–50.

Piazzesi, M. (2005). Bond yields and the federal reserve. *Journal of Political Economy 113*(2).

Protter, P. (1990). *Stochastic Integration and Differential Equations*. Springer-Verlag.

Renò, R. (2004). Nonparametric estimation of the diffusion coefficient via Fourier analysis, with an application to short interest rates. Working Paper, Università di Siena.

Revuz, D. and M. Yor (1998). *Continuous Martingales and Brownian Motion*. Springer-Verlag.

Scott, D. (1992). *Multivariate Density Estimation*. John Wiley & Sons.

Stanton, R. (1997). A nonparametric model of term structure dynamics and the market price of interest rate risk. *Journal of Finance 52*, 1973–2002.

Woerner, J. (2003). Estimation of integrated volatility in stochastic volatility models. *Applied Stochastic Models in Business and Industry*. Forthcoming.