数理解析研究所講究録
第 1620 巻 2009 年 1-10

High frequency statistics with irregularly spaced observations

Takaki Hayashi,[*] Jean Jacod,[†] and Nakahiro Yoshida [‡]

October 27, 2008

The approximation of the quadratic variation by the "realized" quadratic variation plays a very important role in practice, especially for financial data, since such data are necessarily reported at discrete times. Other quantities of interest, based on discrete observations as well, have been considered in many recent papers: these quantities are sums of functions of the successive increments of the process, usually suitable powers or absolute powers of those increments. They are used to estimates some characteristics of the jumps of the observed process, or the volatility of the continuous part, or for various testing problems about jumps for example.

However, if the behavior of the realized quadratic variation and of other similar functionals is well known when the observations come in regularly, this is no longer the case when the observation times are irregularly spaced, and even worse, when they are random. Relatively few papers are so far available in that case: see [14], [1] and [15] for deterministic observation times, and [4] for some special random times like hitting times, and [7], [8], [9], [6] and [2] when the process is multidimensional and the observation times exhibit some sort of relatively restricted randomness, or are random but independent of the observed process. In the five last papers, the situation is quite complex because the various components are observed at different times. All those papers are concerned with a continuous underlying process. One may also quote related works dealing with estimation of various parameters with random sampling, like [5] for diffusions and [3] for Markov processes.

In this short presentation – without proof – we consider a restricted problem, for the following reasons:

1 - The underlying process is 1-dimensional. This avoid considering two or more components which are observed at different times.

2 - The test functions which we use are absolute power functions. Those are the most

*Keio University, Graduate School of Business Administration, 4-1-1 Hiyoshi, Yokohama 223-8526, Japan

†Institut de mathématiques de Jussieu, Université P. et M. Curie (Paris-6), 4 place Jussieu, 75252 Paris, France (CNRS UMR 7586)

‡University of Tokyo, Graduate School of Mathematics, 3-8-9 Komaba, Tokyo 153-8914, Japan

useful ones in practice, but an extension to more general test functions would be (relatively) easy to do.

3 - The underlying process is *continuous.*

On the other hand, we do *not* restrict ourselves to the realized quadratic variation, although the integrated volatility is of course the archetypal and also the most interesting quantity to estimate.

1 - The observed process.

The underlying process X is a 1-dimensional continuous semimartingale X on a filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, \mathbb{P})$, which is of Itô type, that is of the form

$$X_t = X_0 + \int_0^t b_s ds + \int_0^t \sigma_s dW_s, \tag{1}$$

where W is a Wiener process and b and σ are progressively measurable processes. Our aim is to estimate the following "integrated" quantities:

$$V(p)_t = m_p \int_0^t |\sigma_s|^p ds \tag{2}$$

(here m_p is the pth absolute moment of the standard normal law $\mathcal{N}(0,1)$), when $p \geq 2$, together with "feasible" estimators for the variance or conditional variance of these estimators, so as to be able to construct confidence intervals for example.

Apart from being as in (1), we make two different assumptions on X, depending on the results we want to prove:

Assumption (A): We have (1), and the process b is locally bounded, and the process σ is càdlàg (= right continuous with left limits). □

Assumption (B): We have (1), and the process σ is also a (possibly discontinuous) Itô semimartingale, which can be written as

$$\sigma_t = \sigma_0 + \int_0^t \widetilde{b}_s ds + \int_0^t \widetilde{\sigma}_s dW_s + M_t + \sum_{s\leq t} \Delta\sigma_s \, 1_{\{|\Delta\sigma_s|>1\}},$$

where M is a local martingale with $|\Delta M_t| \leq 1$, orthogonal to W, and its predictable quadratic covariation process is $\langle M, M\rangle_t = \int_0^t a'_s ds$, and the predictable compensator of $\sum_{s\leq t} 1_{\{|\Delta\sigma_s|>1\}}$ is $\int_0^t a_s ds$, and the processes $\widetilde{b}$, a and a' are locally bounded, and the processes $\widetilde{\sigma}$ and b are left continuous with right limits. □

Note that M may have jumps, and it may also have a non-vanishing continuous martingale part, which then must be a stochastic integral with respect to another Brownian motion independent of W.

2 - The sampling scheme.

At stage n the process X is observed along a strictly increasing sequence of – possibly random – finite times $T(n,i), i \geq 0$, starting at $T(n,0) = 0$, and we use the notation

$$\left.\begin{array}{rclrcl} \Delta(n,i) &=& T(n,i) - T(n,i-1), & I(n,i) &=& (T(n,i-1), T(n,i)], \\ N_t^n &=& \inf(i : T(n,i) > t) - 1, & \pi_t^n &=& \sup_{i=1,\cdots,N_t^n+1} \Delta(n,i) \end{array}\right\}$$

(π_t^n is the "mesh" up to time t, by convention $\inf(\emptyset) = \infty$ and $\sup(\emptyset) = 0$). Also, for any process Y we write

$$\Delta_i^n Y = Y_{T(n,i)} - Y_{T(n,i-1)}.$$

First, we always assume the following minimal requirements:

$$\left.\begin{array}{lcl} n \geq 1 & \Rightarrow & T(n,i) \longrightarrow \infty \quad \mathbb{P}-\text{a.s., as } i \to \infty \\ t \geq 0 & \Rightarrow & \pi_t^n \xrightarrow{\mathbb{P}} 0 \quad \text{as } n \to \infty. \end{array}\right\} \tag{3}$$

Next, we have a structural assumption:

Assumption (C): There is a sub-filtration $(\mathcal{F}_t^0)_{t\geq 0}$ of $(\mathcal{F}_t)_{t\geq 0}$, with respect to which W and b and σ are adapted, and such that any $(\mathcal{F}_t^0)$ martingale is also an $(\mathcal{F}_t)$-martingale, and such that for all $n, i \geq 1$ the variable $T(n,i)$ is an $(\mathcal{F}_t)$-stopping time which, conditionally on $\mathcal{F}_{T(n,i-1)}$, is independent of the σ-field $\mathcal{F}^0 = \vee_{t>0}\mathcal{F}_t^0$.

This assumption is satisfied when the $T(n,i)$'s are non-random (*deterministic schemes*), and when the $T(n,i)$'s are independent of the processes (W, X, b, σ) (*independent schemes*), but it includes many other cases as well. However it excludes some *a priori* interesting situations: when $\mathcal{F}_t^0 = \mathcal{F}_t$, then (C) amounts to saying that those stopping times are "strongly predictable" in the sense that $T(n,i)$ is $\mathcal{F}_{T(n,i-1)}$-measurable, and this is quite restrictive; for example it excludes the case where the $T(n,i)$'s are the successive hitting times of a spatial grid by X, a case considered in [4] under some restrictive assumptions on X.

Apart from (3) and (C), the sampling scheme should also be not too wildly scattered, asymptotically speaking. Also, its meshes π_t^n should converge to 0 at some deterministic rate, the same for all t. This rate is expressed through a sequence $r_n \to \infty$ of positive – non random – numbers. The assumptions below all involve this sequence r_n, in an implicit way.

Before giving the assumptions, and for any $q \geq 0$, we introduce the processes

$$A(q)_t^n = r_n^{q-1} \sum_{i=1}^{N_t^n} \Delta(n,i)^q.$$

The normalization r_n^{q-1} is motivated by the regular schemes $T(n,i) = i\Delta_n$, for which $A(q)_t^n = \Delta_n[t/\Delta_n]$ (with the choice $r_n = 1/\Delta_n$) converges towards t for any $q \geq 0$. Note also that $A(0)^n_t = N^n_t/r_n$, and since $t-\pi^n_t \leq A(1)^n_t \leq t$ we deduce that

$$A(1)^n_t \stackrel{\mathbb{P}}{\longrightarrow} t$$

as soon as (3) holds. Then, with $q \geq 0$, we set:

Assumption (D(q)): We have (3) and (C), and there is a (necessarily nonnegative) $(\mathcal{F}^0_t)$-optional process $a(q)$, such that for all t we have

$$A(q)^n_t \stackrel{\mathbb{P}}{\longrightarrow} \int_0^t a(q)_s\, ds. \tag{4}$$

Note that (D(q)) for some sequence r_n implies (D(q)) for any other sequence r'_n such that $r'_n/r_n \to \alpha \in [0,\infty)$, and the new limit in (4) is then $\alpha^{q-1}a(q)$ if $\alpha > 0$ or if $q \geq 1$, and in particular vanishes when $r'_n/r_n \to 0$: the forthcoming theorems which explicitly involve r_n are true but "empty" when the limit in (4) vanishes identically. Regular sampling schemes with lag Δ_n satisfy (D(q)) for all $q \geq 0$, with $r_n = 1/\Delta_n$ and $a(q)_t = 1$.

Let us state some important connections between these assumptions: If $0 \leq q < p < q'$ we have for all $0 \leq s < t$ we have

$$A(p)^n_t - A(p)^n_s \;\leq\; \Big(A(q)^n_t - A(q)^n_s\Big)^{\frac{q'-p}{q'-q}} \Big(A(q')^n_t - A(q')^n_s\Big)^{\frac{p-q}{q'-q}}.$$

Then if (D(q)) holds for some $q \neq 1$ and if p is strictly between 1 and q, from any subsequence one may extract a further subsequence which satisfies (D(p)), and we have versions of $a(q)$ and $a(p)$ satisfying $a(p)_t \leq a(q)_t^{(p-1)/(q-1)}$.

Another interesting property is that (4) for all t implies

$$r_n^{q-1} \sum_{i=1}^{N^n_t} H_{T(n,i)} \Delta(n,i)^q \;\stackrel{\text{u.c.p.}}{\longrightarrow}\; \int_0^t H_s a(q)_s ds$$

as soon as H is càdlàg. However in some cases this convergence should hold at a rate faster than $1/\sqrt{r_n}$, and we express this in the following assumption.

Assumption (D'(q)): We have (D(q)), and further for all $t \geq 0$ and all càdlàg $(\mathcal{F}^0_t)$-adapted processes H we have

$$\sqrt{r_n}\Big(r_n^{q-1} \sum_{i=1}^{N^n_t} H_{T(n,i)} \Delta(n,i)^q - \int_0^t H_s a(q)_s ds\Big) \;\stackrel{\text{u.c.p.}}{\longrightarrow}\; 0.$$

This assumption is indeed very strong: for example if we have an independent scheme for which the $\Delta(n,i)$'s are i.i.d. when i varies, then this assumption is *never* satisfied unless $r_n\Delta(n,1)$ converges in law to a (necessarily non 0) constant, as $n \to \infty$.

For a deterministic scheme, (D(q)) may or may not be satisfied, but there is no simple criterion to ensure that it holds. For a random scheme, it may be useful to describe conditions on the laws or on the conditional laws of the lags $\Delta(n,i)$ which ensure (D(q)). For each n and each $q \geq 0$ we choose an $(\mathcal{F}_t)$-optional $(0,\infty]$-valued process $G(q)^n$ such that

$$G(q)^n_{T(n,i-1)} \;=\; r_n^q\, \mathbb{E}(\Delta(n,i)^q \mid \mathcal{F}_{T(n,i-1)}). \tag{5}$$

This specifies $G(q)^n_t$ only at the times $t = T(n,i)$, so there are many such processes $G(q)^n$. A simple choice consists in taking $G(q)^n_t$ to be equal to the right side of (5) when $T(n,i-1) \leq t < T(n,i)$ (a piecewise constant process). But other choices are possible, and perhaps more appropriate in view of the forthcoming assumption. We can obviously take $G(0)^n_t = 1$, and by Hölder's inequality, we can and will choose processes $G(q)^n$ which satisfy

$$0 \leq p \leq q \quad \Rightarrow \quad G(p)^n \;\leq\; (G(q)^n)^{p/q}.$$

Then we set, with $q > 1$:

Assumption (E(q)): We have (C) and (3), and for each $p \in [0,q]$ there is a càdlàg process $G(p)$, adapted to $(\mathcal{F}^0_t)$, and further $G(1)$ and $G(1)_-$ do not vanish, such that for an appropriate choice of $G(p)^n$ we have

$$G(p)^n \;\stackrel{\text{u.c.p.}}{\longrightarrow}\; G(p). \tag{6}$$

Note that (E(q)) for some $q > 1$ yields (D(p)) for all $p \in [0,q)$, with

$$a(p)_t \;=\; \frac{G(p)_t}{G(1)_t},$$

and in particular

$$\frac{1}{r_n}\, N^n_t \;\stackrel{\mathbb{P}}{\longrightarrow}\; \int_0^t \frac{1}{G(1)_s}\, ds.$$

Sometimes we also need a rate of convergence in (6), which is expressed as follows:

Assumption (E'(q)): We have (E(q)), and for each $p \in [0,q]$ we have $\sqrt{r_n}\,(G(p)^n - G(p)) \stackrel{\text{u.c.p.}}{\longrightarrow} 0$.

Finally as an example we introduce a kind of sampling schemes which are somehow restrictive but accommodates *many practical applications*, and are called *mixed renewal schemes*. These schemes are constructed as follows: we consider a filtration $(\mathcal{F}^0_t)$ as in (C), and a double sequence $(\varepsilon(n,i) : i,n \geq 1)$ of i.i.d. positive variables on $(\Omega,\mathcal{F},\mathbb{P})$, independent of $\mathcal{F}^0$, with moments

$$m'_q \;=\; \mathbb{E}(\varepsilon(n,i)^q).$$

We may have $m'_q = \infty$ for $q > 1$, but we assume that $m'_1 < \infty$. We consider a sequence v^n of positive $(\mathcal{F}_t)$-adapted processes, and we define $T(n,i)$ by induction on i as follows:

$$T(n,0) \ = \ 0, \qquad T(n,i+1) \ = \ T(n,i) + \frac{1}{r_n}\, v^n_{T(n,i)}\varepsilon(n,i+1).$$

Then $(\mathcal{F}_t)$ is any filtration containing $(\mathcal{F}^0_t)$ and such that each $T(n,i)$ is a stopping time. In this situation, a natural choice for the processes $G(q)^n$ of (5) is $G(q)^n_t = m'_q(v^n_t)^q$.

Any mixed renewal scheme satisfies (C), and as soon as $v^n \stackrel{\text{u.c.p.}}{\longrightarrow} v$ (convergence in probability, locally uniformly in time), where v is an $(\mathcal{F}^0_t)$-adapted càdlàg process v such that both v and v_- do not vanish, then we have (3) and (E(q)) and (D(q)) for any $q \geq 0$ for which $m'_q < \infty$, and in this case

$$G(q)_t = m'_q(v_t)^q, \qquad a(q)_t = \frac{m'_q}{m'_1}\,(v_t)^{q-1}. \tag{7}$$

And if furthermore $\sqrt{\Delta_n}\,(v_n - v) \stackrel{\text{u.c.p.}}{\longrightarrow} 0$, we also have (E'($q$)).

3 - The estimators.

A natural estimator for $V(p)_t$ in (2) is, at stage n, the variable

$$V^n(p)_t \ = \ \sum_{i=1}^{N^n_t} \Delta(n,i)^{1-p/2}\,|\Delta^n_i X|^p. \tag{8}$$

The upper limit of the sum is such that $V^n(p)_t$ involves exactly the observations actually occurring up to time t only. These estimators are natural because we have:

Theorem 1 *Under (A) and (C) and (3), we have*

$$V^n(p)_t \ \stackrel{\text{u.c.p.}}{\longrightarrow} \ V(p)_t \ = \ m_p \int_0^t |\sigma_s|^p\,ds. \tag{9}$$

This is of course well known for a regular sampling scheme $T(n,i) = i\Delta_n$ for some time lag Δ_n going to 0, because then $V^n(p)_t = \Delta_n^{1-p/2}\sum_{i=1}^{N^n_t}|\Delta^n_i X|^p$. Note that when $p = 2$ this holds under no condition at all on b and σ, other than the fact that (1) makes sense; end this is a very hold result. Note also that the factor $\Delta(n,i)\Delta(n,i+1)^{-p/2}$ which goes out of the sum in the regular case does not in general, and the idea to place this factor inside the sum is due to [1].

As we will see, the behavior of $V^n(p)$ is not enough for our purposes, and we need to establish the convergence in probability for more general processes. For $p > 0$ and $q \geq 0$ we set

$$V^n(p,q)_t \ = \ \sum_{i=1}^{N^n_t} \Delta(n,i)^{q+1-p/2}\,|\Delta^n_i X|^p, \tag{10}$$

so in particular $V^n(p) = V^n(p,0)$. In the regular sampling case $\Delta(n,i) = \Delta_n$ these processes all convey the same information, since $V^n(p,q) = \Delta_n^q V^n(p)$, but this is no longer the case in the irregular sampling case.

Theorem 2 *Let $p \geq 1$ and $q \geq 0$. Assume (A) and (D($q+1$)), and also (D($q+1+\varepsilon$)) for some $\varepsilon > 0$ when $q > 0$. Then*

$$r_n^q V^n(p,q)_t \overset{u.c.p.}{\longrightarrow} V(p,q)_t := m_p \int_0^t |\sigma_s|^p \, a(q+1)_s \, ds. \tag{11}$$

4 - The Central Limit Theorem.

The previous consistency results are undoubtedly useful, but to make full use of them we need some associated CLT (central limit theorem).

For a proper statement of the CLT, we recall the notion of $\mathcal{F}^0$-stable convergence in law for a sequence of random variables (or processes) Y_n defined on $(\Omega, \mathcal{F}, \mathbb{P})$, see [10] for more details. We say that Y_n *converge $\mathcal{F}^0$-stably in law* to Y, where Y is a variable defined on an extension $(\widetilde{\Omega}, \widetilde{\mathcal{F}}, \widetilde{\mathbb{P}})$ of $(\Omega, \mathcal{F}, \mathbb{P})$, if we have

$$\mathbb{E}(Z\, h(Y_n)) \rightarrow \widetilde{\mathbb{E}}(Z\, h(Y)) : \quad Z \text{ bounded } \mathcal{F}^0\text{-measurable, } h \text{ continuous bounded.}$$

There are in fact two versions for the CLT. The first is associated with Theorem 1, and is thus the most useful in practice, and it also holds for (11) when $q > 0$ under a strong additional assumption:

Theorem 3 *Let $p \geq 2$, and assume (A) when $p = 2$ and (B) when $p > 2$. Let $q \geq 0$ and assume one of the following two sets of hypotheses:*

(i) $q = 0$ and (D(2));

(ii) $q > 0$ and (D($q+1$)) and (D($2q+2$)) and (D'($q+1$)).

Then the processes $\sqrt{r_n}\,\left(r_n^q V^n(p,q) - V(p,q)\right)$ converge $\mathcal{F}^0$-stably in law to

$$\overline{V}(p,q)_t = \sqrt{m_{2p} - m_p^2} \int_0^t |\sigma_s|^p \sqrt{a(2q+2)_s} \, dW'_s,$$

where W' is a standard Brownian motion, defined on an extension of $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, \mathbb{P})$ and independent of $\mathcal{F}$. Moreover, conditionally on $\mathcal{F}^0$, the process $\overline{V}(p,q)$ is a continuous centered Gaussian martingale with (conditional) variance at time t:

$$(m_{2p} - m_p^2) \int_0^t |\sigma_s|^{2p} \, a(2q+2)_s \, ds \tag{12}$$

As mentioned before, (D'$(q+1)$) is a very strong assumption, which for example is never satisfied by mixed renewal schemes unless the variables $\varepsilon(n,i)$ are constant. So when $q>0$ the previous CLT hardly applies.

In practice we want to estimate $V(p)_t$, so the above seems enough. However we need also an estimate for the conditional variance (12). For this, we may use $(1-\frac{m_p^2}{m_{2p}})r_n V^n(2p,1)_t$ by virtue of (11): this is why we have introduced the processes $V^n(p,q)$ for $q>0$. But now, for asserting the quality of the latter estimator, we need a CLT for the processes $V^n(p,q)$ under reasonable assumptions, weaker than (ii) above. This is achieved in the following result:

Theorem 4 *Let $p \geq 2$ and $q \geq 0$. Assume (B) and that the sampling scheme satisfies (E'$(2q+2+\varepsilon)$) for some $\varepsilon>0$, and that the $(\mathcal{F}_t^0)$-adapted processes $G(p)$ for $p\in[1,2q+2]$ are Itô semimartingales with the same properties as the process σ in Assumption (B). Then the processes $\sqrt{r_n}\,(r_n^q V^n(p,q)-V(p,q))$ converge $\mathcal{F}^0$-stably in law to*

$$\overline{V}(p,q)_t \;=\; \int_0^t |\sigma_s|^p \sqrt{a(2q+2)_s\, w(p,q)_s}\; dW'_s,$$

where W' is as in Theorem 3 and

$$w(p,q)_t \;=\; m_{2p}-m_p^2\,\frac{2G(q+2)_t\,G(q+1)_t\,G(1)_t-(G(q+1)_t)^2\,G(2)_t}{G(2q+2)_t\;(G(1)_t)^2}.$$

One may check that

- $w(p,q)_t \geq m_{2p}-m_p^2$ always;

- when the assumptions of both theorems above are satisfied, then $w(p,q)_t=m_{2p}-m_p^2$, so the two limits are the same.

Also, if we have an independent scheme with, say $\Delta(n,i)=\varepsilon(n,i)/r_n$ where the $\varepsilon(n,i)$ are all i.i.d. with law η and $m'_1=1$ (this is a special mixed renewal scheme, for which $v^n=v=1$ and so (7) gives $G(q)_t=m'_q$ and $a(q)_t=m'_q$). On the one hand, the average number of sampling times inside $[0,t]$, at stage n, is tr_n. On the other hand we have $a(2q+2)_t>1$ and $w(p,q)_t \geq m_{2p}-m_p^2$, unless $\varepsilon(n,i)=1$ a.s., in which case those are equalities. So the asymptotic conditional variance of the estimator for this independent scheme is always *strictly bigger* than the asymptotic conditional variance for the regular scheme having the same average number of observations.

5 - Estimation of the integrated volatility and other powers.

Suppose we want to estimate the variable $V(p)_t$ at some time $t>0$. We can "standardize" our estimator $V^n(p)_t$ by considering the variable

$$T_t^n \;=\; \sqrt{\frac{m_{2p}}{(m_{2p}-m_p^2)\,V^n(2p,1)_t}}\;(V(p)_t^n-V(p)_t).$$

Then the following holds:

Theorem 5 *Let $p \geq 2$. Assume (D(2)) and (D(2 + ε)) for some $\varepsilon > 0$. Assume also (A) when $p = 2$ and (B) otherwise. Then for any $t > 0$ such that $\int_0^t |\sigma_s|^{2p} a(2)_s ds > 0$ a.s., the sequence T_t^n converges in law (and even $\mathcal{F}^0$-stably in law) to $\mathcal{N}(0,1)$.*

All ingredients in the definition of T_t^n are known to the statistician, except of course the quantity $V(p)_t$ to be estimated. Therefore we can derive (asymptotic) confidence intervals for $V(p)_t$, or tests, in a straightforward way.

An interesting - and crucial - feature of this result is that the properties of the observation scheme are not showing explicitly in the result itself, and in particular the knowledge of the process $a(2)$ and even of the rates r_n is not necessary to apply it. This is a good thing because those are generally unknown, whereas it is also dangerous because one might be tempted to use the property that T_t^n is (approximately) $\mathcal{N}(0,1)$ without checking that the assumptions on the sampling scheme are satisfied. When they are satisfied, and although the rates r_n do not explicitly show up, these rates still govern the "true" rate of convergence.

Once more, r_n is unknown, but N_t^n is of course known. Then as soon as we have also (D(0)) (for example when (E(2)) holds), then $N_t^n / r_n \xrightarrow{\mathbb{P}} \int_0^t a(0)_s ds$, and so the actual rate of convergence for the estimators is also $1/\sqrt{N_t^n}$, as it should be.

References

[1] Barndorff-Nielsen, O.E. and N. Shephard (2005): Power variation and time change. *T. Probab. and its Appl.*, **50**, 1-15.

[2] Barndorff-Nielsen, O.E., P.R. Hansen, A. Lunde and N. Shephard (2008): Multivariate realised kernels: consistent positive semi-definite estimators of the covariation of equity prices with noise and non-synchronous trading. Working paper.

[3] Duffie, D. and P. Glynn (2004): Estimation of continuous-time Markov processes sampled at random time intervals. *Econometrica* **72**, 1773-1808.

[4] Fukasawa, M. (2008): Central limit theorem for the realized volatility based on tick time sampling. Preprint.

[5] Genon-Catalot, V. and J. Jacod (1994): Estimation of the diffusion coefficient for diffusion processes: random sampling. *Scand. J. of Statist.* **21**(3), 193-221.

[6] Griffin, J.E. and R.C. Oomen (2006): Measurement in the presence of non-synchronous trading and market microstructure noise. Working paper (CIREQ Conf. on realized volatility).

[7] Hayashi, T. and N. Yoshida (2005): On covariance estimation of nonsynchronously observed diffusion processes. *Bernoulli*, **11**, 359-379.

[8] Hayashi, T. and N. Yoshida (2008): Asymptotic normality of a covariance estimator for nonsynchronously observed diffusion processes. *Annals of the Institute of Statistical Mathematics*, **60**, 357-396.

[9] Hayashi, T. and N. Yoshida (2006): Nonsynchronous covariance estimator and limit theorem. Preprint.

[10] Jacod, J. and A. Shiryaev (2003): *Limit Theorems for Stochastic Processes*, 2d ed., Springer-Verlag: Berlin.

[11] Jacod, J. and Protter, P. (1998): Asymptotic error distributions for the Euler method for stochastic differential equations. *Ann. Probab.*, **26**, 267-307.

[12] Jacod, J. (2006): Asymptotic properties of realized power variations and related functionals of semimartingales. To appear in *Stoch. Processes and their Appl.*

[13] Jacod, J. (2007): Statistics and high frequency data. SEMSTAT Seminar. To appear.

[14] Malliavin, P. and M.E. Mancino (2002): Fourier series methods for measurement of multivariate volatilities. *Finance Stoch.* **6**, 49-61.

[15] Mykland, P.A. and L. Zhang (2006): ANOVA for diffusions and Itô processes. *The Annals of Statistics*, **34**, 1931-1963.