Estimation of integrated volatility: putting everything together

Jean Jacod – Viktor Todorov

UPMC (Paris 6) – Northwestern University

The aim

Estimate the integrated volatility $C_t = \int_0^t \sigma_s^2\, ds$ of the 1-dimensional process

$$X_t = X_0 + \int_0^t b_s ds + \int_0^t \sigma_s dW_s + \text{jumps}$$

observed at discrete times within the *fixed* time interval $[0,t]$, with a mesh going to 0 (*high frequency setting*).

Three features:

- X has jumps, possibly with *high-activity* (= BG index ≥ 1)
- the sampling times $0 = T(n,0) < T(n,1) < \cdots < T(n,i) < \cdots$ may be irregular, possibly random (*regular sampling* $\Leftrightarrow$ $T(n,i+1) - T(n,i) = \Delta_n$)
- there is a microstructure noise: instead of $X_{T(n,i)}$ we observe

$$Y_i^n = X_{T(n,i)} + \chi_i^n.$$

Notation

$$\Delta(n,i) = T(n,i) - T(n,i-1)$$

$$\Delta_i^n V = V_{T(n,i)} - V_{T(n,i-1)} \qquad V: \text{ any process}$$

Regular sampling means $\Delta(n,i) = \Delta_n$.

Spot Lévy measures of X: the compensator ν of the jump measure of X is assumed to have the factorization

$$\nu(\omega, dt, dx) = dt\, F_{\omega,t}(dx)$$

(this is the "Itô semimartingale property"). The measures $F_t = F_{\omega,t}$ are the spot Lévy measure, and $\int (x^2 \wedge 1)\, F_t(dx) < \infty$.

Problems with jumps *(no-noise and regular sampling cases)*

• *X continuous:* the "optimal" estimator for C_t (in the sense of LAN or LAMN properties)

$$\widehat{C}_t^n = \sum_{i=1}^{[t/\Delta_n]} (\Delta_i^n X)^2$$

the rate is $\frac{1}{\sqrt{\Delta_n}}$, the (conditional) asymptotic variance is $2\int_0^t \sigma_s^4\, ds$.

• *X discontinuous:* $\widehat{C}_t^n$ no longer consistent, and one has 2 main methods:

$$\textit{truncation:} \qquad \widehat{C}_t'^n = \sum_{i=1}^{[t/\Delta_n]} (\Delta_i^n X)^2\, 1_{\{|\Delta_i^n X| \leq u_n\}}$$

$$\textit{multipowers:} \qquad \widehat{C}_t''^n = \alpha_p \sum_{i=1}^{[t/\Delta_n]-k+1} \prod_{j=0}^{k-1} |\Delta_{i+j}^n X|^{2/k}$$

Both these are consistent, and have a CLT with rate $\frac{1}{\sqrt{\Delta_n}}$ and the optimal variance for truncations and a (slightly) bigger variance for multipowers, under mild assumptions on b_t, σ_t (stronger for multipowers, though, for example $1/\sigma_t^2$ should be locally bounded), PLUS the (strong) additional assumption

$$\int (|x|^r \wedge 1)\, F_t(dx) \quad \text{is locally bounded} \tag{1}$$

for some $r < 1$.

More: if a sequence S_n is such that $v_n(S_n - C_t)$ is tight for some sequence $v_n \to \infty$, uniformly for *all* X for which b_t, σ_t and (1) are *uniformly bounded*, then the minimax rate v_n satisfies

$$r \leq 1 \quad \Rightarrow \quad v_n \preceq \frac{1}{\sqrt{\Delta_n}}, \qquad r \leq 1 \quad \Rightarrow \quad v_n \preceq \Big(\frac{\log(1/\Delta_n)}{\Delta_n}\Big)^{(2-r)/2}$$

However: when $r > 1$ and the jumps of X are “close enough” to those of a stable process, then it becomes possible to get the $\frac{1}{\sqrt{\Delta_n}}$ rate (no contradiction: we have switched from a non-parametric situation to a semi-parametric one).

Assumptions on X

$$X_t = X_0 + \int_0^t b_s\,ds + \int_0^t \sigma_s\,dW_s + \int_0^t \int_E \delta(s,z)(\underline{p} - \underline{q}\,)(ds,dz) + \int_0^t \int_E \delta'(s,z)\underline{p}(ds,dz)$$

$$\sigma_t = \sigma_0 + \int_0^t b_s^\sigma\,ds + \int_0^t H_s^\sigma\,dW_s' + \int_0^t \int_E \delta^\sigma(s,z)(\underline{p} - \underline{q}\,)(ds,dz) + \int_0^t \int_E \delta'^\sigma(s,z)\underline{p}(ds,dz)$$

• W and W' are two correlated Brownian motions and $\underline{p}$ is a Poisson measure on $\mathbb{R}_+ \times E$ with (deterministic) compensator $\underline{q}\,(dt,dz) = dt \otimes \eta(dz)$ (η is a σ-finite measure on the Polish space E).

• *up to some localization*:

$b_t, b_t^\sigma, H_t^\sigma$ are optional bounded

$\delta, \delta', \delta^\sigma, \delta'^\sigma$ are predictable, $|\delta|, |\delta^\sigma| \leq 1$ and, for some η-integrable function J,

$$|\delta(t,z)|^{r'},\ |\delta'(t,z)|^r \wedge 1,\ |\delta^\sigma(t,z)|^2,\ |\delta'^\sigma(t,z)| \wedge 1| \leq J(z), \quad \text{with} \quad r < 1,\ r' < 2$$

and the processes $V_t = b_t,\ H_t^\sigma,\ \delta(t,z)^{r'}/J(z)$ for all z satisfy for all finite stopping times $T \leq S$:

$$\mathbb{E}(\sup_{s \in [T,S]} |V_s - V_T|^2 \mid \mathcal{F}_T) \ \leq \ K\mathbb{E}(S - T \mid \mathcal{F}_T) \tag{2}$$

$\cdots/\ldots$

We also need a structural assumption on the high-activity jumps of X, expressed in terms of the BG (Blumenthal-Getoor) index, or successive BG indices:

There is an integer $M \geq 0$, a finite family $2 > \beta_1 > \cdots \beta_M > 0$ of numbers, and M nonnegative predictable càdlàg $a_t^1, \ldots, a_t^M$, such that each $(a_t^m)^{1/\beta_m}$ satisfies (2), and the tail (random) functions $\overline{F}_t(x) = F_t([-x, x]^c)$ have

$$\left| \overline{F}_t(x) - \sum_{m=1}^{M} \frac{a_t^m}{x^{\beta_m}} \right| \leq \frac{K}{x^r}$$

Example:

$$X_t = X_0 + \int_0^t b_s\, ds + \int_0^t \sigma_s\, dW_s + \sum_{m=1}^{M} \int_0^t \gamma_{s-}^m\, dY_s^m + \int_0^t \int_E \delta'(s, z) \underline{p}(ds, dz)$$

with Y^m independent pure jumps Lévy processes (with arbitrary dependencies with $\underline{p}$) and γ^m's are càdlàg adapted processes satisfying (2), provided the Lévy measure of Y^m satisfies $|\overline{F}^m(x) - 1/x^{\beta_m}| \leq K/x^r$ (e.g., Y^m is β_m-stable or tempered stable).

We then have $a_t^m = |\gamma_t^m|^{\beta_m}$.

Assumptions on the observation scheme

Assumption: Each $T(n,i)$ is a stopping time, and

$$\Delta(n,i+1) = \Delta_n \lambda_{T(n,i)} \Phi^n_{i+1}$$

- λ_t positive càdlàg adapted, satisfying (2) and $1/K \le \lambda_t \le K$ (up to localization).
- The variables Φ^n_i are positive, $\mathbb{E}(\Phi^n_{i+1} \mid \mathcal{F}_{T(n,i)}) = 1$ and $\mathbb{E}(|\Phi^n_i|^q \mid \mathcal{F}_{T(n,i)}) \le K_q$, and conditionally on $\mathcal{F}_{T(n,i)}$ the variable Φ^n_{i+1} is independent of the σ-field $\mathcal{F}'_\infty = \bigvee(X_s : s \ge 0)$.

This implies in particular that $N^n_t = \sum_{i\ge1} 1_{\{T(n,i)\le t\}}$ satisfies

$$\Delta_n N^n_t \overset{\text{u.c.p.}}{\Longrightarrow} \Lambda_t := \int_0^t \frac{1}{\lambda_s}\, ds.$$

Examples:

Regular schemes.

Poisson schemes with parameter $1/\Delta_n$ and independent of $\mathcal{F}'_\infty$.

Modulated random walk schemes, where the Φ^n_i's are i.i.d. independent of X.

Assumptions on the noise

Assumption: We observe $Y_i^n = X_{T(n,i)} + \chi_i^n$, where for each n the variables $(\chi_i^n : i \geq 0)$ are independent, conditionally on $\mathcal{F}_\infty$, and satisfy

$$\mathbb{E}(\chi_i^n \,|\mathcal{F}_\infty) = 0, \qquad \mathbb{E}((\chi_i^n)^2 \,|\mathcal{F}_\infty) = \gamma_{T(n,i)}, \qquad \mathbb{E}((\chi_i^n)^3 \,|\mathcal{F}_\infty) = \gamma'_{T(n,i)}$$
$$T(n,i) \leq \tau_m \;\Rightarrow\; \mathbb{E}((\chi_i^n)^8 \mid \mathcal{F}_\infty) \leq K$$

with two càdlàg adapted processes $\gamma_t \geq 0$ and γ'_t satisfying (2) (up to localization)

A centered white noise (with the χ_i^n i.i.d. and independent of $\mathcal{F}_\infty$ satisfies this; a modulated white noise also satisfies this.

Another important example. Let $\rho_t^j \geq 0$ be càdlàg adapted nonnegative with $\sum_{j\in\mathbb{Z}} \rho_t^j = 1$ and $\rho_t^j = \rho_t^{-j}$ and $\sum_{j\in\mathbb{Z}} \rho_t^j \,|j|^8 \leq K$. For each n let $(Z_i^n : i \geq 1)$ be i.i.d. conditionally on $\mathcal{F}_\infty$, with density $x \mapsto \sum_{j\in\mathbb{Z}} \rho^j_{T(n,j)} 1_{[j,j+1)}(x)$. The observation at time $T(n,i)$ is $Y_i^n = [X_{T(n,i)} + Z_i^n]$, so we have a additive (modulated) white noise *plus rounding.*

Remark: If we have "pure rounding", i.e. if we observe $[X_{T(n,i)}]$ (or $[X_{T(n,i)}] + \frac{1}{2}$ to "center" the noise), then no consistent estimator for C_t exists.

Pre-averaging

The de-noising method is pre-averaging, but other methods could probably be used as well. Take a weight (or, kernel) function g on $\mathbb{R}$ with

$$\begin{array}{l} g \text{ is continuous, piecewise } C^1 \text{ with a piecewise Lipschitz derivative } g', \\ s \notin (0,1) \ \Rightarrow \ g(s) = 0, \qquad \int_0^1 g(s)^2 ds \ > \ 0, \end{array}$$

for example $g(x) = (x \wedge (1-x))\, 1_{[0,1]}(x)$. With a sequence $h_n \to \infty$ of integers, set

$$\begin{array}{ll} g_i^n = g(i/h_n), & \overline{g}_i^n \ = \ g_{i+1}^n - g_i^n \\ \phi_j^n = \frac{1}{h_n} \sum_{i \in \mathbb{Z}} g_i^n \, g_{i-j}^n, & \overline{\phi}_j^n = h_n \sum_{i \in \mathbb{Z}} \overline{g}_i^n \, \overline{g}_{i-j}^n \\ \phi(s) = \int g(u) g(u-s)\, du, & \overline{\phi}(s) = \int g'(u) g'(u-s)\, du. \end{array}$$

The *pre-averaged returns* of the observed values Y_i^n are

$$\widetilde{Y}_i^n = \sum_{j=1}^{h_n - 1} g_j^n \, (Y_{i+j}^n - Y_{i+j-1}^n) = - \sum_{j=0}^{h_n - 1} \overline{g}_j^n \, Y_{i+j}^n.$$

For h_n, the (rate)-optimal choice is

$$h_n = \frac{\theta}{\sqrt{\Delta_n}} + \mathrm{o}(\Delta_n^{1/4}), \qquad \text{for some } \theta > 0. \tag{3}$$

Initial estimators

We need anther sequence $k_n \geq 1$ of integers and a sequence u_n of positive reals, such that

$$u_n \to 0, \qquad k_n^2 / h_n u_n^8 \to 0, \qquad k_n^{2+\varepsilon} / h_n \to \infty \quad \forall \varepsilon > 0 \tag{4}$$

We set $w_n = 2h_n k_n$ and, for any $y > 0$,

$$L(y)_j^n = \frac{1}{k_n} \sum_{l=0}^{k_n - 1} \cos\big(y u_n \sqrt{h_n}\,(\widetilde{Y}^n_{j w_n + 2l h_n} - \widetilde{Y}^n_{j w_n + (2l+1) h_n})\big)$$

(a proxy for the real part of the empirical characteristic function of the returns, over a window of length $2w_n$). Taking a difference above allows us to "symmetrize" the problem.

Then, a natural estimator for the integrated volatility over the time interval $[T(n, jw_n), T(n, (j+1)w_n)]$ is then

$$\widehat{c}(y)_j^n = -\frac{1}{y^2 u_n^2 \phi_0^n} \log\Big(L(y)_j^n \bigvee \frac{1}{h_n}\Big).$$

We need to de-bias these estimators, to account for the noise, and also for some intrinsic distortion present even when there is no noise. So, the initial estimators for the integrated volatility are then, with $\sinh(x) = \frac{1}{2}\left(e^x - e^{-x}\right)$ the hyperbolic sine:

$$\widehat{C}(y)^n_t = \frac{2k_n}{h_n} \sum_{j=0}^{[N^n_t/w_n]-1} \left(\widehat{c}(y)^n_j - \frac{\left(\sinh(y^2 u_n^2 \widehat{c}(y)^n_j)\right)^2}{y^2 u_n^2 k_n \phi^n_0} - \frac{\overline{\phi}^n_0}{2w_n \phi^n_0} \sum_{l=1}^{w_n} (\Delta^n_{i+l} Y^n)^2\right)$$

We will see that $\widehat{C}(y)_T$ converges to C_T, and there is an associated Central Limit Theorem with the convergence rate $1/\Delta_n^{1/4}$. However, this CLT exhibits a non-negligible bias, and is in fact about the processes (where $\chi(\beta) = \int_0^\infty \frac{\sin y}{y^\beta}\, dy$):

$$\begin{aligned} Z(y)^n_t &= \tfrac{1}{\Delta_n^{1/4}} \Big(\widehat{C}(y)^n_t - C_t - \textstyle\sum\limits_{m=1}^{M} A^m(y)^n_t\Big), \qquad \text{where} \\ A^m(y)^n_t &= 2\chi(\beta_m)\, \tfrac{\widetilde{\phi}^{n,\beta_m}}{\phi^n_0}\, (yu_n)^{\beta_m - 2} \Delta_n h_n^{\beta_m/2-1}\, A^m_t. \end{aligned}$$

The next theorem presents the CLT for $Z(y)^n_t$, and also for the differences $Z(y)^n_t - Z(1)^n_t$ when $y > 0$. The reason for giving a CLT for these differences is that they will play a key role in the de-biasing procedure developed later on.

Theorem. *With h_n, k_n, u_n as above, for each $t > 0$ we have the $\mathcal{F}_\infty$-stable convergence in law:*

$$\Big(Z(1)^n_t,\ \big(\frac{1}{u_n^2}(Z(y)^n_t - Z(1)^n_t)\big)_{y\in\mathcal{Y}}\Big) \stackrel{\mathcal{L}-s}{\Longrightarrow} \big(Z_t, ((y^2-1)\overline{Z}_t)_{y\in\mathcal{Y}}\big),$$

where the limit is defined on an extension $(\widetilde{\Omega}, \widetilde{\mathcal{F}}, (\widetilde{\mathcal{F}}_t)_{t\geq 0}, \widetilde{\mathbb{P}})$ of the original space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, \mathbb{P})$ and can be written as

$$Z_t = \frac{2}{\theta^{3/2}} \int_0^t \frac{\theta^2 c_s\lambda_s + \frac{\overline{\phi}(0)}{\phi(0)}\,\gamma_s}{\sqrt{\lambda_s}}\, dW^{(1)}_s, \qquad \overline{Z}_t = \frac{2\phi(0)}{\sqrt{3}\,\theta^{3/2}} \int_0^t \frac{(\theta^2 c_s\lambda_s + \frac{\overline{\phi}(0)}{\phi(0)}\,\gamma_s)^2}{\sqrt{\lambda_s}}\, dW^{(2)}_s.$$

where $W^{(1)}$ and $W^{(2)}$ are two independent Brownian motions, independent of $\mathcal{F}$.

Z_t is $\mathcal{F}$-conditionally centered Gaussian, with conditional variance

$$\widetilde{\mathbb{E}}((Z_t)^2 \mid \mathcal{F}) = \frac{4}{\theta^3} \int_0^t \frac{(\theta^2 c_s\lambda_s + \frac{\overline{\phi}(0)}{\phi(0)}\gamma_s)^2}{\lambda_s}\, ds$$

The case $M = 0$

$M = 0$ is equivalent (upon reformulating the basic equation) to the case where

$$X_t = X_0 + \int_0^t b_s\, ds + \int_0^t \sigma_s\, dW_s + \int_0^t \int_E \delta(s,z)\delta'(s,z)\underline{p}(ds,dz)$$

and the last term is pure jump with finite variation. The bias term disappears, so

$$\frac{1}{\Delta_n^{1/4}}\,(\widehat{C}(y)_t^n - C_t) \stackrel{\mathcal{L}-\mathrm{s}}{\longrightarrow} Z_t$$

For a feasible CLT we need to have consistent estimators for the conditional variance. Among many possible choices, one can take

$$\widehat{V}_t^n = \frac{8k_n}{h_n^2} \sum_{j=0}^{[N_t^n/w_n]-1} (\widehat{c}(1)_j^n)^2$$

For all $t>0$ we have the following convergence in probability:

$$\frac{1}{\sqrt{\Delta_n}}\,\widehat{V}_t^n \xrightarrow{\mathbb{P}} \frac{4}{\theta^3}\int_0^t \frac{(\theta^2 c_s\lambda_s + \frac{\overline{\phi}}{\phi}\gamma_s)^2}{\lambda_s}\,ds$$

Therefore, because of the stable convergence in law we get

Theorem. *If $M=0$, for any $t>0$ the variables $(\widehat{C}(1)_t^n - C_t)/\sqrt{\widehat{V}_t^n}$ converge stably in law to an $\mathcal{N}(0,1)$ random variable, in restriction to the set $\{C_t + \int_0^t \gamma_s\,ds > 0\}$.*

The case $M = 1$

When $M = 1$ we have

$$\begin{aligned} \widehat{C}(y)_t^n &= C_t + A^1(y)_t^n + \Delta_n^{1/4} Z(y)_t^n + \mathrm{o}_P(\Delta_n^{1/4}) \\ &= C_t + 2\chi(\beta_1)\,\frac{\widetilde{\phi}^{n,\beta_1}}{\phi_0^n}\,(yu_n)^{\beta_1-2}\Delta_n h_n^{\beta_1/2-1}\,A_t^1 + \Delta_n^{1/4} Z(y)_t^n + \mathrm{o}_P(\Delta_n^{1/4}) \end{aligned}$$

and the bias term has an order of magnitude bigger than $\Delta_n^{1/4}$. However $\widehat{C}(y)_t^n - \widehat{C}(y')_t^n$ is an estimator of this bias term, so we can set for some $\zeta > 1$:

$$\widehat{C}(y,\zeta)_t^n = \widehat{C}(y)_t^n - \frac{(\widehat{C}(\zeta u)_t^n - \widehat{C}(y)_t^n)^2}{\widehat{C}(\zeta^2 y)_t^n - 2\widehat{C}(\zeta y)_t^n + \widehat{C}(y)_t^n}$$

Theorem. *If $M = 0$, for any $t > 0$ the variables $(\widehat{C}(1,\zeta)_t^n - C_t)/\sqrt{\widehat{V}_t^n}$ converge stably in law to an $\mathcal{N}(0,1)$ random variable, in restriction to the set $\{C_t + \int_0^t \gamma_s\,ds > 0\}$..*

The case $M \geq 2$

When $m \geq 2$ we have

$$\widehat{C}(y)^n_t = C_t + \sum_{m=1}^{M} A^m(y)^n_t + \Delta_n^{1/4} Z(y)^n_t + \mathrm{o}_P(\Delta_n^{1/4})$$

and the $A^m(y)^n_t$ have an order of magnitude decreasing with m. The previous de-biasing procedure removes only the biggest part $A^1(y)^n_t$, so in the presence of several terms the procedure has to be iterated.

Iteration will work under some additional structure:

Assumption: *The numbers $2-\beta_m$ all belong to the set $\{j\rho :\ j = 1, 2, \cdots\}$, for some unknown constant $\rho \in (0,1)$ (so necessarily $M \leq [2/\rho]$).*

Iterative procedure

Step 1 - initialization: Choose a real $\zeta > 1$ and an integer $k \geq 1$, and put $\widehat{C}(y,\zeta,0)_T^n = \widehat{C}(y)_t^n$.

Step 2 - iteration: Assuming $\widehat{C}_n(y,\zeta,j-1)$ known for some integer j between 1 and k, define $\widehat{C}_n(y,\zeta,j)$ as

$$\begin{aligned}\widehat{C}(y,\zeta,j)_t^n &= \widehat{C}_n(y,\theta,j-1)_t^n + y^2/(N_t^n)^{1/4} \\ &\quad + \frac{(\widehat{C}_n(\zeta y,\theta,j-1)_t^n - \widehat{C}(y,\zeta,j-1)_t^n)^2}{\widehat{C}(\zeta^2 y,\zeta,j-1)_t^n - 2\widehat{C}(\theta y,\zeta,j-1)_t^n + \widehat{C}(y,\theta,j-1)_t^n}.\end{aligned}$$

Step 3: The final estimator is set to be $\widehat{C}(y,\zeta,N)_T^n$.

The asymptotic result for $\widehat{C}(u_n,\zeta,N)_T^n$ is given in the following theorem.

Theorem. *If we know that $\rho \geq \rho_0$ and if $N \geq 1/\rho_0$, the variables $(\widehat{C}(y,\zeta,N)_t^n - C_t)/\sqrt{\widehat{V}_t^n}$ converge stably in law to an $\mathcal{N}(0,1)$ random variable, in restriction to the set $\{C_t + \int_0^t \gamma_s\, ds > 0\}$.*

Some problems:

1 - A stopping rule for the iteration above (almost done...?)

2 - Optimality

3 - The multi-dimensional case