Integrated noise scale estimation for stochastic integrals

Arturo Jaramillo Gil

Centro de Investigación en Matemáticas (CIMAT)
Joint work with Jorge Cázares

Collaborators

Joint work with Jorge Cazares

Key works by Vladimir Fomichov, Jorge Gonzalez Cazares, Jevgenijs Ivanovs

Objective

Consider a process of the form

$$\xi_t = \int_0^t \eta_s \circledast dV_s + L_t.$$

Objective

Consider a process of the form

$$\xi_t = \int_0^t \eta_s \circledast dV_s + L_t.$$

Goal: recover, only from discrete observations of ξ,

$$I[\eta](t) = \int_0^t \eta_s \, ds.$$

Objective

Consider a process of the form

$$\xi_t = \int_0^t \eta_s \circledast dV_s + L_t.$$

- V is Gaussian with $Var[V_{s+h} - V_s]$ behaving like h^{2H}.
- $\eta > 0$.
- $\circledast$ is a notion of integration.
- L has low activity:

$$\|L\|_p^p := \sup_{n \geq 1} \left(|\Delta_{1,n} L|^p + \cdots + |\Delta_{n,n} L|^p \right) < \infty.$$

Goal: recover, only from discrete observations of ξ,

$$I[\eta](t) = \int_0^t \eta_s \, ds.$$

Canonical example

A classical high-frequency model is

$$X_t = X_0 + I[b](t) + \int_0^t \sigma_s \, dW_s + J_t .$$

Canonical example

A classical high-frequency model is

$$X_t = X_0 + I[b](t) + \int_0^t \sigma_s \, dW_s + J_t.$$

- $I[b]$ is a finite-variation drift term.
- σ is a square integrable continuous adapted process.
- J an integral of a Levy, with finite $(2-\varepsilon)$-variation.

We observe X on a regular grid, and want to recover $I[\sigma^2]$

Thresholding strategies

At the level of increments,

$$\Delta_k^n X = \text{Brownian fluctuation} + \text{drift} + \text{jumps}.$$

Thresholding strategies

At the level of increments,

$$\Delta_k^n X = \text{Brownian fluctuation} + \text{drift} + \text{jumps}.$$

Heuristic

- Brownian increments are typically small.
- Jump increments are typically large.
- Removing large increments typically isolate the continuous martingale contribution.

Thresholding strategies

At the level of increments,

$$\Delta_k^n X = \text{Brownian fluctuation} + \text{drift} + \text{jumps}.$$

Heuristic

- Brownian increments are typically small.
- Jump increments are typically large.
- Removing large increments typically isolate the continuous martingale contribution.

One keeps only increments satisfying

$$|\Delta_k^n X| \leq u_n,$$

Thresholding strategies

At the level of increments,

$$\Delta_k^n X = \text{Brownian fluctuation} + \text{drift} + \text{jumps}.$$

Heuristic

- Brownian increments are typically small.
- Jump increments are typically large.
- Removing large increments typically isolate the continuous martingale contribution.

One keeps only increments satisfying

$$|\Delta_k^n X| \leq u_n,$$

and computes a truncated quadratic variation.

Heuristic considerations

The signal of interest looks like

$$d\xi_t = \eta_t \circledast dV_t + dL_t,$$

Heuristic considerations

The signal of interest looks like

$$d\xi_t = \eta_t \circledast dV_t + dL_t,$$

Discretizing and rescaling by n^H:

$$X_{n,k} = \sigma_{n,k} Z_{n,k} + Y^o_{n,k},$$

where $Z_{n,k} := n^H \Delta_{k,n} V$, and

- $X_{n,k} := n^H \Delta_{n,k} \xi$
- $\sigma_{n,k} Z_{n,k} =$ Riemann sum approximation
- $Y^o_{n,k} =$ Increment of $n^H L$ + scaled Riemann sum error.

Heuristic considerations

The signal of interest looks like

$$d\xi_t = \eta_t \circledast dV_t + dL_t,$$

Discretizing and rescaling by n^H:

$$X_{n,k} = \sigma_{n,k} Z_{n,k} + Y^o_{n,k},$$

where $Z_{n,k} := n^H \Delta_{k,n} V$, and

- $X_{n,k} := n^H \Delta_{n,k} \xi$
- $\sigma_{n,k} Z_{n,k} =$ Riemann sum approximation
- $Y^o_{n,k} =$ Increment of $n^H L$ + scaled Riemann sum error.

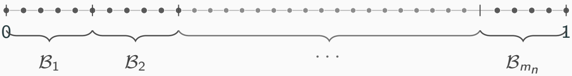
Precise definitions for later

Big blocks-small blocks methodology

We use blocks $\mathcal{B}_i$ to pack indices.

Big blocks-small blocks methodology

We use blocks $\mathcal{B}_i$ to pack indices.



In each block $\mathcal{B}_\ell$, we use the local approximation

$$X_{n,k} = \sigma Z_{n,k} + Y_{n,k},$$

for k in $\mathcal{B}_\ell$.

- σ is a **constant** noise scale inside the block.
- $Y_{n,k}$ absorbs both the term $Y^o_{n,k}$ and the approximation $\sigma_{n,k}$ with the blockwise constant.

A change of coordinates?

Apply a transformation (change of coordinates) $\mathcal{T}$ to the block

$$\boldsymbol{X}_n^{[\ell]} = (X_{n,k};\, k \in \mathcal{B}_\ell).$$

We want two design $\mathcal{T}$ with the properties:

- The vector $\mathcal{T}(\boldsymbol{X}_n^{[\ell]})$ is more predictable
- The equation $\boldsymbol{X}_n^{[\ell]} = \sigma \boldsymbol{Z}_n^{[\ell]} + \boldsymbol{Y}_n^{[\ell]}$, induces a tractable equation

$$\mathfrak{T}(\mathcal{T}(\boldsymbol{X}_n^{[\ell]}), \boldsymbol{Z}_n^{[\ell]}, \boldsymbol{Y}_n^{[\ell]}) = 0 \tag{1}$$

with $\boldsymbol{Y}_n^{[\ell]}$ not being of big influence in (1).

Detour: scale noise estimate in time series

For a vector $(\Lambda_{m,k};\ k = 1, \ldots, m)$, we denote by $\Lambda^{\uparrow}_{k,m}$ the k-th order statistic.

Detour: scale noise estimate in time series

For a vector $(\Lambda_{m,k};\ k = 1, \ldots, m)$, we denote by $\Lambda^{\uparrow}_{k,m}$ the k-th order statistic.

Lemma
Let $p \geq 1$ and let $A, B, C \in \mathbb{R}^n$ satisfy $A_k = B_k + C_k$. Then

$$\|A^{\uparrow} - B^{\uparrow}\|_p \leq \|C\|_p,$$

where $\|\cdot\|_p$ denotes p-norm.

Detour: scale noise estimate in time series

For a vector $(\Lambda_{m,k};\ k = 1, \ldots, m)$, we denote by $\Lambda^{\uparrow}_{k,m}$ the k-th order statistic.

Lemma
Let $p \geq 1$ and let $A, B, C \in \mathbb{R}^n$ satisfy $A_k = B_k + C_k$. Then

$$\|A^{\uparrow} - B^{\uparrow}\|_p \leq \|C\|_p,$$

where $\|\cdot\|_p$ denotes p-norm.

We then obtain the key identity **(Speaker note)**

$$X^{\uparrow}_{k,n} = \sigma Z^{\uparrow}_{k,n} + \widetilde{Y}_{k,n},$$

with

$$\|\widetilde{\boldsymbol{Y}}^{[\ell]}_n\|_p \leq \|\boldsymbol{Y}^{[\ell]}_n\|_p.$$

Heuristic considerations

The term $X^{\uparrow}_{k,n} \approx \sigma Z^{\uparrow}_{k,n}$ is easier to predict since

Heuristic considerations

The term $X^{\uparrow}_{k,n} \approx \sigma Z^{\uparrow}_{k,n}$ is easier to predict since

$$Z^{\uparrow}_{k,n} \approx \psi_{n,k}.$$

where

$$\psi_{n,k} := \Psi^{-1}\big(k/(n+1)\big),$$

for Ψ denoting the standard Gaussian law.

Heuristic considerations

The term $X^{\uparrow}_{k,n} \approx \sigma Z^{\uparrow}_{k,n}$ is easier to predict since

$$Z^{\uparrow}_{k,n} \approx \psi_{n,k}.$$

where

$$\psi_{n,k} := \Psi^{-1}\big(k/(n+1)\big),$$

for Ψ denoting the standard Gaussian law. The approximation $X^{\uparrow}_{k,n} \approx \sigma\psi_{n,k}$ suggest to choose σ so that minimizes the loss **(Speaker note)**

$$\theta(\varsigma) = \sum_{k=1}^{n} \left| X^{\uparrow}_{k,n} - \varsigma\psi_{n,k} \right|.$$

We define our estimator for σ within a single block as

$$\Sigma^{r}(\boldsymbol{X}_n^{[\ell]}) := \arg\min_{\varsigma>0} \theta(\varsigma).$$

Detour for constant volatility estimate

For concreteness: $Z_{n,k}$ are the increments of a fBm (V is a fractional Brownian motion).

Detour for constant volatility estimate

For concreteness: $Z_{n,k}$ are the increments of a fBm (V is a fractional Brownian motion).

Theorem (González Cázares and Jaramillo)
If η is constant equal to $\sigma > 0$ and we have one single block, then,
(Speaker note)

$$\mathbb{E}\left[\left|\frac{\Sigma(\boldsymbol{X}_n)}{\sigma} - 1\right|\right] \leq C\left(n^{-1/p}\mathbb{E}\left[\frac{\|\boldsymbol{Y}_n\|_p}{\sigma}\right] + n^{-((1-H)\wedge 1/2)}\right).$$

Interest connections with robustness

Ingredients of the proof

First step

$$|\Sigma(\boldsymbol{X}_n) - \sigma| \, \|\psi_n\|_1 = \|\Sigma(\boldsymbol{X}_n)\psi_n - \sigma\psi_n\|_1$$

Ingredients of the proof

First step

$$
\begin{aligned}
|\Sigma(\boldsymbol{X}_n) - \sigma| \, \|\psi_n\|_1 &= \|\Sigma(\boldsymbol{X}_n)\psi_n - \sigma\psi_n\|_1 \\
&\leq \|\Sigma(\boldsymbol{X}_n)\psi_n - X_n^{\uparrow}\|_1 + \|X_n^{\uparrow} - \sigma\psi_n\|_1
\end{aligned}
$$

Ingredients of the proof

First step

$$
\begin{aligned}
|\Sigma(\boldsymbol{X}_n) - \sigma| \, \|\psi_n\|_1 &= \|\Sigma(\boldsymbol{X}_n)\psi_n - \sigma\psi_n\|_1 \\
&\leq \|\Sigma(\boldsymbol{X}_n)\psi_n - X_n^{\uparrow}\|_1 + \|X_n^{\uparrow} - \sigma\psi_n\|_1 \\
&\leq 2\|X_n^{\uparrow} - \sigma\psi_n\|_1
\end{aligned}
$$

Ingredients of the proof

First step

$$\begin{aligned}
|\Sigma(\boldsymbol{X}_n) - \sigma|\, \|\psi_n\|_1 &= \|\Sigma(\boldsymbol{X}_n)\psi_n - \sigma\psi_n\|_1 \\
&\leq \|\Sigma(\boldsymbol{X}_n)\psi_n - X_n^{\uparrow}\|_1 + \|X_n^{\uparrow} - \sigma\psi_n\|_1 \\
&\leq 2\|X_n^{\uparrow} - \sigma\psi_n\|_1 \\
&\leq 2\|\sigma Z_n^{\uparrow} - \sigma\psi_n\|_1 + 2\|\widetilde{\boldsymbol{Y}}_n\|_1.
\end{aligned}$$

Ingredients of the proof

First step

$$\begin{aligned}|\Sigma(\boldsymbol{X}_n) - \sigma|\,\|\psi_n\|_1 &= \|\Sigma(\boldsymbol{X}_n)\psi_n - \sigma\psi_n\|_1 \\ &\leq \|\Sigma(\boldsymbol{X}_n)\psi_n - X_n^{\uparrow}\|_1 + \|X_n^{\uparrow} - \sigma\psi_n\|_1 \\ &\leq 2\|X_n^{\uparrow} - \sigma\psi_n\|_1 \\ &\leq 2\|\sigma Z_n^{\uparrow} - \sigma\psi_n\|_1 + 2\|\widetilde{\boldsymbol{Y}}_n\|_1.\end{aligned}$$

Second step

$$\|Z_n^{\uparrow} - \psi_n\| = n d_{W^1}(\iota[Z_n], \iota[\psi_n]),$$

where ι denotes taking the empirical distribution. **(speaker note)**

Details on the choices of the time series

The model:

$$\xi_t = \int_0^t \eta_s \circledast dV_s + L_t.$$

Define

$$s_{k,n}^2 := \mathrm{Var}(\Delta_{k,n} V), \qquad X_{k,n} \quad := s_{k,n}^{-1} \Delta_{k,n} \xi, \qquad Z_{k,n} := s_{k,n}^{-1} \Delta_{k,n} V.$$

Then

$$X_{k,n} = \eta_{t_{k-1}} Z_{k,n} + Y_{k,n},$$

where

$$Y_{k,n} = s_{k,n}^{-1} \Delta_{k,n} L + s_{k,n}^{-1} \left(\int_{t_{k-1}}^{t_k} \eta_s \circledast dV_s - \eta_{t_{k-1}} \Delta_{k,n} V \right)$$

$$\sigma_{k,n} = \eta_{t_{k-1}}.$$

Details on the choices of the blocks

We divide $\{1, \ldots, n\}$ into $M_n = \lfloor n/m_n \rfloor$ consecutive blocks of size m_n:

$$\mathcal{B}_\ell = \{(\ell - 1)m_n + 1, \ldots, \ell m_n\}.$$

For each block, recall the notation

$$\boldsymbol{X}_{m_n}^{[\ell]} = (X_{k,n} : k \in \mathcal{B}_\ell)$$

Details on the choices of the blocks

We divide $\{1, \ldots, n\}$ into $M_n = \lfloor n/m_n \rfloor$ consecutive blocks of size m_n:

$$\mathcal{B}_\ell = \{(\ell - 1)m_n + 1, \ldots, \ell m_n\}.$$

For each block, recall the notation

$$\boldsymbol{X}_{m_n}^{[\ell]} = (X_{k,n} : k \in \mathcal{B}_\ell)$$

Integrated scale estimator

$$\mathcal{IV}^{(m_n)}(\boldsymbol{X}_n) := \frac{T}{M_n} \sum_{\ell=1}^{M_n} \Sigma(\boldsymbol{X}_{m_n}^{[\ell]}).$$

Assumptions for convergence

We assume low activity

$$\sup_{n\to\infty}\sum_{k=1}^{n}\mathbb{E}[|L_{t_k}-L_{t_{k-1}}|^p]<\infty.$$

The dependence of the noise is measured by

$$\kappa_n(r) := \sup_{a\geq 0}\sum_{i,j=1}^{r}\left|\mathrm{Cov}\left(Z_{a+i,n}, Z_{a+j,n}\right)\right|.$$

We assume that the integral satisfies the local approximation

$$\left\|\int_s^t U_u \circledast dV_u - U_s(V_t - V_s)\right\|_{L^2(\Omega)} \leq C|t-s|^{H+\gamma-\varepsilon}.$$

for parameters $\gamma, \varepsilon > 0$.

Theoretical guarantees

Theorem (González Cázares-Jaramillo)
Under the previous assumptions,

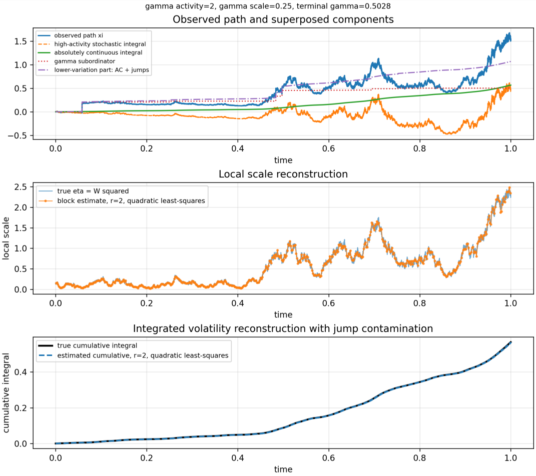
$$\left\| \mathcal{IV}^{(m_n)}(\boldsymbol{X}_n) - \int_0^T \eta_s \, ds \right\|_{L^1(\Omega)} \leq C \left(n^{H - \frac{\gamma(1-p(H-\varepsilon))}{1+\gamma p}} + n^{-\frac{H-\varepsilon+\gamma}{1+\gamma p}} \kappa_n(m_n)^{1/(2p)} \right).$$

where

$$m_n := n^{\frac{H-\varepsilon+\gamma}{\gamma+1/p}}.$$

Numerical illustration

$$\xi_t = \int_0^t W_s^2 \, dW_s + \int_0^t W_s^2 \, ds + J_t,$$



Open problems and conjectures

- Extend the method to noises with α-stable tails.
- Formulate multidimensional versions of the estimator.
- Study random fields, not only stochastic processes.
- To what degree, can the scheme play similar roles as the quadratic variation?

Acknowledgements

Thank you!

Questions and comments are welcome.

References and related work

J. I. González Cázares and A. Jaramillo Gil.
***Robust scale estimation in additive noise via weighted order statistics*.**
Work in progress.

J. I. González Cázares and A. Jaramillo Gil.
***Non-parametric scaling noise estimator for stochastic integrals via order statistics*.**
Work in progress.

V. Fomichov, J. González Cázares and J. Ivanovs.
***Implementable coupling of Lévy process and Brownian motion*.**
Stochastic Processes and their Applications 142, 407–431.