LAMN property for approximation scheme of SDE driven by stable Lévy processes

Laurent DENIS

LMM - Le Mans Université

Tokyo - August 25 th, 2023

$\hookrightarrow$ Joint work with A. Brouste, E. Clément and T. Ngô.

The model

We consider the stochastic differential equation driven by a stable Lévy process as follows

$$X_t = x_0 + \int_0^t b(X_s, \mu)ds + \int_0^t a(X_{s-}, \sigma)dJ_s, \quad t \geq 0, \qquad (1)$$

where $(J_t)_{t\geq 0}$ denotes the standard symmetric β-stable Lévy process characterized by its characteristic function

$$\mathbb{E}(e^{iuJ_1}) = e^{-|u|^\beta}, \quad u \in \mathbb{R}.$$

The unknown parameter is denoted by $\boxed{\theta = (\mu, \sigma, \beta)} \in \Theta$ where Θ is an open subset of $\mathbb{R} \times (0, \infty) \times K$, with some compact $K \subset (0, 2)$.

Our aim

- Prove the LAMN property for approximation scheme associated to the equation and deduce asymptotic properties for high frequency observations of (X_t).
- Build an asymptotically efficient estimator of the parameters with the one-step procedure.
- Illustrate our results with numerical computations.

References

- E. Clement and A. Gloter (SPA,2015) proved the LAMN property for parameter μ in the case

$$X_t = x_0 + \int_0^t b(X_s, \mu)ds + J_t.$$

- A. Brouste and H. Masuda (2018) proved the LAN property for $\theta = (\mu, \sigma, \beta)$ in the case

$$X_t = \mu t + \sigma J_t.$$

- E. Clement, A. Gloter and H. Nguyen (2019) proved LAMN property for $\theta = (\mu, \sigma)$ in the case

$$X_t = x_0 + \int_0^t b(X_s, \mu)ds + \sigma J_t.$$

Assumptions **(A)**

- Let $V_\mu \times V_\sigma$ be a neighborhood of (μ, σ). We assume that $x \mapsto a(x, \sigma)$ is $\mathcal{C}^3$ on $\mathbb{R}$, b is $\mathcal{C}^3$ on $\mathbb{R} \times V_\mu$ and

$$\sup_x \Big(\sup_{\hat{\mu} \in V_\mu} |\partial_x b(x, \hat{\mu})| + |\partial_x a(x, \sigma)| \Big) \le C,$$

$$|\partial_x^2 b(x, \mu)| + |\partial_x^2 a(x, \sigma)| \le C(1 + |x|^p), \quad \text{for some } p > 0,$$

$$\forall x \in \mathbb{R}, \forall \hat{\sigma} \in V_\sigma, \quad a(x, \hat{\sigma}) > 0 \quad \text{and} \quad \sup_{\hat{\sigma} \in V_\sigma} \frac{1}{a(x, \hat{\sigma})} \le C(1 + |x|^p).$$

- We also assume that for any $x \in \mathbb{R}$, $\mu \mapsto b(x, \mu)$ and $\sigma \mapsto a(x, \sigma)$ are $\mathcal{C}^4$ and

$$\sup_{(\hat{\mu}, \hat{\sigma}) \in V_\mu \times V_\sigma} \max_{1 \le \ell \le 4} \big(|\partial_\mu^\ell b(x, \hat{\mu})| + |\partial_\sigma^\ell a(x, \hat{\sigma})| \big) \le C(1 + |x|^p),$$

$$\sup_{\hat{\mu} \in V_\mu} |\partial_x \partial_\mu b(x, \hat{\mu})| \le C(1 + |x|^p).$$

Discretization and approximation Scheme

- We work on time interval $[0,1]$ and consider

$$\forall t \in [0,1],\ X_t = x_0 + \int_0^t b(X_s,\mu)ds + \int_0^t a(X_{s-},\sigma)dJ_s.$$

- For any $n \in \mathbb{N}^*$, we consider the discretization of $[0,1]$: $t_i^n = i/n$ for $i \in \{0,1,\ldots,n\}$.
- To analyse its local asymptotic properties, we are interested in the alternative scheme proposed by E. Clément et A. Gloter (IHP, 2018) on the time points $t_i^n = i/n$ for $i \in \{0,1,\ldots,n\}$

$$\overline{X}_{t_{i+1}^n} = \xi_{t_{i+1}^n - t_i^n}(\overline{X}_{t_i^n},\mu) + a(\overline{X}_{t_i^n},\sigma)(J_{t_{i+1}^n} - J_{t_i^n})$$

where $(\xi_t(x,\mu))_{t\geq 0}$ solves the ODE

$$\xi_t(x,\mu) = x + \int_0^t b(\xi_s(x,\mu),\mu)ds, \quad t \geq 0.$$

Log likehood function

We recall that

$$\overline{X}_{t_{i+1}^n} = \xi_{t_{i+1}^n - t_i^n}(\overline{X}_{t_i^n}, \mu) + a(\overline{X}_{t_i^n}, \sigma)(J_{t_{i+1}^n} - J_{t_i^n})$$

where $(\xi_t(x,\mu))_{t\geq 0}$ solves the ODE

$$\xi_t(x,\mu) = x + \int_0^t b(\xi_s(x,\mu),\mu)ds, \quad t \geq 0.$$

We consider the log-likelihood function based on observing $(\overline{X}_{t_i^n})_{0\leq i\leq n}$ which has an "explicit" expression:

$$\ell_n(\theta) = \sum_{i=0}^{n-1} \log(n^{1/\beta} a(\overline{X}_{t_i^n}, \sigma)^{-1} \phi_\beta(z_n(\overline{X}_{t_i^n}, \overline{X}_{t_{i+1}^n}, \theta))),$$

where ϕ_β denotes the density of $\mathcal{L}(J_1)$ and

$$z_n(x,y,\theta) = \frac{y - \xi_{1/n}(x,\mu)}{n^{-1/\beta} a(x,\sigma)}.$$

Taylor's expansion

For arbitrary bounded $(u_n) \in \mathbb{R}^3$, there exists $0 < \varepsilon < 1$ such that

$$\begin{aligned}\ell_n(\theta + \varphi_n(\theta)u_n) - \ell_n(\theta) &= u_n^\top \Delta_n(\theta) - \frac{1}{2} u_n^\top \mathbb{J}_n(\theta) u + \\ &\frac{1}{3!}((\varphi_n(\theta)u_n)^\top \cdot \partial_\theta \mathbb{I}_n(\theta + \varepsilon\varphi_n(\theta)u_n) \cdot (\varphi_n(\theta)u_n))^\top (\varphi_n(\theta)u_n)\end{aligned}$$

where

$$\begin{aligned}\Delta_n(\theta) &:= \varphi_n(\theta)^\top \partial_\theta \ell_n(\theta), \\ \mathbb{I}_n(\theta) &:= -\partial_\theta^2 \ell_n(\theta) \quad \text{and} \quad \mathbb{J}_n(\theta) := \varphi_n(\theta)^\top \mathbb{I}_n(\theta) \varphi_n(\theta),\end{aligned}$$

and for any $h \in \mathbb{R}^3$ and $\widetilde{\theta} \in \Theta$, we denote

$$h^\top \cdot \partial_\theta \mathbb{I}_n(\widetilde{\theta}) \cdot h = \begin{pmatrix} h^\top \partial_\mu \mathbb{I}_n(\widetilde{\theta}) h \\ h^\top \partial_\sigma \mathbb{I}_n(\widetilde{\theta}) h \\ h^\top \partial_\beta \mathbb{I}_n(\widetilde{\theta}) h \end{pmatrix} \in \mathbb{R}^3.$$

NDNM case: Non degeneracy in the non multiplicative case

In that case:

- $s \mapsto \frac{\partial_\sigma a}{a}(X_s, \sigma)$ is almost surely non constant.
- Almost surely, $\exists t \in (0,1)$, such that $\partial_\mu b(X_t, \mu) \neq 0$.

We introduce the following functions:

$$\begin{aligned} h_\beta(z) =& (\partial_z \phi_\beta / \phi_\beta)(z) \\ k_\beta(z) =& 1 + z h_\beta(z) \\ f_\beta(z) =& (\partial_\beta \phi_\beta / \phi_\beta)(z), \end{aligned}$$

where ϕ_β denotes the density of J_1 (with parameter β).

For this case, we take $\varphi_n(\theta) = \begin{pmatrix} n^{1/2-1/\beta} & 0 & 0 \\ 0 & n^{-1/2} & 0 \\ 0 & 0 & \frac{1}{\log n\sqrt{n}} \end{pmatrix}$.

The LAMN property in that case

Under **(A)** with further $\|a'\|_\infty > 0$: The LAMN property holds:

$$\sup_{u \in \overline{K}} \left| \ell_n(\theta + \varphi_n(\theta)u) - \ell_n(\theta) - \left(u^\top \Delta_n(\theta) - \frac{1}{2} u^\top \mathbb{I}_n(\theta) u \right) \right| \xrightarrow{\mathbb{P}} 0$$

for any compact set $\overline{K} \subset \mathbb{R}^3$, where

$$(\Delta_n(\theta), \mathbb{I}_n(\theta)) \overset{stably}{\Longrightarrow} (\mathbb{I}(\theta)^{1/2}\mathcal{N}, \mathbb{I}(\theta))$$

where $\mathcal{N}$ is a Gaussian variable independent of $\mathbb{I}(\theta)$ and

$$\mathbb{I}(\theta) = \begin{pmatrix} \int_0^1 \frac{\partial_\mu b(X_s,\mu)^2}{a(X_s,\sigma)^2} ds \mathbb{E}(h_\beta^2(J_1)) & 0 \\ 0 & \overline{\mathbb{I}}_{\sigma,\beta} \end{pmatrix}$$

with

$$\overline{\mathbb{I}}_{\sigma,\beta} = \begin{pmatrix} \int_0^1 \frac{\partial_\sigma a(X_s,\sigma)^2}{a(X_s,\sigma)^2} ds \mathbb{E}(k_\beta^2(J_1)) & \frac{1}{\beta^2} \int_0^1 \frac{\partial_\sigma a(X_s,\sigma)}{a(X_s,\sigma)} ds \mathbb{E}(k_\beta^2(J_1)) \\ \frac{1}{\beta^2} \int_0^1 \frac{\partial_\sigma a(X_s,\sigma)}{a(X_s,\sigma)} ds \mathbb{E}(k_\beta^2(J_1)) & \frac{1}{\beta^4} \mathbb{E}(k_\beta^2(J_1)) \end{pmatrix}.$$

There exists a local maximum $\widehat{\theta}_n$ of ℓ_n with probability tending to 1, for which

$$\varphi_n(\theta)^{-1}(\widehat{\theta}_n - \theta) \overset{stably}{\Longrightarrow} \mathbb{I}(\theta)^{-1/2}\mathcal{N}.$$

Remark

Note that the matrix $\mathbb{I}(\theta)$ is invertible a.s. since from the assumption NDNM

$$\frac{1}{\beta^4}\mathbb{E}(k_\beta^2(J_1))\left(\int_0^1 \frac{\partial_\sigma a(X_s,\sigma)^2}{a(X_s,\sigma)^2}ds - \left(\int_0^1 \frac{\partial_\sigma a(X_s,\sigma)}{a(X_s,\sigma)}ds\right)^2\right) > 0, \quad a.s.$$

NDM case (Non degeneracy in the multiplicative case)

In that case:

- $a(x,\sigma) = \sigma \bar{a}(x)$.
- Almost surely, $\exists t \in (0,1)$, such that $\partial_\mu b(X_t,\mu) \neq 0$.

For this case, we take

$$\varphi_n(\theta) = \frac{1}{\sqrt{n}} \begin{pmatrix} n^{1-1/\beta} & 0 & 0 \\ 0 & \varphi_{11,n}(\theta) & \varphi_{12,n}(\theta) \\ 0 & \varphi_{21,n}(\theta) & \varphi_{22,n}(\theta) \end{pmatrix}, \text{ where}$$

$$\begin{cases} \varphi_{11,n}(\theta)\frac{1}{\sigma} + \varphi_{21,n}(\theta)\frac{\log n}{\beta^2} \to \overline{\varphi}_{11}, \quad \varphi_{12,n}(\theta)\frac{1}{\sigma} + \varphi_{22,n}(\theta)\frac{\log n}{\beta^2} \to \overline{\varphi}_{12}, \\ \varphi_{21,n}(\theta) \to \overline{\varphi}_{21}, \quad \varphi_{22,n}(\theta) \to \overline{\varphi}_{22}, \\ \overline{\varphi}_{11}\overline{\varphi}_{22} - \overline{\varphi}_{12}\overline{\varphi}_{21} > 0 \end{cases}$$

Example

One can take $\varphi_{11,n}(\theta) = 1$, $\varphi_{12,n}(\theta) = -\beta^{-2}\sigma\log(n)$, $\varphi_{22,n}(\theta) = 1$ *and* $\varphi_{21,n}(\theta) = 0$ *which yield* $\overline{\varphi}_{11} = \sigma^{-1}$, $\overline{\varphi}_{12} = \overline{\varphi}_{21} = 0$ *and* $\overline{\varphi}_{22} = 1$.

Theorem

1. Assume **(A)** and $\|a'\|_\infty > 0$ then the LAMN property holds:

$$\sup_{u\in\overline{K}}\left|\ell_n(\theta+\varphi_n(\theta)u)-\ell_n(\theta)-\left(u^\top\Delta_n(\theta)-\frac{1}{2}u^\top\mathbb{I}_n(\theta)u\right)\right|\xrightarrow{\mathbb{P}}0,$$

where $(\Delta_n(\theta),\mathbb{I}_n(\theta))\overset{stably}{\Longrightarrow}(\mathbb{I}(\theta)^{1/2}\mathcal{N},\mathbb{I}(\theta))$ with $\mathcal{N}$ is a standard Gaussian variable independent of $\mathbb{I}(\theta)$ and

$$\mathbb{I}(\theta)=\begin{pmatrix}\int_0^1\frac{\partial_\mu b(X_s,\mu)^2}{a(X_s,\sigma)^2}ds\mathbb{E}(h_\beta^2(J_1)) & 0\\ 0 & \overline{\varphi}^\top\overline{\mathbb{I}}_{\sigma,\beta}\overline{\varphi}\end{pmatrix}\quad\text{with}$$

$$\overline{\varphi}=\begin{pmatrix}\overline{\varphi}_{11} & \overline{\varphi}_{12}\\ \overline{\varphi}_{21} & \overline{\varphi}_{22}\end{pmatrix},\ \overline{\mathbb{I}}_{\sigma,\beta}=\begin{pmatrix}\mathbb{E}(k_\beta^2(J_1)) & -\mathbb{E}((k_\beta f_\beta)(J_1))\\ -\mathbb{E}((k_\beta f_\beta)(J_1)) & \mathbb{E}(f_\beta^2(J_1))\end{pmatrix}.$$

2. There exists a local maximum $\widehat{\theta}_n$ of ℓ_n with probability tending to 1, for which

$$\varphi_n(\theta)^{-1}(\widehat{\theta}_n-\theta)\overset{stably}{\Longrightarrow}\mathbb{I}(\theta)^{-1/2}\mathcal{N}.$$

Asymptotic equivalence of two experiments

Recall that the total variation between two probabilities measures P and Q on $(\Omega, \mathcal{F})$ dominated by ν is defined by

$$d_{TV}(P,Q) = \sup_{A \in \mathcal{F}} |P(A) - Q(A)| = \frac{1}{2} \int \left| \frac{dP}{d\nu} - \frac{dQ}{d\nu} \right| d\nu.$$

Let us consider two experiments as follows.

- Experiment $\mathcal{E}^n$ with P_n^θ distribution of $(X_{t_i^n})_{i \in \{0,\dots,n-1\}}$.
- Experiment $\overline{\mathcal{E}}^n$ with $\overline{P}_n^\theta$ distribution of $(\overline{X}_{t_i^n})_{i \in \{0,\dots,n-1\}}$.

In the cas where coefficient a is constant we have:

Theorem (E. Clément, AAP, 2022)

Assume **(A)** *and that* $a(x,\sigma) = \sigma a$ *with* $a > 0$, *we have*

$$d_{TV}(P_n^\theta, \overline{P}_n^\theta) \leq C(a, b, \beta, \sigma) \max \left\{ \frac{1}{\sqrt{n}}, \frac{1}{n^{\frac{4\beta}{\beta+2}}} \right\}.$$

Consequences

The Le Cam distance is bounded by

$$\Delta(\mathcal{E}^n, \overline{\mathcal{E}}^n) \leq \sup_{\theta \in \Theta} d_{TV}(P_n^\theta, \overline{P}_n^\theta).$$

Therefore, the following Corollary holds :

Corollary

Assume **(A)**, *that* $a(x, \sigma) = \sigma a$ *with* $a > 0$ *and that* β *belongs to any compact subset of* $(0, 2)$ *then the experiments are asymptotically equivalent in Le Cam sense i.e.*

$$\lim_{n \to \infty} \Delta(\mathcal{E}^n, \overline{\mathcal{E}}^n) = 0.$$

This means that statistical inference in experiment $\mathcal{E}^n$ *inherits the same asymptotic properties as in experiment* $\overline{\mathcal{E}}^n$.

One step estimation : initial estimator

We follow Todorov (SPA 2013): denoting $\Delta_i X = X_{t_i^n} - X_{t_{i-1}^n}$,

- $V_n^1(p,X) = \sum_{i=2}^n |\Delta_i^n X - \Delta_{i-1}^n X|^p$,
- $V_n^2(p,X) = \sum_{i=4}^n |\Delta_i^n X - \Delta_{i-1}^n X + \Delta_{i-2}^n X - \Delta_{i-3}^n X|^p$,
- $\widehat{\beta}_n^0 = \frac{p \log 2}{\log(V_n^2(p,X)/V_n^1(p,X))} \mathbb{1}_{\{V_n^2(p,X) \neq V_n^1(p,X)\}}$
- $\widehat{\sigma}_n^0$ satisfies $n^{\frac{p}{\widehat{\beta}_n^0}-1} V_n^1(p,X) = \mu_p(\widehat{\beta}_n^0) \int_0^1 |a(X_s, \widehat{\sigma}_n^0)|^p ds$ where $\mu_p(\widehat{\beta}_n^0) = 2^{p/\widehat{\beta}_n^0} \frac{2^p \Gamma(\frac{p+1}{2}) \Gamma(1-p/\widehat{\beta}_n^0)}{\sqrt{\pi}\Gamma(1-p/2)}$
- μ is estimated by maximizing the loglikelihood function $\ell_n\left(\mu, \widehat{\sigma}_n^0, \widehat{\beta}_n^0\right)$ with respect to μ.

Therefore, $\hat{\theta}_n^0 = (\widehat{\mu}_n^0, \widehat{\sigma}_n^0, \widehat{\beta}_n^0)$.

Remark

In the NDM case , $a(X_s, \widehat{\sigma}_n^0) = \widehat{\sigma}_n^0 \, \overline{a}(X_s)$, *we get*

$$\widehat{\sigma}_n^0 = \left[n^{\frac{p}{\widehat{\beta}_n^0}-1} V_n^1(p, \overline{X}) \left(\mu_p(\widehat{\beta}_n^0) \int_0^1 |\overline{a}(X_s)|^p ds \right)^{-1} \right]^{1/p}$$

One step MLE

The one-step MLE $\hat{\theta}_n^1$ is defined by

$$\hat{\theta}_n^1 = \hat{\theta}_n^0 + (\varphi_n(\hat{\theta}_n^0)^{-1}{}^\top \mathbb{I}(\hat{\theta}_n^0)\varphi_n(\hat{\theta}_n^0)^{-1})^{-1}\partial_\theta \ell_n(\hat{\theta}_n^0).$$

Theorem

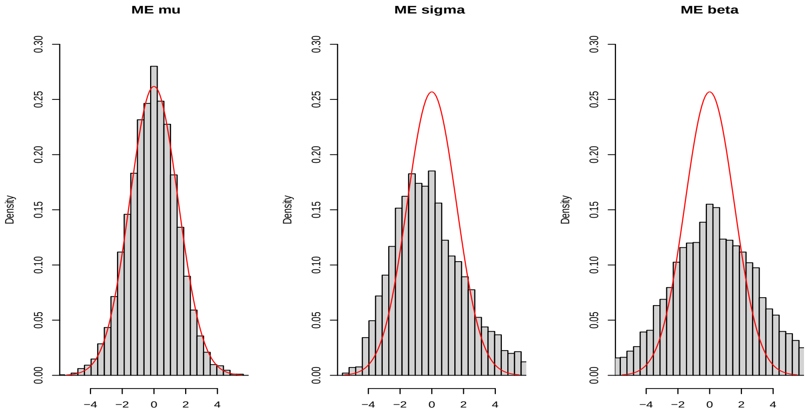
Assume $\beta > 2/3$ then if $p \in \left(\frac{|\beta-1|}{2(\beta\wedge 1)}, \frac{\beta}{2}\right)$, under the LAMN property and sufficiently regular Fisher information matrix, the sequence $(\hat{\theta}_n^1, n \geq 1)$ is asymptotically efficient.

First example: the Ornstein-Uhlenbeck case

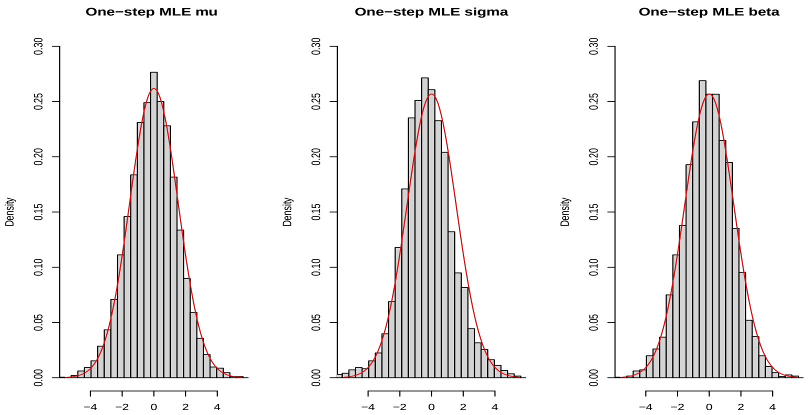
$$X_t = \int_0^t \mu X_s ds + \sigma J_t^{\beta}, \quad t \in [0,1].$$

We simulated 5000 Monte-Carlo samples, $n = 2^{11}$, $\mu = -0.7$, $\sigma = 1$ and $\beta = 1.3$, $p = 0.55$.

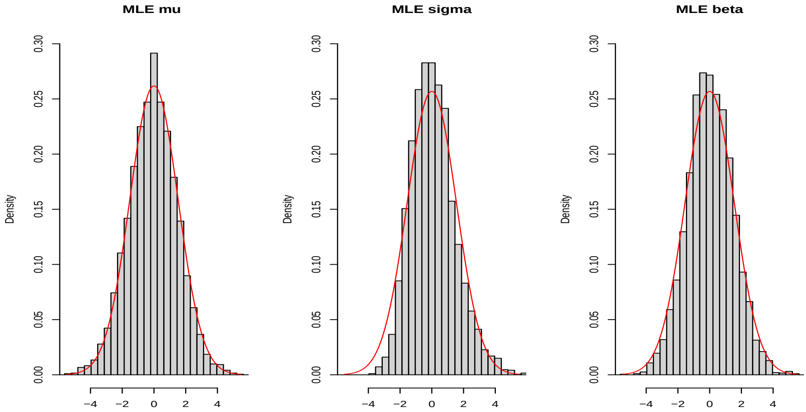
Re-scaled errors for the initial Estimator



Re-scaled errors for the one-step MLE



Re-scaled errors for the MLE



Computational times

Here, for this O-U model, the computational time of MLE is 274028.2 minutes while the one by one-step estimation is 8168.976 minutes. This means that the estimation by one-step procedure is about 33 times faster than the maximum likelihood estimation.

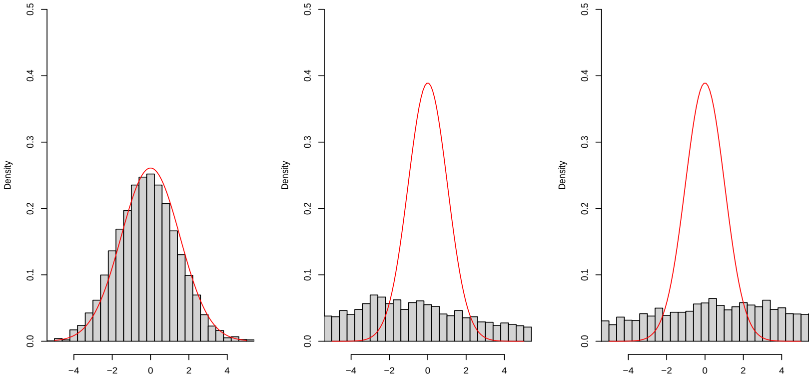
Another example : NDNM case

We consider the square-root model

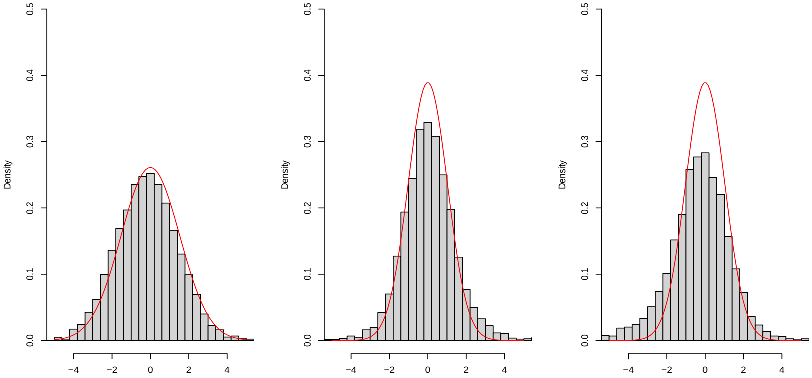
$$X_t = \int_0^t \mu X_s ds + \sigma \int_0^t \exp(\sigma \sin(X_s))\, dJ_s^\beta, \quad t \in [0,1].$$

We simulated 5000 Monte-Carlo samples with $n = 2^{12}$, $\mu = -0.5$, $\sigma = 1$ and $\beta = 1.5$, $p = 0.7$.

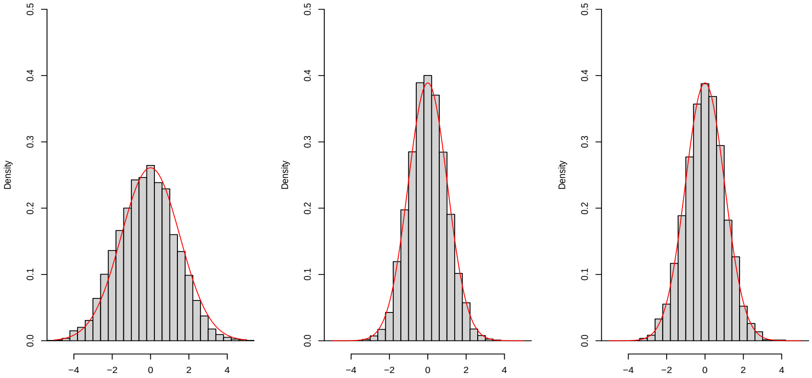
Re-scaled errors for the Initial Estimator



Re-scaled errors for the one-step MLE



Re-scaled errors for the MLE



A few comments concerning the computational times

Here, for this mean-square model, the computational time of MLE is 48960.93 minutes while the one by one-step estimation is 3583.805 minutes. This means that the estimation for one-step procedure is about 13 times faster than the maximum likelihood estimation, but gives similar approximation.