

## SUPPLEMENT TO "JUMP FILTERING AND EFFICIENT DRIFT ESTIMATION FOR LÉVY-DRIVEN SDE'S"

BY ARNAUD GLOTER<sup>\*</sup>, DASHA LOUKIANOVA<sup>\*</sup> AND HILMAR MAI<sup>†</sup>

*Université d'Evry Val d'Essonne* <sup>\*</sup> and *ENSAE-ParisTech* <sup>†</sup>

### CONTENTS

7	Examples and numerical results . . . . .	1
	7.1 Finite activity . . . . .	1
	7.1.1 Ornstein-Uhlenbeck-type processes . . . . .	2
	7.1.2 Hyperbolic diffusions with jumps . . . . .	3
	7.2 Infinite activity . . . . .	5
8	Proofs of auxiliary results . . . . .	7
	8.1 Proof of Lemma 2.1 . . . . .	7
	8.2 Proof of Theorem 5.3 . . . . .	8
	8.3 Proof of Lemma 6.1 . . . . .	9
	8.4 Proof of Lemma 6.3 . . . . .	9
	8.5 Proof of Lemma 6.4 . . . . .	11
	8.6 Equivalent form of the identifiability Assumption 6 . . . . .	11
	References . . . . .	12
	Author's addresses . . . . .	12

**7. Examples and numerical results.** In this section we consider concrete applications of the drift estimator in popular jump diffusion models and investigate the numerical performance in finite sample studies. We consider both examples with finite and infinite jump activity.

In the first part, we give explicit drift estimator for Ornstein-Uhlenbeck and Hyperbolic diffusion and study numerically the performance of the estimator for a finite activity jump process. In the second part, we consider the case of infinite activity jump process, as the Lévy process is chosen to be a stable process.

We consider here for convenience only linear models in the drift parameter that lead to explicit maximum likelihood estimators in order to avoid the need for numerical maximization techniques. Note that the method developed in this work applies equally well to non-linear models by using standard maximization methods on the discretized and jump-filtered likelihood function (3.4).

*7.1. Finite activity.* In this section we consider two different jump diffusion models with finite activity jumps. The first model will consist of Ornstein-Uhlenbeck-type processes that

recently became popular in financial modeling (cf. for example [1]), and the second model will be a 'Hyperbolic diffusion'.

The jump process  $L$  is of compound Poisson type in the case of finite activity such that it can be written as

$$(7.1) \quad L_t = \sum_{i=1}^{N_t} Z_i, \text{ for } t \geq 0,$$

where  $(N_t)_{t \geq 0}$  is a Poisson process with intensity  $\lambda$  and  $(Z_i)_{i \in \mathbb{N}}$  are i.i.d. real random variables independent of  $N$ , with distribution  $\nu/\lambda$ .

7.1.1. *Ornstein-Uhlenbeck-type processes.* Suppose that we have given a discrete sample

$$(7.2) \quad X_{t_0}, \dots, X_{t_n} \quad \text{for } t_i = i\Delta_n \text{ and } i = 0, \dots, n,$$

of an Ornstein-Uhlenbeck-type (OU) process  $(X_t)_{t \geq 0}$  that is defined as a solution of the stochastic differential equation

$$(7.3) \quad dX_t = (\theta_2 - \theta_1 X_t) dt + \sigma dW_t + dL_t \quad X_0 = x,$$

where  $(W_t)_{t \geq 0}$  is a standard Brownian motion and  $(L_t)_{t \geq 0}$  a pure jump Lévy process. Our goal is to estimate the unknown drift parameter  $\theta = (\theta_1, \theta_2) \in \mathbb{R}^2$ . The volatility parameter  $\sigma > 0$  might be unknown and can be seen as a nuisance parameter. The jump component  $(L_t)_{t \geq 0}$  will be of compound Poisson type, i.e. it can be written as in (7.1) with intensity  $\lambda$  and the jump heights  $Z_i$  are supposed to be i.i.d. with Gaussian distribution  $\mathcal{N}(0, 1)$ .

From (3.4) and (3.5) we find that the FMLE for  $\theta$  is the solution  $\hat{\theta}_n^{\text{OU}} = (\hat{\theta}_{1,n}^{\text{OU}}, \hat{\theta}_{2,n}^{\text{OU}})$  to the following set of linear equations in  $\theta_1$  and  $\theta_2$ .

$$(7.4) \quad \begin{aligned} \theta_1 &= \frac{\theta_2 I_n(X, 1) - \sum_{i=1}^n X_{t_{i-1}} \Delta_i^n X \mathbf{1}_{|\Delta_i^n X| \leq v_n^i}}{I_n(X, 2)}, \\ \theta_2 &= \frac{\sum_{i=1}^n \Delta_i^n X \mathbf{1}_{|\Delta_i^n X| \leq v_n^i} + \theta_1 I_n(X, 1)}{t_n}, \end{aligned}$$

where we introduced the functional

$$(7.5) \quad I_n(X, p) := \sum_{i=1}^n X_{t_{i-1}}^p \Delta_i^n Id \quad \text{for } p \in \mathbb{R}.$$

The FLME for the first component of  $\theta$  results in

$$\hat{\theta}_{1,n}^{\text{OU}} = \left( 1 - \frac{I_n(X, 1)^2}{t_n I_n(X, 2)} \right)^{-1} \frac{I_n(X, 1) \sum_{i=1}^n \Delta_i^n X \mathbf{1}_{|\Delta_i^n X| \leq v_n^i} - t_n \sum_{i=1}^n X_{t_{i-1}} \Delta_i^n X \mathbf{1}_{|\Delta_i^n X| \leq v_n^i}}{t_n I_n(X, 2)}.$$

The second component  $\hat{\theta}_{2,n}^{\text{OU}}$  follows now easily by plugging  $\hat{\theta}_{1,n}^{\text{OU}}$  into (7.4).

We give simulation results for  $\hat{\theta}_{1,n}^{\text{OU}}$  in the situation where  $\theta_2 = 0$ . The given mean and standard deviation are each based on 5000 Monte Carlo samples of  $\hat{\theta}_{1,n}^{\text{OU}}$ . We also give the average number of jumps that were detected by the jump filter.

We choose  $v_n = \Delta_n^{1/2-\varepsilon} = \Delta_n^{0.49}$  and consider first the choice of constant weights  $a_n^i = 5$  (recall (3.3)). The results of the simulations are given in columns 3–5 of Table 1 for  $\sigma = 1$ . It appears that the estimator performs well as soon as the discretization distance  $\Delta_n$  is sufficiently small. On the other hand, in the case where  $\sigma = 3$ , the same estimator appears almost useless due to a large bias even for small  $\Delta_n$  (see columns 3–4 of Table 2). This comes from the fact that many increments of the Brownian part  $\sigma(W_{t_i} - W_{t_{i-1}})$  are larger than the threshold  $v_n^i = 5 \times \Delta_n^{0.49}$  in the situation  $\sigma = 3$ . Hence, it is important for finite sample properties of the estimator to take into account the volatility of  $X^c$  for the choice of the jump threshold. As it is better that the estimator does not depend on the knowledge of the value of  $\sigma$ , we introduce the threshold  $v_n^i = 5 \times \widehat{\sigma}_n^i \times \Delta_n^{0.49}$ , where  $(\widehat{\sigma}_n^i)^2$  is an estimation of the quadratic variation of the process on  $K = 30$  past observations,

$$(7.6) \quad (\widehat{\sigma}_n^i)^2 = \frac{1}{\Delta_n K} \sum_{l=1}^K (\Delta_{i-l}^n X)^2$$

and for convenience we set  $(\widehat{\sigma}_n^i)^2 = (\widehat{\sigma}_n^{K+1})^2$  for  $1 \leq i \leq K$ . The estimation of  $\sigma$  by  $\widehat{\sigma}_n^i$  might be upward biased if a jump occurs on  $(t_{i-K}, t_{i-1}]$ . However, as the jumps of the process are isolated, it is asymptotically unlikely that a jump has to be filtered in the interval  $(t_{i-1}, t_i]$  while the estimation of the volatility on the  $(t_{i-K}, t_{i-1}]$  is biased by another jump. We see in columns 6–8 of Tables 1–2, that with this choice, the estimator performs now well, for both values of  $\sigma$ .

Let us remark that the number of filtered jumps is much smaller than the true expected number of jumps (see column 5 of Table 1 and column 8 of Table 1–2). Since, the aim is to estimate the drift parameter, and not the number of jumps, this strong bias is unimportant here. However the number of jump is the main quantity used to estimate the jump intensity and is thus a relevant quantity in many context. For instance, in [5] the author exhibit an asymptotic expansion for the bias when estimating the intensity from the number of jumps, and proposes to calibrate the threshold by some minimization of this bias.

As the number of 'filtered jumps' is a decreasing function of the threshold, it is possible to find the threshold  $v_n^i = a(\Delta_n)^{0.49}$ ,  $a > 0$ , such as the average number of 'filtered jumps' is equal to the expected number of jumps  $\lambda t_n$ . In Table 3 we report the values of  $a$ , numerically obtained by successive guesses, and show how the estimator behaves with this threshold. It appears that the estimator has a higher bias than when the number of jumps was underestimated. Hence, it seems preferable, in some situations, to filter less jumps than the true number of jumps.

*7.1.2. Hyperbolic diffusions with jumps.* In this section we apply the drift estimator to hyperbolic diffusion processes with jumps. They are defined as solutions  $(X_t)_{t \geq 0}$  of the

$t_n$	$n$	$a_n^i = 5$			$a_n^i = 5 \times \widehat{\sigma}_n^i$		
		mean	std dev	jumps filt	mean	std dev	jumps filt
10	100	1.84	0.47	1.23	1.90	0.52	0.43
	400	2.00	0.45	4.16	2.05	0.52	3.04
	1000	2.06	0.45	5.98	2.08	0.49	5.28
50	500	1.77	0.22	6.19	1.80	0.24	3.00
	2000	1.95	0.20	20.8	1.95	0.23	16.3
	5000	1.99	0.20	30.0	2.00	0.22	27.0
100	3000	1.91	0.14	34.9	1.93	0.17	25.9
	$10^4$	1.98	0.14	60.1	1.98	0.16	54.0
	$3 \times 10^4$	2.00	0.14	76.2	2.00	0.15	73.3

TABLE 1

Monte Carlo estimates of mean and standard deviation from 5000 samples of  $\hat{\theta}_{1,n}^{OU}$  for an OU process with compound Poisson jumps with intensity  $\lambda = 1$ ,  $\sigma = 1$  and true parameter  $\theta_1 = 2$ .

$t_n$	$n$	$a_n^i = 5$			$a_n^i = 5 \times \widehat{\sigma}_n^i$		
		mean	std dev	jumps filt	mean	std dev	jumps filt
10	100	1.33	0.51	9.16	1.94	0.59	$8.4 \times 10^{-3}$
	400	1.41	0.53	35.8	2.10	0.64	0.27
	1000	1.44	0.54	85.0	2.14	0.65	1.31
50	500	1.26	0.23	45.4	1.80	0.24	0.06
	2000	1.33	0.30	180	1.98	0.28	1.53
	5000	1.35	0.23	425.0	2.01	0.28	6.63
100	3000	1.30	0.16	273	1.95	0.19	0.11
	$10^4$	1.34	0.17	850	1.99	0.19	13.2
	$3 \times 10^4$	1.36	0.16	2386	2.01	0.19	36.4

TABLE 2

Monte Carlo estimates of mean and standard deviation from 5000 samples of  $\hat{\theta}_{1,n}^{OU}$  for an OU process with compound Poisson jumps with intensity  $\lambda = 1$ ,  $\sigma = 3$  and true parameter  $\theta_1 = 2$ .

$t_n$	$n$	$a$	$v_n^i = a\Delta_n^{0.49}$		
			mean	std dev	jumps filt
10	400	2.551	1.88	0.41	10.0
50	5000	2.897	1.94	0.19	50.0
100	$10^4$	2.9	1.93	0.14	100

TABLE 3

Monte Carlo estimates of mean and standard deviation from 5000 samples of  $\hat{\theta}_{1,n}^{OU}$  for an OU process with compound Poisson jumps with intensity  $\lambda = 1$ ,  $\sigma = 1$ , true parameter  $\theta_1 = 2$ , and with threshold such as the estimated number of jumps is unbiased.

$t_n$	$n$	$a_n^i = 5$			$a_n^i = 5 \times \widehat{\sigma}_n^i$		
		mean	std dev	jumps filt	mean	std dev	jumps filt
10	100	1.06	0.61	1.22	1.12	0.67	0.42
	400	1.10	0.57	4.10	1.11	0.62	3.04
	1000	1.12	0.56	5.95	1.12	0.56	5.31
50	500	0.97	0.26	6.18	0.93	0.27	2.86
	2000	1.01	0.23	20.8	1.01	0.24	16.1
	5000	1.02	0.22	30.0	1.02	0.23	27.0
100	3000	0.99	0.16	35.0	0.99	0.17	25.7
	$10^4$	1.01	0.15	60.0	1.01	0.16	54.0
	$3 \times 10^4$	1.01	0.15	75.6	1.01	0.15	73.1

TABLE 4

Monte Carlo estimates of mean and standard deviation from 5000 samples of  $\hat{\theta}_{1,n}^{\text{hyp}}$  for an Hyperbolic process with compound Poisson jumps with intensity  $\lambda = 1$ ,  $\sigma = 1$  and true parameter  $\theta_1 = 1$ .

following SDE:

$$dX_t = -\frac{\theta X_t}{(1 + X_t^2)^{1/2}} dt + \sigma dW_t + dL_t, \quad X_0 = x.$$

Here, the drift parameter  $\theta > 0$  and the diffusion coefficient  $\sigma > 0$  are unknown and we aim at estimating  $\theta$  from discrete observations  $X_{t_0}, \dots, X_{t_n}$  of  $X$ , where  $t_i = i\Delta_n$  for  $\Delta_n > 0$  and  $i = 0, \dots, n$ . The driving Lévy process  $(L_t)_{t \geq 0}$  is again a Compound Poisson Process with intensity  $\lambda$  and the jump heights  $Z_i$  are supposed to be i.i.d. with Gaussian distribution  $\mathcal{N}(0, 1)$ . From (3.4) we obtain an explicit expression for the estimator,

$$\hat{\theta}_n^{\text{hyp}} = -\sum_{i=1}^n \frac{X_{t_{i-1}}}{(1 + X_{t_{i-1}}^2)^{1/2}} \Delta_i^n X \mathbf{1}_{|\Delta_i^n X| \leq v_n} \left( \sum_{i=1}^n \frac{X_{t_{i-1}}^2 \Delta_i^n Id}{(1 + X_{t_{i-1}}^2)} \right)^{-1}.$$

The Hyperbolic model has a less mean reverting drift than the Ornstein Ullhenbeck model. However, the simulation results are close to the one obtained from the O.U. model (see Table 4).

**7.2. Infinite activity.** In this section we investigate estimation of the drift when the driving Lévy process is of infinite jump activity. This is of course a more challenging problem with regards to the approximation of the continuous martingale part i.e. the jump filtering problem, since we have to distinguish a diffusion component from a process that jumps infinitely often in finite time intervals.

We consider again the O.U. model (7.3), with  $\theta_2 = 0$ , and where the driving Lévy process  $(L_t)_{t \geq 0}$  is an  $\alpha$ -stable process such that  $E[e^{iuL_1}] = e^{-|u|^\alpha}$ . It is known that its Lévy-Khintchine triplet is  $(0, 0, \nu)$ , with a Lévy measure of the form  $\nu(dx) = \kappa_\alpha dx/|x|^{1+\alpha}$ , for some constant  $\kappa_\alpha > 0$ . In Table 5 we show the estimation results for  $\sigma = 1$  and  $\alpha = 0.9$ . We see that the estimator based on the threshold  $v_n^i = 5 \times \Delta_n^{0.49}$  has a serious bias which decay slowly as  $\Delta_n \rightarrow 0$  (see columns 3–5 in table 5). The estimator with threshold

$t_n$	$n$	$a_n^i = 5$			$a_n^i = 5 \times \widehat{\sigma}_n^i$			$a_n^i = 5 \times \widetilde{\sigma}_n^i$		
		mean	std dev	j filt	mean	std dev	j filt	mean	std dev	j filt
10	100	1.62	0.41	8.32	1.85	0.44	1.42	1.79	0.35	2.46
	400	1.88	0.30	13.2	2.00	0.46	4.80	1.95	0.29	6.85
	1000	1.97	0.26	12.7	2.04	0.43	8.85	1.99	0.29	11.5
50	500	1.60	0.19	42.2	1.81	0.17	8.61	1.79	0.12	13.8
	2000	1.82	0.097	66.9	1.96	0.20	25.0	1.94	0.084	34.5
	5000	1.93	0.081	63.7	1.99	0.18	45	1.97	0.084	57.5
100	3000	1.78	0.080	128	1.94	0.13	41.1	1.92	0.059	58.3
	$10^4$	1.93	0.053	128	1.98	0.12	90.1	1.97	0.059	114
	$3 \times 10^4$	1.98	0.082	192	1.99	0.11	166	1.98	0.084	201

TABLE 5

Monte Carlo estimates of mean and standard deviation from 5000 samples of  $\hat{\theta}_{1,n}^{OU}$  for an OU process with  $\alpha$ -stable jumps,  $\alpha = 0.9$ ,  $\sigma = 1$ , and true parameter  $\theta_1 = 2$ .

$t_n$	$n$	$a_n^i = 5$			$a_n^i = 5 \times \widehat{\sigma}_n^i$			$a_n^i = 5 \times \widetilde{\sigma}_n^i$		
		mean	std dev	j filt	mean	std dev	j filt	mean	std dev	j filt
10	100	1.36	0.57	24.5	1.81	0.41	2.0	1.78	0.27	3.47
	400	1.65	0.46	60.7	1.96	0.84	5.28	1.94	0.17	7.76
	1000	1.82	0.29	85.8	2.00	0.38	8.11	1.98	0.13	10.7
50	500	1.36	0.46	131	1.82	0.15	11.1	1.81	0.094	19.0
	2000	1.69	0.30	317	1.96	0.18	26.7	1.94	0.034	38.9
	5000	1.81	0.17	447	1.99	0.14	40.6	1.98	0.029	53.4
100	3000	1.64	0.32	266	1.94	0.11	46.1	1.93	0.031	68.9
	$10^4$	1.82	0.15	892	1.98	0.16	81.4	1.98	0.017	107
	$3 \times 10^4$	1.89	0.082	797	1.99	0.074	121	1.99	0.013	140

TABLE 6

Monte Carlo estimates of mean and standard deviation from 5000 samples of  $\hat{\theta}_{1,n}^{OU}$  for an OU process with  $\alpha$ -stable jumps,  $\alpha = 0.5$ ,  $\sigma = 1$ , and true parameter  $\theta_1 = 2$ .

depending on local estimation of the quadratic variation (7.6) presents a much reduced bias, with a slightly higher standard deviation (see columns 6–8 in table 5). In Table 6 we show the results for  $\alpha = 0.5$ , and the situation is rather similar to  $\alpha = 0.9$ .

Inspecting the behaviour of the estimator, we realized that  $\widehat{\sigma}_n^i$  tends to overestimate  $\sigma$  due to the presence of the infinite number of jumps of the stable process, and may take large values. We propose to reduce the contribution of the stable process in the estimation of the local volatility by removing in the sum (7.6) the contribution of the biggest increment  $\max_{l \in \{i-K, \dots, i-1\}} |\Delta_l^n X|^2$ . This tends to suppress the contribution of the largest jump of the stable process and considerably reduces the upward bias for the estimation of the local volatility. We note  $\widetilde{\sigma}_n^i$  this correction of the quantity  $\widehat{\sigma}_n^i$  and show in the last three columns of Tables 5–6 the behavior of the corresponding estimator. It appears that this estimator works well, with a small bias and a much reduced standard deviation.

**8. Proofs of auxiliary results.** In this section we give the proofs of several results stated and used in Sections 2, 5 and 6 of the paper. We end the section by proving that the Assumption 6 may be rewritten in a simpler form.

First, we give the proof of the ergodicity results and moment bounds of Lemma 2.1. The proof is based on [3].

8.1. *Proof of Lemma 2.1.* Denote

$$\mathcal{G}f(x) = \frac{1}{2}\sigma^2(x)f''(x) + b(\theta, x)f'(x),$$

$$\mathcal{J}f(x) = \int_{\mathbb{R}} (f(x + z\gamma(x)) - f(x))\nu(dz),$$

for any  $f$  such that the two previous expressions are defined and set

$$\mathcal{A} = \mathcal{G} + \mathcal{J}.$$

Let  $q > 2$ ,  $q$  even and  $f^*(x) = |x|^q$ . We show that  $f^*$  satisfies the drift condition

$$\mathcal{A}f^* \leq -c_1f^* + c_2,$$

where  $c_1 > 0, c_2 > 0$ .

Using Taylor's formula together with Assumptions 3 (i), (iii) and 4 we can write

$$|\mathcal{J}f^*(x)| \leq \int_{\mathbb{R}} |z\gamma(x)| \sup_{u \in [x, x+z\gamma(x)]} |f^{*\prime}(u)| \nu(dz) \leq C|\gamma(x)||x|^{q-1} \int_{\mathbb{R}} |z|(1+|z|)^{q-1} \nu(dz) = o(|x|^q)$$

as  $x \rightarrow \infty$ . Using Assumption 3 (ii) and (iv) we get

$$\mathcal{G}f^*(x) = \frac{1}{2}\sigma^2(x)q(q-1)x^{q-2} + b(\theta, x)xqx^{q-2} \leq -C|x|^2qx^{q-2} + o(|x|^q) \leq -Cqf^*(x) + o(|x|^q),$$

for some  $C > 0$ . As  $\mathcal{A}f^*(x)$  is locally bounded, using two previous displays we can choose  $c_2 > 0$  and  $c_1 > 0$  such that for all  $x \in \mathbb{R}$ ,

$$\mathcal{A}f^*(x) \leq -c_1f^*(x) + c_2.$$

Hence, Assumption 3\* from [3] holds and using Theorem 2.2 from [3] we get then

$$(8.1) \quad \sup_{s \geq 0} \mathbf{E}[|X_s^\theta|^q] < \infty$$

and using Fatou's lemma results in

$$\sup_{s \geq 0} \mathbf{E}[|X_{s-}^\theta|^q] < \infty.$$

Hence we proved the assertion (3). Using Assumption 2 and the Theorem 2.1 from [3] we get for all  $\theta \in \Theta$  that  $X^\theta$  admits the unique invariant distribution  $\pi^\theta$ ,  $f^* \in \mathcal{L}^1(\pi^\theta)$  and the ergodic theorem holds. We proved (1) and (2). We continue with the proof of (4). Using ergodic theorem, for all  $q > 0$ ,

$$\lim_{t \rightarrow \infty} \frac{1}{t} \int_0^t |X_s^\theta|^q ds = \pi^\theta(|x|^q), \quad P - a.s.$$

Moreover, using Jensen's inequality and the bound (8.1) we get the uniform integrability of the family  $\{\frac{1}{t} \int_0^t |X_s^\theta|^q ds, t > 0\}$ :

$$E \left( \frac{1}{t} \int_0^t |X_s^\theta|^q ds \right)^{1+\varepsilon} \leq \frac{1}{t} \int_0^t [E |X_s^\theta|^{q(1+\varepsilon)}] ds \leq C,$$

where  $C > 0$ , and hence

$$\lim_{t \rightarrow \infty} \frac{1}{t} \int_0^t E |X_s^\theta|^q ds = \pi^\theta(|x|^q).$$

□

8.2. *Proof of Theorem 5.3.* From (3.1)–(3.2),

$$\begin{aligned} \ell_t(\theta^* + \frac{h}{\sqrt{t}}) - \ell_t(\theta^*) &= -\frac{1}{2} \int_0^t \frac{(b(\theta^* + \frac{h}{\sqrt{t}}, X_s) - b(\theta^*, X_s))^2 ds}{\sigma^2(X_s)} \\ &+ \int_0^t \frac{(b(\theta^* + \frac{h}{\sqrt{t}}, X_s) - b(\theta^*, X_s))}{\sigma(X_s)} dW_s \\ &= -\frac{1}{2} \int_0^1 \int_0^1 \left( \frac{1}{t} \int_0^t \frac{h^\top (\nabla b(\theta^* + \frac{hu}{\sqrt{t}}, X_s) \nabla b^\top(\theta^* + \frac{hu'}{\sqrt{t}}, X_s) h)}{\sigma^2(X_s)} ds \right) dud' \\ &+ \frac{1}{\sqrt{t}} \int_0^t \frac{\nabla b^\top(\theta^*, X_s) h}{\sigma(X_s)} dW_s + R_t \end{aligned}$$

where

$$R_t := \int_0^t \frac{(b(\theta^* + \frac{h}{\sqrt{t}}, X_s) - b(\theta^*, X_s))}{\sigma(X_s)} dW_s - \frac{1}{\sqrt{t}} \int_0^t \frac{\nabla b^\top(\theta^*, X_s) h}{\sigma(X_s)} dW_s.$$

Using Assumption 5, 7, 8 and Lemma 2.1, for all fixed  $r > 0, r' > 0$  such that  $\theta^* + r \in \Theta, \theta^* + r' \in \Theta$  we obtain

$$\lim_{t \rightarrow \infty} \frac{1}{t} \int_0^t \frac{h^\top \nabla b(\theta^* + r, X_s) \nabla b^\top(\theta^* + r', X_s) h}{\sigma^2(X_s)} ds = \int_{\mathbb{R}} \frac{h^\top \nabla b(\theta^* + r, x) \nabla b^\top(\theta^* + r', x) h}{\sigma^2(x)} d\pi(x)$$



$P$ -a.s. and see that this last limit is finite. Moreover, with Assumption 5, 7, 8 and Lemma 2.1 it can be shown that this convergence is uniform with respect to  $r$  and  $r'$ . Hence for  $hu/\sqrt{t} \rightarrow 0$ ,  $hu'/\sqrt{t} \rightarrow 0$ , it gives that  $P - a.s.$

$$(8.2) \quad \lim_{t \rightarrow \infty} \int_0^1 du \int_0^1 du' \frac{1}{t} \int_0^t \frac{h^\top \nabla b(\theta^* + \frac{hu}{\sqrt{t}}, X_s) \nabla b^\top(\theta^* + \frac{hu'}{\sqrt{t}}, X_s) h}{\sigma^2(X_s)} ds \\ = \int_{\mathbb{R}} \frac{h^\top \nabla b(\theta^*, x) \nabla b^\top(\theta^*, x) h}{\sigma^2(X_s)} d\pi(x) = h^\top I(\theta^*) h.$$

Using Markov inequality

$$(8.3) \quad P(|R_t| \geq \varepsilon) \leq \frac{\text{Var} R_t}{\varepsilon^2} \leq \frac{\|h\|^2}{\varepsilon^2} \frac{1}{t} \int_0^t \left( \frac{\|h\|}{\sqrt{t}} \right)^{2\kappa} E \left( \frac{K_1^2(X_s)}{\sigma^2(X_s)} \right) ds,$$

where  $K_1$  is a Holder constant of  $\nabla b$  and is supposed to be at most of polynomial growth. Using ergodic theorem in mean, we obtain  $R_t \rightarrow 0$  in  $P$  probability.

Due to the CLT for martingales in [2]

$$\frac{1}{\sqrt{t}} \int_0^t \frac{\nabla b^\top(\theta^*, X_s) h}{\sigma(X_s)} dW_s \rightarrow \mathcal{N}(0, h^\top I(\theta^*) h)$$

in distribution. Combining the latter equation with (8.2)–(8.3), we obtain (5.14). This implies together with Theorem 5.2 that  $\bar{\theta}_t$  is asymptotically efficient in the sense of the Hájek-Le Cam convolution theorem.  $\square$

8.3. *Proof of Lemma 6.1.* The first claim follows easily from the two lemmas and theorem 66 on p. 339 in [4] together with the fact that  $\sup_{s \geq 0} E[|X_s|^p] < \infty$ , for all  $p \geq 1$ , by Lemma 2.1. The second claim follows from proposition 3.1 in [6]. The third claim follows easily from Burkholder's Inequality (see [4] p.222) and Lemma 2.1 again.  $\square$

8.4. *Proof of Lemma 6.3.* We need to introduce some notations. For  $z > 0$ , we define  $U_z = \int_{t_{i-1}}^{t_i} \int_{|y| \geq 1/z} \mu(ds, dy)$  the number of jumps of  $(L_s)$ ,  $s \in (t_{i-1}, t_i]$ , with a size greater than  $1/z$ , and we set  $U_0 = 0$ . It is clear that  $(U_z)_{z \geq 0}$  is a process whose increments are independent and distributed with Poisson laws. Hence, it is a Poisson process, and by a simple computation we can show that it has a jump intensity equal to  $\bar{U}(z) := (t_i - t_{i-1})z^{-2}(\nu(z^{-1}) + \nu(-z^{-1}))$ , where  $\nu(z) = \nu(dz)/dz$  exists by Assumption 4 (iii).

We define the filtration generated by the process  $(U_z)_{z \geq 0}$ , by setting for all  $z \geq 0$ ,  $\mathcal{G}_z = \sigma\{U_y; y \leq z\}$ . We note  $Z_1^*$  the first jump time of the process  $U$ , which is a stopping time. By construction, we have that  $1/Z_1^*$  is the size of the biggest jumps of the Lévy process  $L$  on  $(t_{i-1}, t_i]$ , or with the notations of Lemma 4.2 that,  $1/Z_1^* = |\Delta L_{T_i^*}|$ , where  $|\Delta L_{T_i^*}| = \max\{|\Delta L_s|; s \in (t_{i-1}, t_i]\}$ .

Moreover, we can write

$$\sum_{t_{i-1} < s \leq t_i; s \neq T_i^*} |\Delta L_s| = \int_{t_{i-1}}^{t_i} \int_{|y| < 1/Z_1^*} |y| \mu(ds, dy) = \int_{(Z_1^*, \infty)} \frac{1}{z} dU_z,$$

where we have used that  $\Delta L_{T_i^*}$  is the only jump with the maximal size  $1/Z_1^*$ . Hence, we have

$$\begin{aligned} P((N_n^i)^c \cap (A_n^i)^c) &= P\left(|\Delta L_{T_i^*}| > \frac{3v_n^i}{\gamma_{min}}; \sum_{t_{i-1} < s \leq t_i; s \neq T_i^*} |\Delta L_s| > \frac{v_n^i}{\gamma_{max}}\right) \\ &= P\left((Z_1^*)^{-1} > \frac{3v_n^i}{\gamma_{min}}; \int_{(Z_1^*, \infty)} z^{-1} dU_z > \frac{v_n^i}{\gamma_{max}}\right) \\ &\leq P\left((Z_1^*)^{-1} > \frac{3\underline{a}v_n}{\gamma_{min}}; \int_{(Z_1^*, \infty)} z^{-1} dU_z > \frac{\underline{a}v_n}{\gamma_{max}}\right) \\ &= E\left[1_{\{(Z_1^*)^{-1} > \frac{3\underline{a}v_n}{\gamma_{min}}\}} P\left(\int_{(Z_1^*, \infty)} z^{-1} dU_z > \frac{\underline{a}v_n}{\gamma_{max}} \mid \mathcal{G}_{Z_1^*}\right)\right] \\ &\leq \frac{\gamma_{max}}{\underline{a}v_n} E\left[1_{\{(Z_1^*)^{-1} > \frac{3\underline{a}v_n}{\gamma_{min}}\}} E\left(\int_{(Z_1^*, \infty)} z^{-1} dU_z \mid \mathcal{G}_{Z_1^*}\right)\right], \end{aligned}$$

where we have used the Markov inequality in the last line. Using now that  $(U_z)_{z \geq 0}$  is a Poisson process with an explicit jump intensity  $\bar{U}(z) = (t_i - t_{i-1})z^{-2}(\nu(z^{-1}) + \nu(-z^{-1}))$ , we deduce,

$$P((N_n^i)^c \cap (A_n^i)^c) \leq \frac{\gamma_{max}}{\underline{a}v_n} E\left[1_{\{(Z_1^*)^{-1} > \frac{3\underline{a}v_n}{\gamma_{min}}\}} E\left(\int_{(Z_1^*, \infty)} z^{-1} \bar{U}(z) dz \mid \mathcal{G}_{Z_1^*}\right)\right].$$

But, by a simple change of variable,  $\int_{(Z_1^*, \infty)} z^{-1} \bar{U}(z) dz = (t_i - t_{i-1}) \int_{|y| < 1/Z_1^*} |y| \nu(y) dy \leq \Delta_n \int_{\mathbb{R}} |y| \nu(y) dy$ . We conclude

$$\begin{aligned} P((N_n^i)^c \cap (A_n^i)^c) &\leq \frac{\gamma_{max}}{\underline{a}v_n} \Delta_n \left(\int_{\mathbb{R}} |y| \nu(y) dy\right) P\left[(Z_1^*)^{-1} > \frac{3\underline{a}v_n}{\gamma_{min}}\right] \\ &\leq C \frac{\Delta_n}{\underline{a}v_n} P\left(\mu((t_{i-1}, t_i] \times [(-\infty, -\frac{3\underline{a}v_n}{\gamma_{min}}) \cup (\frac{3\underline{a}v_n}{\gamma_{min}}, +\infty)]) \geq 1\right) \\ &\leq C \frac{\Delta_n^2}{\underline{a}v_n} \int_{|z| > \frac{3\underline{a}v_n}{\gamma_{min}}} \nu(dz), \end{aligned}$$

where  $C > 0$ . The lemma is proved.  $\square$

8.5. *Proof of Lemma 6.4.* We use the notations introduced in Section 6 and denote the event  $\tilde{A}_n^i = \{L \text{ has at most one jump on } (t_{i-1}, t_i]\}$ , which is well defined since the jump intensity of  $L$  is finite under Assumption 4 (ii'). Then

$$K_n^i \cap (N_n^i)^c \cap \tilde{A}_n^i \subset \{|\Delta_i^n X| \leq v_n^i\} \cap \{|\Delta_i^n X^J| \geq 3v_n^i\} \subset \{|\Delta_i^n X^c| \geq v_n^i\},$$

and by Remark 2 we deduce  $P(K_n^i \cap (N_n^i)^c \cap \tilde{A}_n^i) = O(\Delta_n^2)$ .

Since the process  $L$  has a finite jump activity, it is easy to see that  $P((\tilde{A}_n^i)^c) = O(\Delta_n^2)$ . This implies that  $P(K_n^i \cap (N_n^i)^c) = O(\Delta_n^2)$ .

Let us now control  $E[\sup_{\theta \in \Theta} |B_n^1(\theta)|]$ . Recall the expression of  $B_n^1(\theta)$ ,

$$\begin{aligned} B_n^1(\theta) &= - \sum_{i=1}^n f(\theta, X_{t_{i-1}}) (\Delta_i^n X^J) 1_{K_n^i \cap (N_n^i)^c} \\ &= - \sum_{i=1}^n f(\theta, X_{t_{i-1}}) \left( \Delta_i^n X^c - \Delta_i^n \mathbf{1}_{|\Delta_i^n X| \leq v_n^i} \right) 1_{K_n^i \cap (N_n^i)^c}. \end{aligned}$$

By the Hölder inequality, the  $L^1$  norm of  $\sup_{\theta \in \Theta} |B_n^1(\theta)|$  is upper bounded by

$$\begin{aligned} &E \sup_{\theta \in \Theta} \left| \sum_{i=1}^n f(\theta, X_{t_{i-1}}) \left( \Delta_i^n X^c - \Delta_i^n X \mathbf{1}_{|\Delta_i^n X| \leq v_n^i} \right) 1_{K_n^i \cap (N_n^i)^c} \right| \\ &\leq \sum_{i=1}^n \left( E \sup_{\theta \in \Theta} |f(\theta, X_{t_{i-1}})|^p (|\Delta_i^n X^c| + v_n^i)^p \right)^{1/p} \left( P(K_n^i \cap (N_n^i)^c) \right)^{1/q}, \end{aligned}$$

for  $1/p + 1/q = 1$ . Using Lemma 2.1 (3) and choosing  $1/q = 1 - \varepsilon/2$ , we deduce

$$E[\sup_{\theta \in \Theta} |B_n^1(\theta)|] = O(nv_n \Delta_n^{2-\varepsilon}) = O(n\Delta_n^2).$$

□

### 8.6. Equivalent form of the identifiability Assumption 6.

PROPOSITION 8.1. *Under Assumptions 1 to 4, the Assumption 6 is equivalent to the condition*

$$\forall (\theta, \theta') \in \Theta^2, \quad \text{such that } \theta \neq \theta', \quad b(\theta, \cdot) \neq b(\theta', \cdot).$$

PROOF. It is sufficient to prove that if  $b(\theta, \cdot) \neq b(\theta', \cdot)$  then  $\int_{\mathbb{R}} \frac{(b(\theta, x) - b(\theta', x))^2}{\sigma^2(x)} d\pi^\theta(x) > 0$ . But if  $b(\theta, \cdot) \neq b(\theta', \cdot)$ , the continuous function  $x \mapsto (b(\theta, x) - b(\theta', x))^2 / \sigma^2(x)$  does not vanish on some non empty open set  $\mathcal{O}$ . It remains to show that  $\pi^\theta(\mathcal{O}) > 0$ . It is proved in [3] (see equation (13) p.43) that for all  $\Delta > 0$ ,  $x \in \mathbb{R}$ , and  $\mathcal{O}$  non empty, open set,  $P(X_\Delta^\theta \in \mathcal{O} \mid X_0^\theta = x) > 0$ . From this, we deduce that

$$\pi^\theta(\mathcal{O}) = \int_{\mathbb{R}} P(X_\Delta^\theta \in \mathcal{O} \mid X_0^\theta = x) d\pi^\theta(x) > 0.$$

□

**References.**

- [1] BARNDORFF-NIELSEN, O. E. and SHEPHARD, N. (2001). Non-Gaussian Ornstein-Uhlenbeck-based models and some of their uses in financial economics. *J. R. Stat. Soc., Ser. B, Stat. Methodol.* **63** 167-241.
- [2] KÜCHLER, U. and SØRENSEN, M. (1999). A note on limit theorems for multivariate martingales. *Bernoulli* **5** 483-493.
- [3] MASUDA, H. (2007). Ergodicity and exponential  $\beta$ -mixing bounds for multidimensional diffusions with jumps. *Stochastic Process. Appl.* **117** 35-56.
- [4] PROTTER, P. E. (2004). *Stochastic integration and differential equations. 2nd ed.* Applications of Mathematics 21. Berlin: Springer.
- [5] SHIMIZU, Y. (2008). Some remarks on estimation of diffusion coefficient for jump-diffusions from finite samples. *Bull. Inform. Cybernet.* **40** 51-60.
- [6] SHIMIZU, Y. and YOSHIDA, N. (2006). Estimation of Parameters for Diffusion Processes with Jumps from Discrete Observations. *Statistical Inference for Stochastic Processes* **9** 227-277.

LAMME, UMR CNRS 8071  
UNIVERSITÉ D'EVRY VAL D'ESSONNE  
91037 ÉVRY CEDEX  
FRANCE  
E-MAIL: [arnaud.gloter@univ-evry.fr](mailto:arnaud.gloter@univ-evry.fr)  
[dasha.loukianova@univ-evry.fr](mailto:dasha.loukianova@univ-evry.fr)

CENTRE DE RECHERCHE EN ECONOMIE ET STATISTIQUE  
ENSAE-PARISTECH  
92245 MALAKOFF  
FRANCE  
E-MAIL: [hilmar.mai@ensae.fr](mailto:hilmar.mai@ensae.fr)