ADAPTIVE KERNEL ESTIMATION OF THE LÉVY DENSITY

MÉLINA BEC*, CLAIRE LACOUR**

April 5, 2011

ABSTRACT. This paper is concerned with adaptive kernel estimation of the Lévy density $n(x)$ for pure jump Lévy processes. The sample path is observed at n discrete instants in the "high frequency" context ($\Delta = \Delta_n$ tends to zero while $n\Delta_n$ tends to ∞). We construct a collection of kernel estimators of the function $g(x) = xn(x)$ and propose two methods of local adaptive selection of the bandwidth. The quadratic pointwise risk of the adaptive estimators is studied in both cases. The rate of convergence is proved to be optimal up to a logarithmic factor. We give examples and simulation results for processes fitting in our framework.

KEYWORDS. Adaptive Estimation; High frequency; Pure jump Lévy process; Nonparametric Kernel Estimator.

1. INTRODUCTION

Consider $(L_t, t \geq 0)$ a real-valued Lévy process with characteristic function given by:

$$\psi_t(u) = \mathbb{E}(\exp iuL_t) = \exp\left(t \int_{\mathbb{R}} (e^{iux} - 1)N(dx)\right). \tag{1}$$

We assume that the Lévy measure $N(dx)$ admits a density $n(x)$ and that the function $g(x) = xn(x)$ satisfies:

- (G1): $\int_{\mathbb{R}} |x|n(x)dx = \int_{\mathbb{R}} |g(x)|dx < \infty$.

Under these assumptions, $(L_t, t \geq 0)$ is a pure-jump Lévy process without drift and with finite variation on compact sets. Moreover $\mathbb{E}(|L_t|) < \infty$ (see Bertoin (1996)). Suppose that we have discrete observations $(L_{k\Delta}, k = 1, ..., n)$ with sampling interval Δ. Our aim in this paper is the nonparametric adaptive kernel estimation of the function $g(x) = xn(x)$ based on these observations under the asymptotic framework n tends to ∞. This subject has been recently investigated by several authors. Figueroa-López and Houdré (2006) use a penalized projection method to estimate the Lévy density on a compact set separated from 0. Other authors develop an estimation procedure based on empirical estimations of the characteristic function $\psi_\Delta(u)$ of the increments $(Z_k^\Delta = L_{k\Delta} - L_{(k-1)\Delta}, k = 1, \dots, n)$ and its derivatives followed by a Fourier inversion to recover the Lévy density. For low-frequency data (Δ is fixed), we can quote Watteel and Kulperger (2003), Jongbloed and van der Meulen (2006), Neumann and Reiß (2009) and Comte and Genon-Catalot (2010). In the high frequency context, which is our concern in this paper, the problem is simpler since $\psi_\Delta(u) \to 1$ when $\Delta \to 0$. This implies that $\psi_\Delta(u)$ need not to be estimated and can simply be replaced by 1 in the estimation procedures. This is what is done in Comte and

* UMR CNRS 8145 MAP5, Université Paris Descartes, ** Laboratoire de mathématiques, Université Paris-Sud 11.

Genon-Catalot (2009) and Comte and Genon-Catalot (2010). These authors start from the equality:

$$\mathbb{E}\left[Z_k^\Delta e^{iuZ_k^\Delta}\right] = -i\frac{\psi'_\Delta(u)}{\Delta}, \tag{2}$$

where $g^*(u) = \int e^{iux} g(x)dx$ is the Fourier transform of g well defined under (G1).

Then, as $\psi_\Delta(u) \simeq 1$, equation (2) writes $\mathbb{E}\left[Z_k^\Delta e^{iuZ_k^\Delta}\right] \simeq g^*(u)$ and gives an estimator of $g^*(u)$ as follows:

$$\frac{1}{n\Delta}\sum_{k=1}^n Z_k^\Delta e^{iuZ_k^\Delta}$$

Now, to recover g the authors apply Fourier inversion with cutoff parameter m. Here, we rather introduce a kernel to make inversion possible:

$$\frac{1}{n\Delta}\sum_{k=1}^n Z_k^\Delta K^*(uh) e^{iuZ_k^\Delta}$$

which is in fact the Fourier transform of $1/(nh\Delta)\sum_{k=1}^n Z_k^\Delta K((Z_k^\Delta - x)/h)$. At the end, in the high frequency context, a direct method without Fourier inversion can be applied. Indeed, a consequence of (2) is that the empirical distribution:

$$\hat{\mu}_n(dz) = \frac{1}{n\Delta}\sum_{k=1}^n Z_k^\Delta \delta_{Z_k^\Delta}(dz)$$

weakly converges to $g(z)dz$. This suggests to consider kernel estimators of g of the form

$$\hat{g}_h(x_0) = K_h \star \hat{\mu}_n(x) = \frac{1}{n\Delta}\sum_{k=1}^n Z_k^\Delta K_h(Z_k^\Delta - x) \tag{3}$$

where $K_h(x) = (1/h)K(x/h)$ and $K(-x) = K(x)$ is a symetric kernel such that

$$\int |K(u)|du < \infty \text{ and } \int K(u)du = 1.$$

Below, we study the pointwise $\mathbb{L}^2$-risk of the estimators $(\hat{g}_h(x))$ and evaluate the rate of convergence of this risk as $n \to \infty$, $\Delta = \Delta_n$ tends to 0 and $h = h_n$ tends to 0. This is done under Hölder regularity assumptions for the function g. Note that a pointwise study involving a kernel estimator can be found in van Es et al. (2007) for more specific compound Poisson processes.

In this paper, we study local adaptive bandwidth selection (which the previous authors do not consider). For a given value $x_0 \in \mathbb{R}$, we define two ways of selecting a bandwidth $\hat{h}(x_0)$ such that the resulting adaptive estimator $\hat{g}_{\hat{h}(x_0)}(x_0)$ automatically reaches the optimal rate of convergence corresponding to the unknown regularity of the function g. In the first method (Bandwidth Selection I) we define an estimator $\hat{\beta}(x_0)$ of the unknown regularity β of g and plug in the estimated value in the optimal bandwidth $h_{opt}(\beta)$ that is deduced from the $\mathbb{L}^2$ risk study. The second method (Bandwidth Selection II) follows the scheme developped by Goldenshluger and Lepski (2011) for density estimation. Introducing iterated kernel estimators:

$$\hat{g}_{h,h'}(x_0) = K_{h'} \star \hat{g}_h(x_0) = K_h \star \hat{g}_{h'}(x_0),$$

a data-driven choice $\hat{h}_{(x_0)}$ is defined and the adaptive estimator is given by $\hat{g}_{\hat{h}_{(x_0)}}(x_0)$.

In Section 2, we give notations and assumptions. In Section 3, we study the pointwise mean square error (MSE) of $\hat{g}_h(x_0)$ given in (3) for g belonging to a Hölder class of regularity. We present the two bandwidth selection methods in Section 4 and 5 together with a risk bound for each adaptive estimator. The rate of convergence of the risk is optimal up to a logarithmic loss which is expected in adaptive pointwise context. Examples and simulations in our framework are discussed in Section 6 and some concluding remarks are given in Section 7. Proofs are gathered in Section 8.

2. Notations and assumptions

We present the assumptions on the kernel K and on the function g required to study the estimation given by (3). First, we set some notations. We denote by u^* the Fourier transform of u, $u^*(y) = \int e^{iyx}u(x)dx$. and by $\|u\|$, $< u, v >$, $u \star v$ the quantities

$$\|u\|^2 = \int |u(x)|^2 dx,$$

$$< u, v >= \int u(x)\overline{v}(x)dx \text{ with } z\overline{z} = |z|^2 \text{ and } u \star v(x) = \int u(y)\bar{v}(x-y)dy.$$

Moreover, for any integrable and square-integrable functions u, u_1, u_2 such that $u^* \in \mathbb{L}^1$, we have:

$$(u^*)^*(x) = 2\pi u(-x) \text{ and } \langle u_1, u_2\rangle = (2\pi)^{-1}\langle u_1^*, u_2^*\rangle. \tag{4}$$

2.1. **Kernel Assumptions.**

Let $K : \mathbb{R} \to \mathbb{R}$ be integrable and such that

$$\forall u, \int K(u)du = 1 \text{ and } K(u) = K(-u). \tag{5}$$

We recall that for $l \geq 1$ an integer, $K : \mathbb{R} \to \mathbb{R}$ is a kernel of order l if functions $u \mapsto u^j K(u), j = 0, 1, ..., l$ are integrable and satisfy

$$\int u^j K(u)du = 0, \forall j \in \{1, ..., l\}. \tag{6}$$

The kernels will have to satisfy some of the following assumptions.

- (Ker[1]) K is of order $l = \lfloor \beta \rfloor$ and $\int |x|^\beta |K(x)|dx < +\infty$.
  We define the parameter β later.
- (Ker[2]) $\|K\|_2 < +\infty$.
- (Ker[3]) $K^* \in \mathbb{L}^1$.

Assumptions (Ker[i]), $i = 1, 2, 3$ are standard when working on problems of estimation by kernel methods.

We note that there is a way to build a kernel of order l. Indeed, let u be a symmetric and integrable function such that $u \in \mathbb{L}^2$,$u^* \in \mathbb{L}^1$ and $\int u(y)dy = 1$, and set for any given integer l,

$$K(t) = \sum_{k=1}^{l} \binom{k}{l} (-1)^{k+1} \frac{1}{k} u\left(\frac{t}{k}\right) \tag{7}$$

The kernel K defined by (7) is a kernel of order l which satisfies assumptions (Ker[i]) $i = 1, 2, 3$ (see Kerkyacharian et al. (2001) and Goldenshluger and Lepski (2011)).

2.2. **Assumptions on g.** The target function g has to satisfy the following assumptions.

- (G1) $\|g\|_1 < +\infty$.
- (G2) $\|g\|_2 < +\infty$.

We introduce Hölder classes.

Definition 2.1. *(Hölder class) Let $\beta > 0$, $L > 0$ and let $l = \lfloor \beta \rfloor$ be the largest integer strictly smaller than β. The Hölder class $H(\beta, L)$ on $\mathbb{R}$ is the set of all functions $f : \mathbb{R} \longrightarrow \mathbb{R}$ such that derivative $f^{(l)}$ exists and verifies:*

$$|f^{(l)}(x_0) - f^{(l)}(x_0')| \leq L|x_0 - x_0'|^{\beta - l}, \forall x_0, x_0' \in \mathbb{R}. \tag{8}$$

This allows us to introduce the following assumption

- (G3) $g \in H(\beta, L)$.

(G3) is a classical regularity assumption in nonparametric estimation; it allows to quantify the bias (see Tsybakov (2009)).

- (G4) $M := \|g'\|_\infty < +\infty$.
- (G5) $xg \in \mathbb{L}^1$ and $P := \frac{1}{2\pi} \int |g^{*\prime}(u)| du < +\infty$.

3. Risk bound

In this section, the bandwidth h is fixed, thus we omit the subscript h for the sake of simplicity: we denote $\hat{g} = \hat{g}_h$. Let $x_0 \in \mathbb{R}$. The usual bias variance decomposition of the MSE yields:

$$MSE(x_0, h) := \mathbb{E}[(\widehat{g}(x_0) - g(x_0))^2] = \mathbb{E}[(\widehat{g}(x_0) - \mathbb{E}[\widehat{g}(x_0)])^2] + (\mathbb{E}[\widehat{g}(x_0)] - g(x_0))^2$$

But the bias needs further decomposition:

$$b(x_0)^2 := (\mathbb{E}[\widehat{g}(x_0)] - g(x_0))^2 \quad \leq \quad 2b_1(x_0)^2 + 2b_2(x_0)^2$$

with the usual bias,

$$b_1(x_0) = K_h \star g(x_0) - g(x_0),$$

and the bias resulting from the approximation of $\psi_\Delta(u)$ by 1,

$$b_2(x_0) = \mathbb{E}[\widehat{g}(x_0)] - K_h \star g(x_0).$$

We can provide the following bias bound:

Lemma 3.1. *Under (Ker[1]), (G1), (G3) and (G4) we have:*

$$b(x_0)^2 \quad \leq \quad c_1 h^{2\beta} + c'_1 \Delta^2.$$

with $c_1 = 2\left(L/l! \int |K(v)||v|^\beta dv\right)^2$ and $c'_1 = 2M^2(\|g\|_1 \|K\|_1)^2$, where M is defined in (G4).

Moreover, the variance is controlled as follows:

Lemma 3.2. *Under (Ker[2]), (Ker[3]), (G1), (G2) and (G5), we have*

$$Var[\widehat{g}(x_0)] \leq \frac{1}{nh\Delta}\|K\|_2^2(P+\|g\|_2^2\Delta) \leq c_2\frac{1}{nh\Delta}+c'_2\frac{1}{nh}$$

with $c_2 = P\|K\|_2^2$ *where* P *is defined in (G5), and* $c'_2 = \|g\|_2^2\|K\|_2^2$.

Lemmas 3.1 et 3.2 lead us to the following risk bound:

Proposition 3.1. *Under (Ker[1])-(Ker[3]) , (G1)-(G5) we have*

$$MSE(x_0,h) \leq c_1h^{2\beta}+c_2\frac{1}{nh\Delta}+c'_2\frac{1}{nh}+c'_1\Delta^2. \tag{9}$$

Recall that $\Delta = \Delta_n$ is such that $\lim_{n\to+\infty}\Delta_n = 0$, thus $1/nh$ is negligible compared to $1/nh\Delta$. For Δ^2, we set the following condition:

$$\Delta^3 \leq \frac{1}{nh}, nh \to +\infty. \tag{10}$$

We can then proceed to find the convergence rate, based on the first two terms. It is easily seen that the optimal choice of h is $h_{opt} \propto ((n\Delta)^{-\frac{1}{2\beta+1}})$ and the associated rate has order $O\left((n\Delta)^{-\frac{2\beta}{2\beta+1}}\right)$.

Proposition 3.2. *Under the assumptions of Proposition 3.1 and under condition (10), the choice* $h_{opt} \propto ((n\Delta)^{-\frac{1}{2\beta+1}})$ *minimizes the risk bound (9) and gives* $MSE(x_0,h_{opt}) = O((n\Delta)^{-\frac{2\beta}{2\beta+1}})$.

We stress the fact that this result is new as previous works in estimation (adaptive or not) do not deal with the pointwise risk.

4. Bandwidth selection I

As β is unknown, we need a data-driven selection of the bandwidth. Firstly, we use the method introduced in Lepski (1991).

We set the following additional condition:

$$\frac{1}{n} << \Delta << \frac{1}{n^{1/3}} \tag{11}$$

In other words, the asymptotic context is $n\Delta \to +\infty$ and $n\Delta^3 \to 0$. It allows to ensure condition (10) for all h. It follows from Proposition 3.2 that if $h = h_{opt}(\beta)$ then,

$$\mathbb{E}[\hat{g}_{h_{opt}(\beta)}(x_0)-g(x_0)]^2 \leq C(n\Delta)^{-\frac{2\beta}{2\beta+1}} =: Cu^2_{n,opt}(\beta). \tag{12}$$

We define

$$u_n^2(\beta) = \left(\frac{n\Delta}{\log(n\Delta)}\right)^{-\frac{2\beta}{2\beta+1}} \text{ and } h(\beta) = \left(\frac{n\Delta}{\log(n\Delta)}\right)^{-\frac{1}{2\beta+1}}.$$

We set $\hat{g}_\beta = \hat{g}_{h(\beta)}$.
In the following, we use the letter β for the true parameter. Moreover we assume that β belongs to the set $A = \{\alpha_1,...,\alpha_D\}$ where $\alpha_1 < ... < \alpha_D$ and $D = \lfloor(n\Delta)^\xi\rfloor$ with ξ a fixed positive real. We choose $\hat{\beta}(x_0)$ as follows:

$$\hat{\beta}(x_0) = \max\left\{\alpha \in A,\ \forall\alpha' \in A/\alpha' \leq \alpha,\ |\hat{g}_\alpha(x_0)-\hat{g}_{\alpha'}(x_0)| \leq cu_n(\alpha')\right\}$$

with $c \geq c' + 2N + 2B_1$, $c' = 8\|K\|_2\sqrt{P + \|g\|_2^2}\sqrt{2\xi + 4}$ and $B_1 = M\|g\|_1\|K\|_1$, and

$$N := \frac{L}{l!}\int |K(v)||v|^\beta dv \text{ and } P := \frac{1}{2\pi}\int |g^{*\prime}(v)|dv. \tag{13}$$

Heuristically, $\hat\beta(x_0)$ estimates the regularity index β.

Then, the estimator of interest is: $\hat g_{\hat\beta(x_0)}(x_0)$. We can prove the following bound on the quadratic risk of our adaptive estimator.

Theorem 4.1. *Under the assumptions of Proposition 3.1 and if the $|Z_i|$'s admit a moment of order z with $z > 10 + 4\xi + (5 + 2\xi)/\alpha_1$, we have, under condition (11),*

$$\mathbb{E}[|\hat g_{\hat\beta(x_0)}(x_0) - g(x_0)|] \leq c u_n(\beta). \tag{14}$$

Theorem 4.1 states that our data driven estimator automatically reaches the optimal rate stated in Proposition 3.2, up to a logarithmic factor. Note that for pointwise adaptive density estimation, a logarithmic loss also occurs and has been proved to be nevertheless optimal (see Butucea (2001)).

5. Bandwidth selection II

Now, we follow ideas given in Goldenshluger and Lepski (2010) for density estimation. We set:

$$V(h) = C\frac{\log(n\Delta)}{nh\Delta} \text{ with } C = (c'/2) \times \|K\|_2^2(P + \|g\|_2^2),\ c' \in \mathbb{R}^+, \tag{15}$$

where P is defined in (13). Note that $V(h)$ has the same order as the variance multiplied by $\log(n\Delta)$.

We define $\hat g_{h,h'}(x_0) = K_{h'} \star \hat g_h(x_0) = K_h \star \hat g_{h'}(x_0)$. So we have

$$\hat g_{h,h'}(x_0) = \frac{1}{n\Delta}\sum_{k=1}^n Z_k^\Delta K_{h'} \star K_h(Z_k^\Delta - x_0). \tag{16}$$

Lastly we set

$$A(h, x_0) = \sup_{h'\in H}\{|\hat g_{h,h'}(x_0) - \hat g_{h'}(x_0)|^2 - V(h')\}_+ \text{ with } H = \{\tfrac{j}{M}, 1 \leq j \leq M\}, \tag{17}$$

and M to be specified later.
The adaptive bandwidth h is chosen as follows:

$$\hat h_{(x_0)} = \arg\min_{h\in H}\{A(h, x_0) + V(h)\}.$$

To simplify notations, we set $\hat h = \hat h_{(x_0)}$.

Theorem 5.1. *Let the assumptions of Proposition 3.1 hold and assume that there exists β_0 (known) such that $\beta > \beta_0$. Choose $M = \left[(n\Delta)^{1/(2\beta_0+1)}\right]$ in (17) and take c' in (15) such that $c' \geq 96(1 \vee \|K\|_\infty)$. Then if the $|Z_i|$'s admit a moment of order z such that $z \geq 2(3 + 2/\beta_0)$, we have*

$$\mathbb{E}[|g(x_0) - \hat g_{\hat h}(x_0)|^2] \leq C\left\{\inf_{h\in H}\{\|g - \mathbb{E}[\hat g_h]\|_\infty^2 + V(h)\} + \frac{\log(n\Delta)}{n\Delta}\right\}$$

The comment following Theorem 4.1 also applies to the result of Theorem 5.1.

Let us discuss and compare the two methods. The advantages of Method I are the following. Method I provides an estimator $\hat{\beta}(x_0)$ of β. It is based directly on the differences between the functions $\hat{g}_{h(\alpha)}$ which are easily calculated as empirical means while Method II requires the evaluations of $K_h \star K_{h'}$, sometimes numerically costly.
Nevertheless, the constants required to implement the two methods are not similar. In Method I, N is difficult to estimate (because of L and β). It could be possible to compute a rough estimator. In Method II, we are able to offer empirical natural quantities to replace $\|g\|$ or $\|g'\|_\infty$. Hence, Method II does not require a constant depending on the regularity of the function.

We choose to implement the second method to illustrate our results on various examples.

6. Examples and Simulations

We give some illustrating examples.

Example 1. Compound Poisson processes. Let $L_t = \sum_{i=1}^{N_t} Y_i$, where (N_t) is a Poisson process with constant intensity c and (Y_i) is a sequence of i.i.d random variables with density f independent of the process (N_t). Then, (L_t) is a Lévy process with characteristic function

$$\psi_t(u) = \exp\left(ct\int_{\mathbb{R}}(e^{iux}-1)f(x)dx\right). \tag{18}$$

Its Lévy density is $n(x) = cf(x)$ and thus $g(x) = cxf(x)$.
In the simulations below we have chosen for f the standard Gaussian density. Thus, $g(x) = cxf(x) = cxe^{-x^2/2}/\sqrt{2\pi}$ and $g^*(u) = ciue^{-u^2/2}$. Assumptions (G1), (G2), (G4) and (G5) hold for g. Moreover g belongs to a Hölder class of regularity β for all $\beta > 0$. Thus the rate is close to $(n\Delta/\log(n\Delta))^{-1}$.

Example 2. The Lévy-Gamma process. Let $\alpha > 0$, $\gamma > 0$. The Lévy-Gamma process (L_t) with parameters (γ, α) is such that, for all $t > 0$, L_t has Gamma distribution with parameters $(\gamma t, \alpha)$, i.e the density:

$$\frac{\alpha^{\gamma t}}{\Gamma(\gamma t)} x^{\gamma t-1} e^{-\alpha x} \mathbb{1}_{x\geq 0}.$$

The Lévy density is $n(x) = \gamma x^{-1}e^{-\alpha x}\mathbb{1}_{x>0}$ so that $g(x) = \gamma e^{-\alpha x}\mathbb{1}_{x>0}$ satisfies assumptions (G1), (G2) and (G5). We have $g^*(u) = \gamma/(\alpha - iu)$.
Although g does not satisfy (G3)-(G4), we have implemented the estimation method to study its robustness with regard to this assumption. Note that the simulations results are good in this case also.

Example 3. The bilateral Lévy-Gamma process (Küchler and Tappe (2008)). Consider X, Y two independent random variables, X with distribution $\Gamma(\gamma, \alpha)$ and Y with distribution $\Gamma(\gamma', \alpha')$.Then, $Z = X - Y$ has bilateral gamma distribution with parameters

$(\gamma, \alpha, \gamma', \alpha')$, denoted by $\Gamma(\gamma, \alpha; \gamma', \alpha')$. The characteristic function of Z is equal to:

$$\psi(u) = \left(\frac{\alpha}{\alpha - iu}\right)^{\gamma} \left(\frac{\alpha'}{\alpha' + iu}\right)^{\gamma'} = \exp\left(\int_{\mathbb{R}} (e^{iux} - 1) n(x) dx\right), \tag{19}$$

with

$$n(x) = x^{-1} g(x),$$

and, for $x \in \mathbb{R}$,

$$g(x) = \gamma e^{-\alpha x} 1_{(0,+\infty)(x)} - \gamma' e^{-\alpha' |x|} 1_{(-\infty,0)}(x).$$

The bilateral Gamma process (L_t) has characteristic function $\psi_t(u) = \psi(u)^t$. As the previous example, (G1), (G2) and (G5) hold, and not (G3)-(G4). But simulations are quite satisfactory.

We have implemented the estimation method for processes listed in Examples 1-3 with different kernels (Figure 1) :

1. Epanechnikov Kernel: $K(x) = (3/4)(1 - x^2)\mathbb{1}_{|x| \leq 1}$.
2. Laplace Kernel: $K(x) = (\theta/2)\exp(-|x|/\theta)$ with $\theta = 1$.
3. Kernel K in Section 2 (see (7)) with $l = 2$ and $l = 3$. We choose $u(x) = 1/\sqrt{2\pi}\exp(-x^2/2)$, the Gaussian density.

The first two kernels are of order 1. Kernel [3.] is of order $l = 2$ or $l = 3$. For the implementation, a difficulty is the proper calibration of the constant c' in (15). This is usually done by a large number of previous simulations. We have chosen $c' = 1$ as the adequate value for a variety of models, kernels and number of observations. As expected, when $n\Delta$ increases, the estimation is better. For clarity, we have computed the Mean Integrated Square Error (MISE) of the estimators whose value gets smaller as $n\Delta$ increases. Figure 1 plots ten estimated curves corresponding to our three examples with in the first column $n = 10000$, $\Delta = 0.1$, and in the second $n = 100000$, $\Delta = 0.1$. Globally, it can be seen from Figure 1 that the results are stable even for smaller values of $n\Delta$. For illustration, each example presented in Figure 1 is computed with a specific kernel. We did not observe significant improvement when using the higher order kernel ([3.]). This is surprising because the theoretical part suggests that it should improve the bias term. Lastly, we can see that the estimated curves are not very regular when $n\Delta$ is not large enough, this is due to the pointwise bandwidth selection.

7. Concluding remarks

In this paper, we have investigated in the high frequency framework the nonparametric estimation of the Lévy density $n(.)$ of pure jump Lévy process. This is done through the estimation of the function $g(x) = xn(x)$, and we use kernel estimators. We have studied two methods of bandwidth selection. The rate we obtained are not the optimal ones, but the logarithmic loss which occurs is negligible. Implementation of Method II on simulated data illustrates the quality of the estimation even when some assumptions are not fulfilled.

8. Proofs

For many proofs, we need the two following Propositions (see Proposition 2.1 in Comte and Genon-Catalot (2010) and Proposition 2.1 in Comte and Genon-Catalot (2009)).

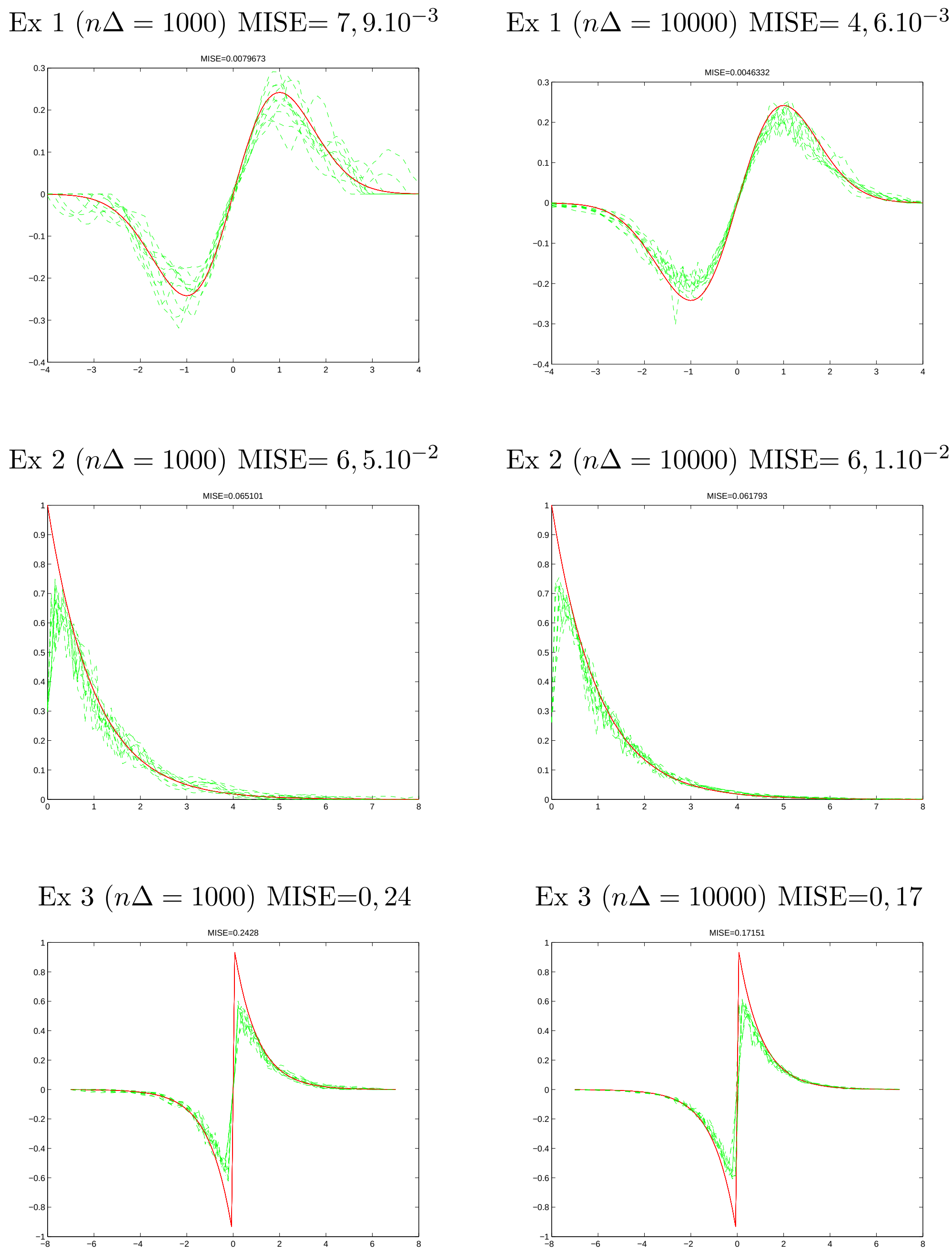

FIGURE 1. Adaptive estimators of g. First line: Example 1 (Compound Poisson and Gaussian density) with $c = 1$, using the Epanechnikov kernel. Second line: Example 2 (Lévy Gamma) with $\alpha = \gamma = 1$ using the Laplace kernel. Third line: Example 3 (Bilateral Lévy Gamma) with $\alpha = \gamma = \alpha' = \gamma' = 1$ using kernel [3.].

Proposition 8.1. *Let $p \geq 1$ an integer such that $\int_{\mathbb{R}} |x|^{p-1}|g(x)|dx < +\infty$ and $\mathbb{E}(|Z_1^\Delta|^p) < +\infty$. Then for $M_k = \int_{\mathbb{R}} x^{k-1}g(x)dx < +\infty$, $k = 1, \dots p$, we have $\mathbb{E}(Z_1^\Delta) = \Delta M_1$, $\mathbb{E}[(Z_1^\Delta)^2] = \Delta M_2 + \Delta^2 M_1$, and more generally $\mathbb{E}[(Z_1^\Delta)^l] = \Delta\ M_l + o(\Delta)$ for all $l = 1, \dots, p$. Moreover, under (G1), $\mathbb{E}(|Z_1^\Delta|) \leq 2\Delta\|g\|_1$.*

Proposition 8.2. *Denote by P_Δ the distribution of Z_1^Δ and define $\mu_\Delta(dx) = \Delta^{-1}xP_\Delta(dx)$ and $\mu(dx) = g(x)dx$. Under (G1), the distribution μ_Δ has a density h_Δ given by*

$$h_\Delta(x) = \int g(x-y)P_\Delta(dy) = \mathbb{E}g(x - Z_1^\Delta).$$

And μ_Δ weakly converges to μ as Δ tends to 0.

8.1. **Proof of bias and variance control.**

8.1.1. *Proof of Lemma 3.1.* First, we study $b_2(x_0)$:

$$\begin{aligned} b_2(x_0) &= \frac{1}{h\Delta}\mathbb{E}\left[Z_1^\Delta K\left(\frac{x_0 - Z_1^\Delta}{h}\right)\right] - \frac{1}{h}\int K\left(\frac{x_0-u}{h}\right)g(u)du \\ &= \frac{1}{h}\int K\left(\frac{x_0-u}{h}\right)\mathbb{E}[g(u - Z_1^\Delta)]du - \frac{1}{h}\int K\left(\frac{x_0-u}{h}\right)g(u)du \\ &= \frac{1}{h}\int K\left(\frac{x_0-u}{h}\right)\mathbb{E}[g(u-Z_1^\Delta) - g(u)]du. \end{aligned}$$

Now, applying the mean value theorem to g, we get

$$|b_2(x_0)| = \left|\frac{1}{h}\int K\left(\frac{x_0-u}{h}\right)\mathbb{E}[-Z_1^\Delta g'(u_{Z_1})]du\right| \text{ with } u_{Z_1} \in [u - Z_1, u\ ]$$

$$\leq \Delta\|g'\|_\infty\|K\|_1\mathbb{E}\left[\frac{|Z_1^\Delta|}{\Delta}\right]. \tag{20}$$

From the results of Proposition 8.1, $\mathbb{E}[|Z_1^\Delta|/\Delta] \leq 2\|g\|_1$.
Now we study $b_1(x_0) = K_h \star g(x_0) - g(x_0)$:

$$\begin{aligned} b_1(x_0) &= \frac{1}{h}\int K\left(\frac{x_0-u}{h}\right)g(u)du - g(x_0), \text{by setting } v = (x_0-u)/h \\ &= \int K(v)g(x_0+vh)dv - g(x_0) \\ &= \int K(v)(g(x_0+vh) - g(x_0))dv, \text{ by (5).} \end{aligned}$$

Using the Taylor formula, there exists $\tau \in [0,1]$ such that

$$g(x_0+vh) - g(x_0) = vhg'(x_0) + ... + \frac{(vh)^l}{l!}g^{(l)}(x_0 + \tau vh).$$

So, under (Ker[1]) using (6),

$$\begin{aligned} b_1(x_0) &= \int K(v)\frac{(vh)^l}{l!}g^{(l)}(x_0+\tau vh)dv \\ &= \int K(v)\frac{(vh)^l}{l!}(g^{(l)}(x_0+\tau vh) - g^{(l)}(x_0))dv. \end{aligned}$$

It follows that

$$\begin{aligned} |b_1(x_0)| &\leq \int |K(v)|\frac{|vh|^l}{l!}|g^{(l)}(x_0+\tau vh)-g^{(l)}(x_0)|dv \\ &\leq \int |K(v)|\frac{|vh|^l}{L}|\tau vh|^{\beta-l}dv, \text{using (8) in (G3)} \\ &\leq \frac{h^\beta L}{l!}\int |K(v)||v|^\beta dv := \sqrt{c_1/2}h^\beta, \qquad (21) \end{aligned}$$

as $\tau\in[0,1]$ and $\int |K(v)||v|^\beta dv<+\infty$ under (Ker[1]).
Gathering (20) and (21), we obtain

$$\begin{aligned} b(x_0)^2 &\leq 2b_1(x_0)^2+2b_2(x_0)^2 \\ &\leq c_1h^{2\beta}+2M^2(\Delta\|g\|_1\|K\|_1)^2. \end{aligned}$$

This completes the proof of Lemma 3.1. □

8.1.2. *Proof of Lemma 3.2.* We have the following relation:

$$\psi_\Delta{}''=i\Delta\psi_\Delta{}'g^*+i\Delta\psi_\Delta g^{*\prime}=-\Delta^2\psi_\Delta g^{*2}+i\Delta\psi_\Delta g^{*\prime}. \qquad (22)$$

As the Z_k^Δ are i.i.d., we have:

$$Var[\widehat{g}(x_0)]=Var\left[\frac{1}{nh\Delta}\sum_{k=1}^n Z_k^\Delta K\left(\frac{Z_k^\Delta-x_0}{h}\right)\right]=\frac{1}{n(h\Delta)^2}Var\left[Z_1^\Delta K\left(\frac{Z_1^\Delta-x_0}{h}\right)\right].$$

Thus,

$$Var[\widehat{g}(x_0)]\leq\frac{1}{n(h\Delta)^2}\mathbb{E}\left[(Z_1^\Delta)^2K^2\left(\frac{Z_1^\Delta-x_0}{h}\right)\right].$$

Writing

$$K^2\left(\frac{Z_1^\Delta-x_0}{h}\right)=\left|\frac{1}{2\pi}\int K^*(u)e^{-i\frac{(x_0-Z_1^\Delta)u}{h}}du\right|^2,$$

we obtain with $v=u/h$

$$\begin{aligned} Var[\widehat{g}(x_0)] &\leq \frac{1}{n\Delta^2}\mathbb{E}\left[(Z_1^\Delta)^2\left|\frac{1}{2\pi}\int K^*(vh)e^{-i(x_0-Z_1^\Delta)v}dv\right|^2\right] \\ &\leq \frac{1}{n\Delta^2(2\pi)^2}\mathbb{E}\left[\int\int Z_1^\Delta e^{iZ_1^\Delta v}K^*(vh)e^{-ix_0v}\overline{Z_1^\Delta e^{iZ_1^\Delta u}K^*(uh)e^{-ix_0u}}dvdu\right]. \end{aligned}$$

Using Fubini (under (Ker[3])), $\mathbb{E}[(Z_1^\Delta)^2e^{iZ_1^\Delta(v-u)}]=-\psi''_\Delta(v-u)$ and formula (22), we find

$$\begin{aligned} Var[\widehat{g}(x_0)] &\leq \frac{1}{n\Delta^2(2\pi)^2}\int\int|-\psi_\Delta{}''(v-u)K^*(vh)K^*(uh)|dvdu \\ &\leq \frac{1}{n\Delta^2(2\pi)^2}\int\int|\Delta^2\psi_\Delta(v-u)(g^*)^2(v-u)K^*(vh)K^*(uh)|dvdu \\ &+ \frac{1}{n\Delta^2(2\pi)^2}\int\int|\Delta\psi_\Delta(v-u)(g^*)'(v-u)K^*(vh)K^*(uh)|dvdu, \text{with (22)} \\ &\leq T_1+T_2 \end{aligned}$$

with

$$T_1 = \frac{1}{n\Delta^2(2\pi)^2}\int\int |\Delta^2\psi_\Delta(v-u)(g^*)^2(v-u)K^*(vh)K^*(uh)|dvdu$$
$$T_2 = \frac{1}{n\Delta^2(2\pi)^2}\int\int |\Delta\psi_\Delta(v-u)(g^*)'(v-u)K^*(vh)K^*(uh)|dvdu.$$

We first bound T_2:

$$\begin{aligned}
T_2 &\leq \frac{1}{n\Delta(2\pi)^2}\sqrt{\int\int |\psi_\Delta(v-u)||(g^*)'(v-u)||K^*(vh)|^2 dvdu} \\
&\qquad \times\sqrt{\int\int |\psi_\Delta(v-u)||(g^*)'(v-u)||K^*(uh)|^2 dvdu} \\
&\leq \frac{1}{n\Delta(2\pi)^2}\int |K^*(vh)|^2 dv \int |\psi_\Delta(z)||(g^*)'(z)|dz \\
&\leq \frac{1}{nh\Delta(2\pi)^2}\int |K^*(u)|^2 du \int |(g^*)'(z)|dz, \text{ setting } u = vh \text{ and because } |\psi_\Delta(z)| \leq 1 \\
&\leq \frac{{\|K\|_2}^2}{2\pi nh\Delta}\int |(g^*)'(z)|dz, \text{where } \|K\|_2 < +\infty \text{ by (Ker[2])} \\
&= \frac{P{\|K\|_2}^2}{nh\Delta}, \text{ by (G5).}
\end{aligned}$$

Following the same line for the study of T_1, we get

$$T_1 \leq \frac{\|K\|_2^2}{(2\pi)nh}\int |(g^*)^2(z)|dz \leq \frac{\|K\|_2^2\|g\|_2^2}{nh},$$

where $\|K\|_2 < +\infty$ by (Ker[2]) and where $\|g\|_2 < +\infty$ by (G2). This completes the proof of Lemma 3.2. □

8.2. **Proof of Theorem 4.1.** We decompose $\mathbb{E}[|\hat{g}_{\hat{\beta}(x_0)}(x_0) - g(x_0)|]$ as follows:

$$\begin{aligned}
\mathbb{E}[|\hat{g}_{\hat{\beta}(x_0)}(x_0) - g(x_0)|] &= \mathbb{E}\left[|\hat{g}_{\hat{\beta}(x_0)}(x_0) - g(x_0)|\mathbb{1}_{\{\hat{\beta}(x_0)\geq\beta\}}\right] \\
&+ \mathbb{E}\left[|\hat{g}_{\hat{\beta}(x_0)}(x_0) - g(x_0)|\mathbb{1}_{\{\hat{\beta}(x_0)<\beta\}}\right].
\end{aligned} \tag{23}$$

We bound the first term by simply using the definition of $\hat{\beta}(x_0)$. We obtain

$$|\hat{g}_\beta(x_0) - \hat{g}_{\hat{\beta}(x_0)}(x_0)|\mathbb{1}_{\{\hat{\beta}(x_0)\geq\beta\}} \leq cu_n(\beta).$$

Also,

$$\begin{aligned}
\mathbb{E}\left[|g(x_0) - \hat{g}_{\hat{\beta}(x_0)}(x_0)|\mathbb{1}_{\{\hat{\beta}(x_0)\geq\beta\}}\right] &\leq \mathbb{E}|g(x_0) - \hat{g}_\beta(x_0)| + \mathbb{E}\left[|\hat{g}_\beta(x_0) - \hat{g}_{\hat{\beta}(x_0)}(x_0)|\mathbb{1}_{\{\hat{\beta}(x_0)\geq\beta\}}\right] \\
&\leq cu_n(\beta)
\end{aligned} \tag{24}$$

which completes the bound of the first term of (23).
We now study the second term. We write

$$\begin{aligned} \mathbb{E}[|\hat g_{\hat\beta(x_0)}(x_0)-g(x_0)|\mathbb{1}_{\{\hat\beta(x_0)<\beta\}}] &\le \left(\mathbb{E}[|\hat g_{\hat\beta(x_0)}(x_0)-g(x_0)|^2]\right)^{1/2}\times \mathbb{P}(\hat\beta(x_0)<\beta)^{1/2} \\ &\le \left(\mathbb{E}[|\hat g_{\hat\beta(x_0)}(x_0)|^2]+\|g\|_\infty^2\right)^{1/2}\times \mathbb{P}(\hat\beta(x_0)<\beta)^{1/2} \quad (25) \end{aligned}$$

Proposition 8.3. *Under the assumptions of Proposition 3.1 and if the $|Z_i|$'s admit a moment of order z with $z>10+4\xi+(5+2\xi)/\alpha_1$, under condition (11), there exists $\kappa>0$ such that*

$$\mathbb{P}(\hat\beta(x_0)<\beta)\le(2+\kappa)(n\Delta)^{-4}.$$

Proposition 8.4. *Under (G1) and (G5), we have*

$$\mathbb{E}[|\hat g_{\hat\beta(x_0)}(x_0)|^2]\le Q\frac{(n\Delta)^2}{\log(n\Delta)^2},$$

with $Q:=\|K\|_\infty^2\left(\mathbb{E}[(Z_1^\Delta)^2/\Delta]+\left[\mathbb{E}[|Z_1^\Delta|/\Delta]\right]^2\right)$.

Propositions 8.3, 8.4 and equation (25) imply that

$$\mathbb{E}[|\hat g_{\hat\beta(x_0)}(x_0)-g(x_0)|\mathbb{1}_{\{\hat\beta(x_0)<\beta\}}] \le \left(Q(n\Delta)^2+\|g\|_\infty^2\right)^{1/2}\sqrt{(2+\kappa)(n\Delta)^{-4}}\le\frac{Q_1}{n\Delta} \quad (26)$$

where Q_1 is a constant which depends on the terms $\mathbb{E}[{Z_1^\Delta}^2/\Delta]+\left[\mathbb{E}[Z_1^\Delta/\Delta]\right]^2$ and $\|g\|_\infty$.
We conclude by gathering (24) and (26):

$$\begin{aligned} \mathbb{E}[|\hat g_{\hat\beta(x_0)}(x_0)-g(x_0)|] &= \mathbb{E}[|\hat g_{\hat\beta(x_0)}(x_0)-g(x_0)|\mathbb{1}_{\{\hat\beta(x_0)\ge\beta\}}]+\mathbb{E}[|\hat g_{\hat\beta(x_0)}(x_0)-g(x_0)|\mathbb{1}_{\{\hat\beta(x_0)\le\beta\}}] \\ &\le cu_n(\beta)+Q_1(n\Delta)^{-1} \\ &\le Cu_n(\beta). \end{aligned}$$

This completes the proof of Theorem 4.1. □

8.2.1. *Proof of Proposition 8.3.* We recall that $g_\alpha(x_0)=g_{h(\alpha)}(x_0)$. First we write that

$$\begin{aligned} \mathbb{P}(\hat\beta(x_0)<\beta) &= \sum_{\alpha_i\in A,\alpha_i<\beta}\mathbb{P}(\hat\beta(x_0)=\alpha_i) \\ &= \sum_{\alpha_i\in A,\alpha_i<\beta}\mathbb{P}(\exists\alpha'\le\alpha_{i+1}|\hat g_{\alpha_{i+1}}(x_0)-\hat g_{\alpha'}(x_0)|>cu_n(\alpha')) \\ &\le \sum_{\alpha_{i+1}\le\beta}\sum_{\alpha'\le\alpha_{i+1}}\mathbb{P}(|\hat g_{\alpha_{i+1}}(x_0)-\hat g_{\alpha'}(x_0)|>cu_n(\alpha')) \\ &\le D^2\sup_{\alpha'\le\alpha\le\beta}\mathbb{P}(|\hat g_\alpha(x_0)-\hat g_{\alpha'}(x_0)|>cu_n(\alpha')). \quad (27) \end{aligned}$$

Now we study $\mathbb{P}(|\hat{g}_\alpha(x_0) - \hat{g}_{\alpha'}(x_0)| > cu_n(\alpha'))$ for $\alpha' \leq \alpha \leq \beta$. We have, with Lemma 3.1

$$\begin{aligned}
|\hat{g}_\alpha(x_0) - \hat{g}_{\alpha'}(x_0)| &\leq |\hat{g}_\alpha(x_0) - g_\alpha(x_0)| + |g_\alpha(x_0) - g(x_0)| + |g(x_0) - g_{\alpha'}(x_0)| \\
&\quad + |g_{\alpha'}(x_0) - \hat{g}_{\alpha'}(x_0)| \qquad (28)\\
&\leq |\hat{g}_\alpha(x_0) - g_\alpha(x_0)| + |g_{\alpha'}(x_0) - \hat{g}_{\alpha'}(x_0)| + \frac{c_1}{2}h(\alpha)^\beta + \frac{c_1}{2}h(\alpha')^\beta + c_1'\Delta \\
&\leq |\hat{g}_\alpha(x_0) - g_\alpha(x_0)| + |g_{\alpha'}(x_0) - \hat{g}_{\alpha'}(x_0)| \\
&\quad + \frac{c_1}{2}\left(\frac{n\Delta}{\log(n\Delta)}\right)^{\frac{-\beta}{2\alpha+1}} + \frac{c_1}{2}\left(\frac{n\Delta}{\log(n\Delta)}\right)^{\frac{-\beta}{2\alpha'+1}} + c_1'\Delta \\
&\leq |\hat{g}_\alpha(x_0) - g_\alpha(x_0)| + |g_{\alpha'}(x_0) - \hat{g}_{\alpha'}(x_0)| + (c_1 + c_1')u_n(\alpha'), \qquad (29)
\end{aligned}$$

where (29) follows by using condition (11). Indeed for $\alpha' \leq \alpha \leq \beta$, we have

$$\left(\frac{n\Delta}{\log(n\Delta)}\right)^{\frac{-\beta}{2\alpha+1}} + \left(\frac{n\Delta}{\log(n\Delta)}\right)^{\frac{-\beta}{2\alpha'+1}} \leq 2\left(\frac{n\Delta}{\log(n\Delta)}\right)^{\frac{-\alpha'}{2\alpha'+1}} = 2u_n(\alpha').$$

Therefore, we get

$$\mathbb{P}(|\hat{g}_\alpha(x_0) - \hat{g}_{\alpha'}(x_0)| > cu_n(\alpha')) \leq \mathbb{P}(|\hat{g}_\alpha(x_0) - g_\alpha(x_0)| + |g_{\alpha'}(x_0) - \hat{g}_{\alpha'}(x_0)| > c'u_n(\alpha')),$$

with $c' = c - c_1 - c_1'$.
Now,

$$\begin{aligned}
\mathbb{P}(|\hat{g}_\alpha(x_0) - g_\alpha(x_0)| + |g_{\alpha'}(x_0) - \hat{g}_{\alpha'}(x_0)| > c'u_n(\alpha')) &\leq \mathbb{P}(|\hat{g}_\alpha(x_0) - g_\alpha(x_0)| > (c'/2)u_n(\alpha')) \\
&+ \mathbb{P}(|\hat{g}_{\alpha'}(x_0) - g_{\alpha'}(x_0)| > (c'/2)u_n(\alpha')),
\end{aligned}$$

and since, $u_n(\alpha') \geq u_n(\alpha)$ we have

$$\mathbb{P}(|\hat{g}_\alpha(x_0) - g_\alpha(x_0)| > (c'/2)u_n(\alpha')) \leq \mathbb{P}(|\hat{g}_\alpha(x_0) - g_\alpha(x_0)| > (c'/2)u_n(\alpha)). \qquad (30)$$

Thus, we just have to bound $2\mathbb{P}(|\hat{g}_\alpha(x_0) - g_\alpha(x_0)| > (c'/2)u_n(\alpha))$ for all $\alpha \leq \beta$.
Note that

$$\hat{g}_\alpha(x_0) - g_\alpha(x_0) = \frac{1}{n}\sum_{k=1}^{n}[Y_k(\alpha) - \mathbb{E}(Y_k(\alpha))] \qquad (31)$$

where

$$Y_k(\alpha) = \frac{{Z_k}^\Delta}{\Delta h(\alpha)}K\left(\frac{{Z_k}^\Delta - x_0}{h(\alpha)}\right). \qquad (32)$$

We truncate the ${Z_k}^\Delta$ and we consider the decomposition:

$$\{|{Z_k}^\Delta| \leq \mu_n\}, \text{ and } \{|{Z_k}^\Delta| > \mu_n\}$$

where

$$\mu_n = k\left(\frac{n\Delta}{\log(n\Delta)}\right)^{\frac{\alpha}{2\alpha+1}} \text{ with } k = \frac{{\|K\|_2}^2(P + \|g\|_2^2)}{c'\|K\|_\infty}. \qquad (33)$$

We decompose (31),

$$\begin{aligned}\hat{g}_\alpha(x_0)-g_\alpha(x_0) &= \frac{1}{n}\sum_{k=1}^n [Y_k(\alpha)\mathbb{1}_{\{|Z_k{}^\Delta|\le\mu_n\}}-\mathbb{E}(Y_k(\alpha)\mathbb{1}_{\{|Z_k{}^\Delta|\le\mu_n\}})]\\ &+ \frac{1}{n}\sum_{k=1}^n [Y_k(\alpha)\mathbb{1}_{\{|Z_k{}^\Delta|>\mu_n\}}-\mathbb{E}(Y_k(\alpha)\mathbb{1}_{\{|Z_k{}^\Delta|>\mu_n\}})]\\ &:= (\hat{g}_\alpha^{(1)}(x_0)-g_\alpha{}^{(1)}(x_0))+(\hat{g}_\alpha^{(2)}(x_0)-g_\alpha{}^{(2)}(x_0)).\end{aligned}$$

So we write

$$\begin{aligned}&\mathbb{P}(|\hat{g}_\alpha(x_0)-g_\alpha(x_0)|>(c'/2)u_n(\alpha))\\ \le\ &\mathbb{P}(|\hat{g}_\alpha^{(1)}(x_0)-g_\alpha{}^{(1)}(x_0)|+|\hat{g}_\alpha^{(2)}(x_0)-g_\alpha{}^{(2)}(x_0)|>(c'/2)u_n(\alpha))\\ \le\ &\mathbb{P}(|\hat{g}_\alpha^{(1)}(x_0)-g_\alpha{}^{(1)}(x_0)|>c'u_n(\alpha)/4)+\mathbb{P}(|\hat{g}_\alpha^{(2)}(x_0)-g_\alpha{}^{(2)}(x_0)|>c'u_n(\alpha)/4).\end{aligned} \tag{34}$$

We need the two following lemmas to conclude the proof.

Lemma 8.1. *Under (Ker[2]) (Ker[3]) , (G1) (G2) et (G5) we have,*

$$\mathbb{P}(|\hat{g}_\alpha^{(1)}(x_0)-g_\alpha{}^{(1)}(x_0)|>c'u_n(\alpha)/4)\le 2(n\Delta)^{-(2\xi+4)}.$$

Lemma 8.2. *If the $|Z_i|$ admits a moment of order z with $z>10+4\xi+(5+2\xi)/\alpha$ there exists $\kappa>0$ such that*

$$\mathbb{P}(|\hat{g}_\alpha^{(2)}(x_0)-g_\alpha{}^{(2)}(x_0)|>c'u_n(\alpha)/4)\le \kappa(n\Delta)^{-(2\xi+4)}.$$

Thus we bound (34) by using Lemmas 8.1 and 8.2 and by gathering this with (27) and (30), we obtain:

$$\begin{aligned}\mathbb{P}(\hat{\beta}(x_0)<\beta) &\le D^2\sup_{\alpha\le\beta}\mathbb{P}(|\hat{g}_\alpha(x_0)-g_\alpha(x_0)|>(c'/2)u_n(\alpha))\\ &\le D^2\sup_{\alpha\le\beta}\mathbb{P}(|\hat{g}_\alpha^{(1)}(x_0)-g_\alpha{}^{(1)}(x_0)|>c'u_n(\alpha)/4)\\ &+ D^2\sup_{\alpha\le\beta}\mathbb{P}(|\hat{g}_\alpha^{(2)}(x_0)-g_\alpha{}^{(2)}(x_0)|>c'u_n(\alpha)/4)\\ &\le 2(n\Delta)^{-4}+\kappa(n\Delta)^{-4}.\end{aligned}$$

This completes the proof of Proposition 8.3. □

8.2.2. *Proof of Proposition 8.4.*

$$\begin{aligned}\mathbb{E}[|\hat{g}_{\hat{\beta}(x_0)}(x_0)|^2] &\le \mathbb{E}\left[\left\{\frac{1}{nh(\hat{\beta}(x_0))\Delta}\sum_{k=1}^n Z_k^\Delta K\left(\frac{Z_k^\Delta-x_0}{h(\hat{\beta}(x_0))}\right)\right\}^2\right]\\ &\le \frac{\|K\|_\infty^2}{(n\Delta)^2}\mathbb{E}\left[\frac{1}{h(\hat{\beta}(x_0))^2}\left\{\sum_{k=1}^n|Z_k^\Delta|\right\}^2\right].\end{aligned}$$

We have, as $\hat{\beta}(x_0)>0$ and thus $\frac{2}{2\hat{\beta}(x_0)+1}<2$,

$$\frac{1}{h(\hat{\beta}(x_0))^2}=\left(\frac{n\Delta}{\log(n\Delta)}\right)^{\frac{2}{2\hat{\beta}(x_0)+1}}<\left(\frac{n\Delta}{\log(n\Delta)}\right)^2.$$

We can write

$$\begin{aligned}\mathbb{E}[|\hat{g}_{\hat{\beta}(x_0)}(x_0)|^2] &\leq \frac{\|K\|_\infty^2}{(n\Delta)^2}\frac{(n\Delta)^2}{\log(n\Delta)^2}\mathbb{E}\left[\left\{\sum_{k=1}^n |Z_k^\Delta|\right\}^2\right] \\ &\leq \frac{\Delta\|K\|_\infty^2}{\log(n\Delta)^2}\mathbb{E}\left[\frac{1}{\Delta}\left\{\sum_{k=1}^n |Z_k^\Delta|\right\}^2\right]. \qquad (35)\end{aligned}$$

Yet,

$$\begin{aligned}\mathbb{E}\left[\left\{\sum_{k=1}^n |Z_k^\Delta|\right\}^2\right] &= Var\left[\sum_{k=1}^n |Z_k^\Delta|\right] + \mathbb{E}\left[\sum_{k=1}^n |Z_k^\Delta|\right]^2 \\ &= nVar\left[|Z_1^\Delta|\right] + \left[n\mathbb{E}[|Z_1^\Delta|]\right]^2 \\ &\leq n\Delta\mathbb{E}[(Z_1^\Delta)^2/\Delta] + (n\Delta)^2\left[\mathbb{E}[|Z_1^\Delta|/\Delta]\right]^2 \\ &\leq (n\Delta)^2\left(\mathbb{E}[(Z_1^\Delta)^2/\Delta] + \left[\mathbb{E}[|Z_1^\Delta|/\Delta]\right]^2\right). \qquad (36)\end{aligned}$$

Therefore, inserting (35) in (36) completes the proof of Proposition 8.4. □

8.2.3. *Proof of lemmas 8.1 and 8.2.* :

Proof of lemma 8.1 :

We apply the Bernstein inequality recalled in Section 9 to $T_i = Y_i(\alpha)\mathbb{1}_{|Z_i^\Delta|\leq\mu_n}$ where $Y_i(\alpha)$ is defined by (32) and μ_n is defined by (33) (see Birgé and Massart (1998), p.366). We choose $\eta = c'u_n(\alpha)/4$. It is easy to see that

$$|T_i| \leq b := \frac{\|K\|_\infty\mu_n}{\Delta h(\alpha)} \text{ and } Var(T_i) \leq \nu^2 := (P + \|g\|_2^2)\frac{\|K\|_2{}^2}{\Delta h(\alpha)}.$$

We find

$$\exp\left(\frac{-n\eta^2}{4\nu^2}\right) = \exp\left(\frac{-n\eta}{4b}\right) = (n\Delta)^{-k_1} \text{ with } k_1 = \frac{c'^2}{4\times 16\|K\|_2{}^2(P+\|g\|_2^2)} = 2\xi + 4.$$

Therefore the Bernstein inequality (45) yields

$$\mathbb{P}(|\hat{g}_\alpha^{(1)}(x_0) - g_\alpha{}^{(1)}(x_0)| > \kappa u_n(\alpha)/2) \leq 2(n\Delta)^{-k_1}.$$

This completes the proof of lemma 8.1. □

Proof of lemma 8.2 :

$$\begin{aligned}\mathbb{P}(|\hat{g}_\alpha^{(2)}(x_0) - g_\alpha{}^{(2)}(x_0)| > c'u_n(\alpha)/4) &\leq 4\frac{\mathbb{E}[|\hat{g}_\alpha^{(2)}(x_0) - g_\alpha{}^{(2)}(x_0)|]}{c'u_n(\alpha)} \\ &\leq \|K\|_\infty\frac{\frac{1}{n}\sum_{k=1}^n 8\mathbb{E}|\frac{Z_k}{\Delta h(\alpha)}\mathbb{1}_{|Z_k|>\mu_n}|}{c'u_n(\alpha)} \\ &\leq \frac{8\|K\|_\infty\mathbb{E}[|Z_1|^{w+1}/\Delta]}{c'h(\alpha)\mu_n{}^w u_n(\alpha)}.\end{aligned}$$

By Proposition 8.1 the quantity $8\|K\|_\infty \mathbb{E}[|Z_1|^{w+1}/\Delta]/c'$ is bounded by a constant.

If the $|Z_i|$'s admit a moment of order z with $\forall \alpha \in A$, $z > 10 + 4\xi + (5+2\xi)/\alpha$ we have

$$\mathbb{P}(|\hat{g}_\alpha^{(2)}(x_0) - {g_\alpha}^{(2)}(x_0)| > \kappa u_n(\alpha)/2) \leq \frac{1}{(n\Delta)^{2\xi+4}}.$$

This completes the proof of lemma 8.2. □

8.3. **Proof of Theorem 5.1.** The goal is to bound $\mathbb{E}[|g(x_0) - \hat{g}_{\hat{h}}(x_0)|^2]$. To do this, we fix $h \in H$. We write

$$|g(x_0) - \hat{g}_{\hat{h}}(x_0)| \leq |\hat{g}_{\hat{h}}(x_0) - \hat{g}_{h,\hat{h}}(x_0)| + |\hat{g}_{h,\hat{h}}(x_0) - \hat{g}_h(x_0)| + |\hat{g}_h(x_0) - g(x_0)|.$$

So we have

$$|g(x_0) - \hat{g}_{\hat{h}}(x_0)|^2 \leq 3|\hat{g}_{\hat{h}}(x_0) - \hat{g}_{h,\hat{h}}(x_0)|^2 + 3|\hat{g}_{h,\hat{h}}(x_0) - \hat{g}_h(x_0)|^2 + 3|\hat{g}_h(x_0) - g(x_0)|^2.$$

Define $B := |\hat{g}_{\hat{h}}(x_0) - \hat{g}_{h,\hat{h}}(x_0)|^2$ and $C := |\hat{g}_{h,\hat{h}}(x_0) - \hat{g}_h(x_0)|^2$.
We have $A(h) \geq |\hat{g}_{\hat{h}}(x_0) - \hat{g}_{h,\hat{h}}(x_0)|^2 - V(\hat{h}) \geq B - V(\hat{h})$. So $B \leq A(h) + V(\hat{h})$.
Moreover, $A(\hat{h}) \geq |\hat{g}_{h,\hat{h}}(x_0) - \hat{g}_h(x_0)|^2 - V(h) \geq C - V(h)$. So $C \leq A(\hat{h}) + V(h)$.
Therefore,

$$|g(x_0) - \hat{g}_{\hat{h}}(x_0)|^2 \leq 3(A(h) + V(\hat{h})) + 3(A(\hat{h}) + V(h)) + 3|\hat{g}_h(x_0) - g(x_0)|^2.$$

Now, by definition of $\hat{h}$, $A(\hat{h}) + V(\hat{h}) \leq A(h) + V(h)$. This allows us to write

$$|g(x_0) - \hat{g}_{\hat{h}}(x_0)|^2 \leq 6A(h) + 6V(h) + 3|\hat{g}_h(x_0) - g(x_0)|^2.$$

Let us denote $b_h(x_0) = \mathbb{E}[\hat{g}_h(x_0)] - g(x_0)$ and $b_{h,2}(x_0) = \mathbb{E}[\hat{g}_h(x_0)] - K_h \star g(x_0)$ (these are the same notations as in Lemma 3.1, but with subscript h). Thus

$$\begin{aligned}\mathbb{E}[|g(x_0) - \hat{g}_{\hat{h}}(x_0)|^2] &\leq 6\mathbb{E}[A(h)] + 6V(h) + 3b_h^2(x_0) + 3Var(\hat{g}_h(x_0)) \\ &\leq 6\mathbb{E}[A(h)] + 3b_h^2(x_0) + 9V(h).\end{aligned}$$

We have to bound $\mathbb{E}[A(h)]$. Let us denote by $g_{h,h'} = \mathbb{E}[\hat{g}_{h,h'}]$ and $g_h = \mathbb{E}[\hat{g}_h]$. We write,

$$\hat{g}_{h,h'} - \hat{g}_{h'} = \hat{g}_{h,h'} - g_{h,h'} - \hat{g}_{h'} + g_{h'} + g_{h,h'} - g_{h'}, \tag{37}$$

and we study the last term of the above decomposition. We have

$$\begin{aligned}|g_{h,h'}(x_0) - g_{h'}(x_0)| &= |\mathbb{E}[\hat{g}_{h,h'}(x_0) - \hat{g}_{h'}(x_0)]| \\ &= |\mathbb{E}[K_{h'} \star \hat{g}_h(x_0) - \hat{g}_{h'}(x_0)]| \\ &= |K_{h'} \star \mathbb{E}[\hat{g}_h(x_0) - g(x_0)] + K_{h'} \star g(x_0) - \mathbb{E}[\hat{g}_{h'}(x_0)]|.\end{aligned}$$

This can be written:

$$\begin{aligned}|g_{h,h'}(x_0) - g_{h'}(x_0)| &= |K_{h'} \star b_h(x_0) + b_{h,2}(x_0)| \\ &\leq |K_{h'} \star b_h(x_0)| + |b_{h,2}(x_0)| \\ &\leq \left| \int K\left(\frac{x_0 - u}{h'}\right) b_h(u) \frac{du}{h'} \right| + |b_{h,2}(x_0)|\end{aligned}$$

Now $|b_{h,2}(x_0)| \leq |b_h(x_0)| \leq \|b_h\|_\infty$ so that

$$\begin{aligned}|g_{h,h'}(x_0) - g_{h'}(x_0)|^2 &\leq 2\|b_h\|_\infty^2 \left(\int |K(v)|dv\right)^2 + 2|b_{h,2}(x_0)|^2 \\ &\leq 2(\|K\|_1^2 + 1)\|b_h\|_\infty^2 \end{aligned} \tag{38}$$

Then by inserting (38) in decomposition (37), we find:

$$\begin{aligned} A(h) &= \sup_{h'} \{|\hat{g}_{h,h'}(x_0) - \hat{g}_{h'}(x_0)|^2 - V(h')\}_+ \\ &\leq 3 \sup_{h'} \{|\hat{g}_{h,h'}(x_0) - g_{h,h'}(x_0)|^2 - V(h')/6\}_+ \\ &\quad + 3 \sup_{h'} \{|\hat{g}_{h'}(x_0) - g_{h'}(x_0)|^2 - V(h')/6\}_+ + 6(\|K\|_1^2 + 1)\|b_h\|_\infty^2. \end{aligned} \tag{39}$$

We can prove the following results:

Proposition 8.5. *Let the assumptions of Proposition 3.1 hold and take c' in (15) such that $c' \geq 96$. Assume that there exists β_0 (known) such that $\beta > \beta_0$. Define $H = \{\frac{j}{M}, 1 \leq j \leq M\}$, with $M = \left[(n\Delta)^{1/(2\beta_0+1)}\right]$. Then if the Z_i's admit a moment of order z such that $z \geq 2(3 + 2/\beta_0)$, we have*

$$\mathbb{E}\left[\sup_{h'} \{|\hat{g}_{h'}(x_0) - g_{h'}(x_0)|^2 - V(h')/6\}_+\right] \leq C\left(\frac{1}{n\Delta} + \frac{\log(n\Delta)}{n\Delta}\right). \tag{40}$$

Proposition 8.6. *Let the assumptions of Proposition 3.1 hold and take c' in (15) such that $c' \geq 96\|K\|_\infty$. Assume that there exists β_0 (known) such that $\beta > \beta_0$. Define $H = \{\frac{j}{M}, 1 \leq j \leq M\}$, with $M = \left[(n\Delta)^{1/(2\beta_0+1)}\right]$. Then if the Z_i's admit a moment of order z such that $z \geq 2(3 + 2/\beta_0)$, we have*

$$\mathbb{E}\left[\sup_{h'} \{|\hat{g}_{h,h'}(x_0) - g_{h,h'}(x_0)|^2 - V(h')/6\}_+\right] \leq C\left(\frac{1}{n\Delta} + \frac{\log(n\Delta)}{n\Delta}\right). \tag{41}$$

Inequalities (40) et (41) together with (39) imply

$$\mathbb{E}[|g(x_0) - \hat{g}_{\hat{h}}(x_0)|^2] \leq C_1\|b_h\|_\infty^2 + C_2 V(h) + C_3\left(\frac{1}{n\Delta} + \frac{\log(n\Delta)}{n\Delta}\right).$$

This completes the proof of Theorem 5.1. □

8.3.1. *Proof of Proposition 8.5.* Note that

$$\hat{g}_{h'}(x_0) - g_{h'}(x_0) = \frac{1}{n}\sum_{k=1}^{n}[Y_k(h') - \mathbb{E}(Y_k(h'))] \tag{42}$$

where

$$Y_k(h') = \frac{Z_k{}^\Delta}{\Delta h'} K\left(\frac{Z_k{}^\Delta - x_0}{h'}\right).$$

The Z_k^Δ's are not bounded, thus we can not apply the Bernstein's inequality directly. As in the first method, we truncate Z_k^Δ and consider the following decomposition:

$$\{|Z_k{}^\Delta| \leq \tilde{\mu}_n\}, \text{and } \{|Z_k{}^\Delta| > \tilde{\mu}_n\}.$$

where

$$\tilde{\mu}_n = \frac{\|K\|_2^2(P + \|g\|_2^2)}{\|K\|_\infty \sqrt{(1/12)V(h')}} \tag{43}$$

We decompose then (42) as follows

$$\begin{aligned}
\hat{g}_{h'}(x_0) - g_{h'}(x_0) &= \frac{1}{n}\sum_{k=1}^{n}[Y_k(h')\mathbb{1}_{\{|Z_k{}^\Delta| \le \tilde{\mu}_n\}} - \mathbb{E}(Y_k(h')\mathbb{1}_{\{|Z_k{}^\Delta| \le \tilde{\mu}_n\}})] \\
&+ \frac{1}{n}\sum_{k=1}^{n}[Y_k(h')\mathbb{1}_{\{|Z_k{}^\Delta| > \tilde{\mu}_n\}} - \mathbb{E}(Y_k(h')\mathbb{1}_{\{|Z_k{}^\Delta| > \tilde{\mu}_n\}})] \\
&:= (\hat{g}_{h'}^{(1)}(x_0) - g_{h'}{}^{(1)}(x_0)) + (\hat{g}_{h'}^{(2)}(x_0) - g_{h'}{}^{(2)}(x_0)).
\end{aligned}$$

We have

$$\begin{aligned}
&\mathbb{E}\left[\sup_{h'}\{|\hat{g}_{h'}(x_0) - g_{h'}(x_0)|^2 - V(h')/6\}_+\right] \\
=\;& \mathbb{E}\left[\sup_{h'}\{|\hat{g}_{h'}^{(1)}(x_0) - g_{h'}^{(1)}(x_0) + \hat{g}_{h'}^{(2)}(x_0) - g_{h'}^{(2)}(x_0)|^2 - V(h')/6\}_+\right] \\
\le\;& \mathbb{E}\left[\sup_{h'}\{2|\hat{g}_{h'}^{(1)}(x_0) - g_{h'}^{(1)}(x_0)|^2 - V(h')/6\}_+\right] + 2\mathbb{E}\left[\sup_{h'}\{|\hat{g}_{h'}^{(2)}(x_0) - g_{h'}^{(2)}(x_0)|^2\}_+\right] \\
\le\;& 2\sum_{h'}\mathbb{E}\left[\{|\hat{g}_{h'}^{(1)}(x_0) - g_{h'}^{(1)}(x_0)|^2 - V(h')/12\}_+\right] + 2\sum_{h'}\mathbb{E}\left[|\hat{g}_{h'}^{(2)}(x_0) - g_{h'}^{(2)}(x_0)|^2\right]. \qquad (44)
\end{aligned}$$

We use the two following lemmas,

Lemma 8.3. *Let the assumptions of Proposition 3.1 hold and and take c' in (15) such that $c' \ge 96$. We have,*

$$E_1 := \mathbb{E}\left[\sup_{h'}\{|\hat{g}_{h'}^{(1)}(x_0) - g_{h'}^{(1)}(x_0)|^2 - V(h')/12\}_+\right] \le C\frac{\log(n\Delta)}{n\Delta}$$

with $C = 16\|K\|_2{}^2(P + \|g\|_2^2)$.

Lemma 8.4. *Let there exists β_0 (known) such that $\beta > \beta_0$. Define $H = \{\frac{j}{M}, 1 \le j \le M\}$, with $M = \left[(n\Delta)^{1/(2\beta_0+1)}\right]$. Then if the Z_i's admit a moment of order z such that $z \ge 2(3 + 2/\beta_0)$, we have*

$$E_2 := \sum_{h'}\mathbb{E}\left[|\hat{g}_{h'}^{(2)}(x_0) - g_{h'}^{(2)}(x_0)|^2\right] \le C'\frac{1}{n\Delta}.$$

Lemmas 8.3 and 8.4 can be used in (44) and yield

$$\mathbb{E}\left[\sup_{h'}\{|\hat{g}_{h'}(x_0) - g_{h'}(x_0)|^2 - V(h')/6\}_+\right] \le C''\left(\frac{1}{n\Delta} + \frac{\log(n\Delta)}{n\Delta}\right)$$

This ends the proof of Proposition 8.5. □

8.3.2. *Proof of Proposition 8.6.* Proposition 8.6 is obtained by following the same lines as for the proof of Proposition 8.5 with $Y_k(h)$ replaced by

$$U_k(h,h') = \frac{{Z_k}^{\Delta}}{\Delta} K_{h'} \star K_h \left({Z_k}^{\Delta} - x_0\right).$$

8.3.3. *Proof of lemmas 8.3 and 8.4.* :

Proof of lemma 8.3:

We apply the Bernstein inequality recalled in Section 9 to $W_i = Y_i(\alpha)\mathbb{1}_{|Z_i^{\Delta}|\leq\tilde{\mu}_n}$ where $Y_i(\alpha)$ is defined by (32) and $\tilde{\mu}_n$ is defined by (43). We choose $\eta = \sqrt{(1/12+y)V(h')}$. It is easy to see that

$$|W_i| \leq b := \frac{\|K\|_\infty \tilde{\mu}_n}{\Delta h'} \text{ and } Var(W_i) \leq \nu^2 := \frac{\|K\|_2^2(P+\|g\|_2^2)}{\Delta h'}.$$

We find

$$\begin{aligned} \exp\left(\frac{-n\eta^2}{4\nu^2}\right) &= \exp\left(-\frac{(1/12)nV(h')\Delta h'}{4{\|K\|_2}^2(P+\|g\|_2^2)}\right) \times \exp\left(-\frac{ynV(h')\Delta h'}{4{\|K\|_2}^2(P+\|g\|_2^2)}\right) \\ &= (n\Delta)^{-c'/96} \times (n\Delta)^{-c'y/8} \end{aligned}$$

and

$$\exp\left(\frac{-n\eta}{4b}\right) \leq (n\Delta)^{-c'/96} \times (n\Delta)^{-c'\sqrt{y}/8}.$$

We can write

$$\begin{aligned} E_1 &:= \mathbb{E}\left[\sup_{h'} \{|\hat{g}_{h'}^{(1)}(x_0) - g_{h'}^{(1)}(x_0)|^2 - V(h')/12\}_+\right] \\ &\leq 4\sum_{h'} V(h') \int_0^\infty (n\Delta)^{-c'/96} \times \max((n\Delta)^{-c'y/8}, (n\Delta)^{-c'\sqrt{y}/8}) dy \\ &\leq 4(n\Delta)^{-c'/96} \sum_{h'} V(h') \max(I_{n\Delta}, J_{n\Delta}) \end{aligned}$$

with $I_{n\Delta} = \int_0^\infty \exp\left(-\frac{c'}{8} y \log(n\Delta)\right) dy$ and $J_{n\Delta} = \int_0^\infty \exp\left(-\frac{c'}{8}\sqrt{y}\log(n\Delta)\right) dy$. Thus if $I_{n\Delta} \leq J_{n\Delta}$,

$$\begin{aligned} E_1 &\leq 4 \times \frac{8(n\Delta)^{-c'/96}}{c'\log(n\Delta)} \sum_{h'} V(h') \\ &\leq 4 \times 4 \times {\|K\|_2}^2(P+\|g\|_2^2)(n\Delta)^{-c'/96} \sum_{h'} \frac{1}{nh'\Delta} \\ &\leq 4 \times 4 \times {\|K\|_2}^2(P+\|g\|_2^2)(n\Delta)^{-(c'/96)-1} \sum_{h'} \frac{1}{h'} \end{aligned}$$

and if $J_{n\Delta} \leq I_{n\Delta}$,

$$\begin{aligned} E_1 &\leq 4 \times \frac{32(n\Delta)^{-c'/96}}{(c'/2)^2 \log(n\Delta)^2} \sum_{h'} V(h') \\ &\leq 4 \times 64 \times \frac{\|K\|_2{}^2 (P + \|g\|_2^2)(n\Delta)^{-c'/96}}{c' \log(n\Delta)} \sum_{h'} \frac{1}{nh'\Delta} \\ &\leq 4 \times 64 \times \frac{\|K\|_2{}^2 (P + \|g\|_2^2)(n\Delta)^{-(c'/96)-1}}{c' \log(n\Delta)} \sum_{h'} \frac{1}{h'}. \end{aligned}$$

Recall that $H = \{\frac{k}{M}, 1 \leq k \leq M\}$. Then

$$\sum_{h'} \frac{1}{h'} = \sum_{k=1}^{M} \frac{M}{k} \leq \log(M)M.$$

So we have

$$\frac{1}{(n\Delta)^{(c'/96)+1} \log(n\Delta)} \sum_{h'} \frac{1}{h'} \leq \frac{1}{(n\Delta)^{(c'/96)+1} \log(n\Delta)} M \log(M),$$

and

$$\frac{1}{(n\Delta)^{(c'/96)+1}} \sum_{h'} \frac{1}{h'} \leq \frac{1}{(n\Delta)^{(c'/96)+1}} M \log(M).$$

As $M = \lfloor (n\Delta)^{1/(2\beta_0+1)} \rfloor \leq n\Delta$ and $c'/96 \geq 1$

$$E_1 \leq 4 \times \|K\|_2{}^2 (P + \|g\|_2^2) \min\left(4\frac{\log(M)}{n\Delta}, 64\frac{\log(M)}{c'(n\Delta)\log(n\Delta)}\right).$$

This completes the proof of lemma 8.3. □

Proof of lemma 8.4:

$$\begin{aligned} \mathbb{E}\left[|\hat{g}_{h'}^{(2)}(x_0) - g_{h'}^{(2)}(x_0)|^2\right] &= Var\left[\frac{1}{n}\sum_{k=1}^{n} \frac{Z_k^\Delta}{\Delta h'} K\left(\frac{Z_k^\Delta - x_0}{h'}\right) \mathbb{1}_{\{|Z_k^\Delta| > \tilde{\mu}_n\}}\right] \\ &\leq \frac{1}{n} \frac{\|K\|_\infty^2}{(\Delta h')^2} \mathbb{E}[(Z_1^\Delta)^2 \mathbb{1}_{\{|Z_1^\Delta| > \tilde{\mu}_n\}}] \\ &\leq \frac{1}{n\Delta} \frac{\|K\|_\infty^2}{h'^2} \frac{\mathbb{E}[|Z_1^\Delta|^{w+2}/\Delta]}{\tilde{\mu}_n^w} \end{aligned}$$

We search conditions for $\sum_{h'} \frac{1}{h'^2 \tilde{\mu}_n^w} \leq$ constant. The following equalities hold up to constants:

$$\sum_{h'} \frac{1}{h'^2 \tilde{\mu}_n^w} = \sum_{h'} \frac{V(h')^{w/2}}{h'^2} = \frac{\log(n\Delta)^{w/2}}{(n\Delta)^{w/2}} \sum_{h'} \frac{1}{h'^{2+w/2}}$$

We set $h' = k/M$, so we have

$$\sum_{h'} \frac{1}{h'^{2+w/2}} = \sum_{k=1}^{M} \left(\frac{M}{k}\right)^{2+w/2} = M^{2+w/2} \sum_{k=1}^{M} \frac{1}{k^{2+w/2}} = O(M^{2+w/2}).$$

Finally, as $M = \lfloor (n\Delta)^{1/(2\beta_0+1)} \rfloor$, we have

$$\sum_{h'} \frac{1}{h'^2 {\mu_n}^w} \leq C \frac{M^{2+w/2} \log(n\Delta)^{w/2}}{(n\Delta)^{w/2}} \leq C \log(n\Delta)^{w/2} (n\Delta)^{\frac{1}{2\beta_0+1}(2+\frac{w}{2})-\frac{w}{2}}$$

We need that $(2+w/2) \times 1/(2\beta_0+1) - w/2 \leq -2$, so we need the Z_i admit a moment of order $z = w+2 \geq 2(3+2/\beta_0)$.

This completes the proof of lemma 8.4. □

9. Appendix

Lemma 9.1. *Let $T_1, ..., T_n$ n independent random variables and $S_n(T) = (1/n)\sum_{i=1}^n [T_i - \mathbb{E}(T_i)]$. Then, for $\eta > 0$,*

$$\begin{aligned} \mathbb{P}(|S_n(T)| \geq \eta) &\leq 2\exp\left(\frac{-n\eta^2/2}{\nu^2 + b\eta}\right) \\ &\leq 2\max\left(\exp\left(\frac{-n\eta^2}{4\nu^2}\right), \exp\left(\frac{-n\eta}{4b}\right)\right), \end{aligned} \tag{45}$$

where $Var(T_1) \leq \nu^2$ and $|T_1| \leq b$.

Acknowledgement

The authors thank Valentine Genon-Catalot and Fabienne Comte for helpful advices.

References

Bertoin, J. (1996). *Lévy processes*, volume 121 of *Cambridge Tracts in Mathematics*. Cambridge University Press, Cambridge.

Birgé, L. and Massart, P. (1998). Minimum contrast estimators on sieves: exponential bounds and rates of convergence. *Bernoulli*, 4(3):329–375.

Butucea, C. (2001). Exact adaptive pointwise estimation on Sobolev classes of densities. *ESAIM Probab. Statist.*, 5:1–31 (electronic).

Comte, F. and Genon-Catalot, V. (2009). Nonparametric estimation for pure jump Lévy processes based on high frequency data. *Stochastic Process. Appl.*, 119(12):4088–4123.

Comte, F. and Genon-Catalot, V. (2010). Nonparametric adaptive estimation for pure jump Lévy processes. *Ann. Inst. Henri Poincaré Probab. Stat.*, 46(3):595–617.

Figueroa-López, J. E. and Houdré, C. (2006). Risk bounds for the non-parametric estimation of Lévy processes. In *High dimensional probability*, volume 51 of *IMS Lecture Notes Monogr. Ser.*, pages 96–116. Inst. Math. Statist., Beachwood, OH.

Goldenshluger, A. and Lepski, O. (2011). Bandwidth selection in kernel density estimation: oracle inequalities and adaptive minimax optimality. *To appear in The Annals of Statistics.*

Jongbloed, G. and van der Meulen, F. H. (2006). Parametric estimation for subordinators and induced OU processes. *Scand. J. Statist.*, 33(4):825–847.

Kerkyacharian, G., Lepski, O., and Picard, D. (2001). Nonlinear estimation in anisotropic multi-index denoising. *Probab. Theory Relat. Fields*, 121:137–170.

Küchler, U. and Tappe, S. (2008). Bilateral gamma distributions and processes in financial mathematics. *Stochastic Process. Appl.*, 118(2):261–283.

Lepski, O. V. (1991). Asymptotically minimax adaptive estimation. I. Upper bounds. Optimally adaptive estimates. *Teor. Veroyatnost. i Primenen.*, 36(4):645–659.

Neumann, M. H. and Reiß, M. (2009). Nonparametric estimation for Lévy processes from low-frequency observations. *Bernoulli*, 15(1):223–248.

Tsybakov, A. B. (2009). *Introduction to nonparametric estimation*. Springer Series in Statistics. Springer, New York. Revised and extended from the 2004 French original, Translated by Vladimir Zaiats.

van Es, B., Gugushvili, S., and Spreij, P. (2007). A kernel type nonparametric density estimator for decompounding. *Bernoulli*, 13(3):672–694.

Watteel, R. N. and Kulperger, R. J. (2003). Nonparametric estimation of the canonical measure for infinitely divisible distributions. *J. Stat. Comput. Simul.*, 73(7):525–542.