KERNEL ESTIMATION FOR LÉVY DRIVEN STOCHASTIC CONVOLUTIONS

F. COMTE[(1)], V. GENON-CATALOT[(1)]

Abstract. We consider a Lévy driven stochastic convolution, also called continuous time Lévy-driven moving average model $X(t) = \int_0^t a(t-s)dZ(s)$ where Z is a Lévy martingale and the kernel $a(.)$ a deterministic function square integrable on $\mathbb{R}^+$. Given N *i.i.d.* continuous time observations $(X_i(t))_{t\in[0,T]}$, $i = 1, \dots, N$, distributed like $(X(t))_{t\in[0,T]}$, we propose two types of nonparametric projection estimators of a^2 under different sets of assumptions. We bound the $\mathbb{L}^2$-risk of the estimators and propose a data-driven procedure to select the dimension of the projection space, illustrated by a short simulation study. July 14, 2021
Mathematical Subject Classification (2010): 62G05-62M09-60G51.

Keywords and phrases: Continuous time moving average. Lévy processes. Model selection. Nonparametric estimation. Projection estimators. Stochastic convolution.

1. Introduction

In this paper, we consider the continuous time moving average (CMA) process, also called stochastic convolution,

$$X(t) = \int_0^t a(t-s)dZ(s) \tag{1}$$

where $(Z(t))_{t\geq 0}$ is a Lévy process such that $\mathbb{E}Z(1) = 0$, $\mathbb{E}Z^2(1) = 1$ and the kernel $a(.) : \mathbb{R}^+ \to \mathbb{R}$ is a deterministic square integrable function. Our aim is the nonparametric estimation of $a^2(.)$ from *i.i.d.* observations $(X_i(t), t \in [0,T], i = 1, \dots, N)$ distributed as $(X(t), t \in [0,T])$.

Thus, we deal with a sample of infinite dimensional data. Such data are often encountered in various fields, e.g. in econometrics (panel data) and more generally in the field of functional data analysis (FDA), see Hsiao (2003), Ramsay *et al.* (2007), Wang *et al.* (2016).

CMA processes have been largely studied in the past decades. Indeed, they provide a large class of stochastic processes including the classical continuous time ARMA (CARMA) processes and also more involved models such as fractional Lévy processes. Generally, stationary versions of $(X(t))_{t\geq 0}$ are investigated, *i.e.* $Y(t) = \int_{-\infty}^{+\infty} a(t-s)dZ(s)$ (see *e.g.* Rajput and Rosinski (1989), Brockwell (2001), Marquart (2006), Brockwell and Lindner (2009), Bender *et al.* (2012), Brockwell *et al.* (2013)). These processes are well fitted to modelling various phenomena in fields such as econometrics and finance (see Comte and Renault (1996)) or electricity prices (see Klüppelberg *et al.* (2010)). Schnurr and Woerner (2011) study the so-called well-balanced Ornstein-Uhlenbeck process and its correlation structure and show that this model can be used

[(1)]: Université de Paris, CNRS, MAP5, UMR 8145, F-75006 Paris, FRANCE,
email: fabienne.comte@parisdescartes.fr,
valentine.genon-catalot@mi.parisdescartes.fr.

as volatility process in stochastic volatility models.
Estimation properties are generally studied from the observation of one sample path in stationary regime (like $(Y(t))_{t\geq 0}$ (see *e.g.* Brockwell *et al.* (2013)). In the same framework, Belomestny *et al.* (2019) are interested in estimation of the Lévy characteristics of $(Z(t))_{t\geq 0}$.
In our contribution, stationarity of the process is not required: T is fixed and N is large. To our knowledge, few papers are concerned with statistical properties in this context. In a previous paper (Comte and Genon-Catalot (2021)), we restrict our attention to Gaussian CMA processes, *i.e.* $Z(t) = W(t)$ is a Wiener process and provide nonparametric projection estimators of the function $a^2(.)$. Proofs, especially for the data-driven procedure, strongly rely on the Gaussian character of $(X(t))_{t\geq 0}$ and cannot be straightforwardly extended to the case where $(Z(t))_{t\geq 0}$ is a Lévy process. The question of this extension is studied here.
In Section 2, we precise the model and the assumptions. In Section 3, we define two collections of projection estimators depending on whether $X(t)$ is a semi-martingale or not. Relying on results of Basse and Pedersen (2009), we establish that the distinction between these two cases is the same as when $Z = W$ is a Brownian motion, *i.e.* when $a(.)$ is continuously differentiable on $[0, +\infty)$ or not. The projection spaces are either, for fixed T, spaces generated by the trigonometric basis of $\mathbb{L}^2([0, T])$ or for large T spaces generated by the Laguerre basis of $\mathbb{L}^2(\mathbb{R}^+)$. Bounds for the $\mathbb{L}^2$-risk of the estimators are provided. A short discussion deals with the impact of discretization of observed paths on estimators' risk bound. In Section 4, we propose a data-driven procedure to select the dimension of the projection space and obtain risk bounds for the resulting estimator proving that it is adaptive in the sense that its risk automatically achieves the compromise between the squared bias and the variance. The findings are illustrated through a short simulation study with Z a compound Poisson process. Proofs, especially of the adaptive result, are completely different from the ones in Comte and Genon-Catalot (2021). Section 5 states some concluding remarks. Section 6 contains proofs. Finally Section 7 gives the necessary recap on Laguerre functions, the Talagrand inequality on which relies our proof of Section 4 and the way to compute or bound moments of $(X(t))_{t\geq 0}$.

2. Lévy driven moving averages

Consider a Lévy process $(Z(t))_{t\geq 0}$ with no Gaussian part and Lévy measure $\nu(dx) = n(x)dx$ satisfying

[H1] $\int_{\mathbb{R}} x^2 n(x)dx < +\infty$ and we assume that $\int_{\mathbb{R}} x^2 n(x)dx = 1$.

The second part of [H1] is an identifiability condition. Without it, we would estimate $\left(\int_{\mathbb{R}} x^2 n(x)dx\right) \times a^2(.)$. Below, we need stronger conditions near infinity for the Lévy density summarized by :

[H2](p) $\quad k_{2p} := \int_{\mathbb{R}} x^{2p} n(x)dx < +\infty$.

We assume that the characteristic function of $Z(t)$ is equal to:

$$\mathbb{E}e^{iuZ(t)} = \exp\,[t \int_{\mathbb{R}} \left(e^{iux} - 1 - iux\right) n(x)dx],$$

so that $\mathbb{E}Z(1) = 0$, $\mathbb{E}Z^2(1) = 1$. Then, $(Z(t))$ is a Lévy martingale which can be written as:

$$Z(t) = \int_{(0,t]} \int_{\mathbb{R}} x(\hat{p}(ds, dx) - dsn(x)dx),$$

where $\hat{p}(ds, dx)$ is the random Poisson measure associated with its jumps. We consider a càdlàg version of the Lévy moving average process:

$$X(t) = \int_0^t a(t-s)dZ(s) \tag{2}$$

where we aim at estimating $g = a^2$ under assumptions of type:

[H3](q) The function $g(t) = a^2(t)$ belongs to $\mathbb{L}^q(\mathbb{R}^+)$, *i.e.* $\int_0^{+\infty} g^q(s)ds = \int_0^{+\infty} a^{2q}(s)ds < +\infty$.

Assumptions [H1] and [H3](1) ensure the existence of (2) (see Section 6.1). Setting

$$G(t) = \int_0^t a^2(s)ds = \int_0^t g(s)ds, \tag{3}$$

we have:

$$\mathbb{E}X^2(t) = \int_0^t \int_{\mathbb{R}} a^2(t-s)dsx^2 n(x)dx = \int_0^t a^2(u)du = G(t).$$

Two cases are to be distinguished:

(1) $X(t)$ is a semi-martingale (more precisely, a $(\mathcal{F}_t^Z)_{t\geq 0}$-semimartingale where $(\mathcal{F}_t^Z)_{t\geq 0}$ is the natural filtration of $(Z_t)_{t\geq 0}$),
(2) $X(t)$ is not a semi-martingale.

In Case (2), we cannot give sense to a stochastic integral $\int_0^t H(s)dX(s)$ for a predictable process $H(s)$. A sufficient condition for case (1) to hold is stated in the following proposition.

Proposition 1. *Assume that $t \mapsto a(t)$ belongs to $C^1([0, +\infty))$. Then,*

$$X(t) = a(0)Z(t) + \int_0^t \left(\int_0^u a'(u-s)dZ(s) \right) du, \quad t \geq 0. \tag{4}$$

3. Projection estimators on a fixed space.

We denote respectively by $\|.\|_T$ (resp. $\langle ., . \rangle_T$) the norm (resp. the scalar product) of $\mathbb{L}^2([0,T])$ and $\|.\|$ (resp. $\langle ., . \rangle$) the norm (resp. the scalar product) of $\mathbb{L}^2(\mathbb{R}^+)$.
As a function of $\mathbb{L}^2([0,T])$ (resp. $\mathbb{L}^2(\mathbb{R}^+)$), when considering an orthonormal basis $(\varphi_{j,T}, j \geq 0)$ (resp. $(\varphi_j, j \geq 0)$ of these spaces, g may be developped into

$$g = \sum_{j\geq 0} \theta_j \varphi_{j,T} \qquad (\text{resp. } g = \sum_{j\geq 0} \theta_j \varphi_j) \tag{5}$$

where $\theta_j = \langle g, \varphi_j \rangle_T$ (resp. $\theta_j = \langle g, \varphi_j \rangle$). The estimation by projection method consists in defining estimators of the coefficients θ_j, say $\hat{\theta}_j$. A collection of projection estimators $(\hat{g}_m, m \geq 0)$ is then by obtained by setting

$$\hat{g}_m = \sum_{j=0}^m \hat{\theta}_j \varphi_j.$$

This requires first the choice of appropriate orthonormal bases, second the choice of an adequate optimal or possibly data-driven m.

In this paragraph, we define our bases and study the $\mathbb{L}^2$-risk of the projection estimators for fixed m. According to the assumptions on the function $a(.)$, different estimators of the coefficients θ_j are proposed. The optimal choice of m may be deduced from the risk bounds.

To build estimators of g, we use two collections of projection spaces.

(1) For fixed T, we estimate g on $[0,T]$. We define the collection $(S_m^{Trig}, m \geq 0)$ of subspaces of $\mathbb{L}^2([0,T])$ where m is odd, generated by the orthonormal trigonometric basis $(\varphi_{j,T})$, $\varphi_{0,T}(t) = \sqrt{1/T}\mathbf{1}_{[0,T]}(t)$, $\varphi_{2j-1,T}(t) = \sqrt{2/T}\cos(2\pi jt/T)\mathbf{1}_{[0,T]}(t)$ and $\varphi_{2j,T}(t) = \sqrt{2/T}\sin(2\pi jt/T)\mathbf{1}_{[0,T]}(t)$ for $j = 1, \ldots, (m-1)/2$. The following properties are useful

$$\sum_{j=0}^{m-1} \varphi_{j,T}^2(t) = \frac{m}{T} \quad \text{and} \quad \int_0^T \varphi_{0,T}(t)dt = \sqrt{T}, \int_0^T \varphi_{j,T}(t)dt = 0 \quad \text{for} \quad j \neq 0.$$

(2) For either T fixed but large enough, or T tending to infinity, we estimate g on $\mathbb{R}^+$. We define the collection of subspaces of $\mathbb{L}^2(\mathbb{R}^+)$, generated by the orthonormal Laguerre basis (see Section 7.1):

$$\ell_j(t) = \sqrt{2}L_j(2t)e^{-t}\mathbf{1}_{t\geq 0}, \quad j \geq 0, \quad L_j(t) = \sum_{k=0}^{j}(-1)^k \binom{j}{k}\frac{t^k}{k!}. \tag{6}$$

We set $S_m^{Lag} = \text{span}\{\ell_j, j = 0, \ldots, m-1\}$, and the following holds

$$\forall t \geq 0, \quad \sum_{j=0}^{m-1} \ell_j^2(t) \leq 2m \quad \text{and} \quad \int_0^{+\infty} \ell_j(t)dt = \sqrt{2}(-1)^j.$$

3.1. **Estimation of $g = a^2$ when $(X(t))_{t\geq 0}$ is a semimartingale.** Here, we assume:

[H4] $t \mapsto a(t)$ belongs to $C^1([0, +\infty))$.

Lemma 1. *Assume* [H1], [H3](1) *and* [H4]. *Denoting by $\theta_j = \langle g, \varphi_j \rangle$, we have*

$$\mathbb{E}\left(\int_0^{+\infty} \varphi_j(s)X(s_-)dX(s)\right) = \frac{1}{2}\left(\theta_j - g(0)\int_0^{+\infty} \varphi_j(s)ds\right), \quad \mathbb{E}\left(\sum_{s\leq T}[\Delta X(s)]^2\right) = Tg(0).$$

Relying on this lemma, we can set:

$$\widehat{\theta}_j = \widehat{\theta}_j(N,T) = 2\left[\frac{1}{N}\sum_{i=1}^{N}\left(\int_0^T \varphi_j(s)X_i(s_-)dX_i(s)\right)\right] + (g(0))^\dagger \int_0^T \varphi_j(s)ds. \tag{7}$$

where $(g(0))^\dagger$ is an estimator of $g(0)$ equal to

$$(g(0))^\dagger = \frac{1}{T}\frac{1}{N}\sum_{i=1}^{N}\sum_{s\leq T}(\Delta X_i(s))^2. \tag{8}$$

The projection estimator of g on a fixed space S_m is given by:

$$\widehat{g}_m = \sum_{j=0}^{m-1} \widehat{\theta}_j \varphi_j.$$

Remark 1. *By the Ito formula with jumps, we have:*

$$-\int_{(0,T]} X^2(s)\varphi_j'(s)ds = 2\int_{(0,T]} \varphi_j(s)X(s_-)dX(s) + \sum_{0<s\leq T} \varphi_j(s)(\Delta X(s))^2 - \varphi_j(T)X_T^2 \tag{9}$$

where:

$$\mathbb{E}\sum_{0<s\leq t} \varphi_j(s)(\Delta X(s))^2 = a^2(0)\mathbb{E}\sum_{0<s\leq T} \varphi_j(s)(\Delta Z(s))^2 = a^2(0)\int_0^T \varphi_j(s)ds.$$

This formula is useful to understand the link between $\widehat{\theta}_j$ defined above and $\widetilde{\theta}_j$ defined in the second strategy below, but it only holds under [H4] *(which is not assumed in the second case).*

The following proposition gives a bound for the $\mathbb{L}^2$-risk of $\widehat{g}_m$ in the case of fixed T and the trigonometric basis.

Proposition 2. *Assume* [H1], [H3](1), [H3](2) *and* [H4]. *When* $(\varphi_j = \varphi_{j,T})$ *is the trigonometric basis,*

$$\mathbb{E}(\|\widehat{g}_m - g\|_T^2) \quad \leq \quad \|g_m - g\|_T^2 + 16g(0)G(T)\frac{m}{N} + 8\mathfrak{C}_{1,T}\frac{T}{N} + 2g^2(0)\frac{k_4}{N} \tag{10}$$

$$\textit{where} \qquad \mathfrak{C}_{1,T} := 3(G^2(T) + G_1^2(T)) + k_4(\|g\|_T^2 + \|g_1\|_T^2)$$

and $k_4 = \int x^4 n(x)dx$, $g_1 = (a')^2, G_1(.) = \int_0^{.} g_1(s)ds$. *Recall that* G *is defined in (3), that* g_m *denotes the orthogonal projection of* g *on* S_m^{Trig} *and that* $\|u\|_T^2 = \int_0^T u^2(s)ds$.

Now, we give risk-bounds in case of an orthonormal basis of $\mathbb{L}^2(\mathbb{R}^+)$ and a special inequality for the Laguerre basis.

Proposition 3. *Assume* [H1], [H3](1), [H3](2) *and* [H4] *and that* $\|g_1\| < +\infty$.
If (φ_j) *is an orthonormal basis of* $\mathbb{L}^2(\mathbb{R}^+)$, *for all* $T \geq 1, N \geq 1, m \geq 0$, *we have*

$$\mathbb{E}(\|\widehat{g}_m - g\|^2) \quad \leq \quad \|g_m - g\|^2 + 16g(0)G(T)\frac{m}{N} + 8\mathfrak{C}_{2,T}\frac{T}{N} + 2g^2(0)\frac{k_4}{N} + \int_T^{+\infty} g^2(s)ds \tag{11}$$

$$\textit{where} \qquad \mathfrak{C}_{2,T} := 3(G^2(T) + G_1^2(T)) + k_4(\|g\|^2 + \|g_1\|^2)]$$

If (φ_j) *is the Laguerre basis of* $\mathbb{L}^2(\mathbb{R}^+)$ *and* $T \geq 6m - 3$, *then*

$$\begin{aligned}\mathbb{E}(\|\widehat{g}_m - g\|^2) \quad \leq \quad & \|g_m - g\|^2 + 8C\mathfrak{C}_{2,T}\frac{m^2}{N} + 16g(0)G(T)\frac{m}{N} + \frac{2}{N}k_4 g^2(0) \\ & + C'\|a\|^2 m \exp(-12\gamma_2 m)\end{aligned} \tag{12}$$

where C, C' *and* γ_2 *are positive constants depending on the basis only.*

The bounds obtained in Propositions 2 and 3 contain three types of terms: the first one is the usual squared bias term $\|g_m - g\|^2$ due to the projection method, decreasing when m increases, the second one is the variance term, increasing with m, and the last ones are residuals.

Let us comment (10) and (11). If $g(0) \neq 0$, the variance order in both cases is m/N. For choosing m, a compromise must be done between the first two terms. If $g(0) = 0$, the variance term vanishes, and m must be chosen as large as possible. Note that this case corresponds to $(X(t))$ derivable.

The difference between (10) and (11) lies in the additional term $\int_T^{+\infty} g^2(s)ds$. In (10), T is fixed and the residual term has negligible order $1/N$. In (11), T must be large enough for the additional term to be small, but not too much because the other residuals terms are of order T/N (see numerical results in Table 1 of Comte and Genon-Catalot (2021)).

The result in (12) is specific to the Laguerre basis with $T \geq 6m - 3$. The variance order m^2/N is larger but the residual terms do not depend on T ($G(T)$ is bounded). The choice of m relies on a compromise between $\|g - g_m\|^2$ and m^2/N. We can consider here the case where $T \to +\infty$: if, in addition to the condition $\|g_1\| < +\infty$, it holds that $(a')^2 \in \mathbb{L}^1(\mathbb{R}^+)$, then $\mathfrak{C}_{2,T}$ is bounded independently of T.

3.2. **Estimation of $g = a^2$ when $(X(t))$ is not a semi-martingale.** In this section, we assume that the basis functions are differentiable on their support. The following Lemma allows to define another estimator.

Lemma 2. *Assume that* [H1], [H3](1) *hold and that* $(\varphi_j)_j$ *is differentiable on* $[0,T]$, *then*

$$\mathbb{E}\left(\int_0^T \varphi_j'(s)X^2(s)ds\right) = \varphi_j(T)G(T) - \int_0^T g(u)\varphi_j(u)du.$$

Therefore, we can set

$$\widetilde{\theta}_j = -\frac{1}{N}\sum_{i=1}^N \left(\int_0^T \varphi_j'(s)X_i^2(s)ds\right) + \varphi_j(T)\widehat{G}(T) \quad \text{and} \quad \widehat{G}(T) = \frac{1}{N}\sum_{i=1}^N X_i^2(T). \tag{13}$$

If $\varphi_j = \varphi_{j,T}$ is the trigonometric basis, then $\varphi_{0,T}(T) = 1/\sqrt{T}, \varphi_{2j-1,T}(T) = \sqrt{2/T}, \varphi_{2j,T}(T) = 0$, $j \geq 1$. Then we define the estimator by

$$\widetilde{g}_m = \sum_{j=0}^{m-1} \widetilde{\theta}_j \varphi_j.$$

We introduce the assumption:

[H5] $\displaystyle\int_0^1 \frac{\|g\|_s^2}{s} ds = c_0 < +\infty$ where we recall that $\|g\|_s^2 = \int_0^s g^2(s)ds$.

Proposition 4. *Assume* [H1] *and* [H3](2).

- *If* $(\varphi_j = \varphi_{j,T})$ *the trigonometric basis, then*

$$\mathbb{E}(\|\widetilde{g}_m - g\|_T^2) \leq \|g_m - g\|_T^2 + \frac{2}{N}(3G^2(T) + k_4\|g\|_T^2)\left(\frac{4\pi^2 m^2}{T} + \frac{m}{T}\right). \tag{14}$$

- *Let* $(\varphi_j = \ell_j)$ *be the Laguerre basis.*
  - *Then, for all* $T \geq 1, N \geq 1, m \geq 0$,

$$\mathbb{E}(\|\widetilde{g}_m - g\|^2) \leq \|g_m - g\|^2 + 4\mathfrak{C}_{3,T}\frac{m}{N} + 4\frac{T}{N}(3G^2(T) + k_4\|g\|_T^2) + \int_T^\infty g^2(s)ds$$

$$\text{with} \qquad \mathfrak{C}_{3,T} := \left[3G^2(T) + k_4\|g\|_T^2 + 2\left(\int_0^T s^{-1}[3G^2(s) + k_4\|g\|_s^2]ds\right)\right]$$

where, if, in addition, [H5] *holds,*

$$\left(\int_0^T s^{-1}[3G^2(s) + k_4\|g\|_s^2]ds\right) \leq (3 + k_4)\left(c_0 + \log(T)\|g\|_T^2\right).$$

  - *If* $T \geq 6(m-1) + 3 = 6m - 3$ *and* (φ_j) *is the Laguerre basis, then*

$$\mathbb{E}(\|\widetilde{g}_m - g\|^2) \leq \|g_m - g\|^2 + c_1(3G^2(T) + k_4\|g\|_T^2)\frac{m^3}{N} + c_2\|a\|^2\frac{m}{N}\exp(-12\gamma_2 m) \tag{15}$$

where c_1, c_2, γ_2 *are constants depending on the basis only.*

Comments on the bounds obtained in Proposition 4 are similar to the comments given after Proposition 2 and 3. Inequality (14) can be compared to (10) and we mainly notice that the variance term increases from m/N to m^2/N. Inequality (15) corresponds to (12) with variance increase from m^2/N to m^3/N. These losses are due to the more general assumptions. In Inequality (15), we can consider $T \to +\infty$.

Moreover, we refer to Section 3 of Comte and Genon-Catalot (2021) for a discussion on optimal theoretical choice of m and on rates of convergence that can be deduced from Propositions 2, 3 and 4, on dedicated function spaces: periodic Fourier-Sobolev spaces for the trigonometric basis and Sobolev-Laguerre spaces for the Laguerre basis.

3.3. **About the impact of discretisations.** It is now commonly admitted that a fine discrete sampling of continuous time processes can be obtained (high frequency data) which is very close to a continuous time record. This justifies our sampling scheme.

However, even if it makes sense to consider the continuous time set-up to build up an estimation theory, it is important to quantify the impact of discretisations on our estimators and this is the aim of the result below.

We restrict our attention to the second type of estimators under assumptions [H0]-[H1].

Suppose we observe $(X_i(k\Delta), k = 1, \dots, n, i = 1, \dots, N)$ with $\Delta = \Delta_n = T/n$ and consider the estimators

$$\widetilde{g}_m^\Delta = \sum_{j=0}^{m-1} \widetilde{\theta}_j^\Delta \varphi_j,$$

where

$$\widetilde{\theta}_j^\Delta = -\frac{1}{N} \sum_{i=1}^{N} \sum_{k=1}^{n} \Delta \varphi_j'(k\Delta) X_i^2(k\Delta) + \varphi_j(T)\widehat{G}(T). \tag{16}$$

Proposition 5. *Assume [H0]-[H1]. Then,*

$$\mathbb{E}\|\widetilde{g}_m^\Delta - g\|^2 \leq \mathbb{E}\|\widetilde{g}_m - g\|^2 + C\Delta^2 G^2(T)(m^3 + m^5) + C\mathbb{E}X^4(T)\frac{1}{N}(\Delta^2 m^5 + \Delta m^\alpha)$$

with $\alpha = 2$ *for the trigonometric basis,* $\alpha = 3$ *for the Laguerre basis.*

Thus, the risk of the discretized estimator is incremented by terms of order of order $\Delta^2 m^5 + \Delta m^2/N$ for the trigonometric basis and of order $\Delta^2 m^5 + \Delta m^3/N$ for the Laguerre basis.
In the case of the trigonometric basis, assume that $m^2 \leq N$ so that the variance term of $\mathbb{E}\|\widetilde{g}_m - g\|^2$ is bounded. Then, if $\Delta \lesssim N^{-7/4}$, $\Delta^2 m^5 + \Delta m^2/N \lesssim 1/N$.
In the case of the Laguerre basis, assume that $m^3 \leq N$ to bound the variance term of $\mathbb{E}\|\widetilde{g}_m - g\|^2$. Then, if $\Delta \lesssim N^{-4/3}$, $\Delta^2 m^5 + \Delta m^3/N \lesssim 1/N$.

4. Adaptation

4.1. **Theoretical result.** Considering the main terms of all risk bounds, we can see that a compromise must be done between the squared bias terms which decrease when m increases while the variance terms increase. In this section, we describe a procedure allowing for a data driven selection of m and we prove that the final estimator reaches an effective tradeoff in term of its integrated $\mathbb{L}^2$-risk bound. For sake of conciseness, we only study the procedure for $\widetilde{g}_m$ and the trigonometric basis.

Let $\mathcal{M}_N = \{m \in \mathbb{N}, m^2 \leq NT\}$ be a collection of models such that the variance of $\widetilde{g}_m$ is bounded and set

$$\widetilde{m} = \arg\min_{m \in \mathcal{M}_N} \left\{ -\|\widetilde{g}_m\|^2 + \text{pen}(m) \right\}, \tag{17}$$

where, for a constant κ precised below,

$$\text{pen}(m) = \kappa \log N \frac{m^2}{NT} \mathbb{E}X^4(T)$$

Theorem 1. *Consider the collection of estimators* $\widetilde{g}_m$ *in the trigonometric basis on* $[0,T]$, *with model selection* $\widetilde{m}$ *given by (17). Assume* $N \geq 3$, [H1], [H2]*(4) and* [H3]*(4). Then, there exists a numerical constant* κ_0 *such that, for all* $\kappa \geq \kappa_0$, *the following holds:*

$$\mathbb{E}\|\widetilde{g}_{\widetilde{m}} - g\|^2 \quad \leq \quad \inf_{m\in\mathcal{M}_N}(3\|g_m - g\|^2 + 4\mathrm{pen}(m)) + C\frac{\log N}{N}.$$

The infimum in the risk bound implies that the $\mathbb{L}^2$-risk of $\widetilde{g}_{\widetilde{m}}$ achieves automatically the best compromise between the square bias term and the variance term.

In practice, we replace the unknown term $\mathbb{E}X^4(T)$ in the penalty by its empirical estimator $N^{-1}\sum_{i=1}^N X_i^4(T)$. Theorem 1 can be extended to this substitution. For the implementation, the constant κ must be fixed. It is standard that the numerical value for κ_0 given in the proof is too large. This is why it must rather be calibrated by preliminary simulation experiments; this is done in Section 5 of Comte and Genon-Catalot (2021), for Z a Brownian motion. More generally, results on simulated data are given in the latter paper especially for examples where $a(t) = t^d \exp(-\alpha t)$ with various values of d. It is worth noting that our assumptions [H3](2) and [H5] hold if $d > -1/4$.

4.2. **Short numerical illustration.** In this section we provide some elements about practical implementation of the method. To that aim, we consider the case where $Z(t) = \sum_{k=1}^{N(t)} \xi_k$ is a compound Poisson process with $(N(t))_{t\geq 0}$ a Poisson process with intensity λ and $(\xi_k, k \geq 1)$ a sequence of *i.i.d.* random variables independent of the Poisson process $(N(t))$. We assume that $\mathbb{E}\xi_1 = 0, \mathbb{E}\xi_1^2 = \sigma^2$ and $\lambda\sigma^2 = 1$.

If $a(.) \in C([0,+\infty))$, then,

$$X(t) = \sum_{n:\tau_n\leq t} a(t-\tau_n)\xi_n$$

where (τ_n) is the sequence of jumps times of $(N(t))$. We have $X(t) = 0$ on $(N(t) = 0)$ and $X(t) = \sum_{k=1}^n a(t-\tau_k)\xi_k$ on $(N(t) = n)$. Thus,

$$\begin{aligned} X(t) &= 0 \text{ for } t \in [0, \tau_1) \\ X(t) &= \sum_{k=1}^{n} a(t-\tau_k)\xi_k \text{ for } t \in [\tau_n, \tau_{n+1}), n \geq 1. \end{aligned}$$

The jump times of X are the sequence $(\tau_n, n \geq 1)$ with $X(\tau_n^-) = \sum_{k=1}^{n-1} a(\tau_n - \tau_k)\xi_k$ and $X(\tau_n) = \sum_{k=1}^n a(\tau_n - \tau_k)\xi_k$. The jump of X at τ_n is $\Delta X(\tau_n) = a(0)\xi_n$. If $a(0) = 0$, the process $(X(t))$ is continuous, see also (4).

In practice, we took the ξ_k's as Gaussian $\mathcal{N}(0,\sigma^2)$, with $\lambda = 8$ and $\sigma = 1/\sqrt{\lambda}$. The observations are generated as

$$X(k\Delta) \equiv \sum_{j,\tau_j\leq k\Delta} a(k\Delta - \tau_j)\xi_j \text{ for } k = 1,\ldots,n$$

with $n^\star = 100$ random variables τ_j in all cases; the parameters are such that $\tau_{n^\star}$ has order (slightly more than) 10 in all cases. Indeed we have $T = 10 = n\Delta$ with $n = 2000$ and $\Delta = 0.1/20$. The number of observations in the results presented here is $N = 4000$. We consider four functions: a function denoted by a_0 and functions a_2, a_3 and a_7 borrowed from [12] (in all cases, recall that $g_i(t) = a_i^2(t)$):

(1) $a_0(t) = (t-5)/\omega_0^{1/2}$, so that $g_0(0) = a_0^2(0) \neq 0$, $\omega_0 = \sqrt{1250}$ is such that $\int_0^{10} g_0^2(u)du = 1$,
(2) $a_2(t) = (\beta(3,3,t/10)/\omega_2^{1/2})^{1/2}$ where $\beta(p,q,x)$ is the density of a $\beta(p,q)$ distribution at point x and $\omega_2 = 14.157$ is such that $\int_{\mathbb{R}^+} g_2^2(u)du \approx 1$.

(3) $a_4(t) = 10b(6t)/(\omega_4)^{0.25}$ with $b(t) = 0.3\Gamma(3,2,t) + 0.7\Gamma(7,4,t)$ where $\Gamma(p,q,x)$ is the density of a $\Gamma(p,q)$ distribution at point x and $\omega_4 = 0.03048$ is such that $\int_{\mathbb{R}^+} g_4^2(u)du \approx 1$.
(4) $a_7(t) = t^{-0.125}e^{-t/5}$, where $\int_{\mathbb{R}^+} g_7^2(u)du \approx 2$.

The estimators are computed in the trigonometric basis, relying on formula (13) for the coefficients $\widetilde{\theta}_j$ of $\widetilde{g}_m = \sum_{j=0}^{m-1} \widetilde{\theta}_j \varphi_{j,T}$ for $m \in \{1, \ldots, 45\}$ where $\widetilde{m}$ selected with (17) and $\kappa = 0.2$ in the penalty $\text{pen}(m)$. Figures 1-4 illustrate the results obtained with the estimation algorithm. Left plots represent one path of $t \mapsto X(t)$ on $[0, 10]$, clearly it has jumps in Figure 1 for a_0, $a_0(0) \neq 0$ while it is continuous for a_2 and a_4 in Figures 2-3 which are such that $a_2(0) = a_4(0) = 0$. Right plots show beams of 25 estimators for each function, with associated MISE given below. The mean of the selected dimensions are also given. They can be compared to the MISE and mean dimension of the best estimator among the collection called "oracle" because it is computed by using the knowledge of the true function. The orders of the MISEs are comparable to the oracles, the selected dimensions seem to be in all cases a little smaller than the oracle. This means that the penalty constant is probably slightly too large, but we kept the choice made in [12]. Slight over-penalization is known to be more safe, at least compared to under-penalization, in term of MISEs orders.

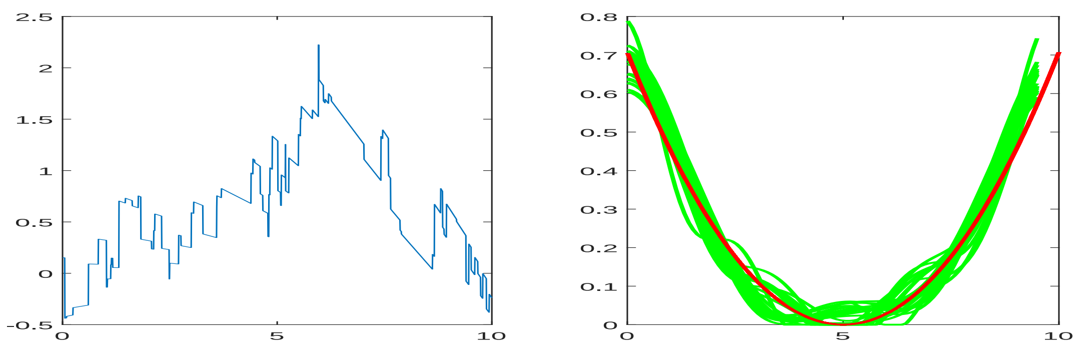

FIGURE 1. Function $g_0(x) = a_0^2(x)$. Left: example of one simulated path. Right: 25 estimated functions. MISE= 0.018 (oracles 0.016), mean of selected dimensions: 4.9 (of oracles 6.4). $N = 4000$, $T = 10$

5. Concluding remarks

In this paper, we study the nonparametric estimation of a^2 from *i.i.d.* observations $(X_i(t), t \in [0, T]), i = 1, \ldots, N)$ distributed as (1). We proceed by projection method on finite dimensional subspaces of $\mathbb{L}^2(\mathbb{R}^+)$. Two different types of estimators are proposed depending on whether $(X(t))_{t \geq 0}$ is a semi-martingale or not and a data-driven procedure is proposed for the most general type of estimators. In our previous paper (where $Z = W$ a Wiener process, Comte and Genon-Catalot (2021)), proofs relied strongly on the Gaussian character of $(X(t))$. The extension to the Lévy case is not straightforward and relies on the general deviation inequality given in the Appendix.

The case where the driving process (Z_t) is a more general Lévy process having a Brownian component and a jump component is also interesting. But then $Z_t = W_t + L_t$ with (W_t) a Brownian motion and (L_t) a pure-jump Lévy process and $(W_t), (L_t)$ independent. Therefore, the observed process becomes $X(t) = X_W(t) + X_L(t)$ where X_W and X_L are independent. Therefore, the study of the estimators based on $X(t)$ can be deduced without much difficulty of

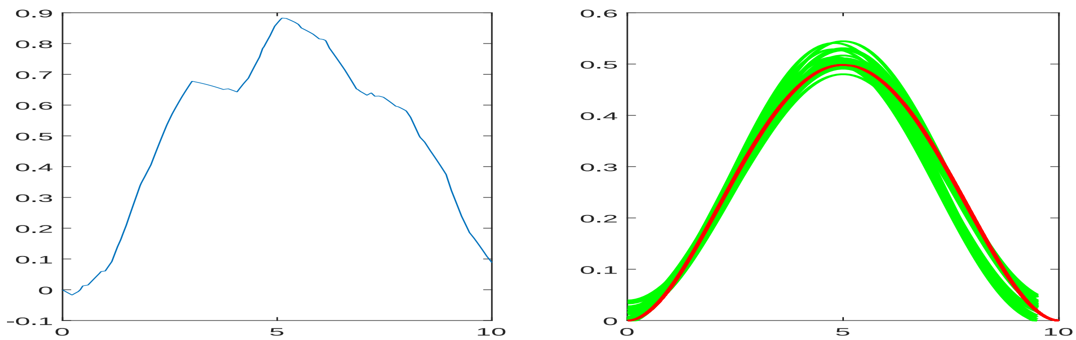

FIGURE 2. Function $g_2(x) = a_2^2(x)$. Left: example of one simulated path. Right: 25 estimated functions. MISE= 0.004 (oracles 0.002), mean of selected dimensions: 2.3 (of oracles 2.6). $N = 4000$, $T = 10$

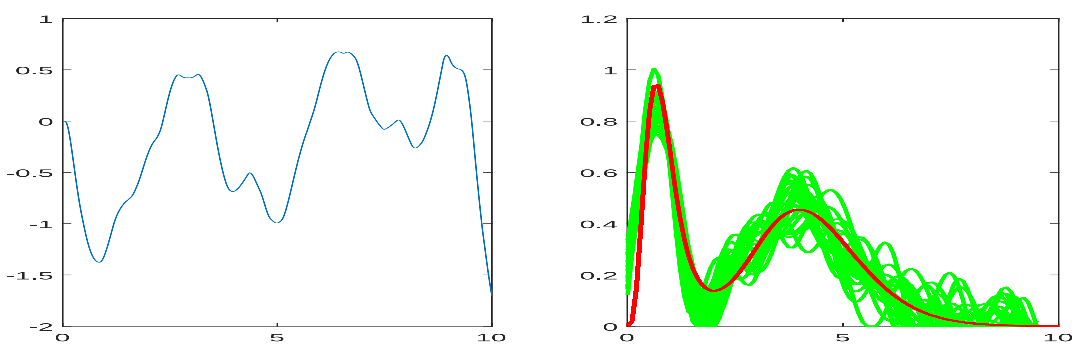

FIGURE 3. Function $g_4(x) = a_4^2(x)$. Left: example of one simulated path. Right: 25 estimated functions. MISE= 0.081 (oracle 0.067), mean of selected dimensions: 13.0 (of oracles 15.3). $N = 4000$, $T = 10$

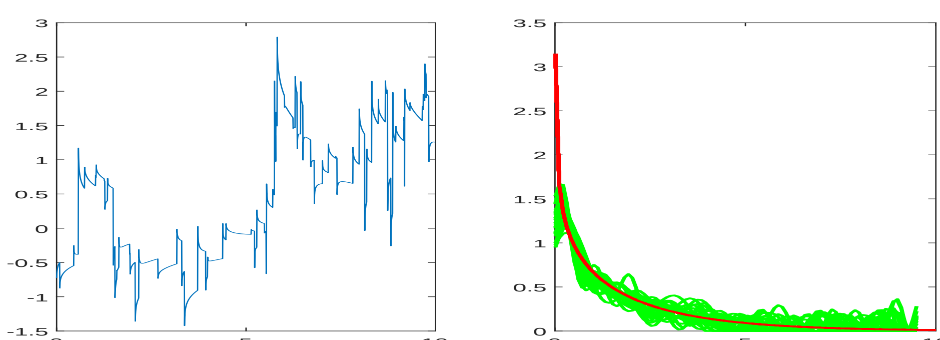

FIGURE 4. Function $g_7(x) = a_7^2(x)$. Left: example of one simulated path. Right: 25 estimated functions. MISE= 0.768 (oracle 0.722), mean of selected dimensions: 16.7 (of oracles 24.6). $N = 4000$, $T = 10$

the separate cases $X = X_W$, $X = X_L$ that we have treated.
From the theoretical and practical points of view, the questions of optimality of our estimators would be worth of investigation.

6. Proofs

6.1. Proof of the existence of (2).

$$\mathbb{E}e^{iuZ(t)} = \exp t[iu\gamma + \int_{\mathbb{R}} \left(e^{iux} - 1 - iux\mathbf{1}_{|x|\leq 1}\right) n(x)dx],$$

where $\gamma = -\int_{\mathbb{R}} x\mathbf{1}_{|x|>1}n(x)dx$ and $\mathbb{E}Z(1) = 0 = \gamma + \int_{\mathbb{R}} x\mathbf{1}_{|x|>1}n(x)dx$. According to Rajput and Rosiński (1989) (Theorem 2.7), see also Basse and Pedersen (2009), the existence of (2) is ensured if and only if, for all t, the following conditions hold:

$$\int_0^t \int_{\mathbb{R}} \left(x^2a^2(s) \wedge 1\right) dsn(x)dx < \infty, \quad \int_0^t \left|a(s)\left(\gamma + \int_{\mathbb{R}} x(\mathbf{1}_{|xa(s)|\leq 1} - \mathbf{1}_{|x|\leq 1})n(x)dx\right)\right| ds < \infty.$$

Note that:

$$\int_0^{+\infty} \int_{\mathbb{R}} \left(x^2a^2(s) \wedge 1\right) dsn(x)dx \leq \int_0^{+\infty} a^2(s)ds \int x^2n(x)dx.$$

For the second one, we have:

$$\begin{aligned} & \int_0^t \left|a(s)\left(\gamma + \int_{\mathbb{R}} x(\mathbf{1}_{|xa(s)|\leq 1} - \mathbf{1}_{|x|\leq 1})n(x)dx\right)\right| ds \\ = & \int_0^t |a(s)\mathbb{E}Z(1)|\, ds + \int_0^t \left|xa(s)\left(\mathbf{1}_{|xa(s)|\leq 1} - 1\right)\right| n(x)dxds \\ = & \int_0^t \left|xa(s)\left(\mathbf{1}_{|xa(s)|>1}\right)\right| n(x)dxds \leq \int_0^{+\infty} a^2(s)ds \int_{\mathbb{R}} x^2n(x)dx. \quad \square \end{aligned}$$

6.2. Proof of Proposition 1.

In Basse and Pedersen (2009) (Theorem 3.1), it is proved that, if $(Z(t))$ is of unbounded variation, $(X(t))$ is an $(\mathcal{F}_t^Z)_{t\geq 0}$-semimartingale if and only if $a(t)$ is absolutely continuous on $\mathbb{R}^+$ with a density a' satisfying, for all $t \geq 0$:

$$\int_0^t \int_{[-1,1]} \left((xa'(s))^2 \wedge |xa'(s)|\right) n(x)dxds < \infty \tag{18}$$

We have under [H1], [H3](1), [H3](2) and [H4]

$$\int_0^t \int_{[-1,1]} \left((xa'(s))^2 \wedge |xa'(s)|\right) n(x)dxds \leq \int_0^t (a'(s))^2 \int_{\mathbb{R}} x^2n(x)dx < \infty$$

So (18) holds. If $(Z(t))$ is of bounded variation (which is equivalent to $\int |x|n(x)dx < \infty$), $(X(t))$ is an $(\mathcal{F}_t^Z)_{t\geq 0}$-semimartingale if and only if it is of bounded variation which is equivalent to a is of bounded variation.
If $(Z(t))$ is of unbounded variation and $(X(t))$ is an $(\mathcal{F}_t^Z)_{t\geq 0}$-semimartingale, it can be decomposed as :

$$X(t) = a(0)Z(t) + \int_0^t \left(\int_0^u a'(u-s)dZ(s)\right) du, \quad t \geq 0,$$

see Proposition 3.2 in Basse and Perdersen (2009). $\square$

6.3. **Proof of Lemma 1.** By (4),

$$\begin{aligned}\int_0^{+\infty}\varphi_j(s)X(s_-)dX(s) &= a(0)\int_0^{+\infty}\varphi_j(s)X(s_-)dZ(s)+\int_0^{+\infty}\varphi_j(s)X(s_-)\int_0^s a'(s-u)dZ(u)ds\\ &= a(0)\int_0^{+\infty}\varphi_j(s)X(s_-)dZ(s)+\int_0^{+\infty}\varphi_j(s)X(s)\int_0^s a'(s-u)dZ(u)ds.\end{aligned}$$

As

$$\begin{aligned}\mathbb{E}\left(\int_0^{+\infty}\varphi_j(s)X(s_-)dZ(s)\right)^2 &= \int_0^{+\infty}\varphi_j^2(s)\mathbb{E}X^2(s)ds\times\int_{\mathbb{R}}x^2n(x)dx\\ &= \int_0^{+\infty}\varphi_j^2(s)G(s)ds\le\|a\|^2<+\infty,\end{aligned}$$

$\mathbb{E}\int_0^{+\infty}\varphi_j(s)X(s_-)dZ(s)=0$ and the first equality follows by:

$$\begin{aligned}\mathbb{E}\int_0^{+\infty}\varphi_j(s)X(s_-)dX(s) &= \int_0^{+\infty}\varphi_j(s)\int_0^s a(s-u)a'(s-u)du\,ds\\ &= \frac{1}{2}\int_0^{+\infty}\varphi_j(s)(a^2(s)-a^2(0))ds.\end{aligned}$$

Using [H3](1) and (4), as $\Delta X(s)=a(0)\Delta Z(s)$,

$$\sum_{s\le T}(\Delta X(s))^2=a^2(0)\sum_{s\le T}(\Delta Z(s))^2<+\infty\quad\text{and}\quad\mathbb{E}\sum_{s\le T}(\Delta Z(s))^2=T.$$

The second equality is proved. □

6.4. **Proof of Proposition 2.** Note that for functions on $S_{m,T}$, the norms $\|.\|_T$ and $\|.\|$ are identical.

When $(\varphi_j)=(\varphi_{j,T})$ is the trigonometric basis on $[0,T]$, $\widehat{\theta}_j$ is an unbiased estimator of θ_j. This implies $\mathbb{E}\|\widehat{g}_m-g\|_T^2=\mathbb{E}\|\widehat{g}_m-\mathbb{E}\widehat{g}_m\|^2+\|g_m-g\|_T^2$. We have, setting $X=X_1$, and using that $\sum_{j=0}^{m-1}\left(\int_0^T\varphi_j(s)ds\right)^2\le T$,

$$\begin{aligned}\mathbb{E}\|\widehat{g}_m-\mathbb{E}\widehat{g}_m\|^2 &\le \frac{2}{N}\sum_{j=0}^{m-1}\mathrm{Var}\left(2\int_0^T\varphi_j(s)X(s_-)dX(s)\right)+\frac{2T}{N}\mathrm{Var}\left(\frac{1}{T}\sum_{s\le T}(\Delta X(s))^2\right)\\ &\le \frac{2}{N}\sum_{j=0}^{m-1}\mathbb{E}\left(2\int_0^T\varphi_j(s)X(s_-)dX(s)\right)^2+\frac{2T}{N}\mathrm{Var}\left(\frac{1}{T}\sum_{s\le T}(\Delta X(s))^2\right).\end{aligned}$$

We have:

$$\left(\int_0^T\varphi_j(s)X(s_-)dX(s)\right)^2\le 2g(0)\left(\int_0^T\varphi_j(s)X(s_-)dZ(s)\right)^2+2\left(\int_0^T\varphi_j(s)X(s)Y(s)ds\right)^2$$

where $Y(s)=\int_0^s a'(s-u)dZ(u)$. Next,

$$\mathbb{E}\left(\int_0^T\varphi_j(s)X(s_-)dZ(s)\right)^2=\int_0^T\varphi_j^2(s)\mathbb{E}(X^2(s))ds\le G(T).$$

Since $(\varphi_j) = (\varphi_{j,T})$ is an orthonormal basis of $\mathbb{L}^2([0,T])$,

$$\sum_{j=0}^{m-1} \mathbb{E}\left(\int_0^T \varphi_j(s)X(s)Y(s)ds\right)^2 = \mathbb{E}\left[\sum_{j=0}^{m-1}\left(\int_0^T \varphi_j(s)X(s)Y(s)ds\right)^2\right] \leq \mathbb{E}\int_0^T X^2(s)Y^2(s)ds.$$

We use that $x^2y^2 \leq (x^4+y^4)/2$ and (see section 7.3)

$$\begin{aligned} \mathbb{E}X^4(s) &= 3\left(\int_0^s a^2(u)du\right)^2 + \int_0^s a^4(u)du\int x^4 n(x)dx \\ &= 3G^2(s) + k_4\int_0^s a^4(u)du. \qquad (19) \end{aligned}$$

Analogously, setting $G_1(s) = \int_0^s (a')^2(u)du$, we obtain:

$$\mathbb{E}Y^4(s) = 3[G_1(s)]^2 + k_4\int_0^s (a'(u))^4 du.$$

It remains to study $\mathbb{E}\left(\frac{1}{T}\sum_{s\leq T}(\Delta X(s))^2\right)^2 = T^{-2}a^4(0)\mathbb{E}\left(\sum_{s\leq T}(\Delta Z(s))^2\right)^2$. By the exponential formula (see *e.g.* Revuz and Yor, 1999, Chap. XII, Prop. 1.12),

$$\mathbb{E}\exp[iu\sum_{s\leq T}(\Delta Z(s))^2] = \exp[T\int_{\mathbb{R}}(e^{iux^2}-1)n(x)dx]. \qquad (20)$$

We deduce: Var $\left(\frac{1}{T}\sum_{s\leq T}(\Delta X(s))^2\right) = k_4a^4(0)/T = k_4g^2(0)/T$. □

6.5. **Proof of Proposition 3.**

Consider a basis (φ_j) of $\mathbb{L}^2(\mathbb{R}^+)$ with arbitrary support. We have $\mathbb{E}\widehat{\theta}_j = \theta_j - \int_T^{+\infty} g(s)\varphi_j(s)ds$ so that $\widehat{g}_m - g = \widehat{g}_m - \mathbb{E}\widehat{g}_m + \mathbb{E}\widehat{g}_m - g_m + g_m - g$ and this implies

$$\mathbb{E}\|\widehat{g}_m - g\|^2 = \|g_m - g\|^2 + \mathbb{E}\|\widehat{g}_m - \mathbb{E}\widehat{g}_m\|^2 + \|\mathbb{E}\widehat{g}_m - g_m\|^2.$$

The first term is the usual bias term due to the projection method. The middle term is a variance term which can be treated as in the previous proposition. The last term is an additional bias term, due to the truncation of the integrals. We have:

$$\|\mathbb{E}\widehat{g}_m - g_m\|^2 = \sum_{j=0}^{m-1}(\mathbb{E}\widehat{\theta}_j - \theta_j)^2 = \sum_{j=0}^{m-1}\left(\int_T^{+\infty} g(s)\varphi_j(s)ds\right)^2 \leq \int_T^{+\infty} g^2(s)ds, \qquad (21)$$

Therefore, we get the first inequality of Proposition 3.
If (φ_j) is the Laguerre basis, we bound the variance term $\mathbb{E}\|\widehat{g}_m - \mathbb{E}\widehat{g}_m\|^2$ and the additional bias term $\|\mathbb{E}\widehat{g}_m - g_m\|^2$ differently. For the variance term, we write:

$$\begin{aligned} \left(\mathbb{E}\int_0^T \varphi_j(s)X(s)Y(s)ds\right)^2 &= \int_{[0,T]^2} \varphi_j(s)\varphi_j(u)\mathbb{E}[X(s)Y(s)X(u)Y(u)]dsdu \\ &\leq \int_{[0,T]^2} |\varphi_j(s)\varphi_j(u)|\left\{\mathbb{E}[(X(s)Y(s))^2]\mathbb{E}[(X(u)Y(u))^2]\right\}^{1/2} dsdu \\ &= \left(\int_0^T |\varphi_j(s)|\left\{\mathbb{E}[(X(s)Y(s))^2)]\right\}^{1/2} ds\right)^2. \qquad (22) \end{aligned}$$

We use the following bound proved in section 6.4:

$$2\mathbb{E}X^2(s)Y^2(s)ds \leq \mathbb{E}X^4(s) + \mathbb{E}Y^4(s) \leq 3(G^2(T) + G_1^2(T)) + k_4(\|g\|_T^2 + \|g_1\|_T^2)$$

There remains to bound $\int_0^T |\varphi_j(s)|ds$. This is done in [12], see Formulae (31)-(32). For $j = 0, \ldots, m-1$ and $T \geq 6(m-1)+3 = 6m-3$, we have

$$\int_0^T |\varphi_j(s)|ds \lesssim j^{1/2} \quad \text{and} \quad \sum_{j=0}^{m-1} \left(\int_0^T |\varphi_j(s)|ds\right)^2 \lesssim m^2 \tag{23}$$

Also by (33) in [12], we have, for the additional bias term,

$$\sum_{j=0}^{m-1} \left[\int_T^{+\infty} \varphi_j(s)g(s)ds\right]^2 \lesssim \|a\|^2 m \, \exp\left(-12\gamma_2 m\right), \tag{24}$$

where γ_2 is a constant depending on the Laguerre basis only, see Section 7. Therefore, the proof of Proposition 3 is complete. □

6.6. **Proof of Lemma 2.** We have

$$\begin{aligned}\mathbb{E}\left(\int_0^T \varphi_j'(s)X^2(s)ds\right) &= \int_0^T \varphi_j'(s)\left(\int_0^s g(s-u)du\right)ds = \int_0^T \varphi_j'(s)G(s)ds \\ &= [\varphi_j(s)G(s)]_0^T - \langle g, \varphi_j\rangle_T = \varphi_j(T)G(T) - \langle g, \varphi_j\rangle_T\end{aligned}$$

which is the result. □

6.7. **Proof of Proposition 4.** Assume that $(\varphi_j = \varphi_{j,T})$ is the trigonometric basis. Then, $\widetilde{\theta}_j$ is an unbiased estimator of θ_j. We only need to study the variance term of the risk.

$$\mathbb{E}\|\widetilde{g}_m - \mathbb{E}\widetilde{g}_m\|_T^2 \leq \frac{2}{N}\left(\sum_{j=0}^{m-1}\mathbb{E}\left(\int_0^T \varphi_{j,T}'(s)X^2(s)ds\right)^2 + \sum_{j=0}^{m-1}\varphi_{j,T}^2(T)\mathbb{E}X^4(T)\right)$$

where $\mathbb{E}X^4(T) = 3(G^2(T) + k_4\|g\|_T^2)$ and $\sum_{j=0}^{m-1}\varphi_j^2(T) = m/T$. We have

$$\varphi_{0,T}'(s) = 0, \;\; \varphi_{2j,T}'(s) = (2\pi j/T)\varphi_{2j-1,T}(s), \;\; \varphi_{2j-1,T}'(s) = -(2\pi j/T)\varphi_{2j,T}(s), \;\; j \geq 1. \tag{25}$$

Using that $(\varphi_{j,T})$ is an orthonormal basis, we obtain, as $\mathbb{E}X^4(s) \leq \mathbb{E}X^4(T)$ (see (19)),

$$\sum_{j=0}^{m-1}\mathbb{E}\left(\int_0^T \varphi_{j,T}'(s)X^2(s)ds\right)^2 \leq \frac{4\pi^2 m^2}{T^2}\mathbb{E}\int_0^T X^4(s)ds \leq (3G^2(T) + k_4\|g\|_T^2)\frac{4\pi^2 m^2}{T}.$$

This gives (14).
Now, assume that $(\varphi_j = \ell_j)$ is the Laguerre basis on $\mathbb{L}^2(\mathbb{R}^+)$ (see Section 7). We still have:

$$\mathbb{E}(\|\widetilde{g}_m - g\|^2) = \mathbb{E}\|\widetilde{g}_m - \mathbb{E}\widetilde{g}_m\|^2 + \|\mathbb{E}\widetilde{g}_m - g_m\|^2 + \|g_m - g\|^2.$$

First,

$$\begin{aligned}\mathbb{E}\|\widetilde{g}_m - \mathbb{E}\tilde{g}_m\|^2 &= \frac{1}{N}\sum_{j=0}^{m-1}\operatorname{Var}\left(\int_0^T \ell_j'(s)X_1^2(s)ds - X_1^2(T)\ell_j(T)\right) \\ &\leq \frac{2}{N}\sum_{j=0}^{m-1}\mathbb{E}\left[\left(\int_0^T \ell_j'(s)X_1^2(s)ds\right)^2\right] + \frac{2}{N}\sum_{j=0}^{m-1}\ell_j^2(T)\mathbb{E}[X_1^4(T)] := \mathbb{T}_1 + \mathbb{T}_2.\end{aligned}$$

Using that $|\ell_j| \leq \sqrt{2}$, we get

$$\mathbb{T}_2 \leq 4(3G^2(T) + k_4\|g\|_T^2)\frac{m}{N}.$$

Next, we use that the Laguerre basis satisfies $\ell_0'(x) = -\ell_0(x)$ and $\ell_j'(x) = -\ell_j(x) - \sqrt{2j/x}\ell_{j-1}^{(1)}(x)$ for $j \geq 1$ where $(\ell_k^{(1)}(x), k \geq 0)$ is the Laguerre basis with index 1 (see section 7) to find

$$\begin{aligned}
\mathbb{T}_1 &\leq \frac{4}{N}\sum_{j=0}^{m-1}\mathbb{E}\left[\left(\int_0^T \ell_j(s)X_1^2(s)ds\right)^2\right] + \frac{4}{N}\sum_{j=1}^{m-1}\mathbb{E}\left[\left(\int_0^T \ell_{j-1}^{(1)}(s)\sqrt{\frac{2j}{s}}X_1^2(s)ds\right)^2\right] \\
&\leq \frac{4}{N}\mathbb{E}\left(\int_0^T X_1^4(s)ds\right) + \frac{8m}{N}\mathbb{E}\left(\int_0^T \frac{X_1^4(s)}{s}ds\right) \\
&\leq \frac{4}{N}T(3G^2(T) + k_4\|g\|_T^2) + \frac{8m}{N}\left(3\int_0^T s^{-1}[G^2(s) + k_4\|g\|_s^2]ds\right)
\end{aligned}$$

where we have used (19). Finally, the variance term is bounded by

$$\begin{aligned}
\mathbb{E}\|\widetilde{g}_m - \mathbb{E}\widetilde{g}_m\|^2 &\leq \frac{4}{N}T(3G^2(T) + k_4\|g\|_T^2) + \frac{8m}{N}\left(3\int_0^T s^{-1}[G^2(s) + k_4\|g\|_s^2]ds\right) \\
&\quad + \frac{4m}{N}(3G^2(T) + k_4\|g\|_T^2)).
\end{aligned}$$

Using [H5] and writing $\int_0^T \cdots = \int_0^1 \cdots + \int_1^T \ldots$, we get

$$3\int_0^T s^{-1}[G^2(s) + k_4\|g\|_s^2]ds \leq (3 + k_4)(c_0 + \log(T)\|g\|_T^2).$$

If [H5] does not hold and $T \geq 6m - 3$, we can bound differently the variance and bias terms. Proceeding as in [12], proof of Proposition 3, we have

$$\sum_{j=0}^{m-1}\mathbb{E}\left(\int_0^T \ell_j'(s)X^2(s)ds\right)^2 \leq (3G^2(T) + k_4\|g\|_T^2)\left(\int_0^T \left[\sum_{j=0}^{m-1}(\ell_j'(s))^2\right]^{1/2} ds\right)^2$$

Still using [12], we have

$$\left(\int_0^T \left[\sum_{j=0}^{m-1}(\ell_j'(s))^2\right]^{1/2} ds\right)^2 \leq 12m^3 + \frac{4m^3}{\gamma_2^2}\exp\left(-(12m-6)\gamma_2\right).$$

Finally, we get

$$\mathbb{E}\|\tilde{g}_m - \mathbb{E}\tilde{g}_m\|^2 \leq \frac{1}{N}(3G^2(T) + k_4\|g\|_T^2)\left(12m^3 + \frac{4m^3}{\gamma_2^2}\exp\left(-(12m-6)\gamma_2\right)\right) \tag{26}$$

So, we have the two variance bounds.
Next, we have $\mathbb{E}\widetilde{\theta}_j = \theta_j - \ell_j(T)G(T) - \int_T^{+\infty}\ell_j(s)g(s)ds$. Therefore

$$\begin{aligned}
\|\mathbb{E}\widetilde{g}_m - g_m\|^2 &= \sum_{j=0}^{m-1}[\mathbb{E}(\widetilde{\theta}_j) - \theta_j]^2 = \sum_{j=0}^{m-1}\left(\ell_j(T)G(T) + \int_T^{+\infty}\ell_j(s)g(s)ds\right)^2 \\
&\leq 2G^2(T)\sum_{j=0}^{m-1}\ell_j^2(T) + 2\sum_{j=0}^{m-1}\left(\int_T^{+\infty}\ell_j(s)g(s)ds\right)^2 \\
&\lesssim \|a\|^2 m\exp(-12\gamma_2 m) + \|a\|^2 m\exp(12\gamma_2 m),
\end{aligned}$$

Indeed $\ell_j(T) \lesssim \exp(-12\gamma_2 m)$ for $T \geq 6m-3$ (first term) and we use (24) (second term). For both, we use $G(T) \leq G(+\infty) = \|a\|^2$. □

6.8. **Proof of Theorem 1.** Note that, as $G(0)=0$, $\langle h, g\rangle_T = h(T)G(T) - \langle h', G\rangle_T$. Let us set:

$$\gamma_{N,T}(h) = \|h\|^2 + \frac{2}{N}\sum_{i=1}^{N}[\int_0^T h'(u)X_i^2(u)du - h(T)X_i^2(T)].$$

We have $\widetilde{g}_m = \arg\min_{h\in S_m}\gamma_{N,T}(h)$, $\gamma_{N,T}(\widetilde{g}_m) = -\|\widetilde{g}_m\|^2$ and

$$\gamma_{N,T}(h) = \|h\|^2 - 2\langle h, g\rangle_T - 2\nu_{N,T}(h) - 2\mu_{N,T}(h)$$

where

$$(27)\quad \nu_{N,T}(h) = -\frac{1}{N}\sum_{i=1}^{N}\int_0^T h'(u)[X_i^2(u) - G(u)]du, \quad \mu_{N,T}(h) = \frac{1}{N}\sum_{i=1}^{N} h(T)(X_i^2(T) - G(T)).$$

Therefore,

$$\gamma_{N,T}(h_1) - \gamma_{N,T}(h_2) = \|h_1 - g\|^2 - \|h_2 - g\|^2 - 2\nu_{N,T}(h_1 - h_2) - 2\mu_{N,T}(h_1 - h_2)$$

Using the definition of $\tilde{m}$, we have for all $g_m \in S_m$,

$$\gamma_{N,T}(\tilde{g}_{\widetilde{m}}) + \text{pen}(\widetilde{m}) \leq \gamma_{N,T}(\widetilde{g}_m) + \text{pen}(m).$$

We deduce, for $\xi_{N,T} = \nu_{N,T} + \mu_{N,T}$,

$$\|\widetilde{g}_{\widetilde{m}} - g\|^2 \quad\leq\quad \|g_m - g\|^2 + 2\xi_{N,T}(\widetilde{g}_{\widetilde{m}} - g_m) + \text{pen}(m) - \text{pen}(\widetilde{m})$$

Let $B_m = \{h \in S_m, \|h\| \leq 1\}$. We use that

$$\begin{aligned} 2\xi_{N,T}(\tilde{g}_{\widetilde{m}} - g_m) &\leq \frac{1}{4}\|\tilde{g}_{\widetilde{m}} - g_m\|^2 + 4\sup_{h\in B_{\widetilde{m}\vee m}}\xi_{N,T}^2(h) \\ &\leq \frac{1}{2}(\|\tilde{g}_{\widetilde{m}} - g\|^2 + \|g - g_m\|^2) + 4\sup_{h\in B_{\widetilde{m}\vee m}}\xi_{N,T}^2(h) \\ &\leq \frac{1}{2}(\|\tilde{g}_{\widetilde{m}} - g\|^2 + \|g - g_m\|^2) + 8\sup_{h\in B_{\widetilde{m}\vee m}}(\nu_{N,T}^2(h) + \mu_{N,T}^2(h)) \end{aligned}$$

Recall that $\mathbb{E}[X_i^2(u)] = G(u)$. For θ a constant to be chosen below, we can split $\nu_{N,T}(h)$ into:

$$(28)\qquad \nu_{N,T,\theta}(h) = -\frac{1}{N}\sum_{i=1}^{N}\int_0^T h'(u)[X_i^2(u)\mathbf{1}_{X_i^2(u)\leq\theta} - \mathbb{E}(X_i^2(u)\mathbf{1}_{X_i^2(u)\leq\theta})]du$$

$$(29)\qquad \nu_{N,T,\theta}^c(h) = -\frac{1}{N}\sum_{i=1}^{N}\int_0^T h'(u)[X_i^2(u)\mathbf{1}_{X_i^2(u)>\theta} - \mathbb{E}(X_i^2(u)\mathbf{1}_{X_i^2(u)>\theta})]du$$

Analogously, we define $\mu_{N,T,\theta}(h)$ and $\mu_{N,T,\theta}^c(h)$ by splitting $\mu_{N,T}(h)$.

Introducing quantities $p_1(m,m')$ and $p_2(m,m')$ to be determined below, we write:

$$\begin{aligned}\frac{1}{2}\|\tilde{g}_{\widetilde{m}}-g\|^2 \quad\leq\quad & \frac{3}{2}\|g_m-g\|^2+16\sup_{h\in B_{\widetilde{m}\vee m}}[\nu^c_{N,T,\theta}(h)]^2+16\sup_{h\in B_{\widetilde{m}\vee m}}[\mu^c_{N,T,\theta}(h)]^2\\ & +\text{pen}(m)-\text{pen}(\widetilde{m})\\ & +16p_1(m,\widetilde{m})+16(\sup_{h\in B_{\widetilde{m}\vee m}}\nu^2_{N,T,\theta}(h)-p_1(m,\widetilde{m}))\\ & +16p_2(m,\widetilde{m})+16(\sup_{h\in B_{\widetilde{m}\vee m}}\mu^2_{N,T,\theta}(h)-p_2(m,\widetilde{m}))\end{aligned}$$

Below, $p_1(m,m')$ and $p_2(m,m')$ are chosen such that, for κ greater than a well chosen constant κ_0, for all m,m', $16(p_1(m,m')+p_2(m,m'))\leq \text{pen}(m)+\text{pen}(m')$ implying that

$$16(p_1(m,\widetilde{m})+p_2(m,\widetilde{m}))+\text{pen}(m)-\text{pen}(\widetilde{m})\leq 2\text{pen}(m).$$

And we bound the expectation of the other terms.

Lemma 3. *Under Assumptions* [H1], [H2]*(2+p) and* [H3]*(2+p), we have:*

$$\mathbb{E}\left(\sup_{h\in S_m,\|h\|\leq 1}[\nu^c_{N,T,\theta}(h)]^2\right)\leq \frac{4\pi^2m^2}{NT}\frac{1}{\theta^p}\left(G^{2+p}(T)+k_{4+2p}\int_0^T a^{4+2p}(u)du\right).$$

$$\mathbb{E}\left(\sup_{h\in S_m,\|h\|\leq 1}[\mu^c_{N,T,\theta}(h)]^2\right)\leq \frac{m}{NT}\frac{1}{\theta^p}\left(G^{2+p}(T)+k_{4+2p}\int_0^T a^{4+2p}(u)du\right).$$

Now, we choose $p=2$ and $\theta=\mathfrak{c}\sqrt{N}/\sqrt{\log(N)}$ where $\mathfrak{c}$ is precised below. As for all $m\in\mathcal{M}_N$, $m^2\leq NT$, we set $M_N=[\sqrt{NT}]$ the largest dimension of the collection, and we have

$$\mathbb{E}\left(\sup_{h\in S_{M_N},\|h\|\leq 1}[\nu^c_{N,T,\theta}(h)]^2\right)+\mathbb{E}\left(\sup_{h\in S_{M_N},\|h\|\leq 1}[\mu^c_{N,T,\theta}(h)]^2\right)\lesssim\frac{\log(N)}{N}. \tag{30}$$

Lemma 4. *Under* [H1], [H2]*(4) and* [H3]*(4), choosing* $\theta=\mathfrak{c}\sqrt{N}/\sqrt{\log(N)}$*,* $\mathfrak{c}=c\mathbb{E}^{1/2}[X^4(T)]$*,* $c=\sqrt{2}/21$*, we have for*

$$\begin{aligned}p_1(m,m') &= 2(1+18\log(N))\mathbb{E}(X^4(T))\frac{4\pi^2(m\vee m')^2}{NT},\\ p_2(m,m') &= 2(1+18\log(N))\mathbb{E}(X^4(T))\frac{m\vee m'}{NT},\end{aligned}$$

$$\mathbb{E}\left[\left(\sup_{h\in B_{\widetilde{m}\vee m}}\nu^2_{N,T,\theta}(h)-p_1(m,\widetilde{m})\right)_+\right]\lesssim\frac{1}{N},\quad \mathbb{E}\left[\left(\sup_{h\in B_{\widetilde{m}\vee m}}\mu^2_{N,T,\theta}(h)-p_2(m,\widetilde{m})\right)_+\right]\lesssim\frac{1}{N}.$$

Therefore, using (30) and Lemma 4, we can conclude that for $\kappa\geq\kappa_0=16\times 2\times 19\times 8\pi^2$, $\text{pen}(m)-\text{pen}(\widetilde{m})+16(p_1(m,\widetilde{m})+p_2(m,\widetilde{m}))\leq 2\text{pen}(m)$, we obtain

$$\mathbb{E}\|\tilde{g}_{\widetilde{m}}-g\|^2\quad\leq\quad 3\|g_m-g\|^2+4\text{pen}(m)+C(\frac{1}{N}+\frac{\log(N)}{N}).$$

The proof of Theorem 1 is now complete.□

Proof of Lemma 3.

$$\begin{aligned}
\mathbb{E}\left(\sup_{h\in S_m,\|h\|\le 1}[\nu^c_{N,T,\theta}(h)]^2\right) &\le \sum_{j=0}^{m-1}\mathbb{E}[\nu^c_{N,T,\theta}(\varphi_{j,T})]^2 = \frac{1}{N}\sum_{j=0}^{m-1} \mathrm{Var}\left(\int_0^T \varphi'_{j,T}(u)X_1^2(u)\mathbf{1}_{X_1^2(u)>\theta}\right)\\
&\le \frac{1}{N}\sum_{j=0}^{m-1}\mathbb{E}\left(\int_0^T \varphi'_{j,T}(u)X_1^2(u)\mathbf{1}_{X_1^2(u)>\theta}du\right)^2\\
&\le \frac{4\pi^2m^2}{NT^2}\mathbb{E}\left(\int_0^T X^4(u)\mathbf{1}_{X_1^2(u)>\theta}du\right),
\end{aligned}$$

see (25). Then, for all $p \ge 1$, we get

$$\begin{aligned}
\mathbb{E}\left(\sup_{h\in S_m,\|h\|\le 1}[\nu^c_{N,T,\theta}(h)]^2\right) &\le \frac{4\pi^2m^2}{NT^2}\frac{1}{\theta^p}\mathbb{E}\left(\int_0^T X^{4+2p}(u)du\right)\\
&\le \frac{4\pi^2m^2}{NT}\frac{1}{\theta^p}\left(G^{2+p}(T)+k_{4+2p}\int_0^T a^{4+2p}(u)du\right),
\end{aligned}$$

where the last bound is obtained by Kunita's Inequality, see Section 7. In the same way, we get

$$\begin{aligned}
\mathbb{E}\left(\sup_{h\in S_m,\|h\|\le 1}[\mu^c_{N,T,\theta}(h)]^2\right) &\le \sum_{j=0}^{m-1}\mathbb{E}[\mu^c_{N,T,\theta}(\varphi_{j,T})]^2\\
&\le \frac{1}{N}\sum_{j=0}^{m-1}\mathbb{E}\left[\left(\varphi_{j,T}(T)X^2(T)\mathbf{1}_{X_1^2(T)>\theta}\right)^2\right] = \frac{m}{NT}\mathbb{E}\left(X^4(T)\mathbf{1}_{X_1^2(T)>\theta}\right)\\
&\le \frac{m}{NT}\frac{1}{\theta^p}\left(G^{2+p}(T)+k_{4+2p}\int_0^T a^{4+2p}(u)du\right).
\end{aligned}$$

This is the second bound of Lemma 3 and the proof of Lemma 3 is complete.□

Proof of Lemma 4. First note that

$$\mathbb{E}\left[\left(\sup_{h\in B_{\widetilde{m}\vee m}}\nu^2_{N,T,\theta}(h)-p_1(m,\widetilde{m})\right)_+\right] \le \sum_{m'\in\mathcal{M}_N}\mathbb{E}\left[\left(\sup_{h\in B_{m\vee m'}}\nu^2_{N,T,\theta}(h)-p_1(m,m')\right)_+\right]$$

To bound each term of the sum above, we apply the Talagrand Inequality recalled in Theorem 2 (Appendix).

Note that $\nu_{N,T,\theta}(h) = \frac{1}{N}\sum_{i=1}^N[f_h(X_i)-\mathbb{E}(f_h(X_i))]$ where, for $x \in D([0,T])$, the space of real valued right-continuous with left-hand limits functions (càdlag) defined on $[0,T]$,

$$f_h(x) = -\int_0^T h'(u)[x^2(u)\mathbf{1}_{x^2(u)\le\theta}]du.$$

Recall that $B_m = \{h \in S_m, \|h\| \le 1\}$. We need bound $\sup_{h\in B_m,x\in D([0,T])}|f_h(x)|$. We have:

$$\sup_{h\in B_m,x\in D([0,T])}|f_h(x)| \le \theta\sup_{h\in B_m}\int_0^T|h'(u)|du \le \theta\sup_{h\in B_m}\sqrt{T}\left(\int_0^T(h'(u))^2du\right)^{1/2}.$$

For $h \in B_m$,

$$\int_0^T(h'(u))^2du \le \int_0^T\left[\sum_{j=0}^{m-1}a_j\frac{2\pi j}{T}\varphi_{j\pm1,T}(u)\right]^2du \le \frac{4\pi^2m^2}{T^2}.$$

So we set

$$M(m) = M = \theta \frac{2\pi m}{\sqrt{T}}.$$

Then, we bound $\sup_{h\in B_m}[\nu_{N,T,\theta}(h)]^2$. We have

$$\begin{aligned}\sup_{h\in B_m}[\nu_{N,T,\theta}(h)]^2 &\le \sum_{j=0}^{m-1}[\nu_{N,T,\theta}(\varphi_{j,T})]^2 = \frac{1}{N}\sum_{j=0}^{m-1} \mathrm{Var}\left(\int_0^T \varphi'_{j,T}(u)X_1^2(u)\mathbf{1}_{X_1^2(u)<\theta}du\right)\\ &\le \frac{1}{N}\sum_{j=0}^{m-1}\mathbb{E}\left(\int_0^T \varphi'_{j,T}(u)X_1^2(u)\mathbf{1}_{X_1^2(u)<\theta}du\right)^2 \le \frac{4\pi^2m^2}{NT}\mathbb{E}(X^4(T))\end{aligned}$$

where $\mathbb{E}(X^4(T)) = 3G^2(T) + k_4\int_0^T a^4(u)du$. We set

$$H^2(m) = H^2 = \frac{4\pi^2m^2}{NT}\mathbb{E}(X^4(T)).$$

Now, we bound

$$\mathbb{V} := \sup_{h\in B_m} \mathrm{Var}\left(\int_0^T h'(u)X_1^2(u)\mathbf{1}_{X_1^2(u)<\theta}du\right) \le \sup_{h\in B_m}\mathbb{E}\left(\int_0^T h'(u)X_1^2(u)\mathbf{1}_{X_1^2(u)<\theta}du\right)^2.$$

Proceeding as previously, we get that the above term is less than

$$\mathbb{V} \le \sup_{h\in B_m}\int_0^T [h'(u)]^2du T\mathbb{E}[X^4(T)] \le \frac{4\pi^2m^2}{T^2}\mathbb{E}[X^4(T)] := v^2(m) = v^2.$$

Therefore, $NH^2 = v^2, NH/M = \sqrt{N}\mathbb{E}^{1/2}(X^4(T))\theta^{-1}$.

First, using the Talagrand inequality with our values of $M(m\vee m'), H(m\vee m'), v^2(m\vee m')$, we bound

$$\mathbb{E}\left[\left(\sup_{h\in B_{m'\vee m}} \nu^2_{N,T,\theta}(h) - p_1(m,m')\right)_+\right].$$

We take

$$\theta = c\mathbb{E}^{1/2}(X^4(T))\frac{\sqrt{N}}{\sqrt{\log(N)}}, \quad \epsilon^2 = \frac{3}{2}\frac{\log(N)}{K_1} = 9\log(N), \quad p_1(m,m') = 2(1+2\epsilon^2)H^2(m\vee m').$$

Thus,

$$\frac{v^2(m\vee m')}{N}\exp\left(-K_1\epsilon\frac{NH^2(m\vee m')}{v^2(m\vee m')}\right) \lesssim \frac{(m\vee m')^2}{N}e^{-(3/2)\log(N)} \lesssim \frac{1}{N^{3/2}}$$

and,

$$\frac{M^2(m\vee m')}{C^2(\epsilon^2)N^2}e^{-\frac{2K_1C(\epsilon^2)\epsilon}{7\sqrt{2}}\frac{NH(m\vee m')}{M(m\vee m')}} \lesssim \frac{1}{\log(N)}\exp(-(3/2)\log(N)) \lesssim \frac{1}{N^{3/2}}$$

for $C(\epsilon^2) = 1$ and $c = \sqrt{2}/21$. Theorem 2 (Appendix) implies

$$\sum_{m'\in\mathcal{M}_N}\mathbb{E}\left[\left(\sup_{h\in S_{m'},\|h\|=1} \nu^2_{N,T,\theta}(h) - p_1(m,m')\right)_+\right] \lesssim \frac{1}{N}.$$

We proceed analogously to bound $\mathbb{E}\left[\left(\sup_{h\in B_{m\vee m'}}\mu^2_{N,T,\theta}(h)-p_2(m,m')\right)_+\right]$. We set $\mu_{N,T,\theta}(h)=\frac{1}{N}\sum_{i=1}^N[g_h(X_i)-\mathbb{E}(g_h(X_i))]$ where, for $x\in D([0,T])$, $g_h(x)=h(T)x^2(T)\mathbf{1}_{x^2(T)\leq\theta}]$. We have

$$\sup_{h\in B_m,x\in D([0,T])}|g_h(x)|\leq\theta\sup_{h\in B_m}|h(T)|\leq\theta(\sum_{j=0}^{m-1}\varphi^2_{j,T}(T))^{1/2}=\theta\sqrt{\frac{m}{T}}:=M(m).$$

Next,

$$\mathbb{E}\left[\sup_{h\in B_m}[\mu_{N,T,\theta}(h)]^2\right]\leq\frac{1}{N}\sum_{j=0}^{m-1}(\varphi_{j,T}(T))^2\mathbb{E}(X^4(T))=\frac{m}{NT}\mathbb{E}(X^4(T)):=H^2(m).$$

Last

$$\sup_{h\in B_m}\text{Var}\left(h(T)X_1^2(T)\mathbf{1}_{X_1^2(T)<\theta}\right)\leq(m/T)\mathbb{E}(X^4(T)):=v^2(m).$$

When choosing $p_2(m,m')=2(1+18\log N)(m\vee m')/NT)\mathbb{E}(X^4(T))$, and proceeding as above with Theorem 2 (see Appendix), we obtain the second part of Lemma 4. □

6.9. **Proof of Proposition 5.** To prove Proposition 5, we need the following Lemma.

Lemma 5. *Under[H0]-[H1],*

$$\mathbb{E}\left(X(t+h)-X(t)\right)^2\leq h^2\|a'\|^2+\|a\|^2_\infty h\leq Ch.$$

Proof of Lemma 5. We write:

$$X(t+h)-X(t)=\int_0^t(a(t+h-u)-a(t-u))dZ(u)+\int_t^{t+h}a(t+h-u)dZ(u).$$

Thus, as $\int x^2n(x)dx=1$,

$$\mathbb{E}\left(X(t+h)-X(t)\right)^2=\int_0^t(a(t+h-u)-a(t-u))^2du+\int_t^{t+h}a^2(t+h-u)du.$$

We have $\int_t^{t+h}a^2(t+h-u)du\leq h\|a\|^2_\infty$ and

$$\begin{aligned}\int_0^t(a(t+h-u)-a(t-u))^2du &= \int_0^t(a(h+v)-a(v))^2dv=h^2\int_0^t[\int_0^1a'(v+\tau h)d\tau]^2dv\\ &\leq h^2\int_0^1d\tau\int_0^t[a'(v+\tau h)]^2dv\leq h^2\|a'\|^2.\end{aligned}$$

□

Proof of Proposition 5. We have $\widetilde{g}^\Delta_m-g=\widetilde{g}^\Delta_m-\widetilde{g}_m+\widetilde{g}_m-g$. Thus,

$$\mathbb{E}\|\widetilde{g}^\Delta_m-g\|^2=\mathbb{E}\|\widetilde{g}_m-g\|^2+\mathbb{E}\|\widetilde{g}^\Delta_m-\widetilde{g}_m\|^2\quad\text{where}$$

$$\mathbb{E}\|\widetilde{g}^\Delta_m-\widetilde{g}_m\|^2=\sum_{j=0}^{m-1}\left[\mathbb{E}(\widetilde{\theta}^\Delta_j-\widetilde{\theta}_j)\right]^2+\sum_{j=0}^{m-1}\text{Var}\ (\widetilde{\theta}^\Delta_j-\widetilde{\theta}_j)=\sum_{j=0}^{m-1}b_j^2+\frac{1}{N}S_j,$$

$b_j=\sum_{k=1}^n\int_{(k-1)\Delta}^{k\Delta}(\varphi'_j(s)G(s)-\varphi'_j(k\Delta)G(k\Delta))ds$ and

$$S_j=\mathbb{E}\left(\sum_{k=1}^n\int_{(k-1)\Delta}^{k\Delta}(\varphi'_j(k\Delta)X^2(k\Delta)-\varphi'_j(s)X^2(s))ds\right)^2.$$

We split $b_j = b_j^{(1)} + b_j^{(2)}$ by splitting

$$\varphi_j'(s)G(s) - \varphi_j'(k\Delta)G(k\Delta) = (\varphi_j'(s) - \varphi_j'(k\Delta))G(s) + \varphi_j'(k\Delta)(G(s) - G(k\Delta)).$$

We have:

$$b_j^{(1)} = -\sum_{k=1}^{n} \int_{(k-1)\Delta}^{k\Delta} G(s) \int_{s}^{k\Delta} \varphi_j''(u)duds = -\sum_{k=1}^{n} \int_{(k-1)\Delta}^{k\Delta} \varphi_j''(u) \int_{(k-1)\Delta}^{u} G(s)dsdu.$$

This yields $|b_j^{(1)}| \leq G(T)\Delta \int_0^T |\varphi_j''(u)|du$. For the Laguerre basis, we apply relation (34) and get, after reordering terms, $\varphi_j'' = \varphi_j + 4\sum_{k=0}^{j-1} \varphi_k + 4\sum_{\ell=0}^{j-2}(j-1-\ell-1)\varphi_\ell$. Using that, for $T \geq 6m-3$, $\varphi_j \leq \exp(-\gamma_2 s)$ yields

$$\int_{6m-3}^{T} |\varphi_j''(u)|du \leq \gamma_2^{-1} 4j^2 \exp(-\gamma_2(6m-3)).$$

And,

$$(\int_0^{6m-3} |\varphi_j''(u)|du)^2 \leq (6m-3)\int_0^{+\infty} [\varphi_j''(u)]^2 du \leq Cj^3(6m-3).$$

Thus, $\sum_{j=0}^{m-1} (b_j^{(1)})^2 \leq CG^2(T)\Delta^2 m^5$. Now, we study the second term:

$$b_j^{(2)} = -\sum_{k=1}^{n} \int_{(k-1)\Delta}^{k\Delta} \varphi_j'(k\Delta) \int_{s}^{k\Delta} g(u)duds = -\sum_{k=1}^{n} \varphi_j'(k\Delta) \int_{(k-1)\Delta}^{k\Delta} g(u)(u - ((k-1)\Delta)du.$$

Thus, $|b_j^{(2)}| \leq C\Delta j \int_0^T g(u)du$. Therefore, $\sum_{j=0}^{m-1} (b_j^{(2)})^2 \leq CG^2(T)\Delta^2 m^3$.

The same bounds hold for the trigonometic basis.

Now, we split S_j into $S_j^{(1)} + S_j^{(2)}$ by splitting

$$\varphi_j'(k\Delta)X^2(k\Delta) - \varphi_j'(s)X^2(s) = (\varphi_j'(k\Delta) - \varphi_j'(s))X^2(k\Delta) + \varphi_j'(s)(X^2(k\Delta) - X^2(s)).$$

We have:

$$\begin{aligned} S_j^{(1)} &= \mathbb{E}\sum_{k,\ell} X^2(k\Delta)X^2(\ell\Delta) \int_{(k-1)\Delta}^{k\Delta} \int_{(\ell-1)\Delta}^{\ell\Delta} \varphi_j''(u)\varphi_j''(v)(u - ((k-1)\Delta)(v - ((k-1)\Delta)dudv \\ &\leq \Delta^2 \left(\sum_{k=1}^{n} [\mathbb{E}(X^4(k\Delta)]^{1/2} \int_{(k-1)\Delta}^{k\Delta} |\varphi_j''(u)|du \right)^2 \leq \Delta^2 \mathbb{E}X^4(T)) \left(\int_0^T |\varphi_j''(u)|du \right)^2. \end{aligned}$$

If $T \geq 6m-3$,

$$\sum_j S_j^{(1)} \leq C\Delta^2 \mathbb{E}X^4(T))m^5.$$

For $S_j^{(2)}$, we can write:

$$\sum_j S_j^{(2)} = \sum_j \mathbb{E}\left(\sum_{k=1}^n \int_{(k-1)\Delta}^{k\Delta} \varphi_j'(s)ds(X^2(k\Delta)-X^2(s))ds\right)^2$$

$$= \sum_{k,\ell=1}^n \int_{(k-1)\Delta}^{k\Delta}\int_{(\ell-1)\Delta}^{\ell\Delta} \sum_j \mathbb{E}\left(\varphi_j'(s)(X^2(k\Delta)-X^2(s))\varphi_j'(u)(X^2(\ell\Delta)-X^2(u))\right) dsdu$$

$$\leq \sum_{k,\ell=1}^n \int_{(k-1)\Delta}^{k\Delta}\int_{(\ell-1)\Delta}^{\ell\Delta} \left[\mathbb{E}\sum_j(\varphi_j'(s))^2(X^2(k\Delta)-X^2(s))^2\mathbb{E}\sum_j(\varphi_j'(u))^2(X^2(\ell\Delta)-X^2(u))^2\right]^{1/2} dsdu$$

$$= \left[\sum_k \int_{(k-1)\Delta}^{k\Delta}\left(\mathbb{E}\sum_j(\varphi_j'(s))^2(X^2(k\Delta)-X^2(s))^2\right)^{1/2} ds\right]^2$$

$$\lesssim 2\Delta\mathbb{E}X^4(T)\left[\int_0^T [\sum_j(\varphi_j'(s))^2]^{1/2}ds\right]^2$$

$$\lesssim 24m^3\Delta\mathbb{E}X^4(T) + o(m^3)$$

a bound obtained previously when (φ_j) is the Laguerre basis (Proposition 3 of Comte and Genon-Catalot (2021)). When (φ_j) is the trigonometric basis,

$$\left[\int_0^T [\sum_j(\varphi_j'(s))^2]^{1/2}ds\right]^2 \leq 4\pi^2m^2. \quad \square$$

7. Appendix

7.1. **Formulae for Laguerre functions.** For this paragraph, we refer to Abramowitz and Stegun (1964) and Comte and Genon-Catalot (2018).

The Laguerre polynomial with index δ, $\delta > -1$, and degree k is given by

$$L_k^{(\delta)}(x) = \frac{1}{k!}e^x x^{-\delta}\frac{d^k}{dx^k}\left(x^{\delta+k}e^{-x}\right) = \sum_{j=0}^k \binom{k+\delta}{k-j}\frac{(-x)^j}{j!}.$$

The following holds:

$$\left(L_k^{(\delta)}(x)\right)' = -L_{k-1}^{(\delta+1)}(x), \quad \text{for } k\geq 1, \text{ and } \int_0^{+\infty}\left(L_k^{(\delta)}(x)\right)^2 x^\delta e^{-x}dx = \frac{\Gamma(k+\alpha+1)}{k!}. \tag{31}$$

We consider the Laguerre functions with index δ, given by

$$\ell_k^{(\delta)}(x) = 2^{(\delta+1)/2}\left(\frac{k!}{\Gamma(k+\delta+1)}\right)^{1/2} L_k^{(\delta)}(2x)e^{-x}x^{\delta/2}. \tag{32}$$

The family $(\ell_k^{(\delta)})_{k\geq 0}$ is an orthonormal basis of $\mathbb{L}^2(\mathbb{R}^+)$.

For $\delta = 0$, we set $L_k^{(0)} = L_k$, $\varphi_k^{(0)} = \ell_k$. Using (31), we obtain for $j \geq 1$:

$$\ell_j'(x) = -\ell_j(x) - \sqrt{\frac{2j}{x}}\,\ell_{j-1}^{(1)}(x). \tag{33}$$

The following properties hold for the ℓ_j's. For all $x \geq 0$,

$$|\ell_j(x)| \leq \sqrt{2}, \quad \int_0^{+\infty} \ell_j(x)dx = \sqrt{2}(-1)^j, \quad j \geq 0,$$

$$\ell_0'(x) = -\ell_0(x),\ \ell_j'(x) = -\ell_j(x) - 2\sum_{k=0}^{j-1} \ell_k(x), j \geq 1. \tag{34}$$

Moreover, the following asymptotic formulae can be found in Askey and Wainger (1965). For $\nu = 4k+2$, and k large enough

$$|\ell_k(x/2)| \leq C \begin{cases} a) \quad 1 & \text{if } 0 \leq x \leq 1/\nu \\ b) \quad (x\nu)^{-1/4} & \text{if } 1/\nu \leq x \leq \nu/2 \\ c) \quad \nu^{-1/4}(\nu - x)^{-1/4} & \text{if } \nu/2 \leq x \leq \nu - \nu^{1/3} \\ d) \quad \nu^{-1/3} & \text{if } \nu - \nu^{1/3} \leq x \leq \nu + \nu^{1/3} \\ e) \quad \nu^{-1/4}(x-\nu)^{-1/4} e^{-\gamma_1 \nu^{-1/2}(x-\nu)^{3/2}} & \text{if } \nu + \nu^{1/3} \leq x \leq 3\nu/2 \\ f) \quad e^{-\gamma_2 x} & \text{if } x \geq 3\nu/2 \end{cases}$$

where γ_1 and γ_2 are positive and fixed constants.

7.2. **A useful inequality.** We recall the Talagrand inequality. The result below follows from the Talagrand concentration inequality given in Klein and Rio (2005) [15] and arguments in Birgé and Massart (1998) [7] (see the proof of their Corollary 2 page 354).

Theorem 2. *(Talagrand Inequality) Let $Y_1, \ldots, Y_n$ be independent random variables with values in a Polish space, let $\nu_{n,Y}(f) = (1/n)\sum_{i=1}^n [f(Y_i) - \mathbb{E}(f(Y_i))]$ and let $\mathcal{F}$ be a countable class of uniformly bounded measurable functions. Then for $\epsilon^2 > 0$*

$$\mathbb{E}\Big[\sup_{f\in\mathcal{F}} |\nu_{n,Y}(f)|^2 - 2(1+2\epsilon^2)H^2\Big]_+$$
$$\leq \quad \frac{4}{K_1}\left(\frac{v^2}{n} e^{-K_1\epsilon^2 \frac{nH^2}{v^2}} + \frac{98M^2}{K_1 n^2 C^2(\epsilon^2)} e^{-\frac{2K_1 C(\epsilon^2)\epsilon}{7\sqrt{2}}\frac{nH}{M}}\right),$$

with $C(\epsilon^2) = (\sqrt{1+\epsilon^2} - 1) \wedge 1$, $K_1 = 1/6$, and

$$\sup_{f\in\mathcal{F}} \|f\|_\infty \leq M, \quad \mathbb{E}\Big[\sup_{f\in\mathcal{F}} |\nu_{n,Y}(f)|\Big] \leq H,\ \sup_{f\in\mathcal{F}} \frac{1}{n}\sum_{k=1}^n \mathrm{Var}(f(Y_k)) \leq v^2.$$

By standard density arguments, this result can be extended to the case where $\mathcal{F}$ is a unit ball of a linear normed space, after checking that $f \mapsto \nu_n(f)$ is continuous and $\mathcal{F}$ contains a countable dense family.

7.3. **Characteristic function and moments of $X(t)$.**

Proposition 6.

$$\mathbb{E}e^{iuX(t)} = \exp\left[\int_0^t \psi(ua(v))\right]dv, \quad \psi(u) = \int_{\mathbb{R}} (e^{iux} - 1 - iux)n(x)dx.$$

Proof of Proposition 6. Consider first the case where $Z(t) = \sum_{i=1}^{N(t)} \xi_i$ is a compound Poisson process with $(N(t))$ a Poisson process with intensity λ and $(\xi_i, i \geq 1)$ a sequence of *i.i.d.* random variables independent of the Poisson process $(N(t))$. Then,

$$X(t) = \sum_{n: T_n \leq t} a(t - T_n)\xi_n$$

where (T_n) is the sequence of jumps times of $(N(t))$. We have $X(t)=0$ on $(N(t)=0)$ and $X(t)=\sum_{k=1}^n a(t-T_k)\xi_k$ on $(N(t)=n)$. Let f denote the common density of the ξ_is and f^* their characteristic function. We use that the conditional distribution of $(T_1,\ldots,T_n)$ given $(N(t)=n)$ is equal to the distribution of $(U_{(k)},k=1,\ldots,n)$ the order statistic of $(U_1,\ldots,U_n)$ n *i.i.d.* random variables with uniform distribution on $[0,t]$:

$$\begin{aligned}\mathbb{E}e^{iuX(t)}\mathbf{1}_{(N(t)=n)} &= \mathbb{E}[\mathbf{1}_{(N(t)=n)}\int e^{iu\sum_{k=1}^n a(t-T_k)x_k} f(x_1)\ldots f(x_k)dx_1\ldots dx_k]\\ &= \mathbb{E}[\mathbf{1}_{(N(t)=n)}\prod_{k=1}^n f^*(ua(t-T_k)) = \mathbb{E}[\mathbf{1}_{(N(t)=n)}\mathbb{E}\prod_{k=1}^n f^*(ua(t-U_{(k)}))\\ &= \mathbb{E}[\mathbf{1}_{(N(t)=n)}\mathbb{E}[\prod_{k=1}^n f^*(ua(t-U_k))] = \mathbb{E}[\mathbf{1}_{(N(t)=n)}\left(\frac{1}{t}\int_0^t f^*(ua(t-v))dv\right)^n\end{aligned}$$

Therefore,

$$\mathbb{E}e^{iuX(t)} = e^{(-\lambda t)}e^{[\lambda t\int_0^t dv f^*(ua(t-v)dv/t]} = \exp\int_0^t dv\psi(ua(t-v))$$

where $\psi(u)=\int_{\mathbb{R}}(e^{iux}-1)\lambda f(x)dx$ is the characteristic exponent of $(Z(t))$. Note that if f is centered, we can write $\psi(u)=\int_{\mathbb{R}}(e^{iux}-1-iux)\lambda f(x)dx$.

In the general case, we have

$$\mathbb{E}e^{iuZ(t)} = \exp[t\int_{\mathbb{R}}\left(e^{iux}-1-iux\right)n(x)dx] = \exp t\psi(u), \quad Z(t)=\int_{(0,t]}\int_{\mathbb{R}} x(\hat{p}(ds,dx)-dsn(x)dx),$$

where $\hat{p}(ds,dx)$ is the random Poisson measure associated with the jumps of $(Z(t))$. We consider for $\varepsilon>0$,

$$Z_\varepsilon(t)=\int_{(0,t]}\int_{|x|>\varepsilon} x(\hat{p}(ds,dx)-dsn(x)dx) = Y_\varepsilon(t) - t\int_{|x|>\varepsilon} xn(x)dx.$$

We have $\mathbb{E}e^{iuZ_\varepsilon(t)} = \exp[t\int_{|x|>\varepsilon}\left(e^{iux}-1-iux\right)n(x)dx] = \exp t\psi_\varepsilon(u)$ and $X_\varepsilon(t)=\int_0^t a(t-v)dZ_\varepsilon(v) = \int_0^t a(t-v)dY_\varepsilon(v) - \int_0^t a(t-v)dv\int_{|x|>\varepsilon} xn(x)dx$. The process $(Y_\varepsilon(t))$ is a compound Poisson process with intensity $\mathbf{1}_{|x|>\varepsilon}n(x)dx$. Therefore,

$$\mathbb{E}e^{iuX_\varepsilon(t)} = \exp[\int_0^t dv\int_{|x|>\varepsilon}(e^{iua(t-v)x}-1)n(x)dx - iu\int_0^t a(t-v)dv\int_{|x|>\varepsilon} xn(x)dx].$$

Thus, $\mathbb{E}e^{iuX_\varepsilon(t)} = \exp[\int_0^t dv\psi_\varepsilon(ua(t-v))]$. Now, it is enough to let ε tend to 0 and the proof is achieved. □

Proposition 7. *We have:* $\mathbb{E}X^4(t) = 3(\int_0^t a^2(u)du)^2 + k_4\int_0^t a^4(u)du.$
For all $p\geq 1$, there exists a constant $D(4+2p)$ depending only on p such that

$$\begin{aligned}\mathbb{E}[X^4(t)\mathbf{1}_{X^2(t)>\theta}] &\leq \frac{1}{\theta^p}\mathbb{E}X^{4+2p}(t)\\ &\leq \frac{D(4+2p)}{\theta^p}\left(G^{2+p}(t) + k_{4+2p}\int_0^t a^{4+2p}(u)du\right).\end{aligned}$$

Proof of Proposition 7. We differentiate four times the characteristic function of $X(t)$ and use that for $\psi(u)=\int(e^{iux}-1-iux)n(x)dx$, $\psi'(0)=0, \psi''(0)=-1, \psi^{(4)}(0)=k_4$.

The second inequality is a direct application of the first Kunita inequality (see Applebaum (2009)) where $D(4+2p)$ is the constant given in this inequality.□

References

[1] Abramowitz, M. and Stegun, I. A. (1964). *Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables.* Dover, New York, ninth dover printing, tenth gpo printing edition.

[2] Applebaum, D. (2009). *Lévy processes and stochastic calculus. Second Edition.* Cambridge University Press. Cambridge.

[3] Askey, R. and Wainger, S. (1965). Mean convergence of expansions in Laguerre and Hermite series. *Amer. J. Math.* **87**, 695-708.

[4] Basse A. and Pedersen, J. (2009). Lévy driven moving averages and semipartingales. *Stoch. Proc. Appl.* **119**, 2970-2991.

[5] Belomestny, D., Panov, V. and Woerner, J. H. C. (2019). Low-frequency estimation of continuous-time moving average Lévy processes. *Bernoulli* **25**, 902-931.

[6] Bender, C., Lindner, A. and Schicks, M. (2012). Finite Variation of Fractional Lévy Processes. *Journal of Theoretical Probability* **25** 594- 612.

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

[8] Brockwell, P. J. (2001) Continuous-time ARMA processes. Stochastic processes: theory and methods, 249-276. *In Handbook of Statist.*, **19**, C.R. Rao and D.N. Shanbhag (eds), North-Holland, Amsterdam.

[9] Brockwell, P. J., Ferrazzano, V. and Klüppelberg, C. (2013) High-frequency sampling and kernel estimation for continuous-time moving average processes. *J. Time Series Anal.* **34**, 385-404.

[10] Brockwell, P. J. and Lindner, A. (2009) Existence and uniqueness of stationary Lévy-driven CARMA processes. *Stochastic Process. Appl.* **119**, 2660-2681.

[11] Comte, F. and Genon-Catalot, V. (2018). Laguerre and Hermite bases for inverse problems. *Journal of the Korean Statistical Society*, **47**, 273-296.

[12] Comte, F. and Genon-Catalot, V. (2021). Non parametric estimation for *i.i.d.* Gaussian continuous time moving average models. *Statistical Inference for Stochastic Processes,* **24**, 149-177.

[13] Comte, F. and Renault, E. (1996). Long memory continuous time models. *Journal of Econometrics*, **73**, 101-149.

[14] Hsiao, C. (2003). *Analysis of panel data.* Cambridge University Press, Second Edition. Cambridge.

[15] Klein, T. and Rio, E. (2005). Concentration around the mean for maxima of empirical processes. *Ann. Probab.* **33** 1060-1077.

[16] Klüppelberg, C., Meyer-Brandis, T. and Schmidt, A. (2010). Electricity spot price modelling with a view towards extreme spike risk. *Quantitative finance* **10** 963-974.

[17] Marquart, T.. Fractional Lévy Processes with an Application to Long Memory Moving Average Processes. (2006). *Bernoulli* **12**, 1099-1126.

[18] Rajput B.S. and Rosiński, J. (1989). Spectral representation of infinitely divisible processes. *Probab. Theorey Related Fields* **82** (3), 451-487.

[19] Ramsay, J.O. and Silverman, B.W. (2007). *Applied functional data analysis: Methods and case studies.* Springer.

[20] Revuz, D. and Yor, M. (1999). *Continuous martingales and Brownian motion.* Third edition. Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences], 293. Springer-Verlag, Berlin.

[21] Schnurr, A. and Woerner, J. H. C. (2011). Well-balanced Lévy driven Ornstein-Uhlenbeck processes. *Stat. Risk Model.* **28**, 343-357.

[22] Wang, J.-L., Chiou, J.-M. and Mueller, H.-G. (2015). Review of functional data analysis, *arKiv preprint*: 1507.05135.