HAL
open science

Nonparametric estimation for i.i.d. Gaussian continuous time moving average models.

Fabienne Comte, Valentine Genon-Catalot

To cite this version:

Fabienne Comte, Valentine Genon-Catalot. Nonparametric estimation for i.i.d. Gaussian continuous time moving average models.. Statistical Inference for Stochastic Processes, 2021, 24 (1), pp.149-177. 10.1007/s11203-020-09228-y . hal-02546705

HAL Id: hal-02546705

https://hal.science/hal-02546705v1

Submitted on 18 Apr 2020

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

NONPARAMETRIC ESTIMATION FOR I.I.D. GAUSSIAN CONTINUOUS TIME MOVING AVERAGE MODELS

FABIENNE COMTE* AND VALENTINE GENON-CATALOT*

Abstract. We consider a Gaussian continuous time moving average model $X(t) = \int_0^t a(t-s)dW(s)$ where W is a standard Brownian motion and $a(.)$ a deterministic function locally square integrable on $\mathbb{R}^+$. Given N *i.i.d.* continuous time observations of $(X_i(t))_{t\in[0,T]}$ on $[0,T]$, for $i = 1,\ldots,N$ distributed like $(X(t))_{t\in[0,T]}$, we propose nonparametric projection estimators of a^2 under different sets of assumptions, which authorize or not fractional models. We study the asymptotics in T, N (depending on the setup) ensuring their consistency, provide their nonparametric rates of convergence on functional regularity spaces. Then, we propose a data-driven method corresponding to each setup, for selecting the dimension of the projection space. The findings are illustrated through a simulation study.

Keywords. Continuous time moving average. Gaussian processes. Model selection. Nonparametric estimation. Projection estimators.
MSC 2010. 62G05-62M09

1. Introduction

Samples of infinite dimensional data, especially of data recorded continuously over a time interval are now a commonly encountered type of data due to the possibilities of modern technology. They arise in many fields of applications, *e.g.* in econometrics where authors rather speak of panel data and supply the field of functional data analysis (FDA) whose scope is no more to be demonstrated (see, for general ideas and lots of examples, Hsiao (2003), Ramsay *et al.* (2007), Wang *et al.* (2016)). Parametric models are most often proposed to deal with FDA. However, nonparametric approaches allow for more flexibility and robustness.

In the present contribution, we consider *i.i.d.* observations $(X_i(t), t \in [0,T], i = 1,\ldots,N)$ of the continuous time moving average (CMA) process

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

where $(W(t), t \geq 0)$ is a Wiener process and $a : \mathbb{R}^+ \to \mathbb{R}$ is a deterministic square integrable function. Our aim is to study the new and challenging question of the nonparametric estimation of the function $g = a^2$ from these observations under very general conditions on the function $a(t)$. Our assumptions include in particular the classical CARMA processes (continuous ARMA) but also more complicated processes such as the continuous time fractionally integrated process of order d (see (3)), defined in Comte and Renault (1996, Definition 2) which is linked with Brownian motion with Hurst index $H = d + (1/2)$.

CMA processes have been the subject of a huge number of contributions concerned with modelling properties. Estimation procedures rely on the observation of a unique sample path on a time interval $[0,T]$ and usually, the stationary version of $(X(t))$, namely

$$Y(t) = \int_{-\infty}^t a(t-s)dW(s), \tag{2}$$

is considered. We refer *e.g.* to Brockwell (2001) for a reference book, Brockwell *et al.* (2012) and the references given therein, where a general Lévy process $(L(t))$ may replace $(W(t))$ (see also *e.g.* Belomestny *et al.* (2019), Schnurr and Woerner (2011)). For what concerns nonparametric estimation, a pointwise estimator of $a(t)$ for mainly Gaussian CARMA(p,q) processes in stationary regime (see formula (2)), is proposed in Brockwell *et al.* (2012) based on the discrete observation of one sample path. Except for this reference, to our knowledge, the nonparametric estimation of $a(t)$ for general CMA processes has not yet been studied.

In the present paper, stationarity of the process is not required. The asymptotic framework will be that either N tends to infinity with fixed T or both $N$and T tend to infinity. We assume that g is square integrable. Considering sequences $(S_m, m \in \mathbb{N})$ of finite dimensional subspaces of $\mathbb{L}^2(\mathbb{R}^+)$, we propose two kinds of projections estimators of g built using the observations $(X_i(t), t \in [0,T], i = 1, \ldots, N)$: *i.e.*, we build estimators of the orthogonal projection g_m of g on S_m by estimating the coefficients of the projection on an orthonormal basis of S_m. The first method relies on the assumption that $a(t)$ belongs to $C^1([0,+\infty))$ which excludes the continuous time fractionally integrated process. In this case, $(X(t))$ is an Itô process with explicit stochastic differential. The second approach which is more general applies without regularity assumptions on $a(t)$. Then, in the general case, we propose a data-driven selection of the dimension leading to an adaptive estimator. For this part, the Gaussian character of the process $(X(t))$ is especially exploited. Proofs which do not rely on this property are possible though longer.

In Section 2, we present assumptions and the collections of models. Two collections are especially investigated. First, we consider for fixed T the collection of spaces generated by the trigonometric basis of $\mathbb{L}^2([0,T])$ and thus we estimate $g_T = g\mathbf{1}_{[0,T]}$. Second, for large T, we consider spaces generated by the Laguerre basis of $\mathbb{L}^2(\mathbb{R}^+)$. This basis has been largely investigated and used in recent references for nonparametric estimation by projection method (see *e.g.* Comte and Genon-Catalot (2018)). The estimators are presented in Sections 2.2 (first method) and 2.3 (second method under more general assumptions). Several risk bounds for the projection estimators on a fixed space are obtained and discussed. In Section 3, we detail the possible rates of convergence that can be deduced from the risk bounds depending on regularity spaces for the unknown function g. Section 4 is concerned with the data-driven choice of the dimension of the projection space. We prove that our estimators are adaptive in the sense that their risk bounds automatically achieve the best compromise between square bias and variance terms (Theorems 1 and 2). Section 5 contains a simulation study. Estimators are implemented on simulated data for various examples of functions g. We give table of risks obtained by Monte-Carlo simulations. In Section 6, some concluding remarks are given. Proofs are gathered in Section 7 and Section 8 contains the necessary definitions and properties of the Laguerre basis.

2. Projection estimators on a fixed space.

2.1. Assumptions and collection of models.

We estimate the function

$$g(t) := a^2(t).$$

Note that the sign of $a(t)$ is not identifiable. Our study will depend on assumptions on the unknown function $a(t)$:

- [H0] The function $g(t) = a^2(t)$ belongs to $\mathbb{L}^1(\mathbb{R}^+) \cap \mathbb{L}^2(\mathbb{R}^+)$
- [H1] The function $a(t)$ belongs to $C^1(\mathbb{R}^+)$, is bounded and $\int_0^{+\infty}(a'(t))^2 dt < +\infty$.

Example 1. Consider the following example: $a(t) = t^d\tilde{a}(t)/\Gamma(d+1)$ where $d > -1/2$ and $\tilde{a} \in C^1(\mathbb{R}^+)$ and $\tilde{a}(0) \neq 0$,

$$X(t) = \int_0^t \frac{(t-s)^d}{\Gamma(d+1)}\tilde{a}(t-s)dW(s). \tag{3}$$

In particular for $\tilde{a}(x) = 1$, this process is the continuous time fractional Brownian motion of order d defined in Comte and Renault (1996, Definition 1) and the general formulation above corresponds to the continuous time fractionally integrated process of order d (Definition 2 therein). The integrability of a^2, a'^2, a^4 near infinity can be ensured by the rate of decrease of $\tilde{a}$ near infinity, for instance if $\tilde{a}(t) = e^{-t}$. The behaviour near 0 depends on d:

- (i) The process $X(t)$ is well defined for any $d > -1/2$ as a is locally square integrable.
- (ii) For $-1/2 < d < 0$, $a(0)$ is not defined.
- (iii) For $d \geq 1$, $a(t)$ belongs to $C^1(\mathbb{R}^+)$ and a' is locally square integrable.
- (iv) As $a(t) \sim (\tilde{a}(0)/\Gamma(d+1))t^d$ at 0, [H0] requires $d > -1/4$.

In other words, fractional processes can be studied only under [H0].

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 $\mathbb{L}^2(\mathbb{R}^+)$. We set

$$G(t) := \int_0^t a^2(s)ds \leq \|a\|^2. \tag{4}$$

Note that $\mathbb{E}(X^2(t)) = G(t)$ is what enables us to estimate g, whereas $\mathbb{E}(Y^2(t)) = \|a\|^2$ would not. To build estimators of g, we use a projection method and consider two settings.

- In the first case, T is fixed and we estimate $g_T = g\mathbf{1}_{[0,T]}$. For this, we consider the collection $(S_m^{Trig}, m \geq 0)$ of subspaces of $\mathbb{L}^2([0,T])$ where S_m^{Trig} has odd dimension m and is generated by the orthonormal trigonometric basis $(\varphi_{j,T})$ where $\varphi_{0,T}(t) = \sqrt{1/T}\mathbf{1}_{[0,T]}(t)$, $\varphi_{2j-1,T}(t) = \sqrt{2/T}\cos(2\pi jt/T)\mathbf{1}_{[0,T]}(t)$ and $\varphi_{2j,T}(t) = \sqrt{2/T}\sin(2\pi jt/T)\mathbf{1}_{[0,T]}(t)$ for $j = 1, \ldots, (m-1)/2$. This basis satisfies
$$\sum_{j=0}^{m-1}\varphi_{j,T}^2(t) = \frac{m}{T} \quad \text{and} \quad \int_0^T \varphi_{0,T}(t)dt = \sqrt{T}, \int_0^T \varphi_{j,T}(t)dt = 0 \quad \text{for} \quad j \neq 0.$$
- In the second case, we may consider that either T is fixed but large enough, or that T tends to infinity. In this case, we estimate g on $\mathbb{R}^+$ and we rather consider a collection of subspaces of $\mathbb{L}^2(\mathbb{R}^+)$, generated by an orthonormal basis. The basis considered here is the Laguerre basis defined by
$$\ell_j(t) = \sqrt{2}L_j(2t)e^{-t}\mathbf{1}_{t\geq 0}, \quad j \geq 0, \quad L_j(t) = \sum_{k=0}^{j}(-1)^k\binom{j}{k}\frac{t^k}{k!}. \tag{5}$$
We set $S_m^{Lag} = \text{span}\{\ell_j, j = 0, \ldots, m-1\}$. We have
$$\forall t \geq 0, \quad \sum_{j=0}^{m-1}\ell_j^2(t) \leq 2m, \quad \text{and} \quad \int_0^{+\infty}\ell_j(t)dt = \sqrt{2}(-1)^j.$$
The second property is obtained by exact computation and the first one comes from the fact that $\forall j, |\ell_j(t)| \leq \sqrt{2}$. Moreover, $\mathcal{L}_j(T) := \int_0^T \ell_j(u)du$ can computed recursively, see (47). All formulae concerning this basis are recalled in Section 8.

Remark 1. *In the case of fixed T, we could also consider the subspaces (S_m^{Hist}) of $\mathbb{L}^2([0,T])$ generated by the histogram basis*

$$\varphi_{j,T}(t) = \sqrt{m/T}\mathbf{1}_{[jT/m,(j+1)T/m[}(t), j = 0, \dots, m-1$$

where $\sum_{j=0}^{m-1} \varphi_{j,T}^2(t) = m/T$ and $\int_0^T \varphi_{j,T}(t)dt = \sqrt{T/m}$. But these basis functions are not differentiable and thus would not be suitable for all our proposals.

For simplicity, in order to use a unique notation, we denote by φ_j either $\varphi_{j,T}$ or ℓ_j and set $S_m = \text{span}\{\varphi_j, j = 0, \dots, m-1\}$. In all cases, under [H0], the function g admits a development

$$g = \sum_{j\geq 0} \theta_j \varphi_j, \text{ with } \theta_j = \int_0^{+\infty} g(s)\varphi_j(s)ds = \langle g, \varphi_j \rangle.$$

We define $g_m(t) = \sum_{j=0}^{m-1} \theta_j \varphi_j(t)$ the orthogonal projection of g on S_m.

2.2. **Estimators under [H0]-[H1].** Under [H1], the stochastic differential of $(X(t))$ satisfies:

$$dX(t) = a(0)dW(t) + \left[\int_0^t a'(t-s)dW(s)\right] dt. \tag{6}$$

(see Comte and Renault, 1996, equation (6)).

Remark 2. *By equation (6), we have, for each trajectory X_i, for $t_k = kT/n$ with fixed T,*

$$\frac{1}{T}\sum_{k=0}^{n-1}(X_i(t_{k+1}) - X_i(t_k))^2 \to_{n\to+\infty} a^2(0) = g(0), \text{ in probability.}$$

Thus, we can assume that $g(0)$ is known, as we have continuous observation of the sample paths.

The construction of our first estimator relies on the following lemma.

Lemma 1. *Under* **[H0]-[H1]**, *denoting by $\theta_j = \langle g, \varphi_j \rangle$, we have*

$$\mathbb{E}\left(\int_0^{+\infty} \varphi_j(s)X(s)dX(s)\right) = \frac{1}{2}\left(\theta_j - g(0)\int_0^{+\infty} \varphi_j(s)ds\right).$$

Obviously, if the basis has support $[0,T]$, integrals are on this interval. Relying on this lemma, we can set:

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

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

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

We refer to Remark 2 concerning the fact that $g(0)$ is known. We mention that here, the histogram basis can be used in the fixed-T setting.

Note that, by the Ito formula and (6), we can write $\hat{\theta}_j$ without stochastic integral, provided that φ_j is differentiable:

$$\hat{\theta}_j = \varphi_j(T)\frac{1}{N}\sum_{i=1}^N X_i^2(T) - \frac{1}{N}\sum_{i=1}^N \int_0^T \varphi_j'(s)X_i^2(s)ds. \tag{8}$$

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

Proposition 1. *Assume* **[H0]**-**[H1]** *and consider that* $(\varphi_j = \varphi_{j,T})$ *is the trigonometric basis. Then*

$$\mathbb{E}(\|\hat{g}_m - g\|_T^2) \leq \|g_m - g\|_T^2 + 8g(0)G(T)\frac{m}{N} + 4\left(2G(T)G_1(T) + g^2(0)\right)\frac{T}{N} + \frac{4\|g\|_T^2}{N} \tag{9}$$

where $G_1(T) = \int_0^T (a'(u))^2 du$. *(Recall that* G *is defined in (4), that* g_m *denotes the orthogonal projection of* g *on* S_m^{Trig} *and that* $\|u\|_T^2 = \int_0^T u^2(s)ds$*.)*
If $g(0) = 0$,

$$\mathbb{E}(\|\hat{g}_m - g\|_T^2) \leq \|g_m - g\|_T^2 + 4G(T)G_1(T)\frac{T}{N} + \frac{\|g\|_T^2}{N} \tag{10}$$

Let us discuss these bounds for fixed T and large N. The bounds involve a standard squared bias term $\|g - g_m\|_T^2$ due to the projection method. For $g(0) \neq 0$, the variance has order m/N and the last two terms are residuals (see (9)). Therefore in this case, for choosing m, the bias-variance compromise can be done between the first two terms.

The case $g(0) = 0$ is different as the process is differentiable, see (6) with $a(0) = 0$, and the bound (10) shows that m must simply be chosen as large as possible.

Proposition 2. *Assume* **[H0]**-**[H1]**.
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*

$$\mathbb{E}(\|\hat{g}_m - g\|^2) \quad \leq \quad \|g_m - g\|^2 + 8g(0)G(T)\frac{m}{N} + c_G\frac{T}{N} + 2\int_T^{+\infty} g^2(s)ds + \frac{4\|g\|^2}{N}. \tag{11}$$

where $c_G = 4\left(2G(T)G_1(T) + g^2(0)\right)$. *If in addition* $g(0) = 0$,

$$\mathbb{E}(\|\hat{g}_m - g\|^2) \leq \|g_m - g\|^2 + 4G(T)G_1(T)\frac{T}{N} + \int_T^{+\infty} g^2(s)ds + \frac{2\|g\|^2}{N}. \tag{12}$$

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

$$\mathbb{E}(\|\hat{g}_m - g\|^2) \leq \|g_m - g\|^2 + C\frac{m^2}{N} + C'\|a\|^2 m\exp\left(-12\gamma_2 m\right) \tag{13}$$

where $C = C'(g(0)^2 + \|g\|_\infty^2 + \|a\|^2\|a'\|^2))$, C' *and* γ_2 *are positive constants depending on the basis only.*

We can discuss these bounds for fixed T or large T. Here again, the bounds involve a standard squared bias term $\|g - g_m\|^2$.
Bounds (11) and (12) may be compared to (9) and (10). In (9), T is fixed so that the variance has order m/N for $g(0) \neq 0$ and the term T/N is a negligible residual. If T can be large, the term T/N may no more be negligible and (11)-(12) involve an additional bias term $\int_T^{+\infty} g^2(s)ds$ which is small for large T. But the order of these terms, depending on T which can not be chosen, are difficult to discuss.
Bound (12) implies as in the trigonometric case that m must be chosen as large as possible. Bound (13) looks more classical: T does not appears, the variance term has order m^3/N and $m\exp\left(-12\gamma_2 m\right)$ is a negligible additional bias term.
Comparing (11) and (13), we see that in the Laguerre case, the variance term is less than

$$\min\left\{8g(0)G(T)\frac{m}{N} + 4\left(2G(T)G_1(T) + g^2(0)\right)\frac{T}{N}, C\frac{m^2}{N}\right\}.$$

Note that the constants $G_1(T)$ and C are difficult to estimate which is a drawback for model selection. In section 5, we propose a practical data-driven choice of m taking into account this difficulty.

2.3. **Estimator under [H0].** In this paragraph, to handle more general processes, including fractional processes, we propose another estimation method. We no longer assume that a belongs to $C^1(\mathbb{R}^+)$. Therefore, the stochastic differential (6), which requires **[H1]**, no more holds. As a counterpart, we consider basis functions that have to be differentiable on their domain, $[0,T]$ or $\mathbb{R}^+$.

The construction of the second estimator is based on the following lemma.

Lemma 2. *Assume that* **[H0]** *holds 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

$$\tilde{\theta}_j = -\frac{1}{N}\sum_{i=1}^N \left(\int_0^T \varphi_j'(s)X_i^2(s)ds\right) + \varphi_j(T)\widehat{G}(T) \quad \text{and} \quad \widehat{G}(T) = \frac{1}{N}\sum_{i=1}^N X_i^2(T). \tag{14}$$

Note that under **[H0]** only, formula (8) no longer holds, this is why we use another notation, $\tilde{\theta}_j$ instead of $\hat{\theta}_j$. 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

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

We introduce the assumption:

- **[H2]** $\displaystyle\int_0^1 \frac{G^2(s)}{s}ds = c_0 < +\infty.$

Actually, **[H2]** is rather weak and allows to consider fractional processes.

Example 1 (*continued*). If we consider, as in example 1, $a(t) = t^d\tilde{a}(t)/\Gamma(d+1)$, where $d > -1/2$ and $\tilde{a} \in C^1(\mathbb{R}^+)$, with $\tilde{a}(0) \neq 0$, then $G^2(s)/s \sim_{s\to 0} s^{4d+1}\tilde{a}^4(0)/\Gamma^4(d+1)$ and **[H2]** holds $(d > -1/2)$. The constraint is weaker than **[H0]**.

The following risk bounds hold for $\tilde{g}_m$.

Proposition 3. *Assume* **[H0]**.

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

$$\mathbb{E}(\|\tilde{g}_m - g\|_T^2) \leq \|g_m - g\|_T^2 + 6G^2(T)\frac{4\pi^2 m^2}{NT} + 6G^2(T)\frac{m}{NT}. \tag{15}$$

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

$$\mathbb{E}(\|\tilde{g}_m - g\|^2) \leq \|g_m - g\|^2 + 12\left(G^2(T) + 2\int_0^T \frac{G^2(u)}{u}du\right)\frac{m}{N} + 12G^2(T)\frac{T}{N} + \int_T^\infty g^2(s)ds, \tag{16}$$

where, if **[H2]** *holds, ,*

$$\int_0^T \frac{G^2(u)}{u}du \leq c_0 + G^2(T)\log(T).$$

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

$$\mathbb{E}(\|\tilde{g}_m - g\|^2) \leq \|g_m - g\|^2 + c_1 G^2(T)\frac{m^3}{N} + c_2\|a\|^2 m \, \exp(-12\gamma_2 m) \tag{17}$$

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

As previously, all the risk bounds involve a squared bias term, $\|g-g_m\|_T^2$ or $\|g-g_m\|^2$. The variance term in (15) can be compared to the one in (9), taking into acount that $G(T) \leq \|a\|^2 < +\infty$, and the order is now $m^2/(NT)$ which for fixed T is larger than m/N obtained for $\hat{g}_m$ with the sama basis. Similarly, the variance term in (17) has order m^3/N, which is larger than m^2/N in (13). This increase is the price of more general assumptions and estimators. As for (16), it is to be compared with (11): the variance order is m/N and there are the two additional terms T/N and $\int_T^\infty g^2(s)ds$, difficult to discuss. We develop a data-driven selection method in section 4, based on (15)-(16), which is implemented on simulated data.

3. Rates of convergence

Rates of convergence can be deduced from Propositions 1 and 3 in the asymptotic framework where N tends to infinity. As it is always the case in nonparametric estimation, we must link the bias term $\|g-g_m\|^2$ with regularity properties of function g, and the regularity spaces depend on the projection spaces.

3.1. **Rates on periodic Fourier-Sobolev spaces for trigonometric basis.** Consider first Inequality (9) and estimators built using the trigonometric basis. Let β be a positive integer, $L > 0$ and define

$$\begin{aligned} W^{per}(\beta, L) \quad = \quad & \{f : [0,T] \to \mathbb{R}, f^{(\beta-1)} \text{ is absolutely continuous,} \\ & \int_0^T [f^{(\beta)}(x)]^2 dx \leq L^2, f^{(j)}(0) = f^{(j)}(T), j = 1, \ldots, \beta-1\}. \end{aligned}$$

By Proposition 1.14 of Tsybakov (2009), a function $f \in W^{per}(\beta, L)$ admits a development $f = \sum_{j=0}^\infty \theta_j \varphi_{j,T}$ such that $\sum_{j\geq 0} \theta_j^2 \tau_j^2 \leq C(L,T)$ where $\tau_j = j^\beta$ for even j, $\tau_j = (j-1)^\beta$ for odd j and $C(L,T) = L^2(T/\pi)^{2\beta}$.

Therefore, if $g \in W^{per}(\beta, L)$, then $\|g-g_m\|^2 \leq C(L,T)m^{-2\beta}$ and Inequality (9) becomes

$$\mathbb{E}(\|\hat{g}_m - g\|_T^2) \leq C(L,T)m^{-2\beta} + C_1 g(0)\frac{m}{N} + C_1\frac{T}{N}.$$

If $g(0) \neq 0$, choosing $m_{\text{opt}} = c_T N^{1/(2\beta+1)}$ yields, for fixed T,

$$\mathbb{E}(\|\hat{g}_{m_{\text{opt}}} - g\|_T^2) \lesssim N^{-2\beta/(2\beta+1)} + \frac{T}{N}.$$

For fixed (not large) T, the estimator $\hat{g}_{m_{\text{opt}}}$ is convergent in MISE when N grows to infinity, with rate $N^{-2\beta/(2\beta+1)}$.
If $g(0) = 0$, we choose m as large as possible and can obtain the rate N^{-1} for fixed T.
On the other hand, if $g \in W^{per}(\beta, L)$, Inequality (15) yields, for a choice $\tilde{m}_{\text{opt}} = \tilde{c}_T N^{1/(2\beta+2)}$ a rate for $\tilde{g}_{\tilde{m}_{\text{opt}}}$ of order $N^{-2\beta/(2\beta+2)} > N^{-2\beta/(2\beta+1)}$. The rate is less good than the one of $\hat{g}_{m_{\text{opt}}}$, but the assumptions are different: the rate of $\hat{g}_{m_{\text{opt}}}$ holds under **[H0]** and **[H1]** while $\tilde{g}_{\tilde{m}_{\text{opt}}}$ only requires **[H0]**.

3.2. **Rates on Sobolev-Laguerre spaces.** Now, look at inequality (13) where $\hat{g}_m$ is computed using the (non compactly supported) Laguerre basis. Assume for consistency that $m^2 \lesssim N$ and $m \leq T/6$. The last term is negligible with respect to the variance term m^2/N and the usual square bias term $\|g - g_m\|^2$. An adequate solution to assess the rate of the bias term is provided by the balls of Sobolev-Laguerre spaces. For $s \geq 0$, let

$$W^s((0,+\infty),K) = \{h : (0,+\infty) \to \mathbb{R}, h \in \mathbb{L}^2((0,+\infty)), \sum_{k\geq 0} k^s \theta_k^2(h) \leq K < +\infty\} \tag{18}$$

where $\theta_k(h) = \int_0^{+\infty} h(u)\varphi_k(u)du$. We set

$$W^s((0,+\infty)) = \{h : (0,+\infty) \to \mathbb{R}, h \in \mathbb{L}^2((0,+\infty)), \sum_{k\geq 0} k^s \theta_k^2(h) \ < +\infty\}$$

for the Sobolev-Laguerre space. The link with regularity properties of functions can be seen for s integer. In this case, if $h : (0,+\infty) \to \mathbb{R}$ belongs to $L^2((0,+\infty))$,

$$\sum_{k\geq 0} k^s (\theta_k(h))^2 < +\infty \tag{19}$$

is equivalent to the property that h admits derivatives up to order $s-1$, with $h^{(s-1)}$ absolutely continuous on $(0,+\infty)$ and for $m = 0, \ldots, s-1$, the functions

$$x^{(m+1)/2}(he^x)^{(m+1)}e^{-x} = x^{(m+1)/2} \sum_{j=0}^{m+1} \binom{m+1}{j} h^{(j)}$$

belong to $\mathbb{L}^2((0,+\infty))$. Moreover, for $m = 0, 1, \ldots, s-1$,

$$\|x^{(m+1)/2}(he^x)^{(m+1)}e^{-x}\|^2 = \sum_{k\geq m+1} k(k-1)\ldots(k-m)\theta_k^2(h).$$

(see Comte and Genon-Catalot, 2015, 2018).

Now, assume that g belongs to $W^s((0,+\infty),K)$ so that $\|g - g_m\|^2 \leq Km^{-s}$. Considering Inequality (13), the minimization of $m^{-s} + m^2/N$ yields $m_{opt} = N^{1/(2+s)}$ and a rate of order $N^{-s/(2+s)}$ for the $\mathbb{L}^2$-risk of $\hat{g}_m$ on a Sobolev-Laguerre ball.
The constraint $m_{opt} = N^{1/(2+s)} \leq T/6$ holds for all s as soon as $T \geq \sqrt{N}$.

As for the first estimator, we can discuss rates of convergence for the $\mathbb{L}^2$-risk of $\tilde{g}_m$, relying on Inequality (17). Assume that $m^3 \lesssim N$ and that g belongs to $W^s((0,+\infty),K)$ so that $\|g - g_m\|^2 \leq Km^{-s}$. By minimizing $(m^3/N) + m^{-s}$, we find $m_{opt} = N^{1/(s+3)}$ and a rate of order $N^{-s/(s+3)}$. The constraint $T > 6m_{opt}$ holds for all s as soon as $T \geq N^{1/3}$.

Inequalities (11) and (16) are appealing: the variance terms are smaller and they require less conditions. However they contain a term $\int_T^{+\infty} g^2(s)ds$: this term is hopefully small for large (not too small) T, but rates of convergence are difficult to discuss. Nevertheless, our model selection procedures rely on these inequalities because the constants $g(0)G(T)$ and $G^2(T)$ are known in theory and possible to estimate in practice.

Example 1 *(continued).* Consider the function $a(t) = t^d \exp(-t)$ with $-1/4 < d < 1/2$, case where $a(0)$ may not be defined and a' is not locally square integrable. Then, g belongs to $W^1((0,+\infty))$ if, moreover, $\sqrt{t}(a^2(t) + 2a(t)a'(t)) \in \mathbb{L}^2((0,+\infty))$, which holds for $0 < d < 1/2$. But for these values of d, we can check that g does not belong to $W^2((0,+\infty))$ as $t(a^2(t) +$

$2(a')^2(t)+(a^2)''(t))$ does not belong to $\mathbb{L}^2((0,+\infty))$. Therefore, the bias term for such a function is of order smaller than m^{-1} but larger than m^{-2}, for $0<d<1/2$.

4. Adaptive procedure under [H0]

As the second estimator can be computed under more general assumptions, we concentrate on this one for finding a data-driven choice of the projection dimension.

The estimator $\tilde{g}_m$ can be obtained as:

$$\tilde{g}_m = \arg\min_{h\in S_m^{(B)}} \gamma_{N,T}(h), \tag{20}$$

for $(B)=(Lag)$ or $(B)=(Trig)$, and where

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

We consider the sets $\mathcal{M}_N^{(Lag)} = \{m\in\mathbb{N}, m\le N/\log(T)\}$ and $\mathcal{M}_N^{(Trig)} = \{m\in\mathbb{N}, m^2\le N\}$. By inequalities (15)-(16), the variance term in the $\mathbb{L}^2$-risk of all $\tilde{g}_m$ with $m\in\mathcal{M}_N^{(B)}$ is bounded, where the superscript (B) indicates the basis: $(B)=(Trig)$ for the trigonometric basis and $(B)=(Lag)$ for the Laguerre basis. Now, we define, for κ a numerical constant,

$$\widetilde{m}^{(B)} := \arg\min_{m\in\mathcal{M}_N^{(B)}} \left(\gamma_{N,T}(\tilde{g}_m) + \mathrm{pen}^{(B)}(m)\right), \tag{21}$$

where

$$\mathrm{pen}^{(Lag)}(m) = \kappa\log(N)\left(G^2(T) + \int_0^T \frac{G^2(u)}{u}du\right)\frac{m}{N},$$

$$\mathrm{pen}^{(Trig)}(m) = \kappa G^2(T)\log(N)\frac{m^2}{NT}.$$

Note that $\gamma_{N,T}(\tilde{g}_m) = -\|\tilde{g}_m\|^2$. Thus, as $\|g-g_m\|^2 = \|g\|^2 - \|g_m\|^2$, $-\|\tilde{g}_m\|^2$ provides an estimation of the squared bias, up to a constant. On the other hand, $\mathrm{pen}^{(B)}(m)$ has the variance order, up to the $\log(N)$ factor. We do not know if this factor is structural or due to technical problems (in the proofs) only. Anyway, the choice of $\widetilde{m}^{(B)}$ mimicks the squared bias-variance compromise. The following risk bound holds.

Theorem 1. *Assume* **[H0]** *and* **[H2]**. *Then, there exists a numerical value* $\kappa_0^{(B)}>0$ *such that* $\forall\kappa\ge\kappa_0^{(B)}$,

$$\mathbb{E}(\|\tilde{g}_{\widetilde{m}^{(B)}} - g\|^2) \le 4\inf_{m\in\mathcal{M}_N^{(B)}}\left(\|g-g_m\|^2 + \mathrm{pen}^{(B)}(m)\right) + C^{(B)}(T,N),$$

$$C^{(Lag)}(T,N) = 32G^2(T)\frac{T\log(N)}{N} + \int_T^{+\infty} g^2(s)ds + \frac{C}{N}\left(\frac{TG^2(T)}{N} + \int_0^T \frac{G^2(u)}{u}du\right)$$

and

$$C^{(Trig)}(T,N) = \frac{C}{N}\left(\frac{1}{T} + \frac{G^2(T)}{T}\frac{1}{N^{1/2}}\right)$$

where C is a numerical constant.

The term $G^2(T)$ is unknown and must be replaced by an estimator. In practical implementation, we set

$$\widehat{G^2(T)} = \frac{1}{3N}\sum_{i=1}^{N} X_i^4(T). \tag{22}$$

Indeed, $\mathbb{E}(X_1^4(T)) = 3G^2(T)$. From theoretical point of view, it can be proved that Theorem 1 still holds with this substitution.

For the implementation of the procedure, we have to fix the constants κ in the penalties (see (21)). The numerical values of $\kappa_0^{(B)}$, given in the proofs, are too large. In this method, finding the minimal value of κ is a difficult problem. This is why the choice of κ in the penalties is standardly calibrated by preliminary simulations.

Theorem 1 shows that the estimator $\tilde{g}_{\widetilde{m}^{(B)}}$ is adaptive in the sense that its $\mathbb{L}^2$-risk automatically achieves the best compromise between squared bias and variance terms, up to remainder terms $C^{(B)}(T,N)$. For $C^{(Trig)}(T,N)$, it is clearly negligible, as $T > 1$ is fixed. As already noticed earlier, the term $C^{(Lag)}(T,N)$ contains T/N and $\int_T^{+\infty} g^2(s)ds$ which are in conflict: T should be large enough for the latter, but not too large for the former. However our risk bounds are valid for any T, N.

Another strategy is possible for Laguerre basis, without **[H2]**, which solves the conflict mentioned above. Let $\mathcal{M}_N^\star = \{m, m^3 \leq N\}$ so that, by inequality (17), the variance term of $\tilde{g}_m$ is bounded and define

$$m^\star := \arg\min_{m\in\mathcal{M}_N^\star} \left(\gamma_{N,T}(\tilde{g}_m) + \mathrm{pen}^\star(m)\right), \text{ where } \mathrm{pen}^\star(m) = \kappa\log(N)G^2(T)\frac{m^3}{N}. \tag{23}$$

Theorem 2. *Assume* **[H0]**. *Consider the Laguerre basis, and $T \geq 6N^{1/3}$. Then, there exists a numerical value $\kappa_0^\star > 0$ such that $\forall\kappa \geq \kappa_0^\star$,*

$$\mathbb{E}(\|\tilde{g}_{m^\star} - g\|^2) \leq 4\inf_{m\in\mathcal{M}_N^\star}\left(\|g - g_m\|^2 + \mathrm{pen}^\star(m)\right) + G^2(T)\frac{C}{N},$$

where C is a constant depending on the basis.

Theorem 2 also shows that the estimator $\tilde{g}_{m^\star}$ is adaptive in the sense that its $\mathbb{L}^2$-risk automatically achieves the best compromise between the squared bias and the variance term of inequality (17). The comments after Theorem 1 apply also here.

5. Simulation study

In this section, we implement the adaptive estimators of the previous sections on simulated data. To simulate an exact discrete sampling of $(X_i(t), i = 1, \ldots, N$ with small sampling interval Δ, we use the property that the vectors $(X_i(k\Delta), k = 1, \ldots, n)'$ with $T = n\Delta$ are *i.i.d.* centered Gaussian vectors with covariance matrix $A = (A_{j,k})$ where for $1 \leq j \leq k$,

$$A_{j,k} = \int_0^{j\Delta} a(j\Delta - u)a(k\Delta - u)du = \int_0^{j\Delta} a(v)a((k-j)\Delta + v)dv.$$

which can be computed exactly or numerically according to the examples. Integrals in the estimators formulae are discretized. The following examples of functions $a(.)$ and thus $g(.)$ are considered.

(1) (Ornstein-Uhlenbeck process) $a_1(t) = \sigma \exp(-\theta t)$,

$$A(j,k) = \frac{\sigma^2 \exp(-\theta k \Delta)}{2\theta}(\exp(\theta j \Delta) - \exp(-\theta j \Delta)),\ k \geq j.$$

We take $\sigma = 0.5, \theta = 0.25$.

(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) $a_3(t) = (\frac{1}{2}\beta(3,3,t/3) + \frac{1}{2}\beta(3,3,t/3-2))^{1/2}$.

(4) $a_4(t) = 10b(6t)/(\omega_4)^{0.25}$ 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 $\int_{\mathbb{R}^+} g_4^2(u)du \approx 1$.

(5) $a_5(t) = t^{1.25}e^{-t/2}$.

(6) $a_6(t) = t^{0.25}e^{-t/3}$.

(7) $a_7(t) = t^{-0.125}e^{-t/5}$.

(8) $a_8(t) = 1/\sqrt{1+t^2}$.

In all cases, recall that $g_i(t) = a_i^2(t)$. The functions a_2 and a_4 are normalized (constants ω_2, ω_4), in order that $\int g_i^2(u)du \approx 1$, $i = 2,4$, while for the other functions, this integral falls between 0.5 and 2.5. In Table 1, we compute the values of residual terms of formula (16): the values of $\int_T^{+\infty} g_i^2(u)du$ are always negligible; but the values of $TG_i^2(T)/N$ are comparable to the risk values obtained in Table 2, and thus not so small.

i	1	2	3	4	5	6	7	8
$\int_T^{+\infty} g_i^2(u)du$	$2.8\ 10^{-6}$	0	0	$5.4\ 10^{-8}$	$1.4\ 10^{-4}$	$1.4\ 10^{-5}$	$1.3\ 10^{-4}$	$3.3\ 10^{-4}$
$\frac{TG_i^2(T)}{N}$	$0.5\ 10^{-2}$	$3.5\ 10^{-2}$	$4.4\ 10^{-2}$	$2.5\ 10^{-2}$	$5.4\ 10^{-2}$	$1.3\ 10^{-2}$	$3.2\ 10^{-2}$	$1.1\ 10^{-2}$

TABLE 1. Order of residual terms, for $T = 10$ and $N = 2000$.

All functions g_i, $i = 1, \ldots, 8$ satisfy **[H0]**. The functions g_2 to g_6 are null at zero, a_6 and a_7 do not satisfy **[H1]**. Thus, the first method (valid under **[H0]-[H1]**) should work for all functions except g_6 and g_7, with parametric rate (and large chosen dimension) for g_2 to g_5. Nevertheless, we implemented both methods for all functions. Note that all functions satisfy **[H2]**.

We also experiment different settings for (N,T): $T = n\Delta = 10$, $n = 400, \Delta = 0.1/4$ with $N = 500, 2000, 8000$.

The estimators are computed via the formulae given in Section 2.2 and 2.3.

More precisely, inspired by Inequalities (9) and (11), we implement a data driven estimator relying on $\hat{g}_m$ given in Section 2.2 with dimension selected as follows: for $(B) = (Lag), (Trig)$,

$$\hat{m}^{(B)} = \arg\min_{1 \leq m \leq D_{\max}} \left\{ -\|\hat{g}_m^{(B)}\|^2 + \kappa_1^{(B)} g^\dagger(0) \widehat{G(T)} \frac{m}{N} \right\}.$$

Note that no theoretical result is given in this case. We compute $(\hat{g}_m^{(Trig)})_{1 \leq m \leq D_{\max}}$ and $(\hat{g}_m^{(Lag)})_{1 \leq m \leq D_{\max}}$ the collections of estimators respectively in trigonometric and Laguerre basis, with coefficients given by (7), with $D_{\max} = 45$. Note that the first term in the curly bracket estimates the squared bias and the second estimates the main variance term. Moreover $\widehat{G}(T) = (1/N)\sum_{i=1}^N X_i^2(T)$ and $g^\dagger(0)$ is computed using the quadratic variation (see Remark 2)

$$g^\dagger(0) = \frac{1}{NT} \sum_{i=1}^N \sum_{k=0}^{n-1} \left[X_i\left(\frac{(k+1)T}{n}\right) - X_i\left(\frac{kT}{n}\right) \right]^2$$

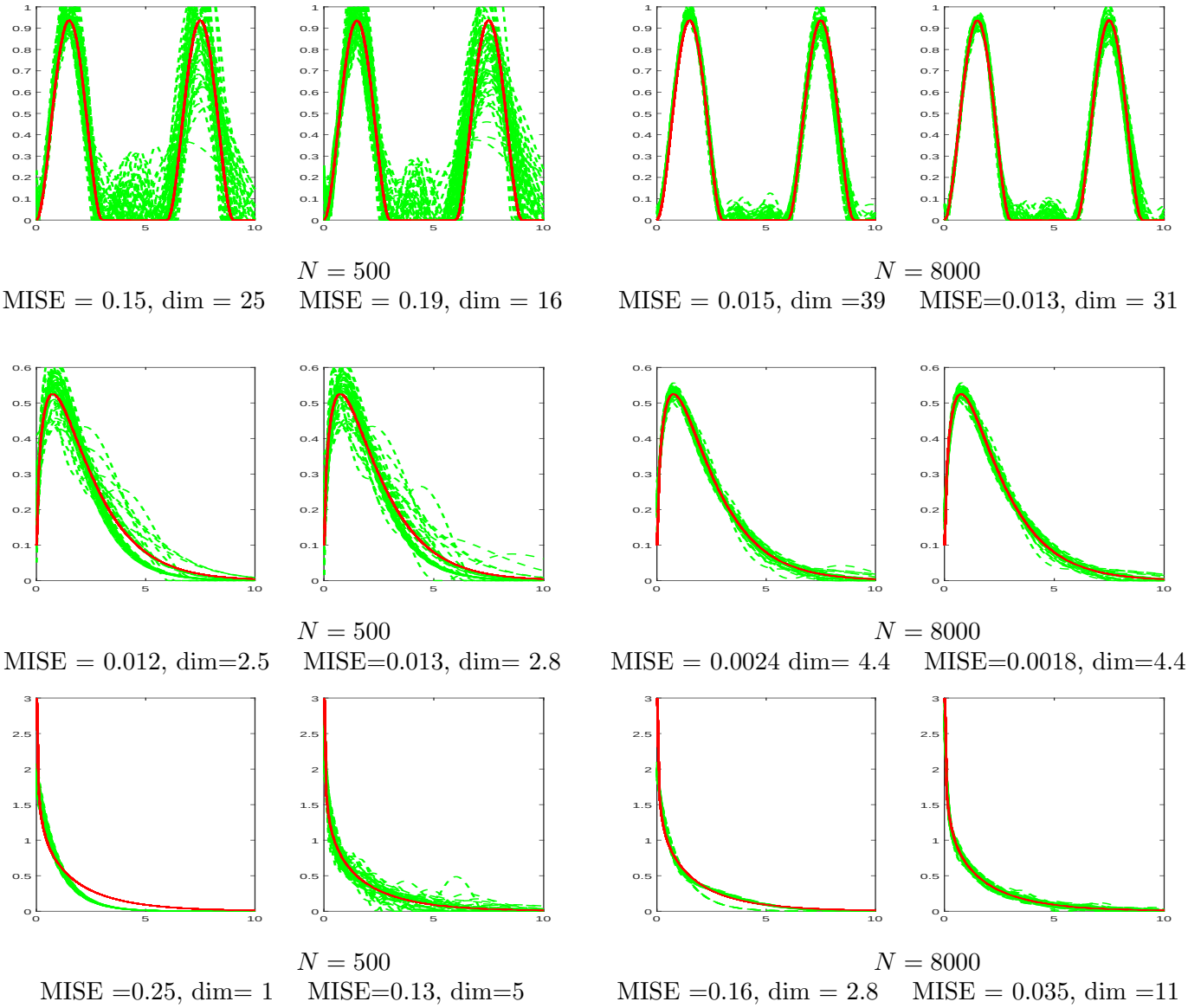

FIGURE 1. Examples of 50 estimated curves (in gray-green) using the Laguerre basis, with the two methods (method 1 left, method 2 right, for each couple of plots) for functions 3 (first line), function 6 (second line) and function 7 (third line), for $N = 500$ (left plots) and $N = 8000$ (right plots). The bold curve is the true function. Under each plot, the MISE over the 50 paths, and the mean of the selected dimensions.

Next, we implement the estimators of Therem 1. We compute $(\tilde{g}_m^{(Trig)})_{1\leq m\leq D_{\max}}$ and $(\tilde{g}_m^{(Lag)})_{1\leq m\leq D_{\max}}$ the collection of estimators in trigonometric and Laguerre basis, with coefficients given by (14). We select

$$\tilde{m}^{(Trig)} = \arg\min_m \left\{ -\|\tilde{g}_m^{(Trig)}\|^2 + \kappa_2^{(Trig)} \widehat{G^2(T)} \log(N) \frac{m^2}{NT} \right\}$$

		$(N,T)=(500,10)$				$(N,T)=(2000,10)$				$(N,T)=(8000,10)$			
a		$\hat{g}_{\hat{m},L}$	$\hat{g}_{\hat{m},T}$	$\tilde{g}_{\tilde{m},L}$	$\tilde{g}_{\tilde{m},T}$	$\hat{g}_{\hat{m},L}$	$\hat{g}_{\hat{m},T}$	$\tilde{g}_{\tilde{m},L}$	$\tilde{g}_{\tilde{m},T}$	$\hat{g}_{\hat{m},L}$	$\hat{g}_{\hat{m},T}$	$\tilde{g}_{\tilde{m},L}$	$\tilde{g}_{\tilde{m},T}$
a_1	MISE	0.07	15.9	1.90	14.7	0.02	9.11	0.28	9.71	0.01	6.19	0.05	8.47
	(std)	(0.1)	(3.5)	(2.4)	(2.2)	(.03)	(0.9)	(0.5)	(0.9)	(.01)	(0.3)	(0.1)	(0.7)
	Or	0.07	12.5	0.07	12.3	0.02	8.21	0.02	8.38	0.01	6.06	0.01	6.92
	dim	1.00	16.4	7.31	11.3	1.00	28.5	3.68	17.4	1.00	41.5	2.63	40.8
	dim Or	1.06	16.1	1.03	16.5	1.04	30.2	1.04	25.7	1.04	42.7	1.03	19.1
a_2	MISE	1.44	1.54	2.02	1.00	0.43	0.44	1.04	0.39	0.12	0.11	0.35	0.19
	(std)	(1.2)	(1.3)	(1.1)	(1.0)	(0.3)	(0.3)	(0.5)	(0.3)	(.09)	(.09)	(.07)	(.07)
	Or	1.24	0.90	1.24	0.90	0.40	0.31	0.39	0.31	0.11	0.09	0.12	0.10
	dim	11.2	10.8	5.08	2.03	14.2	11.4	5.84	2.02	18.3	11.8	7.16	2.02
	dim Or	11.9	2.99	12.0	2.79	16.3	4.17	15.8	3.48	20.5	4.89	16.5	4.08
a_3	MISE	14.5	17.3	20.7	15.1	4.26	4.81	6.93	5.48	1.40	1.47	1.25	3.30
	(std)	(2.5)	(2.6)	(4.1)	(2.3)	(0.8)	(0.8)	(0.8)	(0.9)	(0.2)	(0.2)	(0.1)	(0.2)
	Or	10.3	11.5	9.84	11.4	3.59	4.09	3.12	3.96	1.26	1.38	0.86	1.42
	dim	25.2	20.8	14.9	6.86	32.8	21.1	21.6	7.07	38.7	23.0	31.0	7.45
	dim Or	28.6	8.91	28.4	8.64	36.0	12.9	35.5	10.7	39.3	17.7	38.5	11.5
a_4	MISE	7.46	19.6	4.72	21.06	2.26	5.08	1.69	9.06	0.81	1.49	0.58	3.30
	(std)	(6.2)	(6.8)	(3.0)	(6.3)	(1.2)	(1.6)	(0.6)	(2.8)	(0.3)	(0.5)	(0.2)	(0.8)
	Or	2.77	13.4	2.56	13.21	1.20	4.47	1.00	4.78	0.54	1.45	0.34	2.24
	dim	14.5	29.8	7.16	7.60	15.3	30.7	8.77	10.4	18.9	33.5	11.7	13.7
	dim Or	9.07	13.0	8.96	13.1	13.1	20.8	12.9	16.4	19.7	32.9	18.8	17.8
a_5	MISE	5.26	6.80	2.92	6.99	1.64	2.09	0.95	2.45	0.43	0.55	0.17	1.57
	(std)	(3.9)	(4.1)	(2.4)	(4.3)	(1.1)	(1.2)	(0.9)	(1.0)	(0.3)	(0.3)	(0.2)	(0.3)
	Or	1.53	4.76	1.51	4.70	0.50	1.84	0.47	1.86	0.15	0.53	0.13	0.69
	dim	10.5	16.6	3.41	3.70	11.4	15.7	3.80	4.00	10.6	16.6	4.14	4.13
	dim Or	4.01	6.01	3.99	5.77	4.27	9.06	4.27	7.37	4.62	13.8	4.56	8.32
a_6	MISE	1.18	9.03	1.33	6.42	0.58	2.85	0.51	3.13	0.23	0.95	0.17	1.49
	(std)	(0.9)	(3.0)	(1.2)	(2.1)	(0.2)	(0.8)	(0.2)	(0.8)	(.09)	(0.2)	(.08)	(0.2)
	Or	0.67	4.42	0.63	4.35	0.32	1.95	0.28	2.01	0.15	0.84	0.11	1.16
	dim	2.46	30.4	3.03	5.42	3.13	34.1	3.30	7.53	4.23	38.1	4.23	11.5
	dim Or	3.66	10.5	3.60	10.1	5.62	17.4	5.37	14.0	8.69	29.2	8.19	16.1
a_7	MISE	24.8	115	12.7	101	24.5	80.1	7.27	82.0	14.9	59.9	3.37	73.1
	(std)	(1.6)	(13)	(4.3)	(7.7)	(0.8)	(7.2)	(2.5)	(4.7)	(1.9)	(2.3)	(1.2)	(4.5)
	Or	8.90	87.0	7.15	85.9	4.48	68.6	3.18	71.3	2.07	57.4	1.35	64.9
	dim	1.00	5.16	4.94	8.11	1.00	14.8	6.73	13.7	2.93	36.2	10.9	40.0
	dim Or	6.61	5.29	4.68	5.11	9.29	6.32	6.07	4.59	37.9	44.0	25.3	30.2
a_8	MISE	1.64	20.7	2.55	19.8	1.53	11.0	0.64	12.2	0.53	7.01	0.18	9.57
	(std)	(0.2)	(5.9)	(2.7)	(3.6)	(.05)	(1.5)	(0.6)	(1.4)	(0.2)	(0.4)	(0.2)	(1.7)
	Or	0.68	15.9	0.67	15.7	0.29	9.76	0.28	10.0	0.11	6.78	0.10	7.93
	dim	1.00	15.7	4.87	9.50	1.00	25.7	4.23	14.0	1.96	38.4	5.24	30.7
	dim Or	3.14	15.2	3.18	15.1	4.39	25.4	4.39	21.7	5.36	40.8	5.29	25.4

TABLE 2. Simulation results, 100 MISE one lines MISE, 100 std on lines (std), MISE of the oracle (best unknown choice, using the true), dim = mean of the selected dimensions, dim Or = mean of dimensions associated to the oracle, 200 repetitions.

and

$$\tilde{m}^{(Lag)} = \arg\min_m \left\{ -\|\tilde{g}_m^{(Lag)}\|^2 + \kappa_2^{(Lag)} \widehat{G^2(T)} \log(T) \log(N) \frac{m}{N} \right\}$$

where $\widehat{G^2(T)}$ is defined by (22).

We do not present results using the procedure of Theorem 2, as the method seemed not stable.

Based on preliminary simulations, the constants are calibrated once and for all to the following values $\kappa_1^{(Lag)} = 27$, $\kappa_1^{(Trig)} = 6$, $\kappa_2^{(Lag)} = 0.11$ and $\kappa_2^{(Trig)} = 0.6$.

Table 2 presents the values of the risks of the adaptive estimators computed for the eight functions, following the two methods (method 1: estimators $\hat{g}$, method 2: estimators $\tilde{g}$) and using two bases, Laguerre (index L) and trigonometric (index T). For each function, the first line gives the MISE multipled by 100, over 200 repetitions, with standard deviation multiplied by 100 in parenthesis on the line below. The line "Or" gives the mean of path-by-path minimal integrated error (computed using the true function). The fourth line provides the mean of selected dimensions, and "dim or" the mean of the dimensions associated to the oracle estimators. We can compare lines 1 (MISE)and 3 (Or), and lines 4 (dim) and 5 (dim Or), where MISE and dim should be as close as possible to Or and dim Or.

Naturally, the risk decreases as N increases. Globally, the Laguerre basis performs satisfactorily, and better than the trigonometric one, except for function a_2. Note that the methods are easy to implement and the computation time is quite fast.

To conclude this section, we provide in Figure 1 plots illustrating the behaviour of our estimators following the two strategies in the Laguerre basis, for three of the functions of the list, namely the mixed-beta function 3) and two functions of type $t^d e^{-t/b}$ with $d = 0.25, b = 3$ (function 6)) and $d = -0.25, b = 5$ (function 7)). Each couple of plots corresponds to the representation of 50 estimators computed by the two methods, together with the true function in bold (red). Two values of N are compared, and the MISE are given, to make the orders of Table 2 concrete; the improvement from $N = 500$ to $N = 8000$ is obvious in most cases. We note that the first method seems to still work for function 6), contrary to what was expected from the theory. But it fails for function 7), as expected : the estimator is biased. Method 2 always gives good results.

6. Concluding remarks

In this paper, we consider *i.i.d.* continuous observations of the processes $(X_i(t), t \in [0,T]), i = 1, \ldots, N)$ distributed as the CMA process (1). We build collections of nonparametric estimators of the unknown function $g = a^2$ by projection method on finite-dimensional subspaces of $\mathbb{L}^2(\mathbb{R}^+)$. The subspaces are generated by the trigonometric basis of $\mathbb{L}^2([0,T])$ or the Laguerre basis of $\mathbb{L}^2(\mathbb{R}^+)$. After proving various risk bounds for each estimator, we propose a data-driven selection of the dimension of the projection space and prove that it leads to an adaptive estimator. Our methods are implemented on simulated data and show convincing results in terms of risks and plots with a better performance for the estimators in Laguerre basis.

The consistency of the estimators is ensured for fixed T as N tends to infinity (case of trigonometric basis) or when both $T$and N tend to infinity (case of Laguerre basis) but with T/N not too large. It would be interesting to clarify this point which has an impact on the risk bounds as we noticed on Monte-Carlo simulations.

Our proofs rely on the Gaussian character of (1) especially for the adaptive procedure. The generalization to other processes than the Wiener process in (1) is of interest and left to further work. Clearly, the results could be obtained with more general deviation inequalities, rather than the χ-square deviations specifically used here.

The question of taking into account, from the theoretical point of view, the discretization step used in practice may also be worth investigation. Lastly, there may be some developments about optimality, but the meaning of this in our context would have first to be carefully defined.

7. Proofs

7.1. **Proofs of section 2. Proof of Lemma 1** Using the stochastic differential (6), we write:

$$\int_0^{+\infty} \varphi_j(s)X(s)dX(s) = a(0)\int_0^{+\infty} \varphi_j(s)X(s)dW(s) + \int_0^{+\infty} \varphi_j(s)X(s)\int_0^s a'(s-u)dW(u)ds$$

Therefore, as $\mathbb{E}\int_0^{+\infty} \varphi_j^2(s)X_i^2(s)ds = \int_0^{+\infty} \varphi_j^2(s)G(s)ds \leq \|a\|^2 < +\infty$,

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

which gives the result. □

Proof of Proposition 1. We consider that $(\varphi_j) = (\varphi_{j,T})$ is the trigonometric basis on $[0,T]$. In this case, $\mathbb{E}\hat{\theta}_j = \theta_j$, we can write $\mathbb{E}\|\hat{g}_m - g\|_T^2 = \mathbb{E}\|\hat{g}_m - \mathbb{E}\hat{g}_m\|^2 + \|g_m - g\|_T^2$. We have, setting $X = X_1$,

(24)
$$\mathbb{E}\|\hat{g}_m - \mathbb{E}\hat{g}_m\|^2 = \frac{1}{N}\sum_{j=0}^{m-1} \mathrm{Var}\left(2\int_0^T \varphi_j(s)X(s)dX(s)\right) \leq \frac{1}{N}\sum_{j=0}^{m-1} \mathbb{E}\left(2\int_0^T \varphi_j(s)X(s)dX(s)\right)^2.$$

(Note that for functions on $S_{m,T}$, the norms $\|.\|_T$ and $\|.\|$ are identical). We have:

(25)
$$\mathbb{E}\left(\int_0^T \varphi_j(s)X(s)dX(s)\right)^2 \leq 2g(0)\int_0^T \varphi_j^2(s)\mathbb{E}(X^2(s))ds + 2\mathbb{E}\left[\left(\int_0^T \varphi_j(s)X(s)Y(s)ds\right)^2\right]$$

where $Y(s) = \int_0^s a'(s-u)dW(u)$. We have $\mathbb{E}(X^2(s)) = G(s) \leq G(T) \leq \|a\|^2$. Now, using that $(\varphi_j) = (\varphi_{j,T})$ is an orthonormal basis of $\mathbb{L}^2([0,T])$

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

As $(X(s), Y(s))$ is a Gaussian centered vector, we know that:

$$\mathbb{E}(X^2(s)Y^2(s)) = 2\left[\mathbb{E}(X(s)Y(s))\right]^2 + \mathbb{E}(X^2(s))\mathbb{E}(Y^2(s)) = \frac{1}{4}\left(g(s)-g(0)\right)^2 + G(s)G_1(s), \tag{26}$$

with $G_1(s) = \int_0^s (a'(u))^2du \leq \|a'\|^2$. Therefore, if $g(0) \neq 0$

$$\mathbb{E}\|\hat{g}_m - \mathbb{E}\hat{g}_m\|^2 \leq \frac{4}{N}\left(2g(0)G(T)m + 2TG(T)G_1(T) + Tg^2(0) + \|g\|_T^2\right).$$

Therefore, we obtain (9).

If $g(0) = 0$, (25) becomes

$$\mathbb{E}\left(\int_0^T \varphi_j(s)X(s)dX(s)\right)^2 = \mathbb{E}\left[\left(\int_0^T \varphi_j(s)X(s)Y(s)ds\right)^2\right]$$

and thus

$$\mathbb{E}\|\hat{g}_m - \mathbb{E}\hat{g}_m\|^2 \leq 4\frac{TG(T)G_1(T)}{N} + \frac{\|g\|_T^2}{N},$$

which gives inequality (10). □

Proof of Proposition 2. Now, we look at the case of a basis of $\mathbb{L}^2(\mathbb{R}^+)$. The estimator $\hat{\theta}_j$ is no more unbiased. We write $\hat{g}_m - g = \hat{g}_m - \mathbb{E}\hat{g}_m + \mathbb{E}\hat{g}_m - g_m + g_m - g$ and

$$\mathbb{E}(\|\hat{g}_m - g\|^2) = \|g_m - g\|^2 + \mathbb{E}\|\hat{g}_m - \mathbb{E}\hat{g}_m\|^2 + \|\mathbb{E}\hat{g}_m - g_m\|^2. \tag{27}$$

The first term is the usual square bias term. 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:

$$\|\mathbb{E}\hat{g}_m - g_m\|^2 = \sum_{j=0}^{m-1}(\mathbb{E}\hat{\theta}_j - \theta_j)^2 = \sum_{j=0}^{m-1}\left(\int_T^{+\infty} g(s)\varphi_j(s)ds\right)^2 \leq \int_T^{+\infty} g^2(s)ds, \tag{28}$$

and we obtain inequalities (11) and (12).

Now, consider the Laguerre basis. To get (13), we bound differently (25). We write:

$$\begin{aligned}\left(\mathbb{E}\int_0^T \varphi_j(s)X(s)Y(s)ds\right)^2 &= \int_{[0,T]^2} \varphi_j(s)\varphi_j(u)\mathbb{E}[X(s)Y(s)X(u)Y(u)]dsdu \\ &\leq \int_{[0,T]^2} |\varphi_j(s)\varphi_j(u)| \left\{\mathbb{E}[(X(s)Y(s))^2]\mathbb{E}[(X(u)Y(u))^2]\right\}^{1/2} dsdu \\ &= \left(\int_0^T |\varphi_j(s)| \left\{\mathbb{E}[(X(s)Y(s))^2)]\right\}^{1/2} ds\right)^2. \end{aligned} \tag{29}$$

By (26) and the assumptions, $\mathbb{E}(X(s)Y(s))^2 \leq \frac{1}{2}(g^2(0) + \|g\|_\infty^2) + \|a\|^2\|a'\|^2$ is bounded. Therefore, we need to bound $\int_0^T |\varphi_j(s)|ds$. For this, we split each integral according to the inequalities of Askey and Wainger (1965) recalled in Section 8 (we assume without loss of generality that they hold for all j). We have:

$$\int_0^T |\varphi_j(s)|ds = \int_0^{2T} |\varphi_j(x/2)|dx/2 = \frac{1}{2}(\sum_{\ell=1}^{6} I_\ell)$$

and bound each term. Setting $\nu_j = 4j + 2$,

$$I_1 = \int_0^{1/\nu_j} dx = \nu_j^{-1}, \quad I_2 = \int_{1/\nu_j}^{\nu_j/2} (x\nu_j)^{-1/4}dx \leq \frac{2^{5/4}}{3}\nu_j^{1/2}, \quad I_3 = \int_{\nu_j/2}^{\nu_j - \nu_j^{1/3}} \nu_j^{-1/4}(\nu_j - x)^{-1/4}dx \leq 3/4,$$

$$I_4 = \int_{\nu_j - \nu_j^{1/3}}^{\nu_j + \nu_j^{1/3}} \nu_j^{-1/3}dx = 2, \quad I_5 = \int_{\nu_j + \nu_j^{1/3}}^{3\nu_j/2} \frac{exp(-\gamma_1\nu_j^{-1/2}(x-\nu_j)^{3/2})}{\nu_j^{1/4}(x-\nu_j)^{1/4}}dx \leq \nu_j^{1/6}\frac{\exp(-\gamma_1)}{\gamma_1},$$

$$I_6 = \int_{3\nu_j/2}^T \exp(-\gamma_2 x)dx \leq \frac{\exp(-3\gamma_2\nu_j/2)}{\gamma_2}.$$

Consequently, for $j = 0, \ldots, m-1$ and $T \geq 6(m-1) + 3 = 6m - 3$,

$$\int_0^T |\varphi_j(s)|ds \lesssim j^{1/2}. \tag{30}$$

Finally,

$$\sum_{j=0}^{m-1}\left(\int_0^T |\varphi_j(s)|ds\right)^2 \lesssim m^2 \tag{31}$$

Using again the inequalities of Askey and Wainger (1965) (see Section 8), for $T \geq 6m-3$, for all $j \in \{0,1,\ldots,m-1\}$, $|\varphi_j(x/2)| \leq \exp(-\gamma_2 x)$ and

$$\left|\int_T^{+\infty} \varphi_j(s)g(s)ds\right| \leq \int_T^{+\infty} |\varphi_j(s)|g(s)ds \leq \exp(-2\gamma_2 T)\|a\|^2 \leq \exp(-2\gamma_2(6(m-1)+3))\|a\|^2.$$

The additional bias term (28) is therefore bounded as follows:

$$\sum_{j=0}^{m-1}[\int_T^{+\infty} \varphi_j(s)g(s)ds]^2 \lesssim \|a\|^2 m \ \exp(-12\gamma_2 m). \tag{32}$$

We thus obtain (13) by joining (29), (31) and (32). □

Proof of Lemma 2. We have

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

which is the result. □

Proof of Proposition 3. Assume that $(\varphi_j = \varphi_{j,T})$ is the trigonometric basis. Then, $\tilde{\theta}_j$ is an unbiased estimator of θ_j. We only need to study the variance term of the risk.

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

where $\mathbb{E}X^4(T) = 3\left(\int_0^T a^2(s)ds\right)^2 = 3G^2(T)$ and $\sum_{j=0}^{m-1}\varphi_j^2(T) = m/T$.
We have

$$\varphi_{0,T}'(s) = 0, \ \ \varphi_{2j,T}'(s) = (2\pi j/T)\varphi_{2j-1,T}(s), \ \ \varphi_{2j-1,T}'(s) = -(2\pi j/T)\varphi_{2j,T}(s), \ \ j \geq 1. \tag{33}$$

Proceeding as in Proposition 1 (projection argument), we obtain

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

using that $\mathbb{E}X^4(s) = 3\left(\int_0^s a^2(s)ds\right)^2 \leq 3G^2(T)$. This gives (15).

Now, we look at the case of the Laguerre basis on $\mathbb{L}^2(\mathbb{R}^+)$. We start as above by

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

We get

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

Using that $\mathbb{E}X^4(T) = 3\left(\int_0^T a^2(s)ds\right)^2 = 3G^2(T)$ and $|\ell_j| \leq \sqrt{2}$, we get

$$\mathbb{T}_2 \leq 12G^2(T)\frac{m}{N}.$$

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 8) and we find

$$\begin{aligned}\mathbb{T}_1 &\leq \frac{4}{N}\sum_{j=0}^{m-1}\mathbb{E}\left[\left(\int_0^T \ell_j(s)X_1^2(s)ds\right)^2\right] + \frac{4}{N}\sum_{j=1}^{m-1}\mathbb{E}\left[\left(\int_0^T \ell_{j-1}^{(1)}(s)\sqrt{\frac{2j}{s}}X_1^2(s)ds\right)^2\right] \\ &\leq \frac{4}{N}\mathbb{E}\left(\int_0^T X_1^4(s)ds\right) + \frac{8m}{N}\mathbb{E}\left(\int_0^T \frac{X_1^4(s)}{s}ds\right) = 12TG^2(T)\frac{1}{N} + 24\frac{m}{N}\int_0^T \frac{G^2(s)}{s}ds.\end{aligned}$$

Under **[H2]**, we obtain

$$\mathbb{T}_1 \leq 12TG^2(T)\frac{1}{N} + 24(c_0 + \log(T)G^2(T))\frac{m}{N}.$$

Finally, the variance term is bounded by

$$\mathbb{E}\|\tilde{g}_m - \mathbb{E}\tilde{g}_m\|^2 \leq 12\left(G^2(T)\left(2\log(T)+1\right) + 2c_0\right)\frac{m}{N} + 12G^2(T)\frac{T}{N}.$$

If **[H2]** does not hold and $T \geq 6m-3$, we can bound differently the variance and bias terms.

$$\begin{aligned}\sum_{j=0}^{m-1}\mathbb{E}\left(\int_0^T \ell_j'(s)X^2(s)ds\right)^2 &= \int_{[0,T]^2}\sum_{j=0}^{m-1}\ell_j'(s)\ell_j'(u)\mathbb{E}[X^2(s)X^2(u)]ds\,du \\ &\leq \int_{[0,T]^2}\left[\sum_{j=0}^{m-1}(\ell_j'(s))^2\mathbb{E}(X^4(s))\sum_{j=0}^{m-1}(\ell_j'(u))^2\mathbb{E}(X^4(u))\right]^{1/2} dsdu \\ &= \left(\int_0^T\left[\sum_{j=0}^{m-1}(\ell_j'(s))^2\mathbb{E}(X^4(s))\right]^{1/2} ds\right)^2 \\ &\leq 3G^2(T)\left(\int_0^T\left[\sum_{j=0}^{m-1}(\ell_j'(s))^2\right]^{1/2} ds\right)^2\end{aligned}$$

We decompose the integral to obtain

$$\left(\int_0^T\left[\sum_{j=0}^{m-1}(\ell_j'(s))^2\right]^{1/2} ds\right)^2 \leq 2\left(\int_0^{6m-3}\dots\right)^2 + 2\left(\int_{6m-3}^T\dots\right)^2,$$

and bound each term. Using (46) and again the inequalities of Askey and Wanger (1965)(see Section 8), we get that, for $s \geq 6m-3$, $|\varphi_j'(s)| \leq 2\sum_{k=0}^j|\ell_j(s)| \leq 2(j+1)\exp(-\gamma_2 s)$. Thus, $\sum_{j=0}^{m-1}(\ell_j'(s))^2 \leq 4m^3\exp(-2\gamma_2 s)$. So,

$$\left(\int_{6m-3}^T\left[\sum_{j=0}^{m-1}(\ell_j'(s)^2\right]^{1/2} ds\right)^2 \leq \frac{4m^3}{\gamma_2^2}\exp(-(12m-6)\gamma_2). \tag{34}$$

Now,
(35)
$$\left(\int_0^{6m-3}\left[\sum_{j=0}^{m-1}(\ell_j'(s)^2\right]^{1/2}ds\right)^2 \leq (6m-3)\int_0^{+\infty}\sum_{j=0}^{m-1}(\ell_j'(s))^2\,ds = (6m-3)\sum_{j=0}^{m-1}(1+4j) \leq 12m^3.$$

Finally, we get

$$\mathbb{E}\|\tilde{g}_m - \mathbb{E}\tilde{g}_m\|^2 \leq \frac{3}{N}G^2(T)\left(12m^3 + \frac{4m^3}{\gamma_2^2}\exp\left(-(12m-6)\gamma_2\right)\right) \tag{36}$$

So, we have the two variance bounds.
For the bias term, we have $\mathbb{E}\tilde{\theta}_j = \theta_j - \ell_j(T)G(T) - \int_T^{+\infty}\ell_j(s)g(s)ds$. We have $G(T) \leq G(+\infty) = \|a\|^2$. Moreover, inequality (32) still holds. Joining variance and bias terms, we obtain (16) and (17). □

7.2. **Proof of Theorem 1.** Let us state a preliminary Lemma:

Lemma 3. *Let $V_N = \sum_{i=1}^N (X_i^2 - 1)$ where X_i are i.i.d. standard Gaussian variables. Then for all $\varepsilon \in (0,1]$,*

$$\mathbb{P}(|V_N| \geq N\varepsilon) \leq 2\exp\left(-\frac{N\varepsilon^2}{8}\right).$$

Proof of Lemma 3. By Lemma 1 and Inequalities (4.3)-(4.4) in Laurent and Massart (2000), we have, for any $u > 0$,

$$\mathbb{P}(|V_N| \geq 2\sqrt{Nu} + 2u) \leq 2\exp(-u).$$

Thus, setting $u = Nx$, we have, for any $x > 0$, $\mathbb{P}(|V_N| \geq 2N\sqrt{x} + 2Nx) \leq 2\exp(-Nx)$. Now we set $N\varepsilon = 2N(x + \sqrt{x})$ and using Birgé and Massart (1998), Lemma 8, Inequality (7.14) with $v = \sqrt{2}$ and $c = 2$, we find

$$\mathbb{P}(|V_N| \geq N\varepsilon) \leq 2\exp\left(-\frac{N\varepsilon^2}{4(1+\varepsilon)}\right),$$

and the result follows.□

7.2.1. *Case of Laguerre basis.* Note that, as $G(0) = 0$ and $h(+\infty) = 0$, $\langle h, g\rangle = -\langle h', G\rangle$. Therefore,

$$\gamma_{N,T}(h) = \|h\|^2 - 2\langle h, g\rangle - 2\nu_{N,T}(h) + 2R_T(h)$$

where $\nu_{N,T}(h) = \nu_{N,1}(h) + \nu_{N,2}(h)$,

$$\nu_{N,1}(h) = -\frac{1}{N}\sum_{i=1}^N\int_0^T h'(u)[X_i^2(u) - G(u)]du, \quad \nu_{N,2}(h) = \frac{1}{N}\sum_{i=1}^N h(T)(X_i^2(T) - G(T)), \tag{37}$$

and

$$R_T(h) = \int_T^{+\infty} h(u)g(u)du.$$

Therefore,

$$\gamma_{N,T}(h_1) - \gamma_{N,T}(h_2) = \|h_1 - g\|^2 - \|h_2 - g\|^2 - 2\nu_{N,T}(h_1 - h_2) + 2R_T(h_1 - h_2). \tag{38}$$

Using the definition of $\tilde{m} = \tilde{m}^{(Lag)}$, we have for all $g_m \in S_m$,

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

where for simplicity $\text{pen} = \text{pen}^{(Lag)}$. We deduce

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

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

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

We have $\sup_{h \in B_m} R_T^2(h) \leq \int_T^{+\infty} g^2(u)du$ so that

$$\mathbb{E}\left(\sup_{h \in B_{\widetilde{m}\vee m}} R_T^2(h)\right) \leq \int_T^{+\infty} g^2(u)du.$$

Gathering terms yields

$$\begin{aligned} \frac{1}{2}\|\tilde{g}_{\widetilde{m}} - g\|^2 &\leq \frac{3}{2}\|g_m - g\|^2 + \text{pen}(m) + 8\int_T^{+\infty} g^2(u)du \\ &\quad +8(\sup_{h \in B_{\widetilde{m}\vee m}} \nu_{N,T}^2(h) - p^{(Lag)}(m, \widetilde{m})) + 8p^{(Lag)}(m, \widetilde{m}) - \text{pen}(\widetilde{m}). \end{aligned}$$

where $p^{(Lag)}(m, m') = p_1^{(Lag)}(m, m') + p_2^{(Lag)}(m, m')$,

$$p_1^{(Lag)}(m, m') = 128\log(N)\int_0^T \frac{G^2(u)}{u}du\frac{m \vee m'}{N}, \quad p_2^{(Lag)}(m, m') = 32G^2(T)\log(N)\frac{m \vee m'}{N}.$$

Now we use that $8p^{(Lag)}(m, m') \leq \text{pen}(m) + \text{pen}(m')$ for $\kappa \geq \kappa_0^{(Lag)} = 8 \times 128$, and the result of the following Lemma:

Lemma 4. *Under the Assumptions of Theorem 1, for $\ell = 1, 2$,*

$$\mathbb{E}\left(\sup_{h \in B_{\widetilde{m}\vee m}} \nu_{N,\ell}^2(h) - p_\ell^{(Lag)}(m, \widetilde{m})\right)_+ \leq C_\ell^{(Lag)}(T, N),$$

where

$$C_1^{(Lag)}(T, N) = \frac{C}{N}\left(\frac{TG^2(T)}{N} + \int_0^T \frac{G^2(u)}{u}du\right) + \frac{C\log N}{N}TG^2(T), \quad C_2^{(Lag)}(T, N) = \frac{C}{N}G^2(T)$$

and C is a positive numerical constant.

And we obtain

$$\mathbb{E}(\|\tilde{g}_{\widetilde{m}} - g\|^2) \leq 3\|g_m - g\|^2 + 4\text{pen}(m) + 16\int_T^{+\infty} g^2(u)du + 32(C_1^{(Lag)}(T, N) + C_2^{(Lag)}(T, N)),$$

which ends the proof of Theorem 1 in the Laguerre case. □

Proof of Lemma 4. Let us define

$$Z_N(u) = \frac{1}{N}\sum_{i=1}^N \left(\frac{X_i^2(u)}{G(u)} - 1\right),$$

which is for all u distributed as $(\chi^2(N)-N)/N$, and set

$$A_N(u)=\left\{Z_N^2(u)\leq 16\frac{\log(N)}{N}\right\}.$$

By Lemma 3, $\mathbb{P}(A_N(u)^c)\leq 2N^{-2}$ provided that $16\log(N)/N\leq 1$ i.e. $N\geq 68$.

Now we can write $\nu_{N,1}(h)=-\int_0^T G(u)h'(u)Z_N(u)du$ and split it

$$\begin{aligned}\nu_{N,1}(h) &= -\int_0^T G(u)h'(u)Z_N(u)\mathbf{1}_{A_N(u)}du-\int_0^T G(u)h'(u)Z_N(u)\mathbf{1}_{A_N(u)^c}du \\ &:= \nu_{N,1,1}(h)+\nu_{N,1,2}(h).\end{aligned} \tag{39}$$

Then

$$\left(\sup_{h\in B_{\widetilde{m}\vee m}}\nu_{N,1}^2(h)-p_1^{(Lag)}(m,\widetilde{m})\right)_+\leq\left(2\sup_{h\in B_{\widetilde{m}\vee m}}\nu_{N,1,1}^2(h)-p_1^{(Lag)}(m,\widetilde{m})\right)_++2\sup_{h\in B_{\widetilde{m}\vee m}}\nu_{N,1,2}^2(h).$$

With $B(u):=G(u)Z_N(u)\mathbf{1}_{A_N(u)}$, and by using Formula (45), we have

$$\begin{aligned}\sup_{h\in B_{\widetilde{m}\vee m}}\nu_{N,1,1}^2(h) &\leq \sum_{j=0}^{\widetilde{m}\vee m-1}\left(\int_0^T B(u)\ell_j'(u)du\right)^2 \\ &\leq 2\sum_{j=0}^{\widetilde{m}\vee m-1}\left(\int_0^T B(u)\ell_j(u)du\right)^2+2\sum_{j=1}^{\widetilde{m}\vee m-1}(2j)\left(\int_0^T\frac{B(u)}{\sqrt{u}}\ell_{j-1}^{(1)}(u)du\right)^2 \\ &\leq 2\int_0^T B^2(u)du+4(\widetilde{m}\vee m)\int_0^T\frac{B^2(u)}{u}du.\end{aligned}$$

Now, using the definition of $A_N(u)$,

$$\sup_{h\in B_{\widetilde{m}\vee m}}\nu_{N,1,1}^2(h)\leq 32G^2(T)\frac{T\log(N)}{N}+64\int_0^T\frac{G^2(u)}{u}du\frac{(\widetilde{m}\vee m)\log(N)}{N}.$$

As a consequence

$$\mathbb{E}\left[\left(2\sup_{h\in B_{\widetilde{m}\vee m}}\nu_{N,1,1}^2(h)-p_1^{(Lag)}(m,\widetilde{m})\right)_+\right]\leq 64G^2(T)\frac{T\log(N)}{N}. \tag{40}$$

Similarly, for $C(u):=G(u)Z_N(u)\mathbf{1}_{A_N(u)^c}$, we have

$$\sup_{h\in B_{\widetilde{m}\vee m}}\nu_{N,1,2}^2(h)\leq 2\int_0^T C^2(u)du+4N\int_0^T\frac{C^2(u)}{u}du.$$

Now, by the Rosenthal Inequality (see Hall and Heyde (1980)), $\mathbb{E}(Z_N^4(u))\lesssim N^{-2}$ and thus

$$\mathbb{E}[Z_N^2(u)\mathbf{1}_{A_N(u)^c}]\leq\mathbb{E}^{1/2}[Z_N^4(u)]\mathbb{P}^{1/2}(A_N(u)^c)\lesssim N^{-2}. \tag{41}$$

As a consequence,

$$\mathbb{E}\left(2\sup_{h\in B_{\widetilde{m}\vee m}}\nu_{N,1,2}^2(h)\right)\leq C(\frac{TG^2(T)}{N^2}+\int_0^T\frac{G^2(u)}{u}du\frac{1}{N}). \tag{42}$$

Gathering (40) and (42) implies the result of Lemma 4 for $\ell = 1$ and $(B) = (Lag)$. Now we look at $\nu_{N,2}(h)$ and write

$$\begin{aligned}\sup_{h\in B_{\widetilde{m}\vee m}} \nu_{N,2}^2(h) &= \left(\sup_{h\in B_{\widetilde{m}\vee m}} h^2(T)\right) G^2(T)Z_N^2(T) \leq \sum_{j=0}^{\widetilde{m}\vee m-1} \varphi_j^2(T)G^2(T)Z_N^2(T)\\ &\leq 2(\widetilde{m}\vee m)G^2(T)Z_N^2(T).\end{aligned}$$

Therefore

$$\sup_{h\in B_{\widetilde{m}\vee m}} \nu_{N,2}^2(h)\mathbf{1}_{A_N(T)} \leq 32(\widetilde{m}\vee m)G^2(T)\frac{\log(N)}{N}$$

and using (41),

$$\mathbb{E}\left(\sup_{h\in B_{\widetilde{m}\vee m}} \nu_{N,2}^2(h)\mathbf{1}_{A_N(T)^c}\right) \leq 2NG^2(T)\mathbb{E}(Z_N^2(T)\mathbf{1}_{A_N(T)^c}) \lesssim \frac{G^2(T)}{N}.$$

Finally

$$\begin{aligned}\mathbb{E}\left[(\sup_{h\in B_{\widetilde{m}\vee m}} \nu_{N,2}^2(h) - p_2^{(Lag)}(m,\widetilde{m}))_+\right] &\leq \mathbb{E}\left[(\sup_{h\in B_{\widetilde{m}\vee m}} \nu_{N,2}^2(h)\mathbf{1}_{A_N(T)} - p_2^{(Lag)}(m,\widetilde{m}))_+\right]\\ &\quad +\mathbb{E}\left(\sup_{h\in B_{\widetilde{m}\vee m}} \nu_{N,2}^2(h)\mathbf{1}_{A_N(T)^c}\right)\\ &\leq C\frac{G^2(T)}{N},\end{aligned}$$

where C is a numerical constant. Thus, we obtain Lemma 4 for $\ell = 2$ and $(B) = (Lag)$.

7.2.2. *Case of Trigonometric basis.* We proceed analogously. As now $R_T(h) = 0$, we have, with for simplicity, $\widetilde{m} = \widetilde{m}^{(Trig)}$,

$$\begin{aligned}\|\tilde{g}_{\widetilde{m}} - g\|^2 &\leq 3\|g_m - g\|^2 + 2\mathrm{pen}^{(Trig)}(m)\\ &\quad +16(\sup_{h\in B_{\widetilde{m}\vee m}} \nu_{N,T}^2(h) - p^{(Trig)}(m,\widetilde{m})) + 16p^{(Trig)}(m,\widetilde{m}) - 2\mathrm{pen}^{(Trig)}(\widetilde{m}).\end{aligned}$$

with $p^{(Trig)}(m,m') = p_1^{(Trig)}(m,m') + p_2^{(Trig)}(m,m')$ and we find

$$p_1^{(Trig)}(m,m') = 64\pi^2G^2(T)\log(N)\frac{(m\vee m')^2}{NT}, \quad p_2^{(Trig)}(m,m') = 128G^2(T)\log(N)\frac{m\vee m'}{NT}.$$

In the same way as Lemma 4, the following lemma determines $p_1^{(Trig)}(m,m'), p_2^{(Trig)}(m,m')$.

Lemma 5. *Under the Assumptions of Theorem 1, for $\ell = 1, 2$,*

$$\mathbb{E}\left(\sup_{h\in B_{\widetilde{m}\vee m}} \nu_{N,\ell}^2(h) - p_\ell^{(Trig)}(m,\widetilde{m})\right)_+ \leq C_\ell^{(Trig)}(T,N),$$

where

$$C_\ell^{(Trig)}(T,N) = \frac{C}{NT}, \quad C_2^{(Trig)}(T,N) = C\frac{G^2(T)}{T}\frac{1}{N^{3/2}}$$

and C is a positive numerical constant.

To conclude the proof in the trigonometric basis case, analogously, we use that $8p^{(Trig)}(m,m') \leq \text{pen}^{(Trig)}(m) + \text{pen}^{(Trig)}(m')$ for $\kappa \geq \kappa_0^{(Trig)} = 8 \times 16(8\pi^2+1)$ and obtain

$$\begin{aligned}\mathbb{E}\|\tilde{g}_{\widetilde{m}} - g\|^2 &\leq 3\|g_m - g\|^2 + 4\text{pen}^{(Trig)}(m) \\ &\quad +16(C_1^{(Trig)}(T,N) + C_2^{(Trig)}(T,N))\end{aligned}$$

□

Proof of Lemma 5. Using the properties of the derivatives $\varphi'_{j,T}$ (see (33)) and the definition of $A_N(u)$, the bound for $\nu^2_{N,1,1}(h)$ now writes

$$\begin{aligned}\sup_{h\in B_{\widetilde{m}\vee m}} \nu^2_{N,1,1}(h) &\leq \sum_{j=0}^{\widetilde{m}\vee m-1}\left(\int_0^T B(u)\varphi'_{j,T}(u)du\right)^2 \\ &\leq 4\pi^2\frac{(\widetilde{m}\vee m)^2}{T^2}\int_0^T B^2(u)du = 4\pi^2\frac{(\widetilde{m}\vee m)^2}{T^2}\int_0^T G^2(u)Z_N^2(u)\mathbf{1}_{A_N(u)}du \\ &\leq 16\frac{\log N}{N}4\pi^2\frac{(\widetilde{m}\vee m)^2}{T^2}TG^2(T) = \frac{1}{2}p_1^{(Trig)}(m,\widetilde{m}).\end{aligned}$$

And, using the definition of $\mathcal{M}_N^{(Trig)}$ and (41),

$$\begin{aligned}\sup_{h\in B_{\widetilde{m}\vee m}} \nu^2_{N,1,2}(h) &\leq 4\pi^2\frac{(\widetilde{m}\vee m)^2}{T^2}\int_0^T C^2(u)du = 4\pi^2\frac{(\widetilde{m}\vee m)^2}{T^2}\int_0^T G^2(u)Z_N^2(u)\mathbf{1}_{A_N(u)^c}du \\ &\leq 4\pi^2\frac{N}{T^2}\int_0^T G^2(u)du\frac{C}{N^2} \leq 4\pi^2 C\frac{G^2(T)}{NT} = \frac{1}{2}C_1^{(Trig)}(N,T).\end{aligned}$$

The other term is

$$\begin{aligned}\sup_{h\in B_{\widetilde{m}\vee m}} \nu^2_{N,2}(h) &= \sup_{h\in B_{\widetilde{m}\vee m}} h^2(T)G^2(T)Z_N^2(T)(\mathbf{1}_{A_N(T)} + \mathbf{1}_{A_N(T)^c}) \\ &\leq 16\frac{\widetilde{m}\vee m}{T}G^2(T)\frac{\log N}{N} + \frac{\sqrt{N}}{T}G^2(T)\times\frac{C}{N^2} \\ &\leq p_2^{(Trig)}(m,\widetilde{m}) + C_2^{(Trig)}(N,T).\end{aligned}$$

This implies Lemma 5. □

7.3. **Proof of Theorem 2.** The proof follows the same steps as Theorem 1. We only indicate the changes. Here, we have, proceeding as in Proposition 3:

$$\begin{aligned}\sup_{h\in B_{\widetilde{m}\vee m}} \nu^2_{N,1,1}(h) &\leq \sum_{j=0}^{\widetilde{m}\vee m-1}\left(\int_0^T B(u)\ell'_j(u)du\right)^2 = \int_{[0,T]^2}\left[\sum_{j=0}^{\widetilde{m}\vee m-1}\ell'_j(u))B(u)\ell'_j(v)B(v)\right]dudv \\ &\leq \left[\int_0^T\left(\sum_{j=0}^{\widetilde{m}\vee m-1}(\ell'_j(u)B(u))^2\right)^{1/2}du\right]^2 \\ &\leq 16\frac{\log N}{N}G^2(T)\left[\int_0^T\left(\sum_{j=0}^{\widetilde{m}\vee m-1}(\ell'_j(u))^2\right)^{1/2}du\right]^2 \\ &\leq 16\frac{\log N}{N}G^2(T)(\widetilde{m}\vee m)^3(12+4\gamma_2^{-2}) := \frac{1}{2}p_1^\star(m,\widetilde{m}).\end{aligned}$$

Analogously, for the term with $C(u)$, using that the maximal value in $\mathcal{M}_N^\star$ is bounded by $N^{1/3}$,

$$\sup_{h\in B_{\widetilde{m}\vee m}} \nu_{N,1,2}^2(h) \leq C\frac{G^2(T)}{N^2}\left[\int_0^T \left(\sum_{j=0}^{\widetilde{m}\vee m-1} (\ell_j'(u))^2\right)^{1/2} du\right]^2 \leq C\frac{G^2(T)}{N}.$$

Thus,

$$\mathbb{E}\left(\sup_{h\in B_{\widetilde{m}\vee m}} \nu_{N,1}^2(h) - p_1^\star(m,\widetilde{m})\right)_+ \leq C\frac{G^2(T)}{N}.$$

The study of $\sup_{h\in B_{\widetilde{m}\vee m}} \nu_{N,1,2}^2(h)$ is the same as previously and we can set

$$p_2^\star(m,\widetilde{m}) = p_2^{(Lag)}(m,\widetilde{m}) \leq 32\frac{\log N}{N}G^2(T)(\widetilde{m}\vee m)^3.$$

Then,

$$\mathbb{E}\left(\sup_{h\in B_{\widetilde{m}\vee m}} \nu_{N,2}^2(h) - p_2^\star(m,\widetilde{m})\right)_+ \leq C\frac{G^2(T)}{N}.$$

We set $p^\star(m,m') = p_1^\star(m,m') + p_2^\star(m,m')$ and check that $8p^\star(m,m') \leq \mathrm{pen}^\star(m) + \mathrm{pen}^\star(m')$ for $\kappa \geq \kappa_0^\star = 8\times(16(12+4\gamma_2^{-2})+32)$.

Lastly, we have from the proof of Inequality (17) in Proposition 3 that, for $T \geq 6m-1$, we have

$$\sup_{h\in S_m, \|h\|\leq 1} |R_T(h)| \leq 2\|a\|^2 \frac{m^{3/2}}{\gamma_2}\exp\left(-(6m-3)\gamma_2\right).$$

Therefore, $\sup_{h\in S_{M_N},\|h\|\leq 1} R_T^2(h) \lesssim \frac{1}{N}$. □

8. Appendix

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:

$$\left(L_k^{(\delta)}(x)\right)' = -L_{k-1}^{(\delta+1)}(x), \quad \text{for } k\geq 1, \text{ and } \int_0^{+\infty}\left(L_k^{(\delta)}(x)\right)^2 x^\delta e^{-x}dx = \frac{\Gamma(k+\alpha+1)}{k!}. \tag{43}$$

We consider the Laguerre functions with index δ, given by

$$\ell_k^{(\delta)}(x) = 2^{(\delta+1)/2}\left(\frac{k!}{\Gamma(k+\delta+1)}\right)^{1/2} L_k^{(\delta)}(2x)e^{-x}x^{\delta/2}. \tag{44}$$

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 (43), we obtain for $j \geq 1$:

$$\ell_j'(x) = -\ell_j(x) - \sqrt{\frac{2j}{x}}\,\ell_{j-1}^{(1)}(x). \tag{45}$$

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

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

$$\ell_0'(x) = -\ell_0(x),\ \ell_j'(x) = -\ell_j(x) - 2\sum_{k=0}^{j-1} \ell_k(x), j \geq 1. \tag{46}$$

Then integrating from x to $+\infty$ formula (46) for $j \geq 1$, and setting $\widetilde{\mathcal{L}}_j(x) = \int_x^{+\infty} \ell_j(u)du$, we obtain $\ell_j = \widetilde{\mathcal{L}}_j + 2\sum_{k=0}^{j-1} \widetilde{\mathcal{L}}_k$. Thus, $\widetilde{\mathcal{L}}_j = \ell_j - \ell_{j-1} - \widetilde{\mathcal{L}}_{j-1}$. Using that $\widetilde{\mathcal{L}}_0 = \ell_0$, we obtain by elementary induction $\widetilde{\mathcal{L}}_j = \ell_j + 2\sum_{k=1}^{j}(-1)^k \ell_{j-k}$. Moreover, setting $\mathcal{L}_j(x) = \int_0^x \ell_j(u)du$, we have

$$\mathcal{L}_0(x) = \ell_0(0) - \ell_0(x), \quad \mathcal{L}_j(x) = -\mathcal{L}_{j-1}(x) - \ell_j(x) + \ell_{j-1}(x), j \geq 1. \tag{47}$$

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

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

where γ_1 and γ_2 are positive and fixed constants.

Conflict of interest

On behalf of all authors, the corresponding author states that there is no conflict of interest.

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

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

[4] Birgé, L., and Massart, P. (1998). Minimum contrast estimators on sieves: exponential bounds and rates of convergence. *Bernoulli* **4**, 329-375. MR 1653272

[5] 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.

[6] 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.

[7] Comte, F. and Genon-Catalot, V. (2015). Adaptive Laguerre density estimation for mixed Poisson models. *Electronic Journal of Statistics*, **9**, 1113-1149.

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

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

[10] Hall, P. and Heyde, C. C. (1980). *Martingale limit theory and its application.* Probability and Mathematical Statistics. Academic Press, New York-London.

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

[12] Laurent, B. and Massart, P. (2000) Adaptive estimation of a quadratic functional by model selection. *Ann. Statist.* **28**, 1302-1338.

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

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

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

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

*Laboratoire MAP5, Université Paris Descartes, Paris, France
E-mail address: fabienne.comte@parisdescartes.fr

E-mail address: valentine.genon-catalot@parisdescartes.fr