LAGUERRE AND HERMITE BASES FOR INVERSE PROBLEMS

F. COMTE[(1)] & V. GENON-CATALOT[(2)]

Abstract. We present inverse problems of nonparametric statistics which have a performing and smart solution using projection estimators on bases of functions with non compact support, namely, a Laguerre basis or a Hermite basis. The models are $Y_i = X_i U_i,\ Z_i = X_i + \Sigma_i$, where the X_i's are *i.i.d.* with unknown density f, the Σ_i's are *i.i.d.* with known density f_Σ, the U_i's are *i.i.d.* with uniform density on $[0, 1]$. The sequences $(X_i), (U_i), (\Sigma_i)$ are independent. We define projection estimators of f in the two cases of indirect observations of $(X_1, \ldots, X_n)$, and we give upper bounds for their $\mathbb{L}^2$-risks on specific Sobolev-Laguerre or Sobolev-Hermite spaces. Data-driven procedures are described and proved to perform automatically the bias variance compromise.

(1) Université Paris Descartes, MAP5, UMR CNRS 8145, email: fabienne.comte@parisdescartes.fr
(2) Université Paris Descartes, MAP5, UMR CNRS 8145, valentine.genon-catalot@parisdescartes.fr.

Keywords. Adaptive estimation. Additive noise. Inverse problem. Model selection. Multiplicative censoring. Projection estimator.
MSC2010. 62G05 - 62G07

1. Introduction

The aim of this paper is to demonstrate that some classical inverse problems of nonparametric statistics can be dealt with by using projection estimators either on a Laguerre basis or on a Hermite basis and that this approach provides easy and performing solutions. We make a unified presentation joining novelties and improvements of previous results.

To motivate the framework, first consider $X_1, \ldots, X_n$ n *i.i.d.* random variables with unknown density f. If the X_i's are observed and f is square integrable, nonparametric estimators of f can be built by using a projection method on an orthonormal basis. Many authors use compactly supported bases (see e.g. Massart (2007), Efromovich (1999)). Non compactly supported wavelet bases of $\mathbb{L}^2(\mathbb{R})$ have also been used (see Juditsky and Lambert-Lacroix (2004), and the references therein). The $\mathbb{L}^2(\mathbb{R})$ basis of Hermite functions is another possibility, investigated in *e.g.* Schwarz (1967), Walter (1977) and more recently, in Belomestny *et al.*, 2017. As enlighted in the latter paper, projection estimators on a Hermite basis have the following advantages. First, as the basis is not compactly supported, there is no need to fix an interval for estimation. This is convenient, especially for inverse problems. Second, from the computational point of view, they have a much lower complexity than other estimators based on non compactly supported bases such as deconvolution estimators. Nevertheless they perform asymptotically as well as estimators built from competing methods. If f is $\mathbb{R}^+$-supported, one can use an orthonormal basis of $\mathbb{L}^2(\mathbb{R}^+)$. The Laguerre basis is then well fitted. The qualities of Laguerre projection estimators have been recently investigated (see *e.g.* Mabon (2017), Vareschi (2015), and for actuarial applications Zhang and Su (2017)). If the X_i's are not directly observed, the estimation of f is an inverse problem. The inverse problems considered here are the following ones. First,

Date: October 4, 2017.

we assume that observations are

$$Y_i = X_i U_i, \quad i = 1, \ldots, n \tag{1}$$

where the sequences $(X_i), (U_i)$ are independent and (U_i) are *i.i.d.* with uniform distribution on $[0, 1]$. Model (1) is called multiplicative censoring model and covers several important statistical problems (see *e.g.* Vardi (1989)). Numerous papers deal with the estimation of f for model (1) whether by nonparametric maximum likelihood (Vardi (1989), Vardi and Zhang (1992), Asgharian *et al.* (2012)), by projection methods (Andersen and Hansen (2001), Abbaszadeh *et al.* (2012,2013)) or kernel methods (Brunel *et al.* (2015)). In Belomestny *et al.* (2016), f is supposed to be $\mathbb{R}^+$-supported and estimated by projection estimators on a Laguerre basis. We revisit this problem and obtain an improvement of the risk bound, under a slight additional moment assumption. In the case of real-valued random variables X_i, we investigate the estimation of f from model (1) by using a Hermite basis approach.

Second, we consider observations $Z_1, \ldots, Z_n$ such that

$$Z_i = X_i + \Sigma_i, \quad i = 1, \ldots, n. \tag{2}$$

where (Σ_i) are *i.i.d.* with known density f_Σ and the sequences $(X_i), (\Sigma_i)$ are independent. Density estimation from noisy observations is the subject of a huge number of contributions. This deconvolution problem is classically solved by Fourier methods (see e.g. Comte *et al.* (2006) and the references therein). Recently, when the X_i's are non-negative, the study of one-sided errors, *i.e.* $\Sigma_i \geq 0$, was motivated by applications in the field of finance (see Jirak *et al.* (2014)) or in survival models, (see van Es *et al.* (1998), Jongbloed (1998)). In particular, Mabon (2017) proposes for model (2) projection estimators of f using a Laguerre basis whose properties allow deconvolution of densities on $\mathbb{R}^+$. We detail this approach and provide an improvement of the risk bound of the projection estimators. In the case of real-valued random variables X_i, Σ_i, we study the estimation of f by using a Hermite basis approach.

To our knowledge, the estimation of f from observations (1)-(2) by projection estimators on a Hermite basis, which is investigated here, is new. We compare Hermite estimators to the deconvolution estimators for models (1) and (2). Note that the estimation of f in model (1) by deconvolution is also new.

In each of the above models, we exhibit explicit relations between the projection coefficients of the density of the observed variables and the projection coefficients of the unknown density f. This allows to build projection estimators of f. We provide risk bounds for the estimators. Afterwards, we propose data-driven procedures, including for model (1) a random penalty which avoids a priori knowledge of the variance rate. This is important as the variance order varies in function of moment assumptions.

Laguerre and Hermite bases are related to specific function spaces, respectively the Sobolev-Laguerre spaces (see *e.g.* Shen (2000) and Bongioanni and Torrea (2007)) or the Sobolev-Hermite spaces (Bongioanni and Torrea (2006)). The links between rate of decay of the projection coefficients and regularity properties of functions in these spaces have been described respectively in Comte and Genon-Catalot (2015) and in Belomestny *et al.* (2017). This allows to assess the rate of bias terms of $\mathbb{L}^2$-risks in the projection approach, and to compute upper bounds for the rates of convergence.

In Section 2, we describe the Laguerre and Hermite bases and spaces. In Section 3, for the purpose of comparison with the other models, we study the case of direct observations of $X_1, \ldots, X_n$. Section 4 deals with model (1). Section 5 is concerned with model (2). In Section 6, we give some concluding remarks. Section 7 contains useful formulae for Laguerre and Hermite functions and all proofs. In the Appendix, we recall some properties of Sobolev-Laguerre and Sobolev-Hermite spaces, and the Talagrand inequality used in proofs.

2. About Laguerre and Hermite bases and spaces

2.1. **Laguerre and Hermite bases.** We start by presenting the Laguerre and Hermite bases and the Sobolev-Laguerre and Sobolev-Hermite regularity spaces. More details on Laguerre and Hermite functions are given in Section 7.1.

Below we denote the scalar product and the $\mathbb{L}^2$-norm on $\mathbb{L}^2(\mathbb{R}^+)$ by: $\forall s,t \in \mathbb{L}^2(\mathbb{R}^+)$, $\langle s,t\rangle_+ = \int_0^{+\infty} s(x)t(x)dx$, $\|t\|_+^2 = \int_0^{+\infty} t^2(x)dx$. We also denote by $\|.\|$ the $\mathbb{L}^2$-norm on $\mathbb{R}$ and by $\langle\cdot,\cdot\rangle$ the $\mathbb{L}^2(\mathbb{R})$-scalar product.

Consider the Laguerre polynomials (L_j) and the Laguerre functions (φ_j) given by

$$L_j(x) = \sum_{k=0}^{j}(-1)^k\binom{j}{k}\frac{x^k}{k!}, \qquad \varphi_j(x) = \sqrt{2}L_j(2x)e^{-x}\mathbf{1}_{x\geq 0}, \quad j \geq 0. \tag{3}$$

The collection $(\varphi_j)_{j\geq 0}$ constitutes a complete orthonormal system on $\mathbb{L}^2(\mathbb{R}^+)$, and is such that (see Abramowitz and Stegun (1964)):

$$\forall j \geq 0,\ \ \forall x \in \mathbb{R}^+,\ \ |\varphi_j(x)| \leq \sqrt{2}. \tag{4}$$

For $f \in \mathbb{L}^2(\mathbb{R}^+)$, we can develop f on the Laguerre basis, $f = \sum_{j\geq 0} a_j(f)\varphi_j$, $\quad a_j(f) = \langle f, \varphi_j\rangle_+$. The Hermite polynomial and the Hermite function of order j are given, for $j \geq 0$, by:

$$H_j(x) = (-1)^j e^{x^2}\frac{d^j}{dx^j}(e^{-x^2}), \quad h_j(x) = c_j H_j(x)e^{-x^2/2}, \quad c_j = \left(2^j j!\sqrt{\pi}\right)^{-1/2} \tag{5}$$

The sequence $(h_j, j \geq 0)$ is an orthonormal basis of $\mathbb{L}^2(\mathbb{R})$. When the density f belongs to $\mathbb{L}^2(\mathbb{R})$, it can be developed in the Hermite basis $f = \sum_{j\geq 0} a_j(f)h_j$ where $a_j(f) = \int_{\mathbb{R}} f(x)h_j(x)dx = \langle f, h_j\rangle$. The infinite norm of h_j satisfies (see Abramowitz and Stegun (1964), Szegö (1959) p.242):

$$\|h_j\|_\infty \leq \Phi_0, \qquad \Phi_0 \simeq 1,086435/\pi^{1/4} \simeq 0.8160, \tag{6}$$

and the following more precise bound is available (with C_∞ a constant given in Szegö (1959))

$$\|h_j\|_\infty \leq \frac{C_\infty}{(j+1)^{1/12}}, \quad j = 0, 1, \ldots. \tag{7}$$

We use the notation ψ_j to designate φ_j in the Laguerre case and h_j in the Hermite case, denote by $S_m = \text{span}(\psi_0, \psi_1, \ldots, \psi_{m-1})$ the linear space generated by the m functions $\psi_0, \ldots, \psi_{m-1}$ and by $f_m = \sum_{j=0}^{m-1} a_j(f)\psi_j$ the orthogonal projection of f on S_m, where $a_j(f) = \langle f, \psi_j\rangle$ will mean either $\langle f, \varphi_j\rangle_+$ or $\langle f, h_j\rangle$.

2.2. **Sobolev-Laguerre and Sobolev-Hermite spaces.** For $s \geq 0$, the Sobolev-Laguerre space with index s (see Bongioanni and Torrea (2007)) is defined by:

$$W_L^s = \{\theta : \mathbb{R}^+ \to \mathbb{R},\ \theta \in \mathbb{L}^2(\mathbb{R}^+), |\theta|_s^2 := \sum_{k\geq 0} k^s a_k^2(\theta) < +\infty\}. \tag{8}$$

where $a_k(\theta) = \langle\theta, \varphi_k\rangle_+$. We define the ball $W_L^s(D)$ by

$$W_L^s(D) \doteq \left\{\theta \in W_L^s, |\theta|_s^2 = \sum_{k=0}^{\infty} k^s a_k^2(\theta) \leq D\right\}.$$

Analogously, the Sobolev-Hermite space with regularity s (see Bongioanni and Torrea (2006)) is given by

$$W_H^s = \{\theta : \mathbb{R} \to \mathbb{R},\ \theta \in \mathbb{L}^2(\mathbb{R}), \sum_{k\geq 0} k^s a_k^2(\theta) < +\infty\}. \tag{9}$$

where $a_k(\theta) = \langle \theta, h_k\rangle$, and the Sobolev-Hermite ball

$$W_H^s(D) = \{\theta \in \mathbb{L}^2(\mathbb{R}), \sum_{k\geq 0} k^s a_k^2(\theta) \leq D\}.$$

Thus, for f in $W_L^s(D)$ or in $W_H^s(D)$, we have $\|f - f_m\|^2 \leq Dm^{-s}$.

For details and especially for regularity properties of functions in these spaces, we refer also to Section 7 of Comte and Genon-Catalot (2015), Section 7.2 of Belomestny *et al.* (2016) and Section 4.1 of Belomestny *et al.* (2017), see Appendix A.

3. Projection estimators of f when X_i's are observed

3.1. **Risk bound.** We assume that f belongs to $\mathbb{L}^2(\mathbb{R}^+)$ or $\mathbb{L}^2(\mathbb{R})$ and provide for each $m \geq 1$, a projection estimator of f by estimating the coefficients $a_j(f), j = 0, \ldots, m-1$. In the case where the X_i's are observed, we define the empirical and unbiased estimator of $a_j(f)$ by

$$\hat{a}_j(X) = \frac{1}{n}\sum_{i=1}^n \psi_j(X_i) \quad \text{and the projection estimator} \quad \hat{f}_m^X = \sum_{j=0}^{m-1} \hat{a}_j(X)\psi_j.$$

Clearly, $\hat{f}_m^X$ an unbiased estimator of $f_m = \sum_{j=0}^{m-1} a_j(f)\psi_j$, the orthogonal projection of f on S_m. By the Pythagoras Theorem, we have $\|\hat{f}_m^X - f\|^2 = \|f - f_m\|^2 + \|\hat{f}_m^X - f_m\|^2$. As $(\psi_j)_j$ is orthonormal, we get $\|\hat{f}_m^X - f_m\|^2 = \sum_{j=0}^{m-1}(\hat{a}_j(X) - a_j(f))^2$ and

$$\mathbb{E}[(\hat{a}_j(X) - a_j(f))^2] = \frac{1}{n}\mathrm{Var}(\psi_j(X_1)) = \frac{1}{n}\mathbb{E}(\psi_j^2(X_1)) - \frac{a_j^2(f)}{n}.$$

Let us define

$$V_m^X = \sum_{j=0}^{m-1} \mathbb{E}[\psi_j^2(X_1)], \tag{10}$$

we have

$$\mathbb{E}(\|\hat{f}_m^X - f\|^2) = \|f - f_m\|^2 + \frac{V_m^X}{n} - \frac{\|f_m\|^2}{n}. \tag{11}$$

With (4) or (6), we obtain $V_m^X \leq Cm$, with $C = 2$ in the Laguerre case and $C = \Phi_0^2$ (see (6)) in the Hermite case. In the latter case, using (7), we get $V_m^X \leq Cm^{5/6}$, which is better. However, these bounds on the variance term are not optimal and can be improved under rather weak moment assumptions.

Proposition 3.1. *If* $\mathbb{E}(1/\sqrt{X_1}) < +\infty$ *in the Laguerre case or* $\mathbb{E}(|X_1|^{2/3}) < +\infty$ *in the Hermite case, then for m large enough,* $V_m^X \leq c\sqrt{m}/n$ *where c is a constant, and thus*

$$\mathbb{E}(\|\hat{f}_m^X - f\|^2) \leq \|f - f_m\|^2 + c\frac{\sqrt{m}}{n}. \tag{12}$$

In the Laguerre case, this result improves the variance order obtained in Belomestny *et al.* (2016), under the additional moment assumption $\mathbb{E}(1/\sqrt{X_1}) < +\infty$. For the Hermite case, the order $O(\sqrt{m})$ is obtained in Belomestny *et al.* (2017b), Proposition 2.1, but under the stronger moment condition $\mathbb{E}(|X_1|^5) < +\infty$. Note that even the Cauchy distribution satisfies $\mathbb{E}(|X_1|^{2/3}) < +\infty$.

The risk bound decomposition (12) involves a bias term $\|f - f_m\|^2 = \sum_{j\geq m} a_j^2(f)$ which is decreasing with m and a variance term of order $\sqrt{m}/n$ which is increasing with m. Therefore, to evaluate the rate of convergence, we have to perform a compromise to select relevantly m. More precisely, for f in $W_L^s(D)$ or in $W_H^s(D)$, we have $\|f - f_m\|^2 = \sum_{j\geq m} a_j^2(f) \leq Dm^{-s}$. Choosing $m_{\rm opt} = [n^{1/(s+1/2)}]$ in the r.h.s. of (12) implies

$$\mathbb{E}(\|\hat{f}^X_{m_{\rm opt}} - f\|^2) \leq C_0(s, D) n^{-2s/(2s+1)}$$

where $C_0(s, D)$, is a constant depending on s and D only. This is the usual rate for density estimation for f belonging to a classical Sobolev ball with regularity s: indeed, Schipper (1996) proves that this rate is minimax optimal with exact Pinsker constant on Sobolev balls for an integer s, Efromovich (2002) proves the result for $s < 1/2$ and Rigollet (2006) builds an adaptive deconvolution estimator based on the blockwise Stein method, which reaches the optimal rate with exact constant for any $s > 1/2$.
The upper bound $O(\sqrt{m})$ on the term V_m^X is somehow optimal as we can prove:

Proposition 3.2. *Assume that* $\inf_{a\leq x\leq b} f(x) > 0$ *for some interval* $[a, b]$, *with* $a < b$ *in the Hermite case and* $0 < a < b$ *in the Laguerre case, then, for* m *large enough,* $V_m^X \geq c'\sqrt{m}$ *where* c' *is a constant. Therefore*

$$\mathbb{E}(\|\hat{f}^X_m - f\|^2) \geq \|f - f_m\|^2 + c'\frac{\sqrt{m}}{n} - \frac{\|f\|^2}{n}.$$

Proposition 3.2 in the Hermite case is proved in Belomestny *et al.* (2017), Proposition 2.2.

3.2. **Specific rates under small bias.** For some classes of distributions, the rate of the bias term can be much smaller than polynomial. This implies a much better rate of convergence for the $\mathbb{L}^2$-risk of estimators.
• Laguerre case. Consider the class of mixtures of exponential distributions defined, for $v > 1$, by

$$\mathcal{E}(v) = \left\{ f : f(x) = \int_0^\infty \theta\ \exp(-\theta x) d\Pi(\theta),\ \ \Pi[1/v, v] = 1 \right\}$$

Proposition 3.3. *Let* $f \in \mathcal{E}(v)$ *for some real number* $v > 1$ *and set* $\rho = (v - 1/v + 1)^2 < 1$. *Then, for* $m_{opt} = [\log n/|\log \rho|]$,

$$\mathbb{E}(\|\hat{f}^X_{m_{opt}} - f\|) \lesssim \sqrt{\log n}/n.$$

• Hermite case. In Belomestny *et al.* (2017), Section 4.4, we investigated the risk rate of Hermite estimators on classes of mean mixtures or variance mixtures of Gaussian distributions and obtained the following results. Define, for ϕ the standard Gaussian density,

$$\mathcal{F}(C) = \left\{ f :\ f(x) = \phi \star \Pi(x) = \int \phi(x - u) d\Pi(u),\ \Pi \in \mathcal{P}(C) \right\},$$

$$\mathcal{P}(C) := \left\{ \Pi \in \mathcal{P}(\mathbb{R}), \Pi(|u| > t) \leq C \exp(-t^2/C) \text{ for all positive } t \right\}.$$

For $f \in \mathcal{F}(C)$ and $m_{\rm opt} = [\log(n)(eC + 1/\log(2))]$, we have $\mathbb{E}(\|\hat{f}^X_{m_{\rm opt}} - f\|^2) \lesssim \sqrt{\log(n)}/n$.

Now consider the class of variance mixtures

$$\mathcal{G}(v) = \left\{ f : \ f(x) = \int_0^{+\infty} \frac{\phi(x/u)}{u} d\Pi(u), \Pi([1/\sqrt{v}, \sqrt{v}]) = 1 \right\}, \quad v > 1.$$

For $f \in \mathcal{G}(v)$ and $m_{\rm opt} = [\log(n)/|\log(\rho)|]$, with ρ as above, $\mathbb{E}(\|\hat{f}^X_{m_{\rm opt}} - f\|^2) \lesssim \sqrt{\log(n)}/n$.

3.3. **Adaptive estimation.** The density f being unknown, we do not know what kind of regularity it has, therefore, a data-driven choice of the dimension of the projection space has to be done. The interest of the data-driven procedure is that it allows to realize automatically the finite sample bias-variance compromise and also to *automatically* reach the best possible asymptotic rate without requiring any knowledge on the bias order. The data-driven choice of m mimicks the minimization of the squared bias-variance bound using estimators of the risk bound terms. As $\|f - f_m\|^2 = \|f\|^2 - \|f_m\|^2$, the squared bias is estimated by $-\|\hat{f}^X_m\|^2$, getting rid of $\|f\|^2$ which is unknown but constant. Thus we set, for κ a numerical constant,

$$\hat{m}_X = \arg \min_{m \in \{1,\dots,m_n\}} \left(-\|\hat{f}^X_m\|^2 + \widehat{\rm pen}_X(m) \right), \quad \widehat{\rm pen}_X(m) = \kappa \frac{\widehat{V}^X_m}{n}, \quad \widehat{V}^X_m = \frac{1}{n} \sum_{i=1}^n \sum_{j=0}^{m-1} \psi_j^2(X_i).$$

Clearly, $\widehat{V}^X_m$ is an estimate of $\mathbb{E}(\sum_{j=0}^{m-1} \psi_j^2(X_1))$. We set ${\rm pen}_X(m) := \mathbb{E}(\widehat{\rm pen}_X(m)) = \kappa V^X_m/n$ and prove:

Theorem 3.1. *Assume that $m_n \le (n/\log(n))^\alpha$ with $\alpha = 1$ for Laguerre and $\alpha = 6/5$ for Hermite, then there exists a numerical value κ_0 such that for all $\kappa \ge \kappa_0$,*

$$\mathbb{E}(\|\hat{f}^X_{\hat{m}_X} - f\|^2) \le 3 \inf_{m \in \{1,\dots,n\}} \left(\|f - f_m\|^2 + {\rm pen}_X(m) \right) + \frac{C}{n},$$

where C is a constant depending on $\|f\|$.

The Hermite case is proved in Belomestny *et al.* (2017, Theorem 2.1). We prove here the result in the Laguerre case.

Remark 3.1. Note that the risk bound achieves automatically the bias-variance compromise, up to a negligible term of order $O(1/n)$, and in this sense, $\hat{f}^X_{\hat{m}_X}$ is adaptive.

4. Projection estimator of f when Y_i's are observed

Now, our aim is to build an estimator of f from the observations $Y_1, \dots, Y_n$ given by (1).

4.1. **Preliminary properties and risk bounds.** The estimation of f in model (1) is an inverse problem which can be solved by Hermite or Laguerre projection estimator, depending on the support of the density f. The hidden variables X_i are either *real-valued* or *nonnegative*, with unknown density f. The construction of estimators rely on relations between the density f_Y and survival function $\bar{F}_Y(y) = \mathbb{P}(Y > y)$ of Y_i and those of X_i. We recall these relations:

$$(13) \qquad \forall y \in \mathbb{R}, \quad f_Y(y) = \int_y^{+\infty} \frac{f(x)}{x} dx \, \mathbf{1}(y \ge 0) + \int_{-\infty}^y \frac{f(x)}{|x|} dx \, \mathbf{1}(y < 0),$$

$$(14) \qquad \forall y \in \mathbb{R}, \quad \bar{F}_Y(y) + y f_Y(y) = \bar{F}(y).$$

When f is $\mathbb{R}^+$-supported, (13) reduces to $f_Y(y) = \int_y^{+\infty} (f(u)/u) du \mathbf{1}_{y \ge 0}$. We can prove:

Lemma 4.1. (1) *Let $t:\mathbb{R}\to\mathbb{R}$ be bounded, derivable, then*

$$\mathbb{E}(t(Y_1)+Y_1t'(Y_1))=\mathbb{E}t(X_1). \tag{15}$$

In particular, for all $j\geq 0$,

$$a_j(f)=\langle f,\psi_j\rangle=\langle f_Y,(y\psi_j)'\rangle \tag{16}$$

(2) *Assume that $\mathbb{E}|X_1|<+\infty$. Let $t\in\mathbb{L}^2(\mathbb{R})$ or $\mathbb{L}^2(\mathbb{R}^+)$, then $\mathbb{E}(Y_1^2t^2(Y_1))\leq\mathbb{E}|X_1|\|t\|^2$.*

Equality (15) is the basement of the estimation procedure and leads to (16), which links the coefficients of f and f_Y on the Laguerre or Hermite basis. So, using (16), we get for all $j\geq 0$:

$$a_j(f)=\mathbb{E}(\psi_j(Y_1)+Y_1\psi_j'(Y_1)). \tag{17}$$

Therefore, we define a collection of projection estimator of f based on the observation $(Y_1,\ldots,Y_n)$ by:

$$\hat{f}_m(x)=\sum_{j=1}^{m}\hat{a}_j\psi_j(x),\qquad \text{with } \hat{a}_j=\frac{1}{n}\sum_{i=1}^{n}[Y_i\psi_j'(Y_i)+\psi_j(Y_i)]. \tag{18}$$

Then we can prove the following risk bound.

Proposition 4.1. *Let $\hat{f}_m$ be the estimator defined by (18), then we have*

$$\mathbb{E}(\|\hat{f}_m-f\|^2)=\|f-f_m\|^2+\frac{V_m}{n}-\frac{\|f_m\|^2}{n},\quad V_m=\sum_{j=0}^{m-1}\mathbb{E}\left[Y_1\psi_j'(Y_1)+\psi_j(Y_1)\right]^2 \tag{19}$$

where V_m is such that, for m large enough,

$$V_m\leq cm^{3/2},$$

(1) *in the Laguerre case, if $\mathbb{E}(X_1)<+\infty$ and $\mathbb{E}(1/\sqrt{X_1})<+\infty$,*
(2) *in the Hermite case, if $\mathbb{E}(|X_1|^{2+2/3})<+\infty$,*

where c is a constant which does not depend on m, but depends on the above moments of X_1.

Remark 4.1. In the Laguerre case, with no moment condition, we have, for all $m\geq 1$, $V_m\leq Cm^3$ (see Belomestny *et al.*, 2016).

In the Hermite case, if $\mathbb{E}(|X_1|)<+\infty$, then for all $m\geq 1$, $V_m\leq Cm^2$ and if $\mathbb{E}(X_1^2)<+\infty$, then $V_m\leq Cm^{11/6}$, see the proof in Section 7.10.

We can deduce rates of convergence on Sobolev-Laguerre and Sobolev-Hermite balls.

Proposition 4.2. *Assume that f belongs to $W_L^s(D)$ or to $W_H^s(D)$ (see (8) and (9)). If $\mathbb{E}(X_1)<+\infty$ and $\mathbb{E}(1/\sqrt{X_1})<+\infty$ in the Laguerre case, or if $\mathbb{E}|X_1|^{2/3}<+\infty$ in the Hermite case, then for $m_{\rm opt}=[n^{1/(s+3/2)}]$,*

$$\mathbb{E}(\|\hat{f}_{m_{\rm opt}}-f\|^2)\lesssim n^{-\frac{2s}{2s+3}}.$$

The results of Section 3.2 apply here. For exponential mixtures in the Laguerre case, or mean or variance Gaussian mixtures in the Hermite case, the bias is exponentially decreasing. Thus, the same choices $m_{\rm opt}$ yield a rate of order $[\log(n)]^{3/2}/n$.

Similarly as above, the bound $O(m^{3/2})$ on the variance term cannot be improved:

Proposition 4.3. (1) *Laguerre case: if $\mathbb{E}(1/\sqrt{X_1})<+\infty$ and there exist $0<a<b$, with $\inf_{a\leq x\leq b}f(x)\geq c>0$, then, for m large enough, there exists a constant c_1 such that $V_m\geq c_1m^{3/2}$, where c_1 does not depend on m.*

(2) *Hermite case: If there exist $a < b$, with $\inf_{a\le x\le b} f(x) \ge c > 0$, then, for m large enough, there exists a constant c_2 such that $V_m \ge c_2 m^{3/2}$, where c_2 does not depend on m.*

Consequently in each case, for $c^\star = c_1$ or c_2, we have

$$\mathbb{E}(\|\hat{f}_m - f\|^2) \ge \|f - f_m\|^2 + c^\star \frac{m^{3/2}}{n} - \frac{\|f\|^2}{n}.$$

4.2. **Matrix representations.** Now, we give formulae providing the links between the coefficients of f and those of f_Y on basis $(\psi_j)_{0\le j\le m-1}$ and allowing to build easily the estimators. For g a function, we denote $\vec{a}_k(g) := {}^t(a_0(g), \dots, a_{k-1}(g))$, $k \ge 1$, where $a_j(g) = \langle g, \psi_j \rangle$.

• *Specific Laguerre formula.* Using formula (47) (see Section 7.1), we get $a_0(f) = (1/2)a_0(f_Y) + (1/2)a_1(f_Y)$ and for $j \ge 1$,

$$a_j(f) = -\frac{j}{2} a_{j-1}(f_Y) + \frac{1}{2} a_j(f_Y) + \frac{j+1}{2} a_{j+1}(f_Y).$$

Introducing the matrix $\mathbf{H}_m = ([\mathbf{H}_m]_{k,\ell})_{1\le k,\ell\le m}$ with size $m \times (m+1)$ given by $[\mathbf{H}_m]_{k,\ell} = 0$ if $\ell \ne k-1, k, k+1$ and $[\mathbf{H}_m]_{1,1} = 1/2$, $[\mathbf{H}_m]_{1,2} = 1/2$ and for $k \ge 2$,

$$[\mathbf{H}_m]_{k,k-1} = -\frac{k-1}{2}, \quad [\mathbf{H}_m]_{k,k} = \frac{1}{2}, \quad [\mathbf{H}_m]_{k,k+1} = \frac{k}{2}, \tag{20}$$

yields the linear relation between the vectors of coefficients of f and f_Y: $\vec{a}_{m-1}(f) = \mathbf{H}_m \vec{a}_m(f_Y)$. We thus have for $m \ge 1$, setting $\vec{\hat{a}}_{m-1} = {}^t(\hat{a}_j)_{0\le j\le m-1}$, $\vec{\hat{a}}_m(Y) = {}^t(\hat{a}_j(Y))_{0\le j\le m}$, the following relation which is convenient to compute the estimator

$$\vec{\hat{a}}_{m-1} = \mathbf{H}_m \vec{\hat{a}}_m(Y) \quad \text{with} \quad \hat{a}_j(Y) := \frac{1}{n}\sum_{i=1}^n \varphi_j(Y_i). \tag{21}$$

• *Specific Hermite formula.* From $x f_Y'(x) = -f(x)$ and the relations (52), we get

$$\begin{cases} a_0(f) = (1/2)\left(a_0(f_Y) - \sqrt{2} a_2(f_Y)\right), \quad a_1(f) = (1/2)\left(a_1(f_Y) - \sqrt{6} a_3(f_Y)\right) \\ a_j(f) = (1/2)\left(a_j(f_Y) - \sqrt{(j+1)(j+2)} a_{j+2}(f_Y) + \sqrt{(j-1)j} a_{j-2}(f_Y)\right), \quad j \ge 2. \end{cases}$$

Therefore $\vec{a}_m(f) = \mathbf{A}_{m,m+2}\, \vec{a}_{m+2}(f_Y)$, where $\mathbf{A}_{m,m+2}$ is the $m \times (m+2)$ matrix deduced from the above relations. So, for $m \ge 1$, setting $\vec{\hat{a}}_m = {}^t(\hat{a}_j)_{0\le j\le m-1}$, $\vec{\hat{a}}_{m+2}(Y) = {}^t(\hat{a}_j(Y))_{0\le j\le m+1}$, gives the following relation:

$$\vec{\hat{a}}_m = \mathbf{A}_{m,m+2}\, \vec{\hat{a}}_{m+2}(Y) \quad \text{with} \quad \hat{a}_j(Y) := \frac{1}{n}\sum_{i=1}^n h_j(Y_i). \tag{22}$$

4.3. **Model selection.** Define $\mathcal{M}_n = \{1, \dots, m_n\}$, where m_n is the largest integer of the collection and set

$$\hat{m} = \arg\min_{m\in\mathcal{M}_n} \{-\|\hat{f}_m\|^2 + \widehat{\text{pen}}(m)\}, \quad \widehat{\text{pen}}(m) = \kappa \frac{\widehat{V}_m}{n}, \tag{23}$$

$$\widehat{V}_m = \frac{1}{n}\sum_{i=1}^n \sum_{j=0}^{m-1} \left[Y_i \psi_j'(Y_i) + \psi_j(X_i)\right]^2,$$

where κ is a numerical constant. Note that $\widehat{V}_m$ is an estimate of V_m such that $\mathbb{E}(\widehat{V}_m) = V_m$. We denote $\text{pen}(m) = \mathbb{E}(\widehat{\text{pen}}(m)) = \kappa V_m/n$.

Theorem 4.1. *Assume that, in the Hermite case,* $\mathbb{E}(|X_1|) < +\infty$ *and* $\inf_{a\le x\le b} f(x) > 0$ *for some interval* $[a,b]$ *and in the Laguerre case,* $\mathbb{E}(1/\sqrt{X_1}) < +\infty$ *and* $\inf_{a\le x\le b} f(x) > 0$ *for some interval* $[a,b]$ *with* $0 < a < b$*. Assume also that the collection of models is such that*

$$m_n \le n^{\beta} \text{ with } \quad \beta = 1/3 \text{ for Laguerre case}, \quad \beta = 6/17 \text{ for Hermite case.} \tag{24}$$

Then there exists a numerical constant κ_0 *such that, for* $\kappa \ge \kappa_0$*, the estimator* $\hat{f}_{\hat{m}}$ *where* $\hat{m}$ *is defined by (23) satisfies*

$$\mathbb{E}\left(\|\hat{f}_{\hat{m}} - f\|^2\right) \le C \inf_{m\in\mathcal{M}_n}\left(\|f - f_m\|^2 + \kappa\frac{V_m}{n}\right) + C'\frac{\log(n)}{n},$$

where C *is a numerical constant(*$C = 4$ *suits) and* C' *is a constant depending on* $\|f\|_\infty$*.*

Note that Remark 3.1 applies here also. The restriction (24) implies that the optimal order of m, for f in a Sobolev ball, can be reached only if the function is regular enough i.e. s large enough. More precisely, under the assumptions of Theorem 4.1 and of Proposition 4.2, we have, for $f \in W_L^s(D)$ (Laguerre case) that $m_{\rm opt} = [n^{1/(s+(3/2))}]$ can be reached in the model collection if $s > 3/2$ ($n^{1/(s+3/2)} < n^{1/3}$). For $f \in W_H^s(D)$ (Hermite case), $m_{\rm opt}$ is reached for $s > 4/3$ ($n^{1/(s+3/2)} < n^{6/17}$).

4.4. **Deconvolution estimator based on multiplicative censored observations.** In Belomestny *et al.* (2017), for the case of direct observations $X_1, \dots, X_n$, projection estimators on the Hermite basis are compared to projection estimators on the sine cardinal basis. The comparison is relevant as Sobolev Hermite spaces with regularity index s are included in usual Sobolev spaces with the same regularity index (see Appendix A). For direct observations, the two estimators are proved in Belomestny *et al.* (2017) to be asymptotically equivalent. We prove here that it is also the case in the multiplicative censored case.

Let us recall the definition of the sine cardinal basis. For $u, v \in \mathbb{L}^1(\mathbb{R}) \cap \mathbb{L}^2(\mathbb{R})$, we denote by $u^*(t) = \int e^{itx}u(x)dx$ and $\langle u, v\rangle = \int u\bar{v} = (2\pi)^{-1}\langle u^*, v^*\rangle$. We recall that $u^{**}(\cdot) = 2\pi u(-\cdot)$. Let $\theta(x) = \sin(\pi x)/(\pi x)$ which satisfies $\theta^*(t) = \mathbf{1}_{[-\pi,\pi]}(t)$, where θ^* denotes the Fourier transform of θ. The functions $(\theta_{\ell,j}(x) = \sqrt{\ell}\theta(\ell x - j), j \in \mathbb{Z})$ constitute an orthonormal system in $\mathbb{L}^2(\mathbb{R})$. The space F_ℓ generated by this system is exactly the subspace of $\mathbb{L}^2(\mathbb{R})$ of functions having Fourier transforms with compact support $[-\pi\ell, \pi\ell]$. The orthogonal projection $\bar{f}_\ell$ of f on F_ℓ satisfies $\bar{f}_\ell^* = f^*\mathbf{1}_{[-\pi\ell,\pi\ell]}$ and

$$\bar{f}_\ell(x) = \frac{1}{2\pi}\int_{-\pi\ell}^{\pi\ell} e^{-itx}f^*(t)dt = \sum_{\ell\in\mathbb{Z}} a_{\ell,j}\theta_{\ell,j}, \quad a_{\ell,j} = \langle f, \theta_{\ell,j}\rangle. \tag{25}$$

Note that

$$\|f - \bar{f}_\ell\|^2 = \frac{1}{2\pi}\int_{|t|\ge\pi\ell} |f^*(t)|^2 dt. \tag{26}$$

Now, relation (15) holds for $t(x) = e^{iux}$: $\mathbb{E}(e^{iuY_1} + iuY_1e^{iuY_1}) = \mathbb{E}e^{iuX_1}$. Therefore, we can estimate the Fourier transform $f^*(u) = \int e^{iux}f(x)dx$ of f by

$$\frac{1}{n}\sum_{k=1}^{n}\left(e^{iuY_k} + iuY_ke^{iuY_k}\right).$$

This yields the following new estimator of f, of deconvolution type:

$$\widetilde{f}_\ell(x) = \frac{1}{2\pi}\int_{-\pi\ell}^{\pi\ell} e^{-iux}\frac{1}{n}\sum_{k=1}^{n}\left(e^{iuY_k} + iuY_ke^{iuY_k}\right)du, \tag{27}$$

which is an unbiased estimator of $\bar{f}_\ell$ given in (25). We can integrate and obtain:

$$\widetilde{f}_\ell(x) = \frac{1}{n}\sum_{k=1}^{n}(-x)\frac{\sin(\pi\ell(Y_k-x))}{\pi(Y_k-x)^2} + \frac{Y_k}{Y_k-x}\,\ell\,\cos(\pi\ell(Y_k-x)). \tag{28}$$

Alternatively, using a truncated series expansion gives an approximation of the estimator which may be easier to compute in practice. Let

$$\widetilde{f}_\ell^{(n)}(x) = \sum_{|j|\le L_n}\tilde{a}_{\ell,j}\theta_{\ell,j}(x) \quad \text{where} \quad \tilde{a}_{\ell,j} = \frac{1}{n}\sum_{k=1}^{n}\theta_{\ell,j}(Y_k) + Y_k\theta'_{\ell,j}(Y_k). \tag{29}$$

We can prove:

Proposition 4.4. *Assume that $\mathbb{E}(X_1^2) < +\infty$. The estimator $\widetilde{f}_\ell$ satisfies*

$$\mathbb{E}(\|\widetilde{f}_\ell - f\|^2) \le \|f-\bar{f}_\ell\|^2 + \frac{\pi^2\ell^3\mathbb{E}Y_1^2}{3n} + \frac{\ell}{n}. \tag{30}$$

If moreover $M_2 = \int x^2 f^2(x)dx < +\infty$,

$$\mathbb{E}(\|\widetilde{f}_\ell^{(n)} - f\|^2) \le \|f-\bar{f}_\ell\|^2 + \frac{\ell}{n}\left(1+\mathbb{E}Y_1^2\frac{\pi^2\ell^2}{3}\right) + 4(M_2+1)\frac{\ell^2}{L_n}. \tag{31}$$

As we consider estimators $\widetilde{f}_\ell$ with bounded variance, we will impose $\ell^3 \le n$. Consequently, if $L_n \ge n$, the last residual term ℓ^2/L_n is of order $O(\ell^2/n)$ and thus less than the variance term ℓ^3/n; if $L_n \ge n^{5/3}$, it is of order $O(1/n)$ and thus negligible.

In this context, Sobolev balls are defined by

$$\mathcal{W}^s(L) = \{f \in \mathbb{L}^2(\mathbb{R}),\ \int_{\mathbb{R}}(1+t^{2s})|f^*(t)|^2dt \le L < +\infty\}. \tag{32}$$

It is easy to see that, if $f \in \mathcal{W}^s(L)$, then, from (26), we have $\|f-\bar{f}_\ell\| \le L\ell^{-2s}$. Choosing $\ell_{\rm opt} = cn^{-1/(2s+3)}$, implies, from (30), that $\mathbb{E}(\|\widetilde{f}_{\ell_{\rm opt}} - f\|^2) \le Cn^{-2s/(2s+3)}$. This optimal rate is identical to the rate obtained for $\hat{f}_{m_{\rm opt}}$ under $\mathbb{E}(|X_1|^{2+2/3}) < +\infty$. Consequently, the deconvolution and Hermite estimators are asymptotically equivalent (with $\ell = \sqrt{m}$). However, from the computational efficiency point of view, the Hermite estimator is to be preferred. In Belomestny *et al.* (2017), Section 4.5, a notion of complexity is defined. Hermite estimators are proved to have much lower complexity than deconvolution estimators in the case of direct observations. Analogous computations of complexity in the present indirect observations case lead to the same conclusions.

5. Projection estimator of f when $Z_i = X_i + \Sigma_i$ are observed

5.1. **Estimation strategy in the Laguerre case.** We consider Model (2) where X_i and Σ_i are nonnegative. It holds that

$$f_Z(x) = f\star f_\Sigma(x) = \int_0^x f(u)f_\Sigma(x-u)du.$$

The convolution property of the Laguerre functions $(\varphi_j)_j$ given in formula (49) allows to write

$$\begin{aligned}\sum_{k=0}^{\infty}a_k(f_Z)\varphi_k(x) &= \sum_{j=0}^{+\infty}\sum_{k=0}^{+\infty}a_j(f)a_k(f_\Sigma)\varphi_j\star\varphi_k(x)\\ &= \sum_{k=0}^{\infty}\varphi_k(x)\sum_{\ell=0}^{k}2^{-1/2}\,(a_{k-\ell}(f_\Sigma)-a_{k-\ell-1}(f_\Sigma))a_\ell(f),\end{aligned}$$

with convention $a_k(f) = 0$ if $k < 0$.

Define the $m \times m$ triangular matrix $\mathbf{\Sigma}_m = (\sigma_{i,j})_{0\le i,j\le m-1}$ where

$$\sigma_{i,j} = 2^{-1/2}(\langle f_\Sigma, \varphi_{i-j}\rangle \mathbf{1}_{i-j\ge 0} - \langle f_\Sigma, \varphi_{i-j-1}\rangle \mathbf{1}_{i-j-1\ge 0}). \tag{33}$$

As $\sigma_{i,j} = \sigma(i-j)\mathbf{1}_{i-j\ge 0}$, $\mathbf{\Sigma}_m$ is a Toeplitz triangular matrix with diagonal elements $\sigma_{i,i} = 2^{-1/2}\langle f_\Sigma, \varphi_0\rangle > 0$. It is thus invertible and for all $m \ge 1$,

$$\vec{a}_m(f_Y) = {}^t(a_j(f))_{0\le j\le m-1} = \mathbf{\Sigma}_m^{-1}[(a_j(f_Z))_{0\le j\le m-1}] = \mathbf{\Sigma}_m^{-1}\vec{a}_m(f_Z), \tag{34}$$

The projection estimator of f on S_m based on $(Z_1,\dots,Z_n)$ is given by

$$\tilde{f}_m = \sum_{j=0}^{m-1} \tilde{a}_j\varphi_j, \quad \vec{\tilde{a}}_m = {}^t(\tilde{a}_j)_{0\le j\le m-1} = \mathbf{\Sigma}_m^{-1}\vec{\hat{a}}_m(Z), \qquad m \ge 1 \tag{35}$$

where $\vec{\hat{a}}_m(Z) = [(\hat{a}_j(Z))_{0\le j\le m-1}]$ and $\hat{a}_j(Z)$ is defined by

$$\hat{a}_j(Z) := \frac{1}{n}\sum_{i=1}^{n} \varphi_j(Z_i). \tag{36}$$

This estimator is proposed in Mabon (2017) and inspired from Comte *et al.* (2017). The following risk bound is an improvement of the bound given in Mabon (2017).

Proposition 5.1. *Assume that $\|f_\Sigma\|_\infty < +\infty$ and either $\mathbb{E}(X_1^{-1/2}) < +\infty$ or $\mathbb{E}(\Sigma_1^{-1/2}) < +\infty$. Let $\tilde{f}_m$ be given by (35). Then we have*

$$\mathbb{E}(\|\tilde{f}_m - f\|^2) \le \|f - f_m\|^2 + \frac{[c\sqrt{m}\|\mathbf{\Sigma}_m^{-1}\|_{\text{op}}^2] \wedge [\|f_\Sigma\|_\infty \|\mathbf{\Sigma}_m^{-1}\|_F^2]}{n}$$

*where $\|\mathbf{A}\|_F^2 = \text{Tr}({}^t\mathbf{A}\mathbf{A})$ (*Tr *denotes the trace of the matrix) and $\|\mathbf{A}\|_{\text{op}}^2 = \lambda_{\max}({}^t\mathbf{A}\mathbf{A})$ is the maximal eigenvalue of ${}^t\mathbf{A}\mathbf{A}$.*

The bias term is unchanged. The variance term is increasing in m because of the special form of $\mathbf{\Sigma}_m^{-1}$ (lower triangular and Toeplitz, see Mabon (2017)). If $\Sigma_i = 0$ and thus $\mathbf{\Sigma}_m = Id$, then $\|\mathbf{\Sigma}_m^{-1}\|_{\text{op}}^2 = 1$ and $\|\mathbf{\Sigma}_m^{-1}\|_F^2 = m$ thus the variance term is of order $\sqrt{m}/n$ as expected.

In Comte et *al.* (2017), the order of $\|\mathbf{\Sigma}_m^{-1}\|_F^2$ in function of m is studied. In particular, if Σ_i has a Gamma distribution $\Gamma(r,\lambda)$, $r \in \mathbb{N}$, $r \ge 1$, there exist constants c, C such that

$$cm^{2r} \le \|\mathbf{\Sigma}_m^{-1}\|_{\text{op}}^2 \le \|\mathbf{\Sigma}_m^{-1}\|_F^2 \le Cm^{2r}.$$

Therefore the following corollary holds:

Corollary 5.1. *Assume that $f \in W_L^s(D)$, and that Σ_i has a Gamma distribution $\Gamma(r,\lambda)$, r integer, $r \ge 1$. Then $\tilde{f}_m$ given by (35) satisfies, for $m_{opt} = [n^{2r+s}]$*

$$\mathbb{E}(\|\tilde{f}_{m_{opt}} - f\|^2) \le C(s,D)n^{-s/(2r+s)}.$$

The matrix $\mathbf{\Sigma}_m$ in the case of gamma noise with integer tail parameter can be computed explicitely: we give it for $r = 1, 2$ hereafter.

Remark 5.1.

- For $\Sigma_1 \sim \mathcal{E}(\lambda) = \gamma(1,\lambda)$, we have $[\mathbf{\Sigma}_m]_{i,i} = \lambda/(1+\lambda)$ and

$$[\mathbf{\Sigma}_m]_{i,j} = -2\lambda\frac{(\lambda-1)^{i-j-1}}{(\lambda+1)^{(i-j+1)}} \ \text{ if } j < i \quad \text{ and } [\mathbf{\Sigma}_m]_{i,j} = 0 \text{ otherwise.} \tag{37}$$

  We can compute $[\mathbf{\Sigma}_m^{-1}]_{i,j} = (\lambda+1)/\lambda$ if $i = j$, $2/\lambda$ if $i > j$ and 0 otherwise. Note that

$$\|\mathbf{\Sigma}_m^{-1}\|_F^2 = 2\frac{m^2}{\lambda^2} + m(1 + \frac{2}{\lambda} - \frac{1}{\lambda^2}).$$

- For $\Sigma_1 \sim \Gamma(2,\mu)$, we have $[\mathbf{\Sigma}_m]_{i,i} = (\mu/(1+\mu))^2$, $[\mathbf{\Sigma}_m]_{i+1,i} = -4\mu^2/(1+\mu)^3$ and

$$[\mathbf{\Sigma}_m]_{i,j} = 4(i-j-\mu)\mu^2 \frac{(\mu-1)^{i-j-2}}{(\mu+1)^{i-j+2}} \text{ if } i > j+1 \quad \text{and } [\mathbf{\Sigma}_m]_{i,j} = 0 \text{ otherwise.} \tag{38}$$

We can propose a method to select m automatically, aiming at a data driven bias variance compromise and simplifying Mabon (2017). We define

$$\tilde{m} = \arg\min_{m\in\mathcal{M}_n} \left(-\|\tilde{f}_m\|^2 + \widetilde{\mathrm{pen}}(m)\right) \qquad \text{with } \widetilde{\mathrm{pen}}(m) = \kappa \frac{\log(2+\|\mathbf{\Sigma}_m^{-1}\|_F^2)\|\mathbf{\Sigma}_m^{-1}\|_F^2}{n} \tag{39}$$

where

$$\mathcal{M}_n = \{m \in \mathbb{N}^*,\ m \le n/\log(2+n),\ \|\mathbf{\Sigma}_m^{-1}\|_F^2 \le n\}.$$

Theorem 5.1. *Let $\tilde{f}_m$ be given by (35) and $\tilde{m}$ by (39). There exists a numerical constant κ_0 such that for any $\kappa \ge \kappa_0$, we have*

$$\mathbb{E}(\|\tilde{f}_{\tilde{m}} - f\|^2) \le C_1 \inf_{m\in\mathcal{M}_n} \left(\|f - f_m\|^2 + \widetilde{\mathrm{pen}}(m)\right) + \frac{C_2}{n}.$$

In the penalty, we have chosen $\|\mathbf{\Sigma}_m^{-1}\|_F^2$ rather than $\sqrt{m}\|\mathbf{\Sigma}_m^{-1}\|_{\mathrm{op}}^2$, which is easier to handle.

5.2. **Estimation strategy in the Hermite case.** We consider Model (2) where now $(X_i)_{1\le i\le n}$ and $(\Sigma_i)_{1\le i\le n}$ are real valued. This model is classically dealt with by Fourier deconvolution, and yields the following estimator

$$f_\ell^\diamond(x) := \frac{1}{2\pi}\int_{-\pi\ell}^{\pi\ell} e^{-itx}\frac{\widehat{f_Z^*}(t)}{f_\Sigma^*(-t)}dt \qquad \text{where} \qquad \widehat{f_Z^*}(t) = \frac{1}{n}\sum_{k=1}^n e^{itZ_k}$$

is the empirical characteristic function of Z.

Here we propose projection estimators on the Hermite basis. We use that $h_j^*(t) = \sqrt{2\pi}i^j h_j(t)$ and write that $f_Z = f \star f_\Sigma$ and thus $f_Z^* = f^* f_\Sigma^*$, where $f^* = \sqrt{2\pi}\sum_{j\ge 0} a_j(f) i^j h_j$. Therefore,

$$a_j(f) = \frac{(-i)^k}{\sqrt{2\pi}}\langle f_Z^*, \frac{h_j}{f_\Sigma^*(-.)}\rangle.$$

This leads to define the estimator of $a_j(f)$ as follows:

$$\check{a}_{j,\ell} = \frac{(-i)^j}{\sqrt{2\pi}}\int_{-\pi\ell}^{\pi\ell} \widehat{f_Z^*}(t)\,\frac{h_j(t)}{f_\Sigma^*(t)}dt, \tag{40}$$

and the projection estimator of f as

$$\check{f}_m(x) = \sum_{j=0}^{m-1} \check{a}_{j,\sqrt{m}}\, h_j(x). \tag{41}$$

Note that we have chosen $\ell = \sqrt{m}$ in (40). The following risk bound holds.

Proposition 5.2. *Consider $\check{f}_m$ given by (41) and (40). Then*

$$\mathbb{E}(\|\check{f}_m - f\|^2) \le \sum_{j\ge m} a_j^2(f) + \frac{1}{2\pi}\int_{|t|\ge\pi\sqrt{m}} |f^*(t)|^2 dt + \frac{1}{2\pi\, n}\int_{-\pi\sqrt{m}}^{\pi\sqrt{m}} \frac{1}{|f_\Sigma^*(t)|^2}dt.$$

Thus, if $f \in W_H^s(L) \subset \mathcal{W}^s(L')$ (see (9) and (32) and Appendix A), then the two bias terms in Proposition 5.2 have the same order:

$$\sum_{j\geq m} a_j^2(f) + \frac{1}{2\pi}\int_{|t|\geq \pi\sqrt{m}} |f^*(t)|^2 dt \leq Lm^{-s} + \frac{L'}{2\pi}(\pi\sqrt{m})^{-2s} = Cm^{-s}.$$

For comparison,

$$\mathbb{E}(\|f^{\diamond}_{\sqrt{m}} - f\|^2) \leq \frac{1}{2\pi}\int_{|t|\geq \pi\sqrt{m}} |f^*(t)|^2 dt + \frac{1}{2\pi\, n}\int_{-\pi\sqrt{m}}^{\pi\sqrt{m}} \frac{1}{|f^*_\Sigma(t)|^2} dt.$$

This shows that the two estimators have asymptotically the same rate. The deconvolution estimator has a smaller risk bound. However, in terms of computational efficiency, as described in Belomestny *al.* (2017), the Hermite estimator is to be preferred as it has lower complexity. For sake of brevity, we do not develop the adaptive procedure.

6. Extensions and concluding remarks

In this paper, the use of a Laguerre basis to estimate a function $f \in \mathbb{L}^2(\mathbb{R}^+)$ or a Hermite basis to estimate a function $f \in \mathbb{L}^2(\mathbb{R})$ is illustrated in examples of inverse problems. Projection estimators which are easy to implement are built and studied. Data-driven choices of the projection dimension can be proposed leading to adaptive estimators.

Using formulae of Section 7.1.2, the estimation of a function $f \in \mathbb{L}^2(\mathbb{R}^+)$ by projection estimators on a Laguerre basis $(\varphi_j^{(\delta)})$ for all $\delta > -1$ is possible for direct observations or multiplicative censored observations. Thanks to Lemma 7.2, we can obtain risk bounds of the same type. In the case of multiplicative censoring, setting $a_j^{(\delta)}(f) = \langle f, \varphi_j^{(\delta)}\rangle_+$, and using relation (50), we obtain:

$$a_j^{(\delta)}(f) = \frac{\sqrt{(j+1)(j+\delta+1)}}{2} a_{j+1}^{(\delta)}(f_Y) + \frac{1}{2} a_j^{(\delta)}(f_Y) - \frac{\sqrt{j(j+\delta)}}{2} a_{j-1}^{(\delta)}(f_Y). \tag{42}$$

This allows to define analogously a matrix $\mathbf{H}_m^{(\delta)}$ which helps practical computing of the projection estimator. On the contrary, except for $\delta = 0$, the bases $(\varphi_j^{(\delta)})$ do not seem fitted to the deconvolution setting of densities on $\mathbb{R}^+$.

An extension of the results presented here could be to study the combination of models (1) and (2) and estimate f, the density of X_i if observations are either $(X_i + \Sigma_i)U_i$ or $X_iU_i + \Sigma_i$.

7. Proofs

7.1. **Formulae for Laguerre functions.**

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:

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

The sequence $(\phi_k^{(\delta)}(x) = L_k^{(\delta)}(x)x^{\delta/2}e^{-x/2}\left(\frac{k!}{\Gamma(k+\alpha+1)}\right)^{1/2})$ is an orthonormal basis of $\mathbb{L}^2(\mathbb{R}^+)$. In this paper, we rather consider the Laguerre functions with index δ, given by

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

which are more convenient for computing derivatives or integrals of the basis functions especially when $\delta = 0$. The family $(\varphi_k^{(\delta)})_{k\geq 0}$ is an orthonormal basis of $\mathbb{L}^2(\mathbb{R}^+)$.

7.1.1. *Formulae for $\delta = 0$.* For $\delta = 0$, we set $L_k^{(0)} = L_k$, $\varphi_k^{(0)} = \varphi_k$. Formula (22.7.12) in Abramowitz and Stegun (1964) states that

$$xL_j(x) = -(j+1)L_{j+1}(x) + (2j+1)L_j(x) - jL_{j-1}(x), \quad xL_j'(x) = j(L_j(x) - L_{j-1}(x)). \tag{45}$$

implying

$$x\varphi_j(x) = -\frac{j+1}{2}\varphi_{j+1}(x) + (j+\frac{1}{2})\varphi_j(x) - \frac{j}{2}\varphi_{j-1}(x), \tag{46}$$

$$(x\varphi_j(x))' = \varphi_j(x) + x\varphi_j'(x) = -\frac{j}{2}\varphi_{j-1}(x) + \frac{1}{2}\varphi_j(x) + \frac{j+1}{2}\varphi_{j+1}(x). \tag{47}$$

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

$$\varphi_j'(x) = -\varphi_j(x) - \sqrt{\frac{2j}{x}}\,\varphi_{j-1}^{(1)}(x). \tag{48}$$

The following convolution property (formula 22.13.14 in Abramowitz and Stegun (1964)) makes the Laguerre basis relevant in the deconvolution setting

$$\varphi_k \star \varphi_j(x) = \int_0^x \varphi_k(u)\varphi_j(x-u)du = 2^{-1/2}\left(\varphi_{k+j}(x) - \varphi_{k+j+1}(x)\right) \tag{49}$$

where $\star$ stands for the convolution product.

Lemma 7.1. *For all $x \geq 0$, $\varphi_0'(x) = -\varphi_0(x)$, $\quad \varphi_j'(x) = -\varphi_j(x) - 2\sum_{k=0}^{j-1}\varphi_k(x), j \geq 1$. Moreover, $|\varphi_j'(x)| \leq \sqrt{2}(2j+1) \leq 2\sqrt{2}(j+1)$, $\quad |x\varphi_j'(x) + \varphi_j(x)| \leq \sqrt{2}(j+1)$.*

Proof of Lemma 7.1. The following equality holds $\varphi_j'(x) = -\varphi_j(x) + 2\sqrt{2}e^{-x}L_j'(2x)$ which is a polynomial function of degree j multiplied by e^{-x}. Thus, it can be decomposed as $\varphi_j'(x) = \sum_{k=0}^{j} a_k^{(j)}\varphi_k(x)$ with

$$\begin{aligned} a_k^{(j)} &= \langle \varphi_j', \varphi_k\rangle = \int_0^{+\infty}\varphi_j'(x)\varphi_k(x)dx = [\varphi_j(x)\varphi_k(x)]_0^{+\infty} - \int_0^{+\infty}\varphi_j(x)\varphi_k'(x)dx \\ &= -\varphi_j(0)\varphi_k(0) - \int_0^{+\infty}\varphi_j(x)\varphi_k'(x)dx = -2 - \langle\varphi_j, \varphi_k'\rangle = -2 - a_j^{(k)} \end{aligned}$$

Notice that this formula is also true when $k=j$: $\langle \varphi_j', \varphi_j\rangle = \int_0^{+\infty} \varphi_j'(x)\varphi_j(x)dx = -(1/2)\varphi_j^2(0) = -2/2 = -1$. Thus we obtain:

$$\begin{aligned}\varphi_j'(x) &= \sum_{k=0}^{j} a_k^{(j)}\varphi_k(x) = -2\sum_{k=0}^{j}\varphi_k(x) - \sum_{k=0}^{j}\langle\varphi_j,\varphi_k'\rangle\varphi_k(x)\\ &= -\varphi_j(x) - 2\sum_{k=0}^{j-1}\varphi_k(x) - \sum_{k=0}^{j-1}\langle\varphi_j,\varphi_k'\rangle\varphi_k(x)\end{aligned}$$

Note that the $\langle\varphi_j,\varphi_k'\rangle$ are zero for $k \le j-1$. Thus we obtain the first formula. The bound on $\varphi_j'(x)$ follows from $|\varphi_j(x)| \le \sqrt{2}$ and the bound on $|x\varphi_j'(x)+\varphi_j(x)|$ from (47). □

7.1.2. *Formulae for $\delta > -1$.* The following holds:

$$xL_j^{(\delta)}(x) = -(j+1)L_{j+1}^{(\delta)}(x) + (2j+\delta+1)L_j^{(\delta)}(x) - (j+\delta)L_{j-1}^{(\delta)}(x),$$

$$x(L_j^{(\delta)})'(x) = jL_j^{(\delta)}(x) - (j+\delta)L_{j-1}^{(\delta)}(x),$$

implying

$$x\varphi_j^{(\delta)}(x) = -\frac{\sqrt{(j+1)(j+\delta+1)}}{2}\varphi_{j+1}^{(\delta)}(x) + (j+\frac{\delta+1}{2})\varphi_j^{(\delta)}(x) - \frac{\sqrt{j(j+\delta)}}{2}\varphi_{j-1}^{(\delta)}(x),$$

$$(x\varphi_j^{(\delta)}(x))' = -\frac{\sqrt{j(j+\delta)}}{2}\varphi_{j-1}^{(\delta)}(x) + \frac{1}{2}\varphi_j^{(\delta)}(x) + \frac{\sqrt{(j+1)(j+\delta+1)}}{2}\varphi_{j+1}^{(\delta)}(x). \tag{50}$$

7.1.3. *Asymptotic formulae.* From Askey and Wainger (1965), we have for $\nu = 4k+2\delta+2$, and k large enough

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

where γ_1 and γ_2 are positive and fixed constants. From these estimates, we can prove

Lemma 7.2. *Let p be a nonnegative real number. Assume that a random variable R has density f_R on $\mathbb{R}^+$ and that $\mathbb{E}(R^{p-1/2}) < +\infty$ and $\mathbb{E}(R^p) < +\infty$. For k large enough,*

$$\int_0^{+\infty} x^p[\varphi_k^{(\delta)}(x)]^2 f_R(x)dx \le \frac{c}{\sqrt{k}},$$

where $c = c_p > 0$ is a constant depending on p and $\mathbb{E}(R^{p-1/2})$, $\mathbb{E}(R^p)$.

Proof of Lemma 7.2. We have six terms to compute to find the order of

$$\int_0^{+\infty} x^p[\varphi_k^{(\delta)}(x)]^2 f_R(x)dx = (1/2^{p+1})\int_0^{+\infty} u^p[\varphi_k^{(\delta)}(u/2)]^2 f_R(u/2)du := \sum_{\ell=1}^{6} I_\ell.$$

a) $I_1 \lesssim \frac{1}{2^{p+1}}\int_0^{1/\nu} u^p(u\nu)^\delta f_R(u/2)du \lesssim \|f_R\|\nu^{-(p+1/2)} \lesssim \|f_R\|k^{-(p+1/2)}$.

b) $I_2 \lesssim \nu^{-1/2}\int_{1/\nu}^{\nu/2} f_R(u/2)u^{p-1/2}du \lesssim k^{-1/2}\mathbb{E}(R^{p-1/2})$.

c) $I_3 \lesssim \nu^{-1/2}\nu^{-1/6}\int_{\nu/2}^{\nu-\nu^{1/3}} u^p f_R(u/2)du = o(1/\sqrt{k})$, as $\nu - u \geq \nu^{1/3}$ and $\mathbb{E}(R^p) < +\infty$.

d) $I_4 \lesssim \nu^{-2/3}\int_{\nu-\nu^{1/3}}^{\nu+\nu^{1/3}} u^p f_R(u/2)du = o(1/\sqrt{k})$ as $\mathbb{E}(R^p) < +\infty$.

e) $I_5 \lesssim \nu^{-1/2}\int_{\nu+\nu^{1/3}}^{3\nu/2} u^p(u-\nu)^{-1/2} f_R(u/2)du \lesssim \nu^{-1/2}\nu^{-1/6}\mathbb{E}(R^p) = o(1/\sqrt{k})$,
(exp is bounded by 1, $u - \nu \geq \nu^{1/3}$).

f) $I_6 \lesssim e^{-\gamma_2(3\nu/2)}\mathbb{E}(R^p) = o(1/\sqrt{k})$.
The result of Lemma 7.2 follows from these orders. □

7.2. **Formulae for Hermite functions.** Using (see Abramowitz and Stegun (1964))

$$2xH_j(x) = H_{j+1}(x) + 2jH_{j-1}(x), \quad H_j'(x) = 2jH_{j-1}(x), \quad j \geq 1. \tag{51}$$

we get:

$$\sqrt{2}h_j'(x) = \sqrt{j}\; h_{j-1}(x) - \sqrt{j+1}h_{j+1}(x), 2x\; h_j(x) = \sqrt{2(j+1)}\; h_{j+1}(x) + \sqrt{2j}\; h_{j-1}(x). \tag{52}$$

Moreover, $h_j^*(t) = \sqrt{2\pi} i^j h_j(t)$.

Lemma 7.3. *There exist constants C'_∞, $C"_\infty$ such that, for all $j \geq 0$,*

(1) $\|h_j'\|_\infty \leq C'_\infty (j+1)^{5/12}$,
(2) $\|yh_j' + h_j\|_\infty \leq C"_\infty (j+1)^{11/12}$.

Proof of Lemma 7.3. For (1), we use (52) and the first bound:

$$\|h_j'\|_\infty \leq \frac{C_\infty}{\sqrt{2}}\left(1 + \frac{\sqrt{j+1}}{(j+2)^{1/12}}\right) \leq \sqrt{2}C_\infty (j+1)^{5/12}.$$

Next, (52) implies that

$$\begin{aligned}(yh_j(y))' &= \sqrt{(j+1)/2}h'_{j+1}(y) + \sqrt{j/2}h'_{j-1}(y)\\ &= \frac{1}{2}\left(\sqrt{j(j-1)}h_{j-2}(y) + h_j(y) - \sqrt{(j+1)(j+2)}h_{j+2}(y)\right),\end{aligned}$$

where the last line follows from (52). Thus, $|yh_j'(y) + h_j(y)| \leq 2C_\infty (j+1)^{11/12}$. So we get (2). □

The Hermite polynomials are linked with the Laguerre polynomials as follows (see Abramowitz and Stegun, 1964, p.779, 22.5.40, 22.5.41). For $x \geq 0$,

$$H_{2n}(x) = (-1)^n 2^{2n} n! L_n^{(-1/2)}(x^2), \quad H_{2n+1}(x) = (-1)^n 2^{2n+1} n! x\; L_n^{(1/2)}(x^2). \tag{53}$$

The Hermite polynomials are even for even n and odd for odd n. From this, we deduce the following link between Hermite and Laguerre functions:

Lemma 7.4. *For $x \geq 0$,*

$$h_{2n}(x) = (-1)^n \sqrt{x/2}\; \varphi_n^{(-1/2)}(x^2/2), \quad h_{2n+1}(x) = (-1)^n \sqrt{x/2}\; \varphi_n^{(1/2)}(x^2/2).$$

Proof of Lemma 7.4. We have

$$h_{2n}(x) = c_{2n} e^{-x^2/2} H_{2n}(x) = (-1)^n d_{2n}\; L_n^{(-1/2)}(x^2) e^{-x^2/2} (x^2)^{-1/4} (x^2)^{1/4},$$

with $c_{2n}, d_{2n} > 0$. Thus, $h_{2n}(x) = (-1)^n d_{2n}\ \phi_n^{(-1/2)}(x^2)\ \sqrt{x} = (-1)^n d_{2n}\ \sqrt{x/2}\ \varphi_n^{(-1/2)}(x^2/2)$. Therefore,

$$\int_0^\infty h_{2n}^2(x)dx = 1/2 = d_{2n}^2 \int_0^\infty \left(\phi_n^{(-1/2)}(x^2)\right)^2 x\ dx = d_n^2/2.$$

This shows $d_{2n}^2 = 1$, hence $d_{2n} = 1$. Analogously,

$$h_{2n+1}(x) = c_{2n+1} e^{-x^2/2} H_{2n+1}(x) = (-1)^n d_{2n+1}\ \sqrt{x}\ L_n^{(1/2)}(x^2) e^{-x^2/2} \sqrt{x},$$

with $c_{2n+1}, d_{2n+1} > 0$. We get $h_{2n+1}(x) = (-1)^n d_{2n+1}\ \phi_n^{(1/2)}(x^2)\ \sqrt{x}$. We conclude as above that $d_{2n+1} = 1$. □

7.3. **Asymptotic formulae.** Now, we can use the estimates of Askey and Wainger (1965) for $\varphi_n^{(\delta)}$ to obtain the following result:

Lemma 7.5. *Let p be a nonnegative real number. Assume that a random variable R has density f_R on $\mathbb{R}$ and that $\mathbb{E}(|R|^{p+2/3}) < +\infty$. For k large enough,*

$$\int_0^{+\infty} x^p h_k^2(x) f_R(x) dx \le \frac{c}{\sqrt{k}},$$

where $c > 0$ is a constant which depends on p.

Proof of Lemma 7.5. We start with the even indexes. Again,we have six terms to compute to find the order of

$$\int_0^{+\infty} x^p h_{2k}^2(x) f_R(x) dx = (1/2) \int_0^{+\infty} x^{p+1}\ \left(\varphi_n^{(-1/2)}(x^2/2)\right)^2 f_R(x) dx := \sum_{\ell=1}^{6} J_\ell.$$

We take $\nu = 4k+1$.

$$\begin{aligned}
J_1 &= \int_0^{1/\sqrt{\nu}} x^p h_{2k}^2(x) f_R(x) dx = \frac{1}{2} \int_0^{1/\sqrt{\nu}} x^{p+1}\ \left(\varphi_n^{(-1/2)}(x^2/2)\right)^2 f_R(x) dx \\
&\le C \int_0^{1/\sqrt{\nu}} x^{p+1}\ \left((x^2\nu)^{-1/4}\right)^2 f_R(x) dx \le \frac{C}{\nu^{(p+1)/2}} \int_0^{1/\sqrt{\nu}} f_R(x) dx \le \frac{C}{\nu^{(p+1)/2}}, \\
J_2 &= \int_{1/\sqrt{\nu}}^{\sqrt{\nu/2}} x^p h_{2k}^2(x) f_R(x) dx \le \frac{C}{2} \int_{1/\sqrt{\nu}}^{\sqrt{\nu/2}} x^{p+1} \left(x^2\nu\right)^{-1/2} f_R(x) dx \le \frac{C}{2\sqrt{\nu}} \mathbb{E}(|R|^p), \\
J_3 &= \int_{\sqrt{\nu/2}}^{(\nu-\nu^{1/3})^{1/2}} x^p h_{2k}^2(x) f_R(x) dx \\
&\le \frac{C}{2} \int_{\sqrt{\nu/2}}^{(\nu-\nu^{1/3})^{1/2}} x^{1/3} x^{p+2/3} \nu^{-1/2} (\nu - x^2)^{-1/2} f_R(x) dx \le \frac{C}{2\sqrt{\nu}} \mathbb{E}(|R|^{p+2/3}), \\
J_4 &= \int_{(\nu-\nu^{1/3})^{1/2}}^{(\nu+\nu^{1/3})^{1/2}} x^p h_{2k}^2(x) f_R(x) dx \le \frac{C}{2} \nu^{-1/2} \mathbb{E}(|R|^{p+2/3}), \\
J_5 &= \int_{(\nu+\nu^{1/3})^{1/2}}^{\sqrt{3\nu/2}} x^p h_{2k}^2(x) f_R(x) dx \le \frac{C}{2\sqrt{\nu}} \mathbb{E}(|R|^{p+2/3}), \\
J_6 &= \int_{\sqrt{3\nu/2}}^{+\infty} x^p h_{2k}^2(x) f_R(x) dx \le C' \exp\left(-3\gamma_2\nu/2\right).
\end{aligned}$$

Now, we deal with

$$\int_0^{+\infty} x^p h_{2k+1}^2(x) f_R(x)dx = (1/2)\int_0^{+\infty} x^{p+1} \left(\varphi_n^{(1/2)}(x^2/2)\right)^2 f_R(x)dx := \sum_{\ell=1}^{6} K_\ell,$$

and take $\nu = 4k+3$. The only difference is on the first term:

$$K_1 = \int_0^{1/\sqrt{\nu}} x^p h_{2k+1}^2(x) f_R(x)dx \le \frac{C}{2}\int_0^{1/\sqrt{\nu}} x^p (x^2\nu)^{1/2} f_R(x)dx \le \frac{C}{2\nu^{(p+1)/2}}.$$

This ends the proof. □

7.4. **Proof of Proposition 3.1.** We first prove the result in the Laguerre case. From Lemma 7.2 with $\delta = 0$, $p = 0$, we get that, if $\mathbb{E}(1/\sqrt{X_1}) < +\infty$, then for j large enough,

$$\mathbb{E}(\varphi_j^2(X_1)) = \int_0^{+\infty} \varphi_j^2(x) f(x)dx \le \frac{c}{\sqrt{j}},$$

where $c > 0$ is a constant. As a consequence, for m large enough,

$$\mathbb{E}(\|\hat f_m^X - f_m\|^2) \le \sum_{j=0}^{m-1} \mathbb{E}[\varphi_j^2(X_1)] \lesssim m^{1/2},$$

which gives the announced result.

In the Hermite case, we apply Lemma 7.5 for $p = 0$. Thus for k large enough, $V_m^X \lesssim \Phi_0^2 + \sum_{k=1}^{m-1} k^{-1/2} = O(\sqrt{m})$, and the result follows. □

7.5. **Proof of Proposition 3.2.** We only study the Laguerre case, as the Hermite case is proved in Belomestny *et al.* (2017), Proposition 2.2. By Theorem 8.22.5 in Szego (1959), reminding formula (44), we have: for all $\delta > -1$, and for $\underline{b}/k \le x \le \bar b$, where $\underline{b}, \bar b$ are arbitrary constants,

$$\varphi_k^{(\delta)}(x) = c(kx)^{-1/4}\left(\cos(2\sqrt{2}\sqrt{kx} - \delta\pi/2 - \pi/4) + (kx)^{-1/2}O(1)\right), \tag{54}$$

where $O(1)$ is uniform on $[\underline{b}/k, \bar b]$ and $c = 2^{1/4}/\sqrt{\pi}$.

Take k such that $\underline{b}/k < a < b < \bar b$ and set $d = \inf_{a\le x\le b} f(x)$. Then write

$$\int_0^{+\infty} \varphi_k^2(x) f(x)dx \ge d\int_a^b \varphi_k^2(x)dx.$$

We have,

$$\varphi_k^2(x) = \frac{c^2}{2}(kx)^{-1/2}(1+\sin(4\sqrt{2kx})) + (kx)^{-1}O(1).$$

Therefore, as $\int_a^b \sin(4\sqrt{2kx})dx = 2\int_{\sqrt{a}}^{\sqrt{b}} \sin(4\sqrt{2k}u)udu = O(1/\sqrt{k})$,

$$\int_a^b \varphi_k^2(x)dx \ge \frac{c^2}{2\sqrt{a}} k^{-1/2}\left(b - a + O(\frac{1}{\sqrt{k}})\right) + O(\frac{1}{k}).$$

Consequently, for k large enough, we get

$$\int_0^{+\infty} \varphi_k^2(x) f(x)dx \ge c'/\sqrt{k}. \quad \square$$

7.6. **Proof of Proposition 3.3.** Let X have density $f(x) = \int_0^\infty \theta\ \exp(-\theta x) d\Pi(\theta)$. Then, $X =_{\mathcal{L}} Z/T$ where Z, T are independent, Z has exponential distribution with parameter 1 and T has distribution Π. For fixed $\theta > 0$, elementary computations yield

$$\mathbb{E}\varphi_j(Z/\theta) = \int_0^{+\infty} \theta\ \exp(-\theta x)\varphi_j(x)dx = \sqrt{2}\frac{\theta}{\theta+1}\left(\frac{\theta-1}{\theta+1}\right)^j.$$

Then, we compute

$$\mathbb{E}\varphi_j(X) = \mathbb{E}\left(\mathbb{E}(\varphi_j(Z/T)|T)\right) = \int_0^{+\infty} \sqrt{2}\frac{\theta}{\theta+1}\left(\frac{\theta-1}{\theta+1}\right)^j d\Pi(\theta) = \mathbb{E}\left(\sqrt{2}\frac{T}{T+1}\left(\frac{T-1}{T+1}\right)^j\right)$$

Consequently,

$$\sum_{j\geq m} a_j^2(f) = \sum_{j\geq m} 2\left[\mathbb{E}\frac{T}{T+1}\left(\frac{T-1}{T+1}\right)^j\right]^2 \leq \frac{1}{2}\mathbb{E}T\left(\frac{T-1}{T+1}\right)^{2m} \leq \frac{v}{2}\rho^{2m}.$$

Note that for $Z \sim \mathcal{E}(1)$,

$$\mathbb{E}\left(\frac{1}{\sqrt{X}}\right) \leq \sqrt{v}\mathbb{E}\left(\frac{1}{\sqrt{Z}}\right) < +\infty.$$

Choosing $m_{\text{opt}} = [\log(n)/|\log(\rho)|]$ and using Proposition 3.1, the result follows.$\square$

7.7. **Proof of Theorem 3.1.** The Hermite case is proved in Theorem 2.1 of Belomestny *et al.* (2017b).

The Laguerre case is identical, except for one point. We use that $\sup_x \sum_{j=0}^{m-1} \varphi_j^2(x) \leq 2m$, instead of $\sup_x \sum_{j=0}^{m-1} h_j^2(x) \leq C'_\infty m^{5/6}$. This explains the new bound on m_n. $\square$

7.8. **Proof of formula (13)-(14) and Lemma 4.1.** Equality (13) is elementary. For $y \geq 0$,

$$\begin{aligned}\bar{F}_Y(y) &= \int_y^{+\infty} f_Y(z)dz = \int_y^{+\infty}\int_z^{+\infty}\frac{f(x)}{x}dxdz = \int(\int_y^x dz)\frac{f(x)}{x}\mathbb{1}(y\leq x)dx \\ &= \int_y^{+\infty}(x-y)\frac{f(x)}{x}dx = \int_y^{+\infty} f(x)dx - y\int_y^{+\infty}\frac{f(x)}{x}dx = \bar{F}(y) - yf_Y(y).\end{aligned}$$

For $y \leq 0$,

$$\begin{aligned}F_Y(y) &= \int_{-\infty}^y f_Y(z)dz = \int_{-\infty}^y dz\int_{-\infty}^z \frac{f(x)}{|x|}dx = \int(\int_x^y dz)\frac{f(x)}{|x|}\mathbb{1}(x\leq y)dx \\ &= \int_{-\infty}^y (y-x)\frac{f(x)}{|x|}dx = yf_Y(y) + F(y)\end{aligned}$$

Thus, $\bar{F}_Y(y) = \bar{F}(y) - yf_Y(y)$, which is (14).

For (15), by (13), $yf_Y(y)$ tends to 0 as both y tends to $+\infty$ and $-\infty$. Integrating by parts yields

$$\begin{aligned}&\int_{\mathbb{R}} f_Y(y)(t(y)+yt'(y))dy = -\int_{\mathbb{R}} yt(y)(f_Y(y))'dy \\ &= -[\int_0^{+\infty} yt(y)(-\frac{f(y)}{y})dy + \int_{-\infty}^0 yt(y)\frac{f(y)}{|y|}dy] = \int_{-\infty}^{+\infty} t(y)f(y)dy.\end{aligned}$$

Note that $\mathbb{E}Y_1^2t^2(Y_1) \le \mathbb{E}X_1^2t^2(U_1X_1)$. Then,

$$\begin{aligned}\mathbb{E}X^2t^2(U_1X_1) &= \int_0^{+\infty} xf(x)(\int_0^x t^2(v)dv)dx + \int_{-\infty}^0 |x|f(x)(\int_x^0 t^2(v)dv)dx \\ &\le \mathbb{E}(|X_1|)\|t\|^2. \quad \square\end{aligned}$$

7.9. **Proof of Proposition 4.1.**

• **Laguerre case.** First, we bound V_m in the case $\psi_j = \varphi_j$. We write

$$\mathbb{E}\left[(Y_1\varphi_j'(Y_1) + \varphi_j(Y_1))^2\right] = \mathbb{E}\left[(Y_1\varphi_j'(Y_1))^2\right] + \mathbb{E}\left[2Y_1\varphi_j'(Y_1)\varphi_j(Y_1) + \varphi_j^2(Y_1)\right].$$

Formula (15) applied to $t = \varphi_j^2$ yields

$$\mathbb{E}\left[2Y_1\varphi_j'(Y_1)\varphi_j(Y_1) + \varphi_j^2(Y_1)\right] = \mathbb{E}(\varphi_j^2(X_1)).$$

Therefore

$$\mathbb{E}\left[(Y_1\varphi_j'(Y_1) + \varphi_j(Y_1))^2\right] = \mathbb{E}\left[(Y_1\varphi_j'(Y_1))^2\right] + \mathbb{E}(\varphi_j^2(X_1)),$$

where the second rhs term is the same as in the direct case, and thus $\mathbb{E}(\varphi_j^2(X_1)) \lesssim j^{-1/2}$ for j large enough and if $\mathbb{E}(1/\sqrt{X_1}) < +\infty$.

For the first rhs term, we use formula (48): $y\varphi_j'(y) = -\sqrt{2jy}\varphi_{j-1}^{(1)}(y) - y\varphi_j(y)$. Therefore, for j large enough, we have

$$\int_0^{+\infty}[y\varphi_j'(y)]^2 f_Y(y)dy \lesssim j\int_0^{+\infty} y[\varphi_{j-1}^{(1)}(y)]^2 f_Y(y)dy + \int_0^{+\infty} y^2\varphi_j^2(y)f_Y(y)dy.$$

Now we apply Lemma 7.2 to the first term with $p = 1$ and $\delta = 1$, so that we get that $\int_0^{+\infty} y[\varphi_{j-1}^{(1)}(y)]^2 f_Y(y)dy \lesssim 1/\sqrt{j}$ if $\mathbb{E}(X_1)$ is finite.

For the second term, we use formula (46) and $uv \le (u^2+v^2)/2$ and obtain

$$\begin{aligned}\int_0^{+\infty} y^2\varphi_j^2(y)f_Y(y)dy &= \int_0^{+\infty} y\varphi_j(y)[-\frac{j+1}{2}\varphi_{j+1}(y) + (j+\frac{1}{2})\varphi_j(y) - \frac{j}{2}\varphi_{j-1}(y)]f_Y(y)dy \\ &\le (\frac{3j}{2}+\frac{3}{4})\int_0^{+\infty} y\varphi_j^2(y)f_Y(y)dy + \frac{j+1}{4}\int_0^{+\infty} y\varphi_{j+1}^2(y)f_Y(y)dy \\ &\quad + \frac{j}{4}\int_0^{+\infty} y\varphi_{j-1}^2(y)f_Y(y)dy = O(\sqrt{j})\end{aligned}$$

by Lemma 7.2 with $p = 1$ and $\delta = 0$, if $\mathbb{E}(X_1)$ is finite.

Therefore, we get that, for m large enough, if $\mathbb{E}(1/\sqrt{X_1}) < +\infty$ and $\mathbb{E}(X_1) < +\infty$,

$$V_m = \sum_{j=0}^{m-1}\mathbb{E}\left[(Y_1\varphi_j'(Y_1) + \varphi_j(Y_1))^2\right] \lesssim m^{3/2}.$$

• **Hermite case.** Now, we bound V_m in the case $\psi_j = h_j$. Using (52), we get that

$$xh_j'(x) = \frac{x}{\sqrt{2}}(\sqrt{j}h_{j-1}(x) - \sqrt{j+1}h_{j+1}(x))$$

and the order of $\int_{\mathbb{R}}(yh_j'(y) + h_j(y))^2 f_Y(y)dy$ follows from Lemma 7.5 applied for $p = 2$ and $p = 0$. We obtain a bound of order $\sqrt{j}$ for j large enough, if $\mathbb{E}(|X_1|^{2+2/3}) < +\infty$. Therefore, we get that, for m large enough, if $\mathbb{E}(|X_1|^{2+2/3}) < +\infty$,

$$V_m = \sum_{j=0}^{m-1}\mathbb{E}\left[(Y_1h_j'(Y_1) + h_j(Y_1))^2\right] \lesssim m^{3/2}. \quad \square$$

7.10. **Proof of Remark 4.1.** Using Formula (15) with $t = h_j^2$,

$$V_m = \sum_{j=0}^{m-1} \mathbb{E}(Y_1^2(h_j'(Y_1))^2) + \sum_{j=0}^{m-1} \mathbb{E}h_j^2(X_1) = \sum_{j=0}^{m-1} \mathbb{E}(Y_1^2(h_j'(Y_1))^2) + V_m^X. \tag{55}$$

We know that $V_m^X \leq cm^{5/6}$. For the first term of the r.h.s. above, we can use Lemma 4.1:

$$\mathbb{E}(Y_1^2(h_j'(Y_1))^2) \leq \mathbb{E}|X_1|\|h_j'\|^2 = \mathbb{E}|X_1|(j+\frac{1}{2})$$

as $h_j' = (\sqrt{j}h_{j-1} - \sqrt{j+1}h_{j+1})/\sqrt{2}$ (see (52)). Consequently,

$$\sum_{j=0}^{m-1} \mathbb{E}(Y_1^2(h_j'(Y_1))^2) \leq \frac{m^2}{2}\mathbb{E}|X_1|. \tag{56}$$

Joining (55) and (56) gives the first risk bound.

For the second point, we note that

$$\mathbb{E}(Y_1^2(h_j'(Y_1))^2) \leq \mathbb{E}Y^2\|h_j'\|_\infty^2 = \frac{1}{3}\mathbb{E}X_1^2\|h_j'\|_\infty^2$$

where $\|h_j'\|_\infty \leq C_\infty'(j+1)^{5/12}$ by Lemma 7.3. Therefore,

$$\sum_{j=0}^{m-1} \mathbb{E}(Y_1^2(h_j'(Y_1))^2) \leq \frac{(C_\infty')^2}{3}\mathbb{E}X_1^2\, m^{11/6}. \tag{57}$$

This gives the second risk bound. □

7.11. **Proof of Proposition 4.3.**

• **Laguerre case.** By assumption, there exist $0 < a < b$ and $c > 0$ such that $\inf_{a\leq x\leq b} f(x) > c > 0$. This implies that we can find $0 < a' < b'$ and c' such that $\inf_{a'\leq y\leq b'} yf_Y(y) = c' > 0$. Indeed,

$$\inf_{a/2\leq y\leq a} yf_Y(y) \geq \frac{a}{2}\int_a^b \frac{f(x)}{x}dx \geq \frac{ca}{2}\log(b/a).$$

Thus $\inf_{a'\leq y\leq b'} yf_Y(y) = c' > 0$ holds with $a' = a/2$, $b' = a$, $c' = (ca/2)\log(b/a)$.

Now

$$\begin{aligned}
\mathbb{E}\left[(Y_1\varphi_j'(Y_1) + \varphi_j(Y_1))^2\right] &= \mathbb{E}\left[(Y_1\varphi_j'(Y_1))^2\right] + \mathbb{E}(\varphi_j^2(X_1)) \\
&\geq \mathbb{E}\left\{[\sqrt{2jY_1}\varphi_{j-1}^{(1)}(Y_1) + Y_1\varphi_j(Y_1)]^2\right\} \\
&\geq 2j\int_{a'}^{b'} y[\varphi_{j-1}^{(1)}(y)]^2 f_Y(y)dy + 2\sqrt{2j}\int_{a'}^{b'} y\sqrt{y}\varphi_{j-1}^{(1)}(y)\varphi_j(y)f_Y(y)dy \\
&:= T_1 + T_2.
\end{aligned}$$

We have, by applying Lemma 7.2 for $p = 0$ if $\mathbb{E}(1/\sqrt{X_1}) < +\infty$,

$$|T_2| \leq \sqrt{2j}(b')^{3/2}\int_{a'}^{b'} ([\varphi_{j-1}^{(1)}]^2(y) + \varphi_j^2(y))f_Y(y)dy = O(1).$$

Next, using (54), we proceed as in the proof of Proposition 3.2,

$$T_1 \geq 2jc'\int_{a'}^{b'} [\varphi_{j-1}^{(1)}]^2(y)dy \geq c\sqrt{j}.$$

We conclude that $V_m \geq Cm^{3/2} + O(m)$.

• **Hermite case.** Analogously, we can find $a' < b'$ and c' such that $\inf_{a'\leq y\leq b'} y^2 f_Y(y) = c' > 0$. For instance if $b > 0$, we can assume $a > 0$. Then

$$\inf_{a/2\leq y\leq a} y^2 f_Y(y) \geq \frac{a^2}{4}\int_a^b \frac{f(x)}{x}dx \geq \frac{ca^2}{4}\log(b/a).$$

Thus $\inf_{a'\leq y\leq b'} y^2 f_Y(y) = c' > 0$ holds with $a' = a/2$, $b' = a$, $c' = (ca^2/4)\log(b/a)$.
Here, we can use as in Walter (1977), the following expression for the Hermite function h_j (see Szegö (1959, p.248)):

$$h_j(x) = \lambda_j \cos\left((2j+1)^{1/2}x - \frac{j\pi}{2}\right) + \frac{1}{(2j+1)^{1/2}}\xi_j(x) \tag{58}$$

where $\lambda_j = |h_j(0)|$ if j is even, $\lambda_j = |h_j'(0)|/(2j+1)^{1/2}$ if j is odd and

$$\xi_j(x) = \int_0^x \sin\left[(2j+1)^{1/2}(x-t)\right] t^2 h_j(t)dt. \tag{59}$$

We have

$$\lambda_{2j} = \frac{(2j)!^{1/2}}{2^j j!\pi^{1/4}}, \quad \lambda_{2j+1} = \lambda_{2j}\frac{\sqrt{2j+1}}{\sqrt{2j+3/2}}.$$

By the Stirling formula and its proof, $\lambda_{2j} \sim \pi^{-1/2}j^{-1/4}$, $\lambda_{2j+1} \sim \pi^{-1/2}j^{-1/4}$ and for all j, there exists constants c_1, c_2 such that, for all $j \geq 1$,

$$\frac{c_1}{\pi^{1/2}j^{1/4}} \leq \lambda_j \leq \frac{c_2}{\pi^{1/2}j^{1/4}}. \tag{60}$$

By derivating (58), we get

$$h_j'(x) = -\sqrt{2j+1}\,\lambda_j \sin\left((2j+1)^{1/2}x - \frac{j\pi}{2}\right) + \frac{1}{(2j+1)^{1/2}}\xi_j'(x) \tag{61}$$

with

$$\xi_j'(x) = \sqrt{2j+1}\int_0^x \cos\left[(2j+1)^{1/2}(x-t)\right] t^2 h_j(t)dt.$$

Then we have, using (61),

$$\begin{aligned}\int y^2(h_j'(y))^2 f_Y(y)dy &\geq c'\int_{a'}^{b'}(h_j')^2(y)dy\\ &\geq c'(2j+1)\lambda_j^2\int_{a'}^{b'}\sin^2\left((2j+1)^{1/2}y - \frac{j\pi}{2}\right)dy\\ &\quad -2c'\lambda_j\int_{a'}^{b'}\sin\left((2j+1)^{1/2}y - \frac{j\pi}{2}\right)\xi_j'(y)dy.\end{aligned}$$

We have $j^{-3/4}c_1/\sqrt{\pi} \leq \frac{2\lambda_j}{(2j+1)^{1/2}} \leq j^{-3/4}\sqrt{2/\pi}c_2$ and

$$|\lambda_j\int_a^b \sin\left((2j+1)^{1/2}x - \frac{j\pi}{2}\right)\xi_j'(x)dx| \leq \sqrt{2j+1}\,\lambda_j\int_a^b \frac{|x|^{5/2}}{\sqrt{10}}dx \lesssim j^{1/4}$$

Thus, the second term is lower bounded by $-O(j^{1/4})$. For the first term, $\lambda_j^2 \geq j^{-1/2}c_1^2/\pi$ and

$$(2j+1)\lambda_j^2\int_a^b \sin^2\left((2j+1)^{1/2}x - \frac{j\pi}{2}\right)dx = (2j+1)\lambda_j^2\left\{\frac{b-a}{2} + O(\frac{1}{\sqrt{j}})\right\}.$$

Therefore, as $\lambda_j^2(2j+1) \geq C\sqrt{j}$,

$$\int y^2(h_j'(y))^2 f(y)dy \quad \geq \quad c\left[\sqrt{j}(1+O(j^{-1/2})) - j^{1/4}\right]$$

Consequently, for j large enough, $\int y^2(h_j'(y))^2 f_Y(y)dy \geq c' j^{1/2}$. This implies, as $V_m^X \lesssim m^{5/6}$ $(5/6 < 3/2)$, $V_m = \sum_{j=0}^{m-1} \mathbb{E}[(Y_1 h_j'(Y_1))^2] + V_m^X \geq c m^{3/2}$. $\square$

7.12. **Proof of Theorem 4.1.** Recall that S_m is the space spanned by $\{\psi_0, \dots, \psi_{m-1}\}$ and $B_m = \{t \in S_m, \|t\| = 1\}$. We have $\hat{f}_m = \arg\min_{t \in S_m} \gamma_n(t)$ where

$$\gamma_n(t) = \|t\|^2 - 2n^{-1} \sum_{i=1}^n \phi_t(Y_i), \quad \phi_t(Y_i) := Y_i t'(Y_i) + t(Y_i) \tag{62}$$

and $\gamma_n(\hat{f}_m) = -\|\hat{f}_m\|^2$. Now, we write, for two functions $t, s \in \mathbb{L}^2(\mathbb{R})$,

$$\gamma_n(t) - \gamma_n(s) = \|t - f\|^2 - \|s - f\|^2 - 2\nu_n(t - s)$$

where

$$\nu_n(t) = \frac{1}{n} \sum_{i=1}^n [\phi_t(Y_i) - \langle t, f \rangle], \tag{63}$$

recall that $\langle t, f \rangle = \mathbb{E}(\phi_t(Y_1))$. Then, for any $m \in \mathcal{M}_n = \{1 \leq m \leq m_n\}$, and any $f_m \in S_m$,

$$\gamma_n(\hat{f}_{\hat{m}}) + \widehat{\text{pen}}(\hat{m}) \leq \gamma_n(f_m) + \widehat{\text{pen}}(m).$$

This yields

$$\|\hat{f}_{\hat{m}} - f\|^2 \leq \|f - f_m\|^2 + \widehat{\text{pen}}(m) - \widehat{\text{pen}}(\hat{m}) + 2\nu_n(\hat{f}_{\hat{m}} - f_m).$$

We use that

$$2\nu_n(\hat{f}_{\hat{m}} - f_m) \leq 4 \sup_{t \in B_{m \vee \hat{m}}} [\nu_n(t)]^2 + \frac{1}{4}\|\hat{f}_{\hat{m}} - f_m\|^2,$$

and some classical algebra to obtain:

$$\begin{aligned} \frac{1}{2}\|\hat{f}_{\hat{m}} - f\|^2 \quad &\leq \quad \frac{3}{2}\|f - f_m\|^2 + \widehat{\text{pen}}(m) + 4\left(\sup_{t \in B_{m \vee \hat{m}}} [\nu_n(t)]^2 - p(m \vee \hat{m})\right)_+ \\ &\quad +(4p(m \vee \hat{m}) - \text{pen}(\hat{m})) + (\text{pen}(\hat{m}) - \widehat{\text{pen}}(\hat{m})). \end{aligned} \tag{64}$$

Lemma 7.6. *Assume that $\mathbb{E}(X_1 + (1/\sqrt{X_1})) < +\infty$ in the Laguerre case, $\mathbb{E}(|X_1|) < +\infty$ in the Hermite case, and that $m_n \leq n^\beta$ where $\beta = 1/3$ for Laguerre and $\beta = 6/17$ for Hermite. Then for $p(m) = 4V_m/n$, we have*

$$\mathbb{E}\left(\sup_{t \in B_{m \vee \hat{m}}} [\nu_n(t)]^2 - p(m \vee \hat{m})\right)_+ \leq \frac{c}{n}, \tag{65}$$

$$\textit{and} \quad \mathbb{E}(\text{pen}(\hat{m}) - \widehat{\text{pen}}(\hat{m}))_+ \leq \frac{1}{2}\mathbb{E}(\text{pen}(\hat{m})) + \frac{c' \log(n)}{n}. \tag{66}$$

where c and c' are positive constants.

Taking expectation in (64) and using Lemma 7.6 yields

$$\frac{1}{2}\mathbb{E}(\|\hat f_{\hat m}-f\|^2) \le \frac{3}{2}\|f-f_m\|^2+\text{pen}(m)+\mathbb{E}(4p(m\vee\hat m)-\text{pen}(\hat m))+\frac{1}{2}\mathbb{E}(\text{pen}(\hat m)) + c'\frac{\log(n)}{n}+\frac{4c}{n}.$$

Now we note that, for $\kappa\ge 32:=\kappa_0$, $4p(m\vee\hat m)-\frac{1}{2}\text{pen}(\hat m)\le\frac{1}{2}\text{pen}(m)$. Finally, we get, for all $m\in\mathcal{M}_n$, $\mathbb{E}(\|\hat f_{\hat m}-f\|^2)\le 3\|f-f_m\|^2+3\text{pen}(m)+c''\frac{\log(n)}{n}$, which ends the proof. □

7.13. **Proof of Lemma 7.6.**

7.13.1. *Proof of (65).* Note that

$$\mathbb{E}\left(\sup_{t\in B_{m\vee\hat m}}[\nu_n(t)]^2-p(m\vee\hat m)\right)\le\sum_{m'\in\mathcal{M}_n}\mathbb{E}\left(\sup_{t\in B_{m\vee m'}}[\nu_n(t)]^2-p(m\vee m')\right)_+. \tag{67}$$

We apply the Talagrand Inequality (see Theorem B.1):

$$\mathbb{E}\left(\sup_{t\in B_m}[\nu_n(t)]^2-4H^2\right)_+\le\frac{C_1}{n}\left(v^2e^{-C_2\frac{nH^2}{v^2}}+\frac{M_1^2}{n}e^{-C_3\frac{nH}{M_1}}\right)$$

where H,v,M_1 are such that $\mathbb{E}\left(\sup_{t\in B_m}[\nu_n(t)]^2\right)\le H^2$, $\sup_{t\in B_m}\text{Var}(\phi_t(Y_1))\le v^2$ and $\sup_{t\in B_m}\sup_y|\phi_t(y)|\le M_1$. We have

$$\mathbb{E}\left(\sup_{t\in B_m}[\nu_n(t)]^2\right)\le\sum_{j=0}^{m-1}\mathbb{E}[\nu_n^2(\psi_j)]=\frac{V_m}{n}:=H^2.$$

Using Lemma 4.1, we have

$$\sup_{t\in B_m}\text{Var}(\phi_t(Y_1)) \le \sup_{t\in B_m}\mathbb{E}(\phi_t^2(Y_1))\le\sup_{t\in B_m}\left[\mathbb{E}(Y_1^2[t'(Y_1)]^2)+\mathbb{E}(t^2(X_1))\right].$$

For any $t\in B_m$,

$$\mathbb{E}(t^2(X_1))\le\|t\|_\infty\int|t|f\le\|t\|_\infty\|f\|\lesssim\sqrt{m}\|f\|.$$

We need a specific study for the term $\sup_{t\in B_m}\mathbb{E}(Y_1^2[t'(Y_1)]^2)$.

Laguerre case. From formula (48), we get

$$yt'(y)=-\sum_{j=0}^{m-1}a_j(t)y\varphi_j(y)-\sum_{j=0}^{m-1}a_j(t)\sqrt{2jy}\varphi_{j-1}^{(1)}(y).$$

By (13) and (14), $0\le yf_Y(y)\le 1$ and $0\le y^2f_Y(y)=\int\frac{y^2}{x}\mathbf{1}_{x\ge y}f(x)dx\le\mathbb{E}(X_1)$. Thus, using the orthonormality of $(\varphi_j)_{0\le j\le m-1}$ and $(\varphi_j^{(1)})_{0\le j\le m-1}$, we have, for $t\in B_m$,

$$\begin{aligned}\mathbb{E}[(Y_1t'(Y_1))^2] &\le 2\int_0^{+\infty}(\sum_{j=0}^{m-1}a_j(t)\varphi_j(y))^2y^2f_Y(y)dy+2\int_0^{+\infty}(\sum_{j=0}^{m-1}a_j(t)\sqrt{2j}\varphi_{j-1}^{(1)}(y))^2yf_Y(y)dy\\ &\le 2\mathbb{E}(X_1)+4m.\end{aligned}$$

Thus $\sup_{t\in B_m}\text{Var}(\phi_t(Y_1))\le Cm:=v^2$ where C depends on $\mathbb{E}(X_1)$ and $\|f\|$.

Hermite case. By Lemma 4.1, 2), we have $\sup_{t\in B_m}\mathbb{E}[(Y_1t'(Y_1))^2]\le\sup_{t\in B_m}\mathbb{E}(|X_1|)\|t'\|^2$ and by (52) we easily get that $\|t'\|^2\le 2m$. Thus $v^2=Cm$ where C depends on $\mathbb{E}(|X_1|)$ and $\|f\|$.

Next we note that, using Lemma 7.1 in Laguerre case and Lemma 7.3-(2) in Hermite case, we get, recalling that $\phi_t(y) = yt'(y) + t(y)$,

$$\sup_{t\in B_m}\sup_{y}|\phi_t(y)| \le cm^{\omega} := M_1$$

with $\omega = 3/2$ in Laguerre case and $\omega = 17/12$ in Hermite case. Therefore in both cases, by using condition (24), $M_1^2/n \le 1$ and we obtain

$$\mathbb{E}\left(\sup_{t\in B_m}[\nu_n(t)]^2 - 4\frac{V_m}{n}\right)_+ \le \frac{c_1}{n}\left(me^{-C_2'\frac{V_m}{m}} + e^{-C_3'\frac{\sqrt{nV_m}}{m^{3/2}}}\right).$$

Therefore as $V_m \gtrsim m^{3/2}$ under our assumptions, we get, using that $m^{3/4-3/2} = m^{-3/4} \ge n^{-9/34}$ under $m_n \le n^{6/17}$ (Hermite case, which contains the Laguerre case $m_n \le n^{1/3}$),

$$\mathbb{E}\left(\sup_{t\in B_m}[\nu_n(t)]^2 - 4\frac{V_m}{n}\right)_+ \le \frac{c_1'}{n}\left(me^{-c_2'\sqrt{m}} + e^{-c_3' n^{4/17}}\right).$$

Choosing $p(m) = 4\,V_m/n$ and using Inequality (67) yields (65). □

7.13.2. *Proof of (66).* We proceed in this proof similarly to Massart (2007), chapter 7 (see Theorem 7.7). Let us define

$$Z_i^{(m)} := \sum_{j=0}^{m-1}\phi_{\psi_j}^2(Y_i), \quad \widehat{V}_m = \frac{1}{n}\sum_{i=1}^{n} Z_i^{(m)}.$$

Bernstein's Inequality (see (2.23) in Massart (2007)) writes $\mathbb{P}(|S_n/n| \ge \sqrt{2s^2x/n} + bx/(3n)) \le 2e^{-x}$ for $S_n = \sum_{i=1}^n (U_i - \mathbb{E}(U_i))$, and i.i.d. U_i's, with $\mathrm{Var}(U_1) \le s^2$, $|U_i| \le b$. We consider $U_i = Z_i^{(m)}$ and $x = 2\log(n)$. We have, using Lemma 7.1 in Laguerre case and Lemma 7.3-(2) in Hermite case,

$$\mathrm{Var}(Z_i^{(m)}) \le \mathbb{E}[(Z_i^{(m)})^2] \le V_m\left\|\sum_{j=0}^{m-1}\phi_{\psi_j}^2\right\|_\infty \le Cm^{1/\beta}V_m := s^2,$$

$$\text{and}\quad |Z_i^{(m)}| \le \left\|\sum_{j=0}^{m-1}\phi_{\psi_j}^2\right\|_\infty \le m^{1/\beta} := b$$

Let us define the set

$$\Omega = \left\{\forall m \in \mathcal{M}_n, \quad \frac{1}{n}\left|\sum_{i=1}^{n}(Z_i^{(m)} - \mathbb{E}(Z_i^{(m)}))\right| \le \sqrt{2V_mCm^{1/\beta}\frac{2\log(n)}{n}} + Cm^{1/\beta}\frac{2\log(n)}{3n}\right\}.$$

Now, applying the Bernstein inequality gives $\mathbb{P}(\Omega^c) \le \sum_{m\in\mathcal{M}_n} 2e^{-2\log(n)} \le c/n$. We write

$$\mathbb{E}(\mathrm{pen}(\hat{m}) - \widehat{\mathrm{pen}}(\hat{m}))_+ \le \mathbb{E}\left[(\mathrm{pen}(\hat{m}) - \widehat{\mathrm{pen}}(\hat{m}))_+\mathbf{1}_\Omega\right] + \mathbb{E}\left[(\mathrm{pen}(\hat{m}) - \widehat{\mathrm{pen}}(\hat{m}))_+\mathbf{1}_{\Omega^c}\right].$$

$$\begin{aligned}\text{On } \Omega,\quad |\widehat{V}_{\hat{m}} - V_{\hat{m}}| &\le \sqrt{2V_mCm^{1/\beta}\frac{2\log(n)}{n}} + Cm^{1/\beta}\frac{2\log(n)}{3n}\\ &\le \frac{1}{2}V_{\hat{m}} + \frac{8}{3}C\frac{\hat{m}^{1/\beta}\log(n)}{n},\end{aligned}$$

using that $2xy \le x^2 + y^2$ applied to $\sqrt{2VA} = 2\sqrt{V/2}\sqrt{A} \le V/2 + A$ with $V = V_{\hat m}$ and $A = 2Cm^{1/\beta}\log(n)/n$. Thus, as by definition of $\mathcal{M}_n$, $\hat m^{1/\beta} \le m_n^{1/\beta} \le n$,

$$\mathbb{E}\left[(\mathrm{pen}(\hat m) - \widehat{\mathrm{pen}}(\hat m))_+ \mathbf{1}_\Omega\right] + \le \frac{1}{2}\mathbb{E}(\mathrm{pen}(\hat m)) + c\frac{\log(n)}{n}.$$

On the other hand, on Ω^c, $\mathbb{E}\left[(\mathrm{pen}(\hat m) - \widehat{\mathrm{pen}}(\hat m))_+ \mathbf{1}_{\Omega^c}\right] \le 2\kappa\mathbb{P}(\Omega^c) \le c/n$. Thus (66) follows from the last two inequalities.□

7.14. **Proof of Proposition 4.4.** We have for all $\ell > 0$,

$$\mathbb{E}[\|\tilde f_\ell - f\|^2] = \|f - \bar f_\ell\|^2 + \mathbb{E}[\|\tilde f_\ell - \bar f_\ell\|^2].$$

Next,

$$\begin{aligned}
\mathbb{E}[\|\tilde f_\ell - \bar f_\ell\|^2] &= \frac{1}{2\pi}\mathbb{E}[\|\tilde f_\ell^* - \bar f_\ell^*\|^2] = \frac{1}{2\pi n}\int_{-\pi\ell}^{\pi\ell} \mathrm{Var}\left(e^{iuY_1} + iuY_1 e^{iuY_1}\right) du \\
&= \frac{1}{2\pi n}\int_{-\pi\ell}^{\pi\ell} \mathbb{E}\left|e^{iuY_1} + iuY_1 e^{iuY_1}\right|^2 du - \frac{\|\bar f_\ell\|^2}{n} \\
&= \frac{1}{2\pi n}\int_{-\pi\ell}^{\pi\ell} (1 + u^2\mathbb{E}(Y_1^2))du - \frac{\|\bar f_\ell\|^2}{n} \\
&= \frac{\ell}{n}\left(1 + \frac{\pi^2}{3}\mathbb{E}(Y_1^2)\ell^2\right) - \frac{\|\bar f_\ell\|^2}{n}.
\end{aligned}$$

Gathering the two terms gives Inequality (30). On the other hand, we have for all $\ell > 0$,

$$\mathbb{E}[\|\tilde f_\ell^{(n)} - f\|^2] \le \|f - \bar f_\ell\|^2 + 2\|\bar f_\ell - \mathbb{E}\tilde f_\ell^{(n)}\|^2 + 2\mathbb{E}[\|\tilde f_\ell^{(n)} - \mathbb{E}(\tilde f_\ell^{(n)}\|^2] := T_1 + T_2 + T_3. \quad (68)$$

The term T_1 is the same bias term as before. The term T_2 is bounded in Belomestny *et al.* (2017), Proposition 3.1 and we have the bound $T_2 \le 4\ell^2(M_2 + 1)/L_n$. This term is $O(\ell/n)$ if $\ell \le n$ and $L_n \ge n^2$.

For T_3, using Lemma 4.1 with $t = \theta_{\ell,j}^2$, we write that

$$\begin{aligned}
T_3 &= \sum_{|j|\le L_n} \mathrm{Var}(\tilde{\tilde a}_{\ell,j}) = \frac{1}{n}\sum_{|j|\le L_n} \mathrm{Var}\left[\theta_{\ell,j}(Y_1) + Y_1\theta'_{\ell,j}(Y_1)\right] \\
&\le \frac{1}{n}\sum_{|j|\le L_n} \mathbb{E}\left[\theta_{\ell,j}^2(Y_1) + 2Y_1\theta_{\ell,j}(Y_1)\theta'_{\ell,j}(Y_1) + Y_1^2(\theta'_{\ell,j}(Y_1))^2\right] \\
&= \frac{1}{n}\sum_{|j|\le L_n} \left[\mathbb{E}[\theta_{\ell,j}^2(X_1)] + \mathbb{E}[Y_1^2(\theta'_{\ell,j}(Y_1))^2]\right]
\end{aligned}$$

Next, we know that $\sum_{|j|\le L_n}\theta_{\ell,j}^2 \le \sum_{j\in\mathbb{Z}}\theta_{\ell,j}^2 = \ell$. As moreover

$$\sum_{j\in\mathbb{Z}}\left[\theta'_{\ell,j}(x)\right]^2 = \frac{1}{2\pi}\sum_{j\in\mathbb{Z}}\left[\langle\frac{\theta_{\ell,j}^*}{\sqrt{2\pi}}, te^{-itx}\rangle\right]^2 = \frac{1}{2\pi}\int_{-\pi\ell}^{\pi\ell} t^2 dt = \frac{\pi^2\ell^3}{3},$$

we obtain

$$T_3 \le \frac{\ell}{n}\left(1 + \mathbb{E}(Y_1^2)\frac{\pi^2\ell^2}{3}\right).$$

Plugging this bound in Inequality (68) gives Inequality (31). □

7.15. **Proof of Proposition 5.1.** The risk of the estimator can be written as usual

$$\|\tilde f_m - f\|^2 = \|f - f_m\|^2 + \|\tilde f_m - f_m\|^2$$

where $f_m = \sum_{j=0}^{m-1} a_j(f)\varphi_j$ is the projection of f on $S_m = \text{span}(\varphi_0, \dots, \varphi_{m-1})$ and $\|f - f_m\|^2$ is the square bias term. Next we have

$$\|\tilde f_m - f_m\|^2 = \sum_{j=0}^{m-1} (\tilde a_j - a_j(f))^2 = \|\mathbf{\Sigma}_m^{-1}(\vec{\hat a}(Z)_{m-1} - \mathbb{E}(\vec{\hat a}(Z)_{m-1}))\|_2^2,$$

where $\|\vec x\|_2$ denotes the Euclidean norm of the m-vector $\vec x$. So,

$$\begin{aligned}
\mathbb{E}(\|\tilde f_m - f_m\|^2) &\le \|\mathbf{\Sigma}_m^{-1}\|_{\text{op}}^2 \mathbb{E}(\|\vec{\hat a}(Z)_{m-1} - \mathbb{E}(\vec{\hat a}(Z)_{m-1})\|_2^2) \\
&\le \|\mathbf{\Sigma}_m^{-1}\|_{\text{op}}^2 \sum_{j=0}^{m-1} \text{Var}(\hat a_j(Z)) = \frac{1}{n}\|\mathbf{\Sigma}_m^{-1}\|_{\text{op}}^2 \sum_{j=0}^{m-1} \text{Var}(\varphi_j(Z_1)) \\
&\le \frac{1}{n}\|\mathbf{\Sigma}_m^{-1}\|_{\text{op}}^2 \sum_{j=0}^{m-1} \mathbb{E}(\varphi_j^2(Z_1)) \le \frac{c\sqrt{m}\|\mathbf{\Sigma}_m^{-1}\|_{\text{op}}^2}{n},
\end{aligned}$$

using Lemma 7.2. On the other hand,

$$\begin{aligned}
\mathbb{E}(\|\tilde f_m - f_m\|^2) &= \frac{1}{n}\sum_\ell \text{Var}\left(\sum_j [\mathbf{\Sigma}_m^{-1}]_{\ell,j}\varphi_j(Z_1)\right) \le \frac{1}{n}\sum_\ell \mathbb{E}\left[\left(\sum_j [\mathbf{\Sigma}_m^{-1}]_{\ell,j}\varphi_j(Z_1)\right)^2\right] \\
&\le \frac{\|f_Z\|_\infty}{n}\sum_\ell \int \left(\sum_j [\mathbf{\Sigma}_m^{-1}]_{\ell,j}\varphi_j(z)\right)^2 dz \\
&= \frac{\|f_Z\|_\infty}{n}\sum_\ell \sum_j [\mathbf{\Sigma}_m^{-1}]_{\ell,j}^2 \le \frac{\|f_\Sigma\|_\infty}{n}\|\mathbf{\Sigma}_m^{-1}\|_F^2.
\end{aligned}$$

Combining the previous bounds implies the result. □

7.16. **Proof of Proposition 5.2.** Let us denote by $\check f_{\ell,m}(x) = \sum_{j=0}^{m-1} \check a_{j,\ell}\, h_j(x)$ and $\mathbb{E}(\check f_{\ell,m}) = f_{\ell,m}$ where $f_{\ell,m} = \sum_{j=0}^{m-1} a_{j,\ell} h_j$,

$$a_{j,\ell}(f) = \mathbb{E}(\check a_{j,\ell}) = \frac{(-i)^j}{\sqrt{2\pi}} \int_{-\pi\ell}^{\pi\ell} f^*(t) h_j(t) dt.$$

We have

$$\begin{aligned}
\mathbb{E}(\|\check f_{\ell,m} - f\|^2) &= \|f - f_{\ell,m}\|^2 + \mathbb{E}(\|\check f_{\ell,m} - f_{\ell,m}\|^2) \\
&= \|f - f_m\|^2 + \sum_{j=0}^{m-1} (a_j(f) - a_{j,\ell}(f))^2 + \sum_{j=0}^{m-1} \text{Var}(\check a_{j,\ell}) \\
&= \sum_{j\ge m} a_j^2(f) + \sum_{j=0}^{m-1} (a_j(f) - a_{j,\ell}(f))^2 + \frac{1}{n}\sum_{j=0}^{m-1} \text{Var}\left(\frac{(-i)^j}{\sqrt{2\pi}} \int_{-\pi\ell}^{\pi\ell} e^{itZ_1} \frac{h_j(t)}{f_\Sigma^*(t)} dt\right)
\end{aligned}$$

Then

$$\sum_{j=0}^{m-1} (a_j(f) - a_{j,\ell}(f))^2 = \sum_{j=0}^{m-1} \frac{1}{2\pi}\left|\int f^*(t)\mathbf{1}_{|t|\ge\pi\ell} h_j(t) dt\right|^2 \le \frac{1}{2\pi}\int_{|t|\ge\pi\ell} |f^*(t)|^2 dt.$$

On the other hand, for the variance term, we have

$$\begin{aligned}\frac{1}{n}\sum_{j=0}^{m-1}\mathrm{Var}\left(\frac{(-i)^j}{\sqrt{2\pi}}\int_{-\pi\ell}^{\pi\ell} e^{itZ_1}\frac{h_j(t)}{f_\Sigma^*(t)}dt\right) &\le \frac{1}{n}\sum_{j=0}^{m-1}\mathbb{E}\left(\left|\frac{(-i)^j}{\sqrt{2\pi}}\int_{-\pi\ell}^{\pi\ell} e^{itZ_1}\frac{h_j(t)}{f_\Sigma^*(t)}dt\right|^2\right)\\ &= \frac{1}{2\pi\, n}\mathbb{E}\left(\sum_{j=0}^{m-1}\left|\int \frac{\mathbf{1}_{|t|\le\pi\ell}e^{itZ_1}}{f_\Sigma^*(t)}h_j(t)dt\right|^2\right)\\ &\le \frac{1}{2\pi\, n}\int\frac{\mathbf{1}_{|t|\le\pi\ell}}{|f_\Sigma^*(t)|^2}dt.\end{aligned}$$

Therefore we obtain

$$\mathbb{E}(\|\check{f}_{\ell,m}-f\|^2)\le\sum_{j\ge m}a_j^2(f)+\frac{1}{2\pi}\int_{|t|\ge\pi\ell}|f^*(t)|^2dt+\frac{1}{2\pi\, n}\int\frac{\mathbf{1}_{|t|\le\pi\ell}}{|f_\Sigma^*(t)|^2}dt.$$

Choosing $\ell=\sqrt{m}$ gives the announced result. □

7.17. **Proof of Theorem 5.1.** Let $\mathfrak{M}=\max\mathcal{M}_n$ the maximal element of the collection. We follow the lines of the proof of Theorem 4.1, with (62) replaced by

$$\tilde{\gamma}_n(t)=\|t\|^2-2\langle t,\tilde{f}_{\mathfrak{M}}\rangle,$$

and (63) by $\tilde{\nu}_n(t)=\langle t,\tilde{f}_{\mathfrak{M}}-f_{\mathfrak{M}}\rangle$. Note that for $t\in S_m$, then $\tilde{\nu}_n(t)=\langle t,\tilde{f}_m-f_m\rangle$. Then we get

$$\|\tilde{f}_{\tilde{m}}-f\|^2 \;\le\; 3\|f-f_m\|^2+2\widetilde{\mathrm{pen}}(m)+8(\sup_{t\in\mathcal{B}_{m,\tilde{m}}}\tilde{\nu}_n^2(t)-\tilde{p}(m,\tilde{m}))+8\tilde{p}(m,\tilde{m})-2\widetilde{\mathrm{pen}}(\tilde{m})$$

with $\tilde{p}(m,m')$ satisfying $4\tilde{p}(m,m')\le\widetilde{\mathrm{pen}}(m)+\widetilde{\mathrm{pen}}(m')$ for $\kappa\ge\kappa_0$, where $\tilde{p}(m,m')$ and κ_0 are specified by the next Lemma.

Lemma 7.7. *Under the assumptions of Theorem 5.1, for $m^*=m\vee m'$ and*

$$\tilde{p}(m,m')=2(2\vee\|f_\Sigma\|_\infty)(1+2c\log(2+\|\Sigma_{m^*}^1\|_F^2))\frac{\|\Sigma_{m^*}^{-1}\|_F^2}{n},\quad c\ge\max(3/b,21^2/2b^2)$$

where b is a constant given in Theorem B.1, we have

$$\mathbb{E}\left[\left(\sup_{\vec{t}\in B(\tilde{m},m)}\tilde{\nu}_n^2(\vec{t})-\tilde{p}(m,\tilde{m})\right)_+\right]\le\frac{K}{n} \tag{69}$$

We obtain that $\forall m\in\mathcal{M}_n$,

$$\mathbb{E}(\|\tilde{f}_{\tilde{m}}-f\|^2)\;\le\;3\|f-f_m\|^2+4\widetilde{\mathrm{pen}}(m)+8\frac{K}{n},$$

which ends the proof of Theorem 5.1. □

7.18. **Proof of Lemma 7.7.** The proof of (69) follows the line of the proof of Proposition 7.1 in Mabon (2015). We start as in the proof of Lemma 7.6 and compute the terms H^2, v and M of Theorem B.1. For $t\in B(\tilde{m},m)$ and $m^*=m\vee m'$, we get

$$\mathbb{E}\left(\sup_{t\in B(m',m}\tilde{\nu}_n^2(t)\right)\le\sum_{j=0}^{m^*-1}\mathbb{E}\left(\tilde{\nu}_n^2(\varphi_j)\right)\le\sum_{j=0}^{m^*-1}\mathbb{E}\left(\langle\varphi_j,\tilde{f}_{m^*}-f_{m^*}\rangle^2\right)=\mathbb{E}(\|\tilde{f}_{m^*}-f_{m^*}\|^2).$$

From Proposition 5.1, we deduce $H^2 = (2 \vee \|f_\Sigma\|_\infty)\|\Sigma_{m^*}^{-1}\|_F^2/n$. Clearly, $v = nH^2$. Moreover

$$\tilde{\nu}_n(t) = \frac{1}{n}\sum_{i=1}^n [\psi_t(Z_i) - \mathbb{E}(\psi_t(Z_i))], \quad \psi_t(x) = \sum_{j=0}^{m^*-1} \langle t, \varphi_j\rangle [V_{m^*}^{-1}\vec{\varphi}_{m^*-1}(x)]_j$$

where $\vec{\varphi}_{m-1}(x) = {}^t(\varphi_0(x), \dots, \varphi_{m-1}(x))$ and $[\vec{x}]_j$ denotes the jth coordinate of vector $\vec{x}$. Thus

$$\sup_{t\in B(m',m)} \sup_x |\psi_t(x)| \le \sup_x \|\mathbf{\Sigma}_{m^*}^{-1}\vec{\varphi}_{m^*-1}(x)\|_2 \le \|\mathbf{\Sigma}_{m^*}^{-1}\|_{\rm op}\sqrt{2m^*} := M.$$

Let $\alpha(m^*) = c\log(2 + \|\mathbf{\Sigma}_{m^*}^1\|_F^2)$, and let us apply Theorem B.1

$$\begin{aligned}
&\mathbb{E}\left(\sup_{t\in B(m',m)} \tilde{\nu}_n^2(t) - 2(1+2\alpha(m^*))H^2\right)_+ \\
\le\ & \frac{C}{n}\left(\|\mathbf{\Sigma}_{m^*}^{-1}\|_F^2 \exp(-b\alpha(m^*)) + \frac{m^*\|\mathbf{\Sigma}_{m^*}^{-1}\|_{\rm op}^2}{n}\exp\left(-\frac{\sqrt{2}b}{7}\frac{\sqrt{\alpha(m^*)n}\|\mathbf{\Sigma}_{m^*}^{-1}\|_F}{\sqrt{m^*}\|\mathbf{\Sigma}_{m^*}^{-1}\|_{\rm op}}\right)\right) \\
\le\ & \frac{C}{n}\left(\frac{1}{\|\mathbf{\Sigma}_{m^*}^{-1}\|_F^{2bc-2}} + \|\mathbf{\Sigma}_{m^*}^{-1}\|_F^2 \exp\left(-\frac{\sqrt{2}b}{7}\sqrt{\alpha(m^*)\log(n+2)}\right)\right),
\end{aligned}$$

where we have used that $m^* \le n/\log(n+2)$ and $\|\mathbf{\Sigma}_m^{-1}\|_{\rm op}^2 \le \|\mathbf{\Sigma}_m^{-1}\|^2$. Therefore

$$\mathbb{E}\left(\sup_{t\in B(m',m)} \tilde{\nu}_n^2(t) - 2(1+2\alpha(m^*))H^2\right)_+ \le \frac{C}{n}\left(\frac{1}{\|\mathbf{\Sigma}_{m^*}^{-1}\|_F^{2bc-2}} + \frac{1}{\|\mathbf{\Sigma}_{m^*}^{-1}\|_F^{\sqrt{2c}b/7-2}}\right).$$

For $c \ge \max(3/b, 21^2/2b^2)$ and as $\|\mathbf{\Sigma}_m^{-1}\|_F^2 \ge 2m^*/a_0(f_\Sigma)$, we get

$$\mathbb{E}\left(\sup_{t\in B(m',m)} \tilde{\nu}_n^2(t) - 2(1+2\alpha(m^*))H^2\right)_+ \le \frac{C'}{n}\frac{1}{(m^*)^4}$$

so that

$$\sum_{m'\in\mathcal{M}_n} \mathbb{E}\left(\sup_{t\in B(m',m)} \tilde{\nu}_n^2(t) - 2(1+2\alpha(m^*))H^2\right)_+ \le C"/n.$$

This concludes of Lemma 7.7. □

References

[1] Abbaszadeh, M., Chesneau, C. and Doosti, H. (2012) Nonparametric estimation of density under bias and multiplicative censoring via wavelet methods. *Statist. Probab. Lett.* **82**, 932-941.

[2] Abbaszadeh, M., Chesneau, C. and Doosti, H. (2013) Multiplicative censoring: estimation of a density and its derivatives under the Lp-risk. *REVSTAT* **11**, 255-276.

[3] Abramowitz, M. and Stegun, I. A. (1964) Handbook of mathematical functions with formulas, graphs, and mathematical tables. National Bureau of Standards Applied Mathematics Series, 55 For sale by the Superintendent of Documents, U.S. Government Printing Office, Washington, D.C.

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

[5] Asgharian, M., Carone, M., Fakoor, V. (2012) Large-sample study of the kernel density estimators under multiplicative censoring. *Ann. Statist.* **40**, 159-187.

[6] Andersen, K. E. and Hansen, M. B. (2001) Multiplicative censoring: density estimation by a series expansion approach. *J. Statist. Plann. Inference* **98**, 137-155.

[7] D. Belomestny, F. Comte and V. Genon-Catalot. (2016) Nonparametric Laguerre estimation in the multiplicative censoring model. *Preprint.*

[8] D. Belomestny, F. Comte and V. Genon-Catalot. (2017) Sobolev-Hermite versus Sobolev nonparametric density estimation on R, to appear in *The Annals of the Institute of Mathematical Statistics.*

[9] Bongioanni, B. and Torrea, J.L. (2006). Sobolev spaces associated to the harmonic oscillator. *Proc.Indian Acad. Sci. (math. Sci.)* **116** (3), 337-360.
[10] Bongioanni, B. and Torrea, J.L. (2009). What is a Sobolev space for the Laguerre function system? *Studia Mathematica* **192** (2), 147-172.
[11] Brunel, E., Comte, F. and Genon-Catalot, V. (2016) Nonparametric density and survival function estimation in the multiplicative censoring model. *TEST* **25**, 570-590.
[12] Chesneau, C. (2013) Wavelet estimation of a density in a GARCH-type model. *Comm. Statist. Theory Methods* **42**, 98-117.
[13] Comte, F., Cuenod, C.-A., Pensky, M. and Rozenholc, Y. (2017). Laplace deconvolution and its application to Dynamic Contrast Enhanced imaging. *J. R. Stat. Soc.* Ser. **B**, **79**, 69-94.
[14] Comte, F. and Genon-Catalot, V. (2015) Adaptive Laguerre density estimation for mixed Poisson models. *Electron. J. Stat.* **9**, 1113-1149.
[15] Comte, F., Rozenholc, Y. and Taupin, M.-L. (2006) Penalized contrast estimator for adaptive density deconvolution. *Canad. J. Statist.* **34**, 431-452.
[16] Efromovich, S. (1999). *Nonparametric curve estimation.* Methods, theory, and applications. Springer Series in Statistics. Springer-Verlag, New York.
[17] Efromovich, S. (2009). Lower bound for estimation of Sobolev densities of order less 1/2. *J. Statist. Plann. Inference* **139**, 2261-2268.
[18] van Es, B., Spreij, P. and van Zanten, H. (2003). Nonparametric volatility density estimation. *Bernoulli* **9**, 451-465.
[19] van Es, B., Jongbloed, G. and van Zuijlen, M. (1998) Isotonic inverse estimators for nonparametric deconvolution. *Ann. Statist.* **26**, 2395-2406.
[20] Jirak, M., Meister, A. and Reiss, M. (2014) Adaptive function estimation in nonparametric regression with one-sided errors. *Ann. Statist.* **42**, 1970-2002.
[21] Jongbloed, G. (1998) Exponential deconvolution: two asymptotically equivalent estimators. *Statist. Neerlandica* **52**, 6-17.
[22] Juditsky, A. and Lambert-Lacroix, S. (2004) On minimax density estimation on R. *Bernoulli* **10**, 187-220.
[23] Klein, T. and Rio, E. (2005). Concentration around the mean for maxima of empirical processes. *Ann. Probab.* **33** 1060-1077.
[24] Mabon, G. (2017). Adaptive deconvolution on the nonnegative real line. *Scandinavian Journal of Statistics*, **44**, 707-740.
[25] Massart, P. (2007) *Concentration inequalities and model selection.* Lectures from the 33rd Summer School on Probability Theory held in Saint-Flour, July 6-23, 2003. With a foreword by Jean Picard. Lecture Notes in Mathematics, 1896. Springer, Berlin, 2007.
[26] Rigollet, P. (2006) Adaptive density estimation using the blockwise Stein method. *Bernoulli* **12**, 351-370.
[27] Shen, J. (2000). Stable and efficient spectral methods in unbounded domains using Laguerre functions. *SIAM J. Numer. Anal.* **38**, 1113-1133.
[28] Schipper, M. (1996). Optimal rates and constants in L2-minimax estimation of probability density functions. *Mathematical Methods of Statistics,* **5**, 253-274.
[29] Vardi, Y. (1989) Multiplicative censoring, renewal processes, deconvolution and decreasing density: nonparametric estimation. *Biometrika* **76**, 751-761.
[30] Vardi, Y. and Zhang, C.-H. (1992) Large sample study of empirical distributions in a random-multiplicative censoring model. *Ann. Statist.* **20**, 1022-1039.
[31] Vareschi, T. (2015) Noisy Laplace deconvolution with error in the operator. *J. Statist. Plann. Inference* **157/158**, 16-35.
[32] Zhang, Z. and Su, W. (2017) A new efficient method for estimating the Gerber-Shiu function in the classical risk model, to appear in *Scandinavian Actuarial Journal*, http://dx.doi.org/10.1080/03461238.2017.1371068

Appendix A. Regularity properties of functions in Sobolev-Laguerre and Sobolev-Hermite spaces

For this appendix, we refer to Bongionanni and Torrea (2006,2009), Belomestny *et al.* (2016,2017) and Comte and Genon-Catalot (2015).

Laguerre case. For $a > 0$, consider the functions

$$\varphi_k^{(\delta,a)}(x) = a^{(\delta+1)/2} \left(\frac{k!}{\Gamma(k+\delta+1)} \right)^{1/2} \exp(-ax/2) x^{\delta/2} L_k^{(\delta}(ax), \quad k \geq 0.$$

This sequence is an orthonormal basis of $\mathbb{L}^2(\mathbb{R}^+)$. We set $a_k^{(\delta,a)}(f) = \langle f, \varphi_k^{(\delta,a)}\rangle$ and define the Sobolev-Laguerre space with regularity index s and scale parameter a by:

$$W^s_{\delta,a} = \{f : (0,+\infty) \to \mathbb{R}, f \in L^2((0,+\infty)), \sum_{k\geq 0} k^s (a_k^{(\delta,a)}(f))^2 < +\infty\}.$$

For f a derivable function on $(0,+\infty)$, consider the operator

$$\partial_a^\delta(f)(x) = \sqrt{x} f'(x) + (\frac{a}{2}\sqrt{x} - \frac{\delta}{2\sqrt{x}}) f(x).$$

If $g = \exp(ax/2)x^{-\delta/2} f$ and f is $m+1$ times derivable, then, the following holds:

$$\partial_a^{\delta+m} \ldots \circ \partial_a^{\delta+1} \circ \partial_a^\delta(f)(x) = g^{(m+1)}(x) \exp(-ax/2) x^{(\delta+m+1)/2}.$$

For s integer, the space $W^s_{\delta,a}$ is exactly the space of functions $f : (0,+\infty) \to \mathbb{R}, f \in L^2((0,+\infty))$ such tht f is $s-1$ derivable, $f^{(s-1)}$ is absolutely continuous and $\partial_a^{\delta+m} \ldots \circ \partial_a^{\delta+1} \circ \partial_a^\delta(f)$ belongs to $L^2((0,+\infty))$ for $m = 0, 1, \ldots, s-1$.
The choice $a = 2, \delta = 0$ seems the simplest one, and $W^s_L = W^s_{0,2}$.
Hermite case. For s integer, $f \in W^s_H$ holds if and only if f admits derivatives up to order s which satisfy $f, f', \ldots, f^{(s)}, x^{s-\ell} f^{(\ell)}, \ell = 0, \ldots, s-1$ belong to $\mathbb{L}^2(\mathbb{R})$. The usual Sobolev space with regularity index s is defined by

$$\mathcal{W}^s = \{f \in \mathbb{L}^2(\mathbb{R}), \|f\|^2_{s,sob} = \int_{\mathbb{R}} (1+t^{2s})|f^*(t)|^2 dt < +\infty\} \tag{70}$$

If s is integer, then

$$\begin{aligned}\mathcal{W}^s \;=\; \{&f \in \mathbb{L}^2(\mathbb{R}), f \text{ admits derivatives up to order } s \\ &\text{such that } \||f\||^2_{s,sob} = \|f\|^2 + \|f'\|^2 + \ldots + \|f^{(s)}\|^2 < +\infty\}.\end{aligned}$$

Therefore, for s integer, $W^s_H \subset \mathcal{W}^s$. Morevover, the following properties are proved in Bongioanni and Torrea (2006): for all $s > 0$,

- $W^s_H \subsetneq \mathcal{W}^s$. If $f \in \mathcal{W}^s$ has compact support, then $f \in W^s_H$.
- $f \in W^s_H \Rightarrow x^s f \in \mathbb{L}^2(\mathbb{R})$.

Appendix B. Talagrand's inequality

We recall the Talagrand concentration inequality given in Klein and Rio (2005).

Theorem B.1. *Consider $n \in \mathbb{N}^*$, $\mathcal{F}$ a class at most countable of measurable functions, and $(X_i)_{i\in\{1,\ldots,n\}}$ a family of real independent random variables. Define, for $f \in \mathcal{F}$, $\nu_n(f) = (1/n)\sum_{i=1}^n (f(X_i) - \mathbb{E}[f(X_i)])$, and assume that there are three positive constants M, H and v such that $\sup_{f\in\mathcal{F}} \|f\|_\infty \leq M$, $\mathbb{E}[\sup_{f\in\mathcal{F}} |\nu_n(f)|] \leq H$, and $\sup_{f\in\mathcal{F}} (1/n)\sum_{i=1}^n \mathrm{Var}(f(X_i)) \leq v$. Then for all $\alpha > 0$,*

$$\mathbb{E}\left[\left(\sup_{f\in\mathcal{F}} |\nu_n(f)|^2 - 2(1+2\alpha)H^2\right)_+\right] \;\leq\; \frac{4}{b}\left(\frac{v}{n} e^{-b\alpha\frac{nH^2}{v}} + \frac{49M^2}{bC^2(\alpha)n^2} e^{-\frac{\sqrt{2}bC(\alpha)\sqrt{\alpha}}{7}\frac{nH}{M}}\right)$$

with $C(\alpha) = (\sqrt{1+\alpha} - 1) \wedge 1$, and $b = \frac{1}{6}$.

By 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 $f \to \nu_n(f)$ is continuous and $\mathcal{F}$ contains a countable dense family.