Volume 13, No. 3 (2004), pp. Allerton Press, Inc.

MATHEMATICAL METHODS OF STATISTICS

MINIMAX ESTIMATION OF LINEAR FUNCTIONALS IN THE CONVOLUTION MODEL

C. Matias[1] and M.-L. Taupin[2]

[1]CNRS – Laboratoire Statistique et Génome
Tour Evry 2, 523 pl. des Terrasses de l'Agora, 91 000 Evry, France
E-mail: matias@genopole.cnrs.fr

[2]Laboratoire de Probabilités, Statistique et Modélisation UMR 8628, Université Paris-Sud
Bâtiment 425, 91 405 Orsay Cedex, France
E-mail: Marie-Luce.Taupin@math.u-psud.fr

Consider the convolution model $Y_k = X_k + \varepsilon_k$, $k = 1, \dots, n$, where the (X_k)'s and the $(\varepsilon_k)'s$, are two independent sequences of independent and identically distributed random variables, the (X_k)'s with unknown density g and the $(\varepsilon_k)'s$ having the Gaussian density f_ε with zero mean and unit variance. In this model we aim at estimating, using the observations $Y_1, \dots, Y_n$, some linear functionals of the density g of the form $\Gamma_f(y) = \int f(x)g(x)f_\varepsilon(y-x)\,dx$, where f is a known function, either polynomial or trigonometric. We extend Taupin's results [21] by giving lower bounds for pointwise minimax risk and upper and lower bounds for minimax $\mathbb{L}_p(\mathbb{R})$-risk, when $2 \le p \le \infty$.

Key words: nonparametric estimation, density estimation, linear functionals, convolution model, kernel estimator, Fano's lemma, van Trees inequality, minimax risk.

2000 Mathematics Subject Classification: 62G05, 62G07, 62G20, 62C20.

1. Introduction

Consider the convolution model

$$Y_k = X_k + \varepsilon_k, \qquad k = 1, \dots, n,$$

where the (X_k)'s and the (ε_k)'s are two independent sequences of independent and identically distributed (i.i.d.) real-valued random variables, the (X_k)'s with

[1]CNRS, Laboratoire Statistique et Génome, UMR 8071, Evry.

[2]IUT de Paris V et Université d'Orsay, Laboratoire de Probabilités, Statistique et Modélisation, UMR 8628.

unknown density g with respect to the Lebesgue measure on $\mathbb{R}$ and the (ε_k)'s having the Gaussian density f_ε with zero mean and unit variance. In this model we aim at estimating linear integral functionals of the unknown density g of the form

$$\Gamma_f(y) = \int f(x)g(x)f_\varepsilon(x-y)\,dx \qquad \text{for} \quad y \in \mathbb{R},$$

where f is a known function, using the observations $Y_1, \dots, Y_n$. We focus here on the cases where f is either a polynomial or a trigonometric function with a particular interest to the special case $f \equiv 1$ corresponding to the density h of the observations $Y_1, \dots, Y_n$, which is, due to independence between X_k and ε_k, given by the convolution product

$$h(y) = g * f_\varepsilon(y) = \int g(x)f_\varepsilon(x-y)\,dx \qquad \text{for} \quad y \in \mathbb{R}.$$

Let us first motivate the interest in those functionals. Consider a nonlinear structural errors-in-variables regression model described by the observation of the random variables $(Z_1, Y_1), \dots, (Z_n, Y_n)$ satisfying the relations

$$Z_k = f_{\beta^0}(X_k) + \eta_k, \quad Y_k = X_k + \varepsilon_k, \qquad k = 1, \dots, n,$$

where the function f is known up to a finite-dimensional parameter β^0 and the errors $(\eta_1, \varepsilon_1), \dots, (\eta_n, \varepsilon_n)$ are centered, i.i.d., with respective variances σ_η^2 and $\sigma_\varepsilon^2 = 1$ for the sake of simplicity. We assume furthermore that the errors (ε_k) are normally distributed. The sequence $(X_k)_{1 \le k \le n}$ is not observed and is a sequence of i.i.d. random variables with unknown density g. Moreover, the sequences $(X_k)_{k \ge 0}$, $(\eta_k)_{k \ge 0}$, and $(\varepsilon_k)_{k \ge 0}$ are independent. In this errors-in-variables regression model, the purpose is to estimate the parameter β^0 in the presence of the unknown density g of the unobserved variables considered as a nuisance parameter. In this context, Taupin [21] proposed an estimator of this parameter β^0 based on the criterion

$$\frac{1}{n}\sum_{i=1}^n W(Y_i)[Z_i - \mathbb{E}\left(f_\beta(X_i) \mid Y_i\right)]^2,$$

$W(\cdot)$ being a compactly supported weight function, where the conditional expectation $\mathbb{E}\left(f_\beta(X_i) \mid Y_i\right)$ is replaced by a nonparametric estimator based on the sample $Y_1, \dots, Y_n$. Denoting by X, Y, and ε the generic variables, and using the independence between X and ε, the conditional expectation $\mathbb{E}\left(f_\beta(X) \mid Y\right)$ can be written as

$$\mathbb{E}\left(f_\beta(X) \mid Y\right) = \frac{\int f_\beta(x)g(x)f_\varepsilon(Y-x)\,dx}{\int g(x)f_\varepsilon(Y-x)\,dx} = \frac{\Gamma_{f_\beta}(Y)}{h(Y)}.$$

Taupin [21] proposed an estimator of this conditional expectation obtained by estimating separately the numerator and the denominator. In particular, she constructed an estimator of this functional Γ_f for general functions f and gave the corresponding upper bounds on the pointwise quadratic risk and on uniform-risk for various classes of functions f such as polynomial functions, exponential functions, trigonometric functions, and more generally functions f admitting an analytic continuation into a strip $\mathcal{B}_\gamma = \{x + iy;\ (x, y) \in \mathbb{R}^2,\ |y| \leq \gamma\}$, or functions f belonging to Sobolev classes.

The upper bound for the rate of convergence of the estimator of β^0 may depend on the asymptotic properties of the pointwise quadratic risk and of the $\mathbb{L}_\infty$-risk of the estimator of Γ_f. Therefore it is of interest to study deeply the estimation problem of such functionals.

The estimator. The functional Γ_f is simply estimated by a plug-in deconvolution kernel density estimator as given in Fan [11] or Butucea [4], that is

$$\widehat{\Gamma}_{f,n}(y_0) = \int f(x) f_\varepsilon(y_0 - x)\widehat{g}_n(x)\,dx \tag{1}$$

$$\text{with} \quad \widehat{g}_n(x) = \frac{C_n}{n}\sum_{i=1}^n \widetilde{K_n}(C_n(x - Z_i)),$$

where $\widetilde{K_n}^\star(t) = K^\star(t)/f_\varepsilon^\star(tC_n)$, K being a kernel to be chosen, $(C_n)_{n\geq 0}$ is a sequence increasing to infinity, and $u^\star(t) = \int \exp(itx)u(x)\,dx$ denotes the Fourier transform of u. The choice of the kernel will be adapted to the function f in Γ_f, but has to satisfy the following conditions.

[K1] The Fourier transform $K^\star$ of K has a bounded support and $|K^\star| \leq \mathbf{1}_{[-\tau,\tau]}$.
[K2] The kernel K belongs to $\mathbb{L}_2(\mathbb{R})$ and is an even function.
[K3] $K^\star(t) = 1$ for any t in $[-1, 1]$ and $K^\star$ is nonnegative.

Condition **[K1]** ensures, in particular, the existence of the deconvolution density estimator $\widehat{g}_n$ and consequently also the existence of $\widehat{\Gamma}_{f,n}$. Condition **[K2]** ensures that the Fourier transform of the kernel is an even real-valued function, and Condition **[K3]** allows the control of the bias term since $|1 - K^\star(t)| \leq \mathbf{1}_{|t|\geq 1}$, for all t. The so-called sine kernel defined by

$$S(x) = \sin(x)/(\pi x), \quad x \in \mathbb{R}, \qquad \text{with} \quad S^\star(t) = \mathbf{1}_{[-1;1]}(t), \tag{2}$$

satisfies Conditions **[K1]**–**[K3]**. Since the kernel S is not integrable, we may prefer, for instance, a kernel, called the analogue of the de La Vallée-Poussin kernel (see Nikol'skii [16]), defined for all x in $\mathbb{R}$ by

$$V(x) = \frac{\cos(x) - \cos(2x)}{\pi x^2} \quad \text{with} \quad V^\star(t) = \begin{cases} 1 & \text{if } t \in [-1, 1], \\ 0 & \text{if } |t| \geq 2, \\ (2 - |t|) & \text{if } |t| \in [1, 2], \end{cases} \tag{3}$$

which satisfies Conditions **[K1]**–**[K3]** and belongs to $\mathbb{L}_1(\mathbb{R})$. Subsequently, S will denote the sine kernel and V the de La Vallée-Poussin kernel. It is easy to see that, according to Parseval's Formula, $\widehat{\Gamma}_{f,n}(y_0)$ is the same estimator as the one proposed by Taupin [21]. It is also noteworthy that when $f \equiv 1$ we have $\Gamma_1 = h$, and $\widehat{h}_n(y_0)$ is simply a kernel density estimator based on the kernel K, so that $\widehat{h}_n^{(k)}$, $k = 0, \ldots, \ell$, is also given by

$$\widehat{h}_n^{(k)}(y_0) = \frac{1}{n}\sum_{j=1}^n K_n^{(k)}(y_0 - Y_j) = \frac{C_n^{k+1}}{n}\sum_{j=1}^n K^{(k)}(C_n(y_0 - Y_j)), \tag{4}$$

where $K_n(\cdot) = C_n K(C_n \cdot)$.

Our aim is to provide lower bounds for the minimax risk with respect to various loss functions, and to show that the estimator defined by (1) achieves those optimal rates of convergence when f is either a polynomial function or a trigonometric function. Note that when f is an exponential function, we have the relation $\Gamma_f(y) = \int \exp(\beta x) g(x) f_\varepsilon(x-y)\,dx = \exp(y\beta + \beta^2/2) h(y+\beta)$, and therefore estimating Γ_f corresponds to a special case of estimating h.

Minimax risks. We consider various minimax risks: the pointwise minimax quadratic risk for the estimation of $\Gamma_f(y_0)$ when y_0 is fixed and the minimax $\mathbb{L}_p(\mathbb{R})$-risk ($2 \le p \le \infty$) for the global estimation of Γ_f. Denote by $\mathcal{G}$ the set of probability densities g with respect to the Lebesgue measure on $\mathbb{R}$ and by $\mathcal{H}$ the set of densities written as the convolution of a density g in $\mathcal{G}$ with the standard Gaussian density f_ε, that is

$$\mathcal{G} = \left\{ g;\, g \ge 0,\, \int_{\mathbb{R}} g(x)\,dx = 1 \right\}, \quad \mathcal{H} = \left\{ g * f_\varepsilon;\, g \ge 0,\, \int_{\mathbb{R}} g(x)\,dx = 1 \right\}.$$

For a fixed function f, $\mathcal{G}_f$ denotes the set of densities g for which the linear functional Γ_f exists, that is

$$\mathcal{G}_f = \left\{ g \in \mathcal{G} \text{ such that } \forall y \in \mathbb{R},\, f(\cdot)g(\cdot)f_\varepsilon(\cdot - y) \in \mathbb{L}_1(\mathbb{R}) \right\}.$$

Note that when f is a polynomial function of degree ℓ, the functionals Γ_f are defined for densities g in $\mathcal{G}_f$ having at least ℓ finite moments. This condition relies on the fact that we use the functional Γ_f to estimate $\mathbb{E}\,(f(X) \mid Y)$ and then when f is a polynomial function of degree ℓ, we need that $\mathbb{E}\,(|X|^\ell) < \infty$. When f is a trigonometric function, the functional Γ_f exists for any density g.

The pointwise minimax quadratic risk over the set $\mathcal{G}_f$ is defined for any $y_0 \in \mathbb{R}$ by

$$R_n(f, y_0) = \inf_{T_n} \sup_{g \in \mathcal{G}_f} \mathbb{E}^{\,1/2}[\Gamma_f(y_0) - T_n]^2,$$

where the infimum is taken over all the estimators T_n based on the observations $Y_1, \dots, Y_n$. In the same way, the minimax $\mathbb{L}_p(\mathbb{R})$-risk is defined by

$$R_{n,p}(f) = \inf_{T_n} \sup_{g \in \mathcal{G}_f} \mathbb{E}\,\|\Gamma_f - T_n\|_p, \qquad 2 \le p \le \infty,$$

where the infimum is taken over all the estimators T_n based on the observations $Y_1, \dots, Y_n$, and where $\|u\|_p = \left[\int |u(x)|^p\,dx\right]^{1/p}$ and $\|u\|_\infty = \sup_{x \in \mathbb{R}} |u(x)|$.

Previous results. Let us present previous results starting with the particular case $f \equiv 1$ corresponding to the problem of estimating the density $h = f_\varepsilon * g$. Take $K \equiv S$ and $C_n = \sqrt{\log n}$. Then by classical calculations on kernel estimation (see Ibragimov and Hasminskii [15] and Taupin [21]) the following results hold:

$$\limsup_{n \to \infty} \frac{\sqrt{n}}{(\log n)^{1/4}} \sup_{h \in \mathcal{H}} \mathbb{E}^{\,1/2} \left[\widehat{h}_n(y_0) - h(y_0)\right]^2 < +\infty,$$

and

$$\limsup_{n\to\infty} \frac{\sqrt{n}}{(\log n)^{1/4}\sqrt{\log\log n}} \sup_{h\in\mathcal{H}} \mathbb{E}\, \|\widehat{h}_n - h\|_\infty < +\infty.$$

Note that, due to the structure of the convolution model, the regularity of the density h is determined by the density of ε's, which is supposed to be the standard Gaussian density. Then the density $f_\varepsilon * g$ has obviously strong smoothness properties.

First, the class $\mathcal{H}$ is embedded into the totally bounded (with respect to the $\mathbb{L}_2(\mathbb{R})$-norm) set $\mathcal{A}_\gamma$ of analytic densities on the strip $\mathcal{B}_\gamma = \{x + iy;\, (x, y) \in \mathbb{R}^2,\, |y| \le \gamma\}$ satisfying $\int (\operatorname{Re} h(x + i\gamma))^2\, dx \le C_\gamma$, where $\operatorname{Re}(z)$ denotes the real part of the complex number z. Golubev and Levit [14] proved that the optimal rate of convergence of the quadratic risk over this class is of order $\sqrt{\log n}/\sqrt{n}$, which implies that the order of the minimax quadratic risk on $\mathcal{H}$ is less than or equal to $\sqrt{\log n}/\sqrt{n}$. Secondly, by the fact that $|(f_\varepsilon * g)^\star| \le f_\varepsilon^\star$, the class $\mathcal{H}$ is contained in the even smaller class $\mathcal{B}_{1,1/2}(2)$, where the class $\mathcal{B}_{A,\rho}(r)$ is defined as $\mathcal{B}_{A,\rho}(r) = \big\{\phi \text{ density};\, |\phi^\star(t)| \le Ae^{-\rho|t|^r}\big\}$. These smoothness classes have been considered by Davis [9, 10], and lately by Levit[1] and Artiles Martinez (see [1] and [2]). Since the minimax quadratic risk for estimating a density over the class $\mathcal{B}_{A,\rho}(r)$ is of order $(\log n)^{1/(2r)}/\sqrt{n}$, the minimax quadratic risk on $\mathcal{H}$ is less than or equal to $(\log n)^{1/4}/\sqrt{n}$. Taupin [20] established a lower bound for the minimax quadratic risk over the class $\mathcal{H}$, which shows that the $(\log n)^{1/4}/\sqrt{n}$ rate cannot be further improved.

The results obtained for the estimation of h lead us to make some remarks about the convolution model.

The smoothness properties of h come from the convolution, $h = f_\varepsilon * g$, of a density g with the Gaussian density, without any additional assumption on g. Therefore adaptivity is not of our concern since our density h has a regularity which does not depend on unknown parameters.

Secondly, we would like to stress the difference between our objective and the one in deconvolution problems. In the latter case the purpose is to estimate the density g of X. It is known that the slowest rates of convergence for estimating g are obtained for the smoothest error densities, whereas the faster rates of convergence for estimating $h = g * f_\varepsilon$ are obtained for the smoothest errors densities. We refer, e.g., to Pensky and Vidakovic [17], Cator [6], Butucea [4], Butucea and Tsybakov [5] or Comte and Taupin [7] for recent results on deconvolution density estimation.

We now turn to the problem of estimating Γ_f focusing on polynomial and trigonometric functions f. In the polynomial case, when $f(x) = \sum_{k=0}^{\ell} \beta_k x^k$ with $\ell \ge 1$ and $\beta = (\beta_0, \cdots, \beta_\ell)$ a fixed $(\ell + 1)$-tuple of real numbers, Taupin [21] established the following upper bounds: for any y_0 in $\mathbb{R}$, if $C_n = \sqrt{\log n}$ and $K \equiv S$ in $\widehat{\Gamma}_{f,n}$ defined by (1), then

$$\limsup_{n\to\infty} \frac{\sqrt{n}}{(\log n)^{(2\ell+1)/4}} \mathbb{E}^{1/2}\, [\widehat{\Gamma}_{f,n}(y_0) - \Gamma_f(y_0)]^2 < \infty,$$

[1] May 1996, talk at the Ecole Normale in Paris (ULM). Unpublished

and if Z_C denotes any compact subset of $\mathbb{R}$,

$$\limsup_{n\to\infty} \frac{\sqrt{n}}{(\log n)^{(2\ell+3)/4}} \mathbb{E}\,[\sup_{y\in Z_c} |\widehat{\Gamma}_{f,n}(y) - \Gamma_f(y)|\,] < \infty.$$

Note that, due to the fact that the errors ε are Gaussian, for a polynomial function f the functional Γ_f can be related to the derivatives of h up to order ℓ (see Lemma 2.1). Therefore the problem of estimating Γ_f is related to the problem of estimating the derivatives of the density h up to order ℓ.

When $f(x) = \sum_{k=0}^{\ell} \beta_k \cos(kx)$ or $f(x) = \sum_{k=0}^{\ell} \beta_k \sin(kx)$ with $\ell \geq 1$ and $\beta = (\beta_0, \dots, \beta_\ell)$ a fixed $(\ell+1)$-tuple of real numbers, Taupin [21] proved the following upper bounds: for any y_0 in $\mathbb{R}$, if $C_n = \sqrt{\log n}$ and $K \equiv S$ in $\widehat{\Gamma}_{f,n}$ defined by (1), then

$$\limsup_{n\to\infty} \frac{\sqrt{n}}{\exp\{\ell\sqrt{\log n}\}} \mathbb{E}^{\,1/2}[\widehat{\Gamma}_{f,n}(y_0) - \Gamma_f(y_0)]^2 < \infty,$$

and if Z_C denotes any compact subset of $\mathbb{R}$, then

$$\limsup_{n\to\infty} \frac{\sqrt{n}}{\sqrt{\log n}\exp\{\ell\sqrt{\log n}\}} \mathbb{E}\,[\sup_{y\in Z_C} |\widehat{\Gamma}_{f,n}(y) - \Gamma_f(y)|\,] < \infty.$$

Taupin [21] noticed that for general f, the smoother is f the faster is the rate of convergence for the estimation of Γ_f. In both cases considered (polynomial or trigonometric functions), f admits an analytic continuation in the whole complex plane and so does the functional Γ_f but the rates of convergence are really different. We will show in this paper that these rates cannot be essentially improved (see below).

Results. Our aim here is threefold. First, we improve the existing upper bounds with respect to the uniform norm. As a matter of fact, these previous bounds were obtained for y lying in a compact set and were not optimal. We give new upper bounds for the uniform norm on $\mathbb{R}$ when f is a polynomial or a trigonometric function. Secondly, we extend those results by giving upper bounds for $\mathbb{L}_p(\mathbb{R})$-risks, when $2 \leq p < \infty$. Finally, our main contribution concerns lower bounds. We prove that all these rates of convergence obtained in the polynomial case are optimal in the minimax sense (pointwise and for $\mathbb{L}_p(\mathbb{R})$-risks, $2 \leq p \leq \infty$), and are nearly optimal in the trigonometric case (pointwise and for $\mathbb{L}_\infty(\mathbb{R})$-risks), in the sense that there is a small loss in lower bounds, negligible with respect to the dominating term in the rate of convergence.

This paper is organized as follows. Section 2 presents some elementary properties of Γ_f and $\widehat{\Gamma}_{f,n}$ in the cases we are interested in. In Section 3, we give the results about pointwise estimation, starting by recalling previous results on upper bounds and then giving lower bounds. In Sections 4 and 5 we deal with the problem of estimation with respect to $\mathbb{L}_p(\mathbb{R})$-risk, when $2 \leq p \leq \infty$, starting with upper bounds (Section 4) and finally establishing the optimality properties of our estimator through lower bounds (Section 5). The proofs are presented in Sections 6 and 7 (for the main results) and in Section 8 (for technical lemmas).

2. Some Properties of our Functionals

Subsequently, we will denote by $\mathcal{P}_{\beta,\ell}$ any polynomial function of the form $\mathcal{P}_{\beta,\ell}: x \mapsto \sum_{j=0}^{\ell} \beta_j x^j$ with β_ℓ different from zero. In the same way, we will denote by $\mathcal{C}_{\beta,\ell}$, $\ell \geq 1$, any linear combination of cosine functions of the form $\mathcal{C}_{\beta,\ell}\colon x \mapsto \sum_{j=0}^{\ell} \beta_j \cos(jx)$ and similarly $\mathcal{S}_{\beta,\ell}\colon x \mapsto \sum_{j=0}^{\ell} \beta_j \sin(jx)$, with β_ℓ different from zero. In both cases ℓ is a fixed integer and the parameters $(\beta_j)_{0\leq j\leq \ell}$ are real fixed numbers. These particular forms of the function f considered here imply that the corresponding functional Γ_f satisfies some useful formulae given below.

Lemma 2.1. *For a fixed integer ℓ let f be of type $\mathcal{P}_{\beta,\ell}$, then there exist polynomial functions $\{Q_{\beta,j}\}$ for $1 \leq j \leq \ell$ of degree* $\deg(Q_{\beta,j}) = j$ *such that for all y in $\mathbb{R}$*

$$\Gamma_f(y) = \beta_\ell h^{(\ell)}(y) + \sum_{k=0}^{\ell-1} Q_{\beta,\ell-k}(y) h^{(k)}(y), \qquad \textit{for all} \quad y \in \mathbb{R}. \tag{5}$$

Moreover, the corresponding estimator $\widehat{\Gamma}_{f,n}$ defined by (1) *satisfies*

$$\widehat{\Gamma}_{f,n}(y) = \beta_\ell \widehat{h}_n^{(\ell)} + \sum_{k=0}^{\ell-1} Q_{\beta,\ell-k}(y) \widehat{h}_n^{(k)}(y), \qquad \textit{for all} \quad y \in \mathbb{R}. \tag{6}$$

This lemma follows immediately from the following remarks. Denoting $\Gamma_k(y) = \int x^k f_\varepsilon(x-y) g(x)\,dx$, we have the recurrence formula $\Gamma_k(y) = \int (x-y)^k f_\varepsilon(x-y) g(x)\,dx + \sum_{j=0}^{k-1} \binom{k}{j} y^{k-j} \Gamma_j(x)$. Moreover, there exist coefficients $\{\alpha_j\}$ such that $h^{(k)}(y) = \int (x-y)^k f_\varepsilon(x-y) g(x)\,dx + \sum_{j=0}^{k-1} \alpha_j \int (x-y)^j f_\varepsilon(x-y) g(x)\,dx$. This gives the result for Γ_f and the proof of (6) follows the same lines. Consequently, when f is a polynomial function of degree ℓ, the estimation of the functional Γ_f is a problem equivalent to the estimation of the derivatives $h^{(k)}$ up to order ℓ. We now turn to the trigonometric case.

Lemma 2.2. *For a fixed integer $\ell \geq 1$ let f be of type $\mathcal{C}_{\beta,\ell}$. Consider the functional Γ_f, its estimator $\widehat{\Gamma}_{f,n}$ defined by* (1)*, and $\widehat{h}_n$ defined by* (4)*. Then we have*

$$\Gamma_f(y) = \sum_{j=0}^{\ell} \beta_j \frac{e^{-j^2/2}}{2} \left[e^{ijy} h(y+ij) + e^{-ijy} h(y-ij) \right], \tag{7}$$

$$\widehat{\Gamma}_{f,n}(y) = \sum_{j=0}^{\ell} \beta_j \frac{e^{-j^2/2}}{2} \left[e^{ijy} \widehat{h}_n(y+ij) + e^{-ijy} \widehat{h}_n(y-ij) \right]. \tag{8}$$

Analogously, when f is of type $\mathcal{S}_{\beta,\ell}$, we get

$$\Gamma_f(y) = \sum_{j=0}^{\ell} \beta_j \frac{e^{-j^2/2}}{2i} \left[e^{ijy} h(y+ij) - e^{-ijy} h(y-ij) \right], \tag{9}$$

$$\widehat{\Gamma}_{f,n}(y) = \sum_{j=0}^{\ell} \beta_j \frac{e^{-j^2/2}}{2i} \left[e^{ijy} \widehat{h}_n(y+ij) - e^{-ijy} \widehat{h}_n(y-ij) \right]. \tag{10}$$

This lemma follows from the fact that h is the convolution of a density with the Gaussian density, and hence it admits an analytic continuation on the whole complex plane. Therefore, the results follow by writing that if $f(x) = \cos(jx) = [\exp(ijx) + \exp(-ijx)]/2$, then $\Gamma_f(y)$ can also be written as

$$\begin{aligned}\Gamma_f(y) &= \frac{e^{-j^2/2}}{2\sqrt{2\pi}} \left[\int e^{ijy} e^{-(y+ij-x)^2/2} g(x)\, dx + \int e^{-ijy} e^{-(y-ij-x)^2/2} g(x)\, dx \right] \\ &= \frac{e^{-j^2/2}}{2} \left[e^{ijy} h(y+ij) + e^{-ijy} h(y-ij) \right].\end{aligned}$$

Any kernel K satisfying Conditions **[K1]**–**[K3]** also admits an analytic continuation on the whole complex plane, so that the formulae are also valid for the estimator $\widehat{\Gamma}_{f,n}$.

Remark 1. Formulae (5) to (10) are useful for obtaining lower bounds, especially for constructing sub-families related to the density h.

3. Pointwise Estimation: Upper and Lower Bounds

We first consider polynomial functions, starting by recalling the results by Taupin [21] about upper bounds for the pointwise quadratic risk and then giving the corresponding lower bounds.

Proposition 3.1. (Consequence of Proposition 3.1 in Taupin [21].) *Fix an integer $\ell \geq 1$ and a polynomial function f of type $\mathcal{P}_{\beta,\ell}$. Consider the estimator $\widehat{\Gamma}_{f,n}$ defined by* (1) *with the kernel S defined by* (2) *and the bandwidth $C_n = \sqrt{\log n}$. Then for all y_0 in $\mathbb{R}$,*

$$\limsup_{n\to\infty} \sup_{g\in\mathcal{G}_f} \frac{1}{|\beta_\ell|} \left(\frac{\pi(2\ell+1)}{h(y_0)} \right)^{1/2} \frac{\sqrt{n}}{(\log n)^{(2\ell+1)/4}} \mathbb{E}^{1/2} \left[\widehat{\Gamma}_{f,n}(y_0) - \Gamma_f(y_0) \right]^2 \leq 1.$$

The following theorem states that this rate of convergence is the best achievable one and hence the estimator defined by (1) achieves it.

Theorem 3.1. *Fix an integer $\ell \geq 1$ and a real number y_0. Then there exists a density h_0 in $\mathcal{H}$ such that*

$$\liminf_{n\to\infty} \inf_{T_n} \sup_{h\in\mathcal{H}} \frac{\sqrt{n}}{(\log n)^{(2\ell+1)/4}} \mathbb{E}^{1/2} \left[T_n - h^{(\ell)}(y_0) \right]^2 \geq \left(\frac{h_0(y_0)}{\pi(2\ell+1)} \right)^{1/2}.$$

Furthermore, if f is a polynomial function of type $\mathcal{P}_{\beta,\ell}$, then

$$\liminf_{n\to\infty} \inf_{T_n} \sup_{g\in\mathcal{G}_f} \frac{\sqrt{n}}{(\log n)^{(2\ell+1)/4}} \mathbb{E}^{1/2} \left[T_n - \Gamma_f(y_0) \right]^2 \geq |\beta_\ell| \left(\frac{h_0(y_0)}{\pi(2\ell+1)} \right)^{1/2}.$$

These infima are taken over all estimators T_n based on the observations $Y_1, \ldots, Y_n$.

Remark 2. Note that the rate obtained in Theorem 3.1 is optimal but with a constant different from the one obtained in Proposition 3.1. Nevertheless, it seems that it cannot be further improved for all $g \in \mathcal{G}_f$. Indeed, as noticed above, these bounds hold whatever be the density g, without additional assumptions on g. We could have the same constants in the upper and lower bounds if we restricted g to belong to a smoothness class. Then we would obtain a local minimax asymptotic result. Without smoothness constraints on g, it seems that there is no hope to obtain the constant $h(y_0)/[\pi(2\ell+1)]$ in both bounds.

The lower bound when $\ell = 0$ was obtained by Taupin [20]. The main tool in the proof is the van Trees inequality (see Gill and Levit [13]).

We now turn to trigonometric functions. The upper bound for the pointwise estimation is given in the following theorem.

Theorem 3.2. (Consequence of Proposition 3.1 in Taupin [13].) *For a fixed integer $\ell \geq 1$ let f be a trigonometric function of the form $\mathcal{C}_{\beta,\ell}$ or $\mathcal{S}_{\beta,\ell}$. Let $\widehat{\Gamma}_{f,n}$ be defined by* (1)*, with the kernel S defined by* (2) *and the bandwidth $C_n = \sqrt{\log n - (\frac{1}{2})\log\log n}$. Then, for all y_0 in $\mathbb{R}$,*

$$\limsup_{n\to\infty} \sup_{g\in\mathcal{G}_f} \left(\frac{2\pi}{|\beta_\ell| e^{-\ell^2/2} [h * \varphi_\ell(y_0)]^{1/2}} \right) \frac{\sqrt{n}}{\exp\{\ell\sqrt{\log n}\}} \mathbb{E}^{1/2} [\widehat{\Gamma}_{f,n}(y_0) - \Gamma_f(y_0)]^2 \leq 1,$$

where $\varphi_\ell(y) = 1/(y^2+\ell^2)$.

We give a new proof of this theorem based on Lemma 2.2 (which was not used in [21]). This proof is of interest since it uses techniques specifically related to this link between trigonometric functionals and the density h.

The following theorem states that the estimator defined by (1) or equivalently by (8) is "nearly minimax". As a matter of fact, the upper bound is of order $n^{-1/2}\exp\{\ell\sqrt{\log n}\}$ and the lower bound of order $n^{-1/2}(\log n)^{-1/4}\exp\{\ell\sqrt{\log n}\}$, with a loss of order $(\log n)^{1/4}$ negligible with respect to the considered rate.

Theorem 3.3. *For a fixed integer $\ell \geq 1$ let f be a trigonometric function of type $\mathcal{C}_{\beta,\ell}$ or $\mathcal{S}_{\beta,\ell}$. Then there exists a density $h_0 = g_0 * f_\varepsilon$ in $\mathcal{H}$ such that for all y_0 in $\mathbb{R}$,*

$$\liminf_{n\to\infty} \inf_{T_n} \sup_{g\in\mathcal{G}_f} \frac{\sqrt{n}(\log n)^{1/4}}{\exp\{\ell\sqrt{\log n}\}} \mathbb{E}^{1/2}[T_n - \Gamma_f(y_0)]^2 \geq \frac{\Gamma_0(y_0)}{(4\pi\ell^2 h_0(y_0))^{1/2}},$$

where the infimum is taken over all the estimators T_n based on the observations $Y_1, \ldots, Y_n$, and $\Gamma_0(y_0) = \int f(x) g_0(x) f_\varepsilon(x - y_0)\, dx$.

4. Upper Bounds for the $\mathbb{L}_p(\mathbb{R})$-Risk

Now we come to our main contribution, that is estimating Γ_f and h with respect to $\mathbb{L}_p(\mathbb{R})$-norm, when $2 \leq p \leq \infty$. We start with upper bounds for the rate of convergence of the $\mathbb{L}_p(\mathbb{R})$-risk before showing that these rates are the best achievable and that our estimator defined in (1) with suitable kernel essentially achieves these rates.

4.1. Upper bound of the $\mathbb{L}_p(\mathbb{R})$-risk for polynomial functions. In this subsection, we are interested in upper bounds for the $\mathbb{L}_p(\mathbb{R})$-risk of $\widehat{\Gamma}_{f,n}$ when f

is a polynomial function of type $\mathcal{P}_{\beta,\ell}$, with $\ell \geq 1$. According to Lemma 2.1, we start with the rate of convergence of the estimators of the derivatives $h^{(k)}$, $0 \leq k \leq \ell$.

Theorem 4.1. *Let ℓ be a fixed integer and $\widehat{h}_n^{(\ell)}$ the estimator defined by* (4) *with the kernel V defined by* (3) *and the bandwidth $C_n = \sqrt{\log n}$. If $2 \leq p < \infty$, then*

$$\limsup_{n\to\infty} \sup_{h\in\mathcal{H}} \frac{\sqrt{n}}{(\log n)^{(2\ell+1)/4}} \mathbb{E}\, \|\widehat{h}_n^{(\ell)} - h^{(\ell)}\|_p < +\infty.$$

If $p = \infty$, then

$$\limsup_{n\to\infty} \sup_{h\in\mathcal{H}} \frac{\sqrt{n}}{(\log n)^{(2\ell+1)/4}\sqrt{\log\log n}} \mathbb{E}\, \|\widehat{h}_n^{(\ell)} - h^{(\ell)}\|_\infty < +\infty.$$

The result for $p = \infty$ and $\ell = 0$ was obtained by Taupin [21].

Remark 3. The main tool for the proof of upper bounds is Rosenthal's inequality with optimal constants. This tool does not allow us to give the exact constant in these upper bounds.

Now we turn to a general polynomial function f. We start by presenting some additional conditions needed for the control of the $\mathbb{L}_p(\mathbb{R})$-risks related to the estimation of $x^{\ell-k}h^{(k)}(x)$.

First note that Theorem 4.1 holds for any integrable kernel K satisfying Conditions **[K1]**–**[K3]**. And when f is a polynomial function of degree less or equal to ℓ, by using Lemma 2.1 and according to (6), we have to choose a kernel admitting a more regular Fourier transform than V (which means a kernel admitting more finite moments) in order to estimate the functional Γ_f. This new kernel has to satisfy the following condition.

[K4] For all $0 \leq k \leq \ell - 1$ and for all $p \geq 2$, the integral $\int |x|^{p(\ell-k)} |K^{(k)}(x)|^p\, dx$ is finite.

Classical analysis results ensure the existence of an integrable kernel satisfying Conditions **[K1]**–**[K4]** even if Assumption **[K4]** has to hold for all $k \in \mathbb{N}$. Moreover, Theorem 4.1 still holds with any kernel satisfying Conditions **[K1]**–**[K4]**.

Secondly, for a polynomial function, the density g has to admit a finite moment of order $p\ell$. Consequently, for any $M > 0$ and any $r \geq 1$, we define the set

$$\mathcal{G}_r(M) = \left\{ g \in \mathcal{G};\ \int |y|^r h(y)\, dy \leq M, \text{ where } h = g * f_\varepsilon \right\}.$$

This condition concerns only the densities g since all the moments of the Gaussian variable are bounded.

Corollary 4.1. *For a fixed integer ℓ, let f be a polynomial function of type $\mathcal{P}_{\beta,\ell}$. Consider the estimator $\widehat{\Gamma}_{f,n}$ defined by* (1) *with a kernel K satisfying* **[K1]**–**[K4]** *and the bandwidth $C_n = \sqrt{\log n}$. If $2 \leq p < \infty$, then for any finite $M_{p\ell} > 0$*

$$\limsup_{n\to\infty} \sup_{g\in\mathcal{G}_{p\ell}(M_{p\ell})} \frac{\sqrt{n}}{(\log n)^{(2\ell+1)/4}}\, \mathbb{E}\, \|\widehat{\Gamma}_{f,n} - \Gamma_f\|_p < +\infty.$$

If $p = \infty$, for any infinite sequence $\{M_{p'\ell}\}_{p'}$, $p' \geq 2$, of positive finite numbers satisfying

$$M_{p'\ell} \leq VK^{p'} \exp\left\{(p'/2-1)e^{2p'} + (p')^2/2 - (p'/2)\log p'\right\} \tag{11}$$

with V, K positive constants, we have

$$\limsup_{n\to\infty} \sup_{g\in\bigcup_{p'}\mathcal{G}_{p'\ell}(M_{p'\ell})} \frac{\sqrt{n}}{(\log n)^{(2\ell+1)/4}\sqrt{\log\log n}}\,\mathbb{E}\,\|\widehat{\Gamma}_{f,n} - \Gamma_f\|_\infty < +\infty.$$

Remark 4. Note that the second part of Corollary 4.1 improves the result of Taupin [21] since her result gave an upper bound only for a uniform norm on a compact set and the rate she gave, namely $\sqrt{n}(\log n)^{-(2\ell+1)/4}(\log n)^{-1/2}$, is slower than this one. Also observe the classical loss of order $\sqrt{\log\log n}$ between the rate of convergence of the $\mathbb{L}_p(\mathbb{R})$-risk for $2 \leq p < \infty$ and the rate of the $\mathbb{L}_\infty(\mathbb{R})$-risk.

Our result holds when we consider $\mathbb{L}_p(K)$-norms on a compact set K without any condition on $\mathbb{E}\,(|Y|^{p\ell})$. When we consider $\mathbb{L}_p(\mathbb{R})$-norms, the uniform control of the $p\ell$th moment of g seems unavoidable. Furthermore, the method used for the uniform convergence requires the control of these $p\ell$th moments for all p, which is even stronger. Also note that the condition on the growth of the moments $\mathbb{E}\,|Y|^m$ is satisfied for usual distributions admitting all finite moments. In particular, this condition is fulfilled by distributions satisfying the classical condition $\mathbb{E}\,|Y|^m \leq VK^{m-2}m!$ with V, K positive constants.

4.2. Upper bound of the $\mathbb{L}_p(\mathbb{R})$-risk for trigonometric functions.

Theorem 4.2. *For a given integer $\ell \geq 1$, let f be a fixed trigonometric function of type $\mathcal{C}_{\beta,\ell}$ (resp. $\mathcal{S}_{\beta,\ell}$). Consider $\widehat{\Gamma}_{f,n}$ as defined by* (1) *with the kernel S (defined by* (2)) *and the bandwidth $C_n = \sqrt{\log n - \frac{1}{2}\log\log n}$. If $2 \leq p < \infty$, then we have*

$$\limsup_{n\to\infty} \sup_{g\in\mathcal{G}_f} \frac{\sqrt{n}}{\exp(\ell\sqrt{\log n})}\mathbb{E}\,\|\widehat{\Gamma}_{f,n} - \Gamma_f\|_p < +\infty.$$

If $p = \infty$, then we have

$$\limsup_{n\to\infty} \sup_{g\in\mathcal{G}_f} \frac{\sqrt{n}}{\exp(\ell\sqrt{\log n})\sqrt{\log\log n}}\mathbb{E}\,\|\widehat{\Gamma}_{f,n} - \Gamma_f\|_\infty < +\infty.$$

Even for trigonometric functions, we observe the classical loss of order $\sqrt{\log\log n}$ between the rate of convergence of the $\mathbb{L}_p(\mathbb{R})$-risk for $2 \leq p < \infty$ and the rate of the $\mathbb{L}_\infty(\mathbb{R})$-risk.

5. Lower Bounds

5.1. Lower bounds for the $\mathbb{L}_\infty(\mathbb{R})$-risk. The following theorem gives the lower bound for the minimax uniform risk over $\mathcal{H}$ for the estimation of the density h of the observations and its derivatives $h^{(\ell)}$ when $\ell \geq 1$.

Theorem 5.1. *For any integer $\ell \geq 0$, we have*

$$\liminf_{n\to\infty} \inf_{T_n} \sup_{h\in\mathcal{H}} \frac{\sqrt{n}}{(\log n)^{(2\ell+1)/4}\sqrt{\log\log n}} \|h^{(\ell)} - T_n\|_\infty > 0,$$

with the infimum taken over all estimators T_n based on the observations $Y_1, \dots, Y_n$.

The following corollary to Theorem 5.1, which is a consequence of Lemma 2.1, gives the lower bound for the uniform estimation of Γ_f when f is a polynomial function.

Corollary 5.1. *For an integer $\ell \geq 1$, let f be a polynomial function of type $\mathcal{P}_{\beta,\ell}$. Then we have*

$$\liminf_{n\to\infty} \inf_{T_n} \sup_{g\in\mathcal{G}_f} \frac{\sqrt{n}}{(\log n)^{(2\ell+1)/4}\sqrt{\log\log n}} \|\Gamma_f - T_n\|_\infty > 0,$$

with the infimum taken over all estimators T_n based on the observations $Y_1, \dots, Y_n$.

The next theorem gives the lower bound for the minimax uniform risk when f is a trigonometric function.

Theorem 5.2. *For an integer $\ell \geq 1$, let f be a trigonometric function of type $\mathcal{C}_{\beta,\ell}$ or $\mathcal{S}_{\beta,\ell}$. Then we have*

$$\liminf_{n\to\infty} \inf_{T_n} \sup_{g\in\mathcal{G}_f} \frac{\sqrt{n}(\log n)^{3/4}}{\exp(\ell\sqrt{\log n})\sqrt{\log\log n}} \|\Gamma_f - T_n\|_\infty > 0,$$

with the infimum taken over all estimators T_n based on the observations $Y_1, \dots, Y_n$.

By comparing Theorems 5.2 and 4.2, we observe, as for the pointwise risk, a loss of order $(\log n)^{3/4}$, which is negligible with respect to the considered rate of convergence.

5.2. Lower bounds for the $\mathbb{L}_p(\mathbb{R})$-risk, $2 \leq p < \infty$, for polynomial functions. The following theorem gives a lower bound for the minimax $\mathbb{L}_p(\mathbb{R})$-risk over $\mathcal{H}$, for the estimation of the density h and its derivatives $h^{(\ell)}$ when $\ell \geq 1$.

Theorem 5.3. *For any integer $\ell \geq 0$ and any $p \geq 2$, we have*

$$\liminf_{n\to\infty} \inf_{T_n} \sup_{h\in\mathcal{H}} \frac{\sqrt{n}}{(\log n)^{(2\ell+1)/4}} \|h^{(\ell)} - T_n\|_p > 0,$$

with the infimum taken over all estimators T_n based on the observations $Y_1, \dots, Y_n$.

The following corollary to Theorem 5.3, which is a consequence of Lemma 2.1, gives a lower bound for the estimation of Γ_f when f is a polynomial function. Its proof is a generalization of the methods used in the proof of Theorem 5.3.

Corollary 5.2. *For an integer $\ell \geq 1$, let f be a polynomial function of type $\mathcal{P}_{\beta,\ell}$. Then, for any $2 \leq p < \infty$, we have*

$$\liminf_{n\to\infty} \inf_{T_n} \sup_{g\in\mathcal{G}_f} \frac{\sqrt{n}}{(\log n)^{(2\ell+1)/4}} \|\Gamma_f - T_n\|_p > 0,$$

with the infimum taken over all estimators T_n based on the observations $Y_1, \dots, Y_n$.

6. Proofs of Pointwise Results

6.1. Proof of pointwise results for polynomial functions.

6.1.1. *Proof of Theorem* 3.1 (lower bound for polynomial functions). The first step of the proof consists in constructing a one-dimensional family of densities $\{h_{\theta,\ell}\}_{|\theta|\le\theta_n}$, $\theta_n = \sqrt{\log n}/\sqrt{n}$, contained in $\mathcal{H}$ for n large enough. Then we consider the sub-family $\{\Gamma_{\theta,\ell}\}_{|\theta|\le\theta_n}$ with $\theta_n = \sqrt{\log n}/\sqrt{n}$, where, according to Lemma 2.1, $\Gamma_{\theta,\ell}$ is given by $\Gamma_{\theta,\ell} = \beta_\ell h^{(\ell)}_{\theta,\ell} + \sum_{k=0}^{\ell-1} Q_{\beta,\ell-k} h^{(k)}_{\theta,\ell}$. Therefore, by construction, the sub-family $\{\Gamma_{\theta,\ell}\}_{|\theta|\le\theta_n}$ is contained in $\{\Gamma_f; g\in\mathcal{G}_f\}$ for a polynomial function f of type $\mathcal{P}_{\beta,\ell}$. The second step consists in applying the van Trees inequality (see Gill and Levit [13]).

We start by constructing the family of densities. Consider the kernel $S_n(x) = C_n S(C_n x)$ with $S(x) = \sin(x)/(\pi x)$ and $(C_n)_{n\ge 0}$ a sequence of real positive numbers tending to infinity. Denote by g_2 the probability density $x \mapsto g_2(x) = \sin^2(x)/(\pi x^2)$, and define g_0 and h_0 to be the probability densities

$$g_0 = g_2 * g_2 \qquad \text{and} \qquad h_0 = f_\varepsilon * g_0.$$

Moreover, define the normalizing constant

$$\overline{S}_n^{(\ell)}(y_0) = \int S_n^{(\ell)}(y_0-u)h_0(u)\,du. \tag{12}$$

For some fixed sequence of parameters $(\theta_n)_{n\ge 0}$ decreasing to zero (to be specified later), we define the parametric path $\{h_{\theta,\ell}\}_{|\theta|\le\theta_n}$ by

$$h_{\theta,\ell}(y) = h_0(y)\left[1+\theta\big(S_n^{(\ell)}(y_0-y) - \overline{S}_n^{(\ell)}(y_0)\big)\right], \tag{13}$$

for all $|\theta|\le\theta_n$ and $y\in\mathbb{R}$. Note that the constant $\overline{S}_n^{(\ell)}(y_0)$ ensures that the density $h_{\theta,\ell}$ integrates to one.

Also note that our path involves the derivative $S_n^{(\ell)}$ of the kernel S_n. The rate of convergence in this lower bound will be the same if we use S_n instead of $S_n^{(\ell)}$. But the use of S_n would provide a different constant.

Denote by $I(\theta,\ell)$ the Fisher information for the family of probability densities $\{h_{\theta,\ell}\}_{|\theta|\le\theta_n}$ given by

$$I(\theta,\ell) = \int\left[\frac{\partial \log h_{\theta,\ell}(x)}{\partial\theta}\right]^2 h_{\theta,\ell}(x)\,dx = \int \frac{\big(S_n^{(\ell)}(y_0-x) - \overline{S}_n^{(\ell)}(y_0)\big)^2}{1+\theta\big(S_n^{(\ell)}(y_0-x)-\overline{S}_n^{(\ell)}(y_0)\big)} h_0(x)\,dx.$$

The following lemma is an immediate extension of a result due to Taupin [20] in case $\ell = 0$. It ensures that with an appropriate choice of the parameters $(C_n)_{n\ge0}$ and $(\theta_n)_{n\ge0}$, the family of densities defined by (13) is contained in the set $\mathcal{H}$. Moreover, it gives an evaluation of the Fisher information of this family.

Condition 1. $\theta_n C_n^{\ell+1} e^{(C_n+2)^2/2} \underset{n\to\infty}{\longrightarrow} 0.$

Lemma 6.1. *Under Condition* 1 *and for large enough* n, *the path* $\{h_{\theta,\ell}\}_{|\theta|\leq\theta_n}$ *is contained in* $\mathcal{H}$. *Moreover,*

$$I(\theta,\ell) = C_n^{2\ell+1}\frac{h_0(y_0)}{\pi(2\ell+1)}(1+o(1)) \qquad \textit{for all} \quad |\theta|\leq\theta_n.$$

We now use the van Trees inequality to get a lower bound for the minimax quadratic risk for the estimation of the derivatives $h^{(\ell)}$ and consequently for polynomial functionals Γ_f, by using Lemma 2.1. Let $\theta\mapsto\lambda_0(\theta)$ be a probability density on $[-1,1]$ satisfying $\lambda_0(-1)=\lambda_0(1)=0$ such that λ_0 is continuously differentiable on $]-1;1[$. Its Fisher information is defined by

$$I_0 = \int_{-1}^{1}\frac{{\lambda_0'}^2(\theta)}{\lambda_0(\theta)}\,d\theta.$$

By rescaling this probability density on the interval $[-\theta_n;\theta_n]$, we define the probability density $\lambda(\theta)=\theta_n^{-1}\lambda_0(\theta_n^{-1}\theta)$ with Fisher information $I(\lambda)=\theta_n^{-2}I_0$. For any fixed y_0 in $\mathbb{R}$ we have

$$\begin{aligned}\inf_{\widehat{h}_n}\sup_{h\in\mathcal{H}}\mathbb{E}\,[\widehat{h}_n-h^{(\ell)}(y_0)]^2 &\geq \inf_{\widehat{h}_n}\sup_{|\theta|\leq\theta_n}\mathbb{E}_{h_\theta}[\widehat{h}_n-h^{(\ell)}_{\theta,\ell}(y_0)]^2\\ &\geq \inf_{\widehat{h}_n}\int_{-\theta_n}^{+\theta_n}\mathbb{E}_{h_{\theta,\ell}}[\widehat{h}_n-h^{(\ell)}_{\theta,\ell}(y_0)]^2\lambda(\theta)\,d\theta,\end{aligned}$$

where the infima are taken over all the estimators $\widehat{h}_n$ based on the observations $Y_1,\dots,Y_n$. Applying the van Trees inequality [13], we get

$$\begin{aligned}(14)\qquad &\inf_{\widehat{h}_n}\sup_{h\in\mathcal{H}}\mathbb{E}\,[\widehat{h}_n-h^{(\ell)}(y_0)]^2\\ &\geq\left[\int_{-\theta_n}^{\theta_n}\frac{\partial h^{(\ell)}_{\theta,\ell}(y_0)}{\partial\theta}\lambda(\theta)\,d\theta\right]^2\left[n\int_{-\theta_n}^{\theta_n}I(\theta,\ell)\lambda(\theta)\,d\theta+I(\lambda)\right]^{-1}.\end{aligned}$$

By definition, $h^{(\ell)}_{\theta,\ell}$ satisfies

$$\frac{\partial h^{(\ell)}_{\theta,\ell}(y_0)}{\partial\theta}=\frac{d^\ell}{dy^\ell}[h_0(y)S_n^{(\ell)}(y_0-y)]\big|_{y=y_0}-h_0^{(\ell)}(y_0)\overline{S}_n^{(\ell)}(y_0),$$

and using that $S_n^{(2\ell)}(0)=C_n^{2\ell+1}S^{(2\ell)}(0)=C_n^{2\ell+1}[\pi(2\ell+1)]^{-1}$, we obtain

$$\left[\frac{d^\ell}{dy^\ell}h_0(y)S_n^{(\ell)}(y_0-y)\right]\bigg|_{y=y_0}=C_n^{2\ell+1}h_0(y_0)(-1)^\ell S^{(2\ell)}(0)(1+o(1)).$$

We now use that $\overline{S}_n^{(\ell)}(y_0)=S_n^{(\ell)}*h_0$ and therefore, for all t in $\mathbb{R}$, the following equality holds:

$$\big(S_n^{(\ell)}*h_0-h_0^{(\ell)}\big)^\star(t)=(it)^\ell h_0^\star(t)(S_n^\star(t)-1).$$

Using the properties of the kernel S_n and the fact that the function $|h^*(t)|$ is bounded by $e^{-t^2/2}$, we get

$$\|S_n^{(\ell)} * h_0 - h_0^{(\ell)}\|_\infty \leq 2(\pi)^{-1} C_n^{\ell-1} e^{-C_n^2/2}, \tag{15}$$

which yields

$$\overline{S}_n^{(\ell)}(y_0) = h_0^{(\ell)}(y_0)(1 + o(1)). \tag{16}$$

Consequently

$$\left[\int_{-\theta_n}^{\theta_n} \frac{\partial h_{\theta,\ell}^{(\ell)}(y_0)}{\partial \theta} \lambda(\theta)\, d\theta\right]^2 = \left[h_0(y_0) C_n^{2\ell+1} S^{(2\ell)}(0)\right]^2 (1 + o(1)).$$

Apply Lemma 6.1 to get the following lower bound

$$\inf_{\widehat{h}_n} \sup_{h \in \mathcal{H}} \mathbb{E}\,[\widehat{h}_n - h^{(\ell)}(y_0)]^2 \geq \frac{C_n^{2(2\ell+1)} h_0^2(y_0)[\pi(2\ell+1)]^{-2}(1+o(1))}{nC_n^{2\ell+1} h_0(y_0)[\pi(2\ell+1)]^{-1}(1+o(1)) + \theta_n^{-2} I_0}.$$

Maximize this bound under the constraint given by Condition 1 and choose $C_n = \sqrt{\log n} - 2\sqrt{2}$ and $\theta_n = n^{-1/2}\sqrt{\log n}$ to obtain the desired result for $h^{(\ell)}$ and consequently for Γ_f by using Lemma 2.1. The result for Γ_f follows by noting that the previous calculations hold if we replace $\partial h_{\theta,\ell}^{(\ell)}(y_0)/\partial\theta$ by $\partial \Gamma_{\theta,\ell}(y_0)/\partial\theta$. □

6.2. Proofs of pointwise results for trigonometric functions.

6.2.1. *Proof of Theorem* 3.2 (upper bound for trigonometric functions). For the sake of simplicity, we only give the proof for the function $f\colon x \mapsto \cos(\ell x)$. Using the triangle inequality, the risk $\mathbb{E}\,[\widehat{\Gamma}_{f,n}(y_0) - \Gamma_f(y_0)]^2$ is bounded by the sum of the squared bias term $\mathbb{E}\,[\widehat{\Gamma}_{f,n}(y_0)] - \Gamma_f(y_0)$ and the variance term $\mathrm{Var}(\widehat{\Gamma}_{f,n})$. We start with the bound for the bias $\mathbb{E}\,[\widehat{\Gamma}_{f,n}(y_0)] - \Gamma_f(y_0)$, which is, by Lemma 2.2, given by

$$\frac{e^{-\ell^2/2}}{2}\Big[e^{i\ell y_0}\big(\mathbb{E}\,S_n(y_0+i\ell-Y_1) - h(y_0+i\ell)\big) + e^{-i\ell y_0}\big(\mathbb{E}\,S_n(y_0-i\ell-Y_1) - h(y_0-i\ell)\big)\Big].$$

By using Parseval's identity we infer that for $\delta \in \{-1; +1\}$

$$\int S_n(y_0 + \delta i\ell - x) h(x)\, dx = \frac{1}{2\pi}\int e^{ity_0 - \delta t\ell} h^\star(t) S_n^\star(t)\, dt,$$
$$\text{and} \quad h(y_0 + \delta i\ell) = \frac{1}{2\pi}\int e^{ity_0 - \delta t\ell} h^\star(t)\, dt,$$

which implies the following bound:

$$\big|\mathbb{E}\,S_n(y_0 + \delta i\ell - Y_1) - h(y_0 + \delta i\ell)\big| \leq \frac{1}{2\pi}\int e^{\delta t\ell} |h^\star(t)||S_n^\star(t) - 1|\, dt.$$

Hence, by using that S satisfies Conditions **[K1]**–**[K3]** we get the bound for the bias

$$\sup_{y_0\in\mathbb{R}} \left|\mathbb{E}\,[\widehat{\Gamma}_{f,n}(y_0)] - \Gamma_f(y_0)\right|^2 \leq \frac{1}{4\pi^2}\frac{e^{-(C_n-\ell)^2}}{(C_n-\ell)^2}. \tag{17}$$

We now come to the variance term. Using the independence of the variables and $\mathrm{Var}(X) \leq \mathbb{E}\,[|X|^2]$ we get

$$\mathrm{Var}(\widehat{\Gamma}_{f,n}(y_0)) \leq \frac{e^{-\ell^2}}{2n}\Big[\mathbb{E}\,|S_n(y_0+i\ell-Y_1)|^2 + \mathbb{E}\,|S_n(y_0-i\ell-Y_1)|^2\Big].$$

By definition, setting δ in $\{-1;1\}$, we get

$$\mathbb{E}\,|S_n(y_0+\delta i\ell-Y_1)|^2 = C_n\int |S(u+\delta i\ell C_n)|^2 h(y_0+u/C_n)\,du$$

with

$$|S(u+\delta i\ell C_n)|^2 = \frac{e^{2\ell C_n}+e^{-2\ell C_n}-2\cos(2u)}{4\pi^2(u^2+\ell^2C_n^2)}.$$

Consequently the variance is bounded as follows:

$$\begin{aligned}\mathrm{Var}(\widehat{\Gamma}_{f,n}(y_0)) &\leq \frac{e^{-\ell^2}C_n(e^{2\ell C_n}+3)}{4\pi^2 n}\int\frac{h(y_0+u/C_n)}{u^2+\ell^2C_n^2}\,du\\ &\leq \frac{e^{-\ell^2}(e^{2\ell C_n}+3)}{4\pi^2 n}\int\frac{h(y_0+v)}{v^2+\ell^2}\,dv.\end{aligned}$$

Combine this inequality with (17) and get the result by choosing

$$C_n = \sqrt{\log n - \tfrac{1}{2}\log\log n}. \qquad \square$$

6.2.2. *Proof of Theorem* 3.3 (lower bound for trigonometric functions). We only give the proof for the function $f\colon x\mapsto\cos(\ell x)$. To prove this theorem, we consider the special case $\{h_{\theta,0}\}_{|\theta|\leq\theta_n}$ of the path used in the proof of Theorem 3.1, shorten as $\{h_\theta\}_{|\theta|\leq\theta_n}$, and let

$$\Gamma_\theta(y) = \frac{e^{-\ell^2/2}}{2}\big[e^{i\ell y}h_\theta(y+i\ell) + e^{-i\ell y}h_\theta(y-i\ell)\big]. \tag{18}$$

Under Condition 1, the family $\{h_\theta\}_{|\theta|\leq\theta_n}$ is contained in $\mathcal{H}$, so that by construction and Lemma 2.2, the family $\{\Gamma_\theta\}_{|\theta|\leq\theta_n}$ belongs to $\{\Gamma_f; g\in\mathcal{G}_f\}$. Proceeding as for polynomial functions, we consider a probability density $\lambda_0(\theta)$ on the interval $[-1,1]$ with Fisher information given by

$$I_0 = \int_1^1 \frac{{\lambda_0'}^2(\theta)}{\lambda_0(\theta)}\,d\theta,$$

where $\lambda_0(-1) = \lambda_0(1) = 0$ and λ_0 is continuously differentiable on $]-1;1[$. Applying the van Trees inequality we see that the pointwise minimax quadratic risk is bounded from below as follows:

$$\inf_{\widehat{\Gamma}_n} \sup_{g\in\mathcal{G}_f} \mathbb{E}\,[\widehat{\Gamma}_n - \Gamma_f(y_0)]^2 \geq \left[\int_{-\theta_n}^{\theta_n} \frac{\partial\Gamma_\theta(y_0)}{\partial\theta}\lambda(\theta)\,d\theta\right]^2 \left[n\int_{-\theta_n}^{\theta_n} I(\theta)\lambda(\theta)\,d\theta + I(\lambda)\right]^{-1}.$$

The denominator is studied by applying Lemma 6.1 in the proof of Theorem 3.1 with $\ell = 0$. Thus, it remains to study the behavior of the numerator. Using definitions (13) and (18) we get

$$\frac{\partial\Gamma_\theta(y_0)}{\partial\theta} = \frac{e^{-\ell^2/2}}{2}\Big[e^{i\ell y_0}h_0(y_0+i\ell)S_n(-i\ell)+e^{-i\ell y_0}h_0(y_0-i\ell)S_n(i\ell)\Big]-\Gamma_0(y_0)\overline{S}_n(y_0),$$

where $\Gamma_0(y_0) = \frac{1}{2}e^{-\ell^2/2}\big[e^{i\ell y_0}h_0(y_0+i\ell) + e^{-i\ell y_0}h_0(y_0-i\ell)\big]$. Using that the kernel S is an even function, and applying (16) we get

$$\frac{\partial\Gamma_\theta(y_0)}{\partial\theta} = \Gamma_0(y_0)\big(C_nS(i\ell C_n) - \overline{S}_n(y_0)\big) = \Gamma_0(y_0)C_nS(i\ell C_n)(1+o(1)).$$

The definition of the kernel S leads to

$$\left[\int_{-\theta_n}^{\theta_n} \frac{\partial\Gamma_\theta(y_0)}{\partial\theta}\lambda(\theta)\,d\theta\right]^2 = \frac{\Gamma_0(y_0)^2}{4\pi^2\ell^2}e^{2\ell C_n}[1+o(1)],$$

which, combined with Lemma 6.1, gives that

$$\inf_{\widehat{\Gamma}_n} \sup_{g\in\mathcal{G}_f} \mathbb{E}\,[\widehat{\Gamma}_n - \Gamma_f(y_0)]^2 \geq \frac{(4\pi^2\ell^2)^{-1}\Gamma_0(y_0)^2e^{2\ell C_n}(1+o(1))}{nC_nh_0(y_0)\pi^{-1}(1+o(1)) + \theta_n^{-2}I_0}.$$

Under Condition 1, choose $C_n = \sqrt{\log n} - 2\sqrt{2}$ and $\theta_n = \sqrt{\log n}/\sqrt{n}$ to get the result. $\square$

7. Proofs of results for $\mathbb{L}_p(\mathbb{R})$-risk

7.1. Proofs of upper bounds for the $\mathbb{L}_p(\mathbb{R})$-risk when $2 \leq p \leq \infty$.

6.2.2. *Proof of Theorem* 4.1 (upper bound of $\mathbb{L}_p(\mathbb{R})$-risk for the density and its derivatives). By using the triangle inequality, we get

$$\mathbb{E}\,\|h^{(\ell)} - \widehat{h}_n^{(\ell)}\|_p \leq \|h^{(\ell)} - \mathbb{E}\,\widehat{h}_n^{(\ell)}\|_p + \mathbb{E}\,\|\widehat{h}_n^{(\ell)} - \mathbb{E}\,\widehat{h}_n^{(\ell)}\|_p. \tag{19}$$

The first term in (19) (bias term) $\|\mathbb{E}\,\widehat{h}_n^{(\ell)} - h^{(\ell)}\|_p$, is bounded, for $2 \leq p \leq \infty$, by a method valid for any kernel K satisfying Assumption **[K2]**. The second term $\mathbb{E}\,\|\widehat{h}_n^{(\ell)} - \mathbb{E}\,\widehat{h}_n^{(\ell)}\|_p$ for $2 \leq p < \infty$ is controlled by applying Rosenthal's inequality [19], with optimal constants given in Pinelis [18]. The control of $\mathbb{E}\,\|\widehat{h}_n^{(\ell)} - \mathbb{E}\,\widehat{h}_n^{(\ell)}\|_\infty$ is obtained following Ibragimov and Hasminskii [15].

Let us start with the control of $\|\mathbb{E}\,\widehat{h}_n^{(\ell)} - h^{(\ell)}\|_p$ for $2 \leq p \leq \infty$. Using that the kernel V satisfies Conditions **[K1]**–**[K3]**, we apply the same reasoning as for inequality (15) to obtain

$$\|\mathbb{E}\,\widehat{h}_n^{(\ell)} - h^{(\ell)}\|_\infty = \|V_n^{(\ell)} * h - h^{(\ell)}\|_\infty \leq \frac{2}{\pi} C_n^{\ell-1} e^{-C_n^2/2}.$$

Consider now the case $2 \leq p < \infty$. For any function $u \in \mathbb{L}_2(\mathbb{R})$ with $u^\star \in \mathbb{L}_1(\mathbb{R})$,

$$\|u\|_p^p \leq \frac{1}{(2\pi)^{p-1}} \|u^\star\|_1^{p-2} \|u^\star\|_2^2. \tag{20}$$

Apply this result to the function $V_n^{(\ell)} * h - h^{(\ell)}$ to obtain a bound on $\|\mathbb{E}\,\widehat{h}_n^{(\ell)} - h^{(\ell)}\|_p$ when $2 \leq p < \infty$,

$$\begin{aligned}&\|V_n^{(\ell)} * h - h^{(\ell)}\|_p \\ &\quad \leq \frac{1}{(2\pi)^{(p-1)/p}} \left(\int |t|^\ell e^{-t^2/2} \mathbf{1}_{|t|>C_n}\, dt \right)^{\frac{p-2}{p}} \left(\int |t|^{2\ell} e^{-t^2} \mathbf{1}_{|t|>C_n}\, dt \right)^{1/p}.\end{aligned}$$

Note that

$$\int |t|^\ell e^{-t^2/2} \mathbf{1}_{|t|>C_n}\, dt \leq 2 C_n^{\ell-1} e^{-C_n^2/2}(1+o(1))$$

and

$$\int |t|^{2\ell} e^{-t^2} \mathbf{1}_{|t|>C_n}\, dt \leq C_n^{2\ell-1} e^{-C_n^2}(1+o(1))$$

to get the bound

$$\|V_n^{(\ell)} * h - h^{(\ell)}\|_p \leq \frac{1}{2^{1/p}\pi^{1-1/p}} C_n^{\ell-1+1/p} e^{-C_n^2/2}(1+o(1)). \tag{21}$$

We now aim at bounding the term $\mathbb{E}\,\|\widehat{h}_n^{(\ell)} - \mathbb{E}\,\widehat{h}_n^{(\ell)}\|_p$ for $2 \leq p < \infty$. Write $\widehat{h}_n^{(\ell)} - \mathbb{E}\,\widehat{h}_n^{(\ell)}$ as a sum of independent random variables

$$\big(\widehat{h}_n^{(\ell)} - \mathbb{E}\,\widehat{h}_n^{(\ell)}\big)(x) = \frac{1}{n} \sum_{j=1}^n \mathcal{X}_n^{(\ell)}(x, Y_j),$$

where the random variables $\mathcal{X}_n^{(\ell)}(x, Y_j)$ are centered and defined by $\mathcal{X}_n^{(\ell)}(x, Y_j) = V_n^{(\ell)}(x - Y_j) - \mathbb{E}\,V_n^{(\ell)}(x - Y_j)$. Using the concavity of the function $x \mapsto x^{1/p}$ for $p \geq 1$ and applying Fubini's Theorem we infer that

$$\begin{aligned}\mathbb{E}\,\|\widehat{h}_n^{(\ell)} - \mathbb{E}\,\widehat{h}_n^{(\ell)}\|_p &= \mathbb{E}\left\| \frac{1}{n} \sum_{j=1}^n \mathcal{X}_n^{(\ell)}(\cdot, Y_j) \right\|_p \\ &\leq \left[\int \mathbb{E} \left| \frac{1}{n} \sum_{j=1}^n \mathcal{X}_n^{(\ell)}(x, Y_j) \right|^p dx \right]^{1/p}.\end{aligned} \tag{22}$$

We now apply Rosenthal's inequality, with optimal constants as given in Pinelis [18], to the quantity $\mathbb{E}\left|n^{-1}\sum_{j=1}^{n}\mathcal{X}_n^{(\ell)}(x,Y_j)\right|^p$. We will see later that the optimal constants are useful in the case $p=\infty$.

Lemma 7.1 (Corollary 1 in [18]). *Let $X_1,\dots,X_n$ be n independent centered random variables. Put $S_n = X_1+\cdots+X_n$ and $X_n^\star = \max_{1\le i\le n}|X_i|$. Then there exists an absolute constant κ such that, for all $p\ge 1$,*

$$\mathbb{E}^{1/p}|S_n|^p \le \kappa\big[e\sqrt{p}\,\mathbb{E}^{1/2}|S_n|^2 + p\mathbb{E}^{1/p}|X_n^\star|^p\big].$$

An immediate consequence of this inequality is that

$$\mathbb{E}^{1/p}\left|\frac{S_n}{n}\right|^p \le \frac{e\kappa\sqrt{p}}{\sqrt{n}}\,\mathbb{E}^{1/2}|X_1|^2 + \frac{\kappa p}{n^{1-1/p}}\,\mathbb{E}^{1/p}|X_1|^p. \tag{23}$$

Apply inequality (23) to obtain that there exists an absolute constant κ such that

$$\begin{aligned}\mathbb{E}\left|\frac{1}{n}\sum_{j=1}^{n}\mathcal{X}_n^{(\ell)}(x,Y_j)\right|^p \le{}& 2^{p-1}\left[\frac{e\kappa\sqrt{p}}{\sqrt{n}}\mathbb{E}^{1/2}|\mathcal{X}_n^{(\ell)}(x,Y_1)|^2\right]^p\\ &+2^{p-1}\left[\frac{\kappa p}{n^{1-1/p}}\mathbb{E}^{1/p}|\mathcal{X}_n^{(\ell)}(x,Y_1)|^p\right]^p\end{aligned}$$

and therefore

$$\begin{aligned}\mathbb{E}\,\|\widehat{h}_n^{(\ell)}-\mathbb{E}\,\widehat{h}_n^{(\ell)}\|_p \le \Bigg\{&2^{p-1}\left[\frac{e\kappa\sqrt{p}}{\sqrt{n}}\right]^p\int\mathbb{E}^{p/2}|\mathcal{X}_n^{(\ell)}(x,Y_1)|^2\,dx\\ &+2^{p-1}\left[\frac{\kappa p}{n^{1-1/p}}\right]^p\int\mathbb{E}\,|\mathcal{X}_n^{(\ell)}(x,Y_1)|^p\,dx\Bigg\}^{1/p}.\end{aligned}$$

By using Fubini's Theorem, the last term $\int\mathbb{E}\,|\mathcal{X}_n^{(\ell)}(x,Y_1)|^p dx$ is bounded as follows:

$$\int\mathbb{E}\,|\mathcal{X}_n^{(\ell)}(x,Y_1)|^p\,dx \le 2^p\iint|V_n^{(\ell)}(x-y)|^p h(y)\,dy\,dx = 2^pC_n^{(\ell+1)p-1}\|V^{(\ell)}\|_p^p. \tag{24}$$

It remains now to bound the first and main term $\int\mathbb{E}^{p/2}|\mathcal{X}_n^{(\ell)}(x,Y_1)|^2\,dx$. Note that $\mathbb{E}\,|\mathcal{X}_n^{(\ell)}(x,Y_1)|^2 = \mathrm{Var}(V_n^{(\ell)}(x-Y_1)) \le \mathbb{E}\,|V_n^{(\ell)}(x-Y_1)|^2$, so that

$$\int\mathbb{E}^{p/2}|\mathcal{X}_n^{(\ell)}(x,Y_1)|^2\,dx \le \int\left(\int C_n^{2\ell+1}(V^{(\ell)}(u))^2h(x+u/C_n)\,du\right)^{p/2}dx.$$

Proceed to a Taylor expansion of $u\mapsto h(x+u/C_n)$ (the function h in $\mathcal{H}$ admits an analytic continuation on the whole complex plane) valid since the integral $\int u[V^{(\ell)}(u)]^2\,du$ is finite. Hence we get

$$\int\mathbb{E}^{p/2}|\mathcal{X}_n^{(\ell)}(x,Y_1)|^2\,dx \le C_n^{(2\ell+1)p/2}\,\|V^{(\ell)}\|_2^p\,\|h\|_{p/2}^{p/2}(1+o(1)) \quad\text{as}\quad n\to\infty. \tag{25}$$

Combining (24) and (25) we find that $\mathbb{E}\,\|\widehat{h}_n^{(\ell)} - \mathbb{E}\,\widehat{h}_n^{(\ell)}\|_p$ is bounded by

$$2\kappa\left\{\left[\frac{e\sqrt{p}}{\sqrt{n}}\right]^p C_n^{(2\ell+1)p/2}\,\|V^{(\ell)}\|_2^p\,\|h\|_{p/2}^{p/2}(1+o(1))+2^p\left[\frac{p}{n^{1-1/p}}\right]^p C_n^{(\ell+1)p-1}\|V^{(\ell)}\|_p^p\right\}^{1/p}.$$

Finally for $p > 2$ we have

$$\mathbb{E}\,\|\widehat{h}_n^{(\ell)} - \mathbb{E}\,\widehat{h}_n^{(\ell)}\|_p \le 2\frac{e\kappa\sqrt{p}}{\sqrt{n}} C_n^{(2\ell+1)/2}\,\|V^{(\ell)}\|_2\,\|h\|_{p/2}^{1/2}\,(1+o(1)), \tag{26}$$

and for $p = 2$,

$$\mathbb{E}\,\|\widehat{h}_n^{(\ell)} - \mathbb{E}\,\widehat{h}_n^{(\ell)}\|_2 \le 2\kappa\|V^{(\ell)}\|_2\sqrt{2e^2+16}\,\frac{C_n^{(2\ell+1)/2}}{\sqrt{n}}(1+o(1)).$$

Note that these two bounds are uniformly bounded for $h \in \mathcal{H}$ since $\|h\|_\infty \le (2\pi)^{-1/2}$ and hence $\|h\|_{p/2} \le \|h\|_1^{2/p} = 1$. The final result for $2 \le p < \infty$ follows by gathering (21) and (26) and taking $C_n = \sqrt{\log n}$.

For $p = \infty$, the control of the analogue of the variance term $\mathbb{E}\,\|\widehat{h}_n^{(\ell)} - \mathbb{E}\,\widehat{h}_n^{(\ell)}\|_\infty$ follows arguing as Ibragimov and Hasminskii [15]. More precisely, the bound for $\mathbb{E}\,\|\widehat{h}_n^{(\ell)} - \mathbb{E}\,\widehat{h}_n^{(\ell)}\|_\infty$ is deduced from the following lemma (see Nikol'skii [16]). Denote by $\mathcal{M}_{C,p}$ the set of functions belonging to $\mathbb{L}_p(\mathbb{R})$ and having a Fourier transform compactly supported on $[-C;C]$.

Lemma 7.2 (Nikol'skii [16], p. 150). *If $1 \le p \le p' \le \infty$, then for g in $\mathcal{M}_{C,p}$ we have $\|g\|_{p'} \le 2C^{1/p-1/p'}\|g\|_p$.*

We apply Lemma 7.2 to $\widehat{h}_n^{(\ell)} - \mathbb{E}\,\widehat{h}_n^{(\ell)} \in \mathcal{M}_{2C_n,p}$ and infer that for any $p \le \infty$,

$$\|\widehat{h}_n^{(\ell)} - \mathbb{E}\,\widehat{h}_n^{(\ell)}\|_\infty \le 2(2C_n)^{1/p}\|\widehat{h}_n^{(\ell)} - \mathbb{E}\,\widehat{h}_n^{(\ell)}\|_p.$$

Take $p = \log C_n$ to get that

$$\|\widehat{h}_n^{(\ell)} - \mathbb{E}\,\widehat{h}_n^{(\ell)}\|_\infty \le 2(2C_n)^{1/\log C_n}\|\widehat{h}_n^{(\ell)} - \mathbb{E}\,\widehat{h}_n^{(\ell)}\|_p,$$

and therefore, by applying (26) with $\|h\|_{p/2}^{1/2} \le 1$ for all p, we find

$$\mathbb{E}\,\|\widehat{h}_n^{(\ell)} - \mathbb{E}\,\widehat{h}_n^{(\ell)}\|_\infty \le 4\kappa e^2\,\frac{\sqrt{\log C_n}}{\sqrt{n}}C_n^{(2\ell+1)/2}\|V^{(\ell)}\|_2(1+o(1)),$$

and the result follows for $p = \infty$ again by taking $C_n = \sqrt{\log n}$.

This final argument shows how the optimal constants in Rosenthal's inequality are important to get the result when $p = \infty$. □

7.1.2. *Proof of Corollary* 4.1 (upper bound of $\mathbb{L}_p(\mathbb{R})$-risk for polynomial functions). According to Lemma 2.1, the main point of the proof lies in showing that the rate of convergence of $\widehat{\Gamma}_{f,n}$ is given by the rate of convergence of $\widehat{h}_n^{(\ell)}$. This follows by using Theorem 4.1 and the triangle inequality. More precisely, denoting by Q_s the function $y \mapsto y^s$, it suffices to show that there exist constants C and C' such that for all $0 \le k \le \ell - 1$ and for all $0 \le s \le \ell - k$, we have, when $2 \le p < \infty$,

$$\mathbb{E}\,\|Q_s(\widehat{h}_n^{(k)} - h^{(k)})\|_p \le \frac{C(\log n)^{(2\ell+1)/4}}{\sqrt{n}}(1 + o(1)),$$

and when $p = \infty$ then

$$\mathbb{E}\,\|Q_s(\widehat{h}_n^{(k)} - h^{(k)})\|_\infty \le \frac{C'(\log n)^{(2\ell+1)/4}\sqrt{\log\log n}}{\sqrt{n}}(1 + o(1)).$$

We follow the lines of the proof of Theorem 4.1. The control of the bias term defined by $\|Q_s(h^{(k)} - K_n^{(k)} * h)\|_p$ $(2 \le p \le \infty)$ uses the following identity valid for all $t \in \mathbb{R}$

$$[Q_s(h^{(k)} - K_n^{(k)} * h)]^*(t) = \frac{1}{i^s}\frac{\partial^s}{\partial t^s}[h^{(k)} - K_n^{(k)} * h]^*(t) = i^{k-s}\frac{\partial^s}{\partial t^s}[Q_k h^*(1 - K_n^*)](t).$$

Consequently

$$\|Q_s(h^{(k)} - K_n^{(k)} * h)\|_\infty \le \frac{1}{2\pi}\sum_{j=0}^{s}\binom{s}{j}\int\left|\frac{\partial^j}{\partial t^j}Q_k h^*\right|(t)\left|\frac{\partial^{s-j}}{\partial t^{s-j}}(1 - K_n^*)\right|(t)\,dt.$$

Under Assumptions **[K2]** and **[K4]**, the kernel K has the property that for all $0 \le j \le s$, there exists a constant $\mathcal{C}(s)$ such that

$$\left|\frac{\partial^{s-j}}{\partial t^{s-j}}(1 - K_n^*)\right|(t) \le \mathcal{C}(s)\mathbf{1}_{|t| \ge C_n}.$$

Moreover, the quantity $\partial^j Q_k h^*(t)/\partial t^j$ is a linear combination of powers of t times derivatives of h^* at t. Note that $h^* = f_\varepsilon^* g^*$, so that its derivative is a linear combination of derivatives of f_ε^* times derivatives of g^*. Since $\mathbb{E}\,|X|^k < \infty$ for all $k \le \ell$, the kth derivative of g^* exists and is uniformly bounded. Consequently, we bound the bias term in the following way:

$$\|Q_s(h^{(k)} - K_n^{(k)} * h)\|_\infty \le C\int |t|^{k+s}e^{-t^2/2}\mathbf{1}_{|t| \ge C_n}\,dt \le O(1)C_n^{k+s-1}e^{-C_n^2/2},$$

where $0 \le k \le \ell - 1$ and $0 \le s \le \ell - k$, and finally

$$\|Q_s(h^{(k)} - K_n^{(k)} * h)\|_\infty \le O(1)C_n^{\ell-1}e^{-C_n^2/2}.$$

Arguing as in the proof of Theorem 4.1, by using (20), we bound the bias term for the $\mathbb{L}_p(\mathbb{R})$-risk $(p < \infty)$ by writing that

$$\begin{aligned}\|Q_s(h^{(k)} - K_n^{(k)} * h)\|_p \le O(1)&\left[\int |t|^{k+s}e^{-t^2/2}\mathbf{1}_{|t| \ge C_n}\,dt\right]^{(p-2)/p} \\ &\times\left[\int |t|^{2(k+s)}e^{-t^2}\mathbf{1}_{|t| \ge C_n}\,dt\right]^{1/p} \le O(1)C_n^{\ell-1}e^{-C_n^2/2},\end{aligned}$$

the last inequality being valid since $0 \le k \le \ell - 1$ and $0 \le s \le \ell - k$.

Let us study the analogue of the variance term $\mathbb{E}\,\|Q_s(\widehat{h}_n^{(k)} - K_n^{(k)} * h)\|_p$ for $0 \le k \le \ell - 1$, $0 \le s \le \ell - k$, and $2 \le p < \infty$. Denote

$$Q_s(x)(\widehat{h}_n^{(k)} - K_n^{(k)} * h)(x) = \frac{1}{n}\sum_{j=1}^{n} x^s \mathcal{X}_n^{(k)}(x, Y_j)$$

and apply Rosenthal's inequality (23) to get that

$$\begin{aligned}(27)\quad \mathbb{E}\,\|Q_s(\widehat{h}_n^{(k)} - K_n^{(k)} * h)\|_p \le \Bigg\{ & 2^{p-1}\left[\frac{e\kappa\sqrt{p}}{\sqrt{n}}\right]^p \int \mathbb{E}^{\,p/2}\{|\mathcal{X}_n^{(k)}(x, Y_1)|^2 |x|^{2s}\}\, dx \\ & + 2^{p-1}\left[\frac{\kappa p}{n^{1-1/p}}\right]^p \int \mathbb{E}\,|\mathcal{X}_n^{(k)}(x, Y_1)|^p |x|^{ps}\, dx \Bigg\}^{1/p}.\end{aligned}$$

Arguing as in the proof of Theorem 4.1, we need to control

$$\int \mathbb{E}^{\,p/2}\{|\mathcal{X}_n^{(k)}(x, Y_1)|^2 |x|^{2s}\}\, dx \quad \text{and} \quad \int \mathbb{E}\,|\mathcal{X}_n^{(k)}(x, Y_1)|^p |x|^{ps}\, dx.$$

First, by using that $x^{2s} \le 2^{2s-1}(|x-y|^{2s} + |y|^{2s})$, we get

$$\begin{aligned}\int \mathbb{E}^{\,p/2}\{|\mathcal{X}_n^{(k)}(x, Y_1)|^2 |x|^{2s}\}\, dx \le 2^{(2s-1)p/2} \int \Bigg(& \int |x-y|^{2s} |K_n^{(k)}(x-y)|^2 h(y)\, dy \\ & + \int |K_n^{(k)}(x-y)|^2 |y|^{2s} h(y)\, dy \Bigg)^{p/2} dx.\end{aligned}$$

Setting $u = C_n(y - x)$ in the outer integral we obtain that

$$\begin{aligned}(28)\quad & \int \mathbb{E}^{\,p/2}\{|\mathcal{X}_n^{(k)}(x, Y_1)|^2 |x|^{2s}\}\, dx \\ & \le 2^{sp-1} C_n^{(k+1)p-1}\left(\|Q_s K^{(k)}\|_p^p + \left(\int |y|^{2s} h(y)\, dy\right)^{p/2} \|K^{(k)}\|_p^p \right).\end{aligned}$$

Now, by using Jensens's inequality, we infer that

$$\int \mathbb{E}\,|\mathcal{X}_n^{(k)}(x, Y_1)|^p |x|^{ps}\, dx \le 2^p \iint |x|^{ps} |K_n^{(k)}(x-y)|^p h(y)\, dy\, dx,$$

which is bounded by

$$2^{p(s+1)} \iint |x-y|^{ps} |K_n^{(k)}(x-y)|^p h(y)\, dy\, dx + 2^{p(s+1)} \iint |y|^{ps} |K_n^{(k)}(x-y)|^p h(y)\, dy\, dx,$$

and therefore, for $0 \le s \le \ell - k$,

$$\begin{aligned}(29)\quad & \int \mathbb{E}\,|\mathcal{X}_n^{(k)}(x, Y_1)|^p |x|^{ps}\, dx \\ & \le 2^{p(s+1)}\left[\|Q_s K^{(k)}\|_p^p + \|K^{(k)}\|_p^p \left(\int |y|^{ps} h(y)\, dy\right)\right] C_n^{(k+1)p-1}.\end{aligned}$$

Combining (27)–(29) we obtain that $\mathbb{E}\,\|Q_s(\widehat{h}_n^{(k)} - K_n^{(k)} * h)\|_p$ is bounded by

$$
\begin{aligned}
(30)\quad 2^{s+2}e\kappa\Bigg[\sqrt{\frac{p^p}{n^p}}\big(\|Q_sK^{(k)}\|_p^p + M_{2s}^{p/2}\|K^{(k)}\|_p^p\big)\\
+\frac{p^p}{n^{p-1}}\big(\|Q_sK^{(k)}\|_p^p + \mathbb{E}\,(|Y|^{ps})\|K^{(k)}\|_p^p\big)\Bigg]^{1/p} C_n^{(k+1)-1/p}.
\end{aligned}
$$

Since $k+1\le \ell$, we see that for any density g such that all the moments $\mathbb{E}\,(|Y|^{2s})$ and $\mathbb{E}\,(|Y|^{ps})$ (with $0\le s\le \ell$) are uniformly bounded by $M_{p\ell}$

$$
\mathbb{E}\,\|Q_s(\widehat{h}_n^{(k)} - K_n^{(k)} * h)\|_p \le O(1)\frac{C_n^{\ell-1/p}}{\sqrt{n}} \le O(1)\bigg(\frac{(\log n)^{(2\ell+1)/4}}{\sqrt{n}}\bigg),
$$

and the result of Corollary 4.1 follows for $2\le p<\infty$.

One ends up the proof for $p=\infty$ in the same way as the proof of Theorem 4.1. Namely, we apply Lemma 7.2 with $p=\log C_n$ in (30). Since $C_n=e^p$ and $\log n = C_n^2 = e^{2p}$, under the condition $g\in\mathcal{G}_{p\ell}(M_{p\ell})$ we have

$$
\sup_{g\in\mathcal{G}_{p\ell}(M_{p\ell})} \mathbb{E}\,(|Y|^{p\ell}) \le M_{p\ell} \le VK^p\exp\big((p/2-1)e^{2p} + p^2/2 - (p/2)\log p\big),
$$

and using that $k+1\le\ell$, we ensure the bound

$$
M_{ps} \le \frac{C_n^{p/2+1}n^{p/2-1}}{p^{p/2}}, \qquad \text{i.e.,}\qquad \frac{p}{n^{1-1/p}}M_{ps}^{1/p}C_n^{(k+1)-1/p} \le \frac{\sqrt{p}C_n^{\ell+1/2}}{\sqrt{n}}.
$$

The result for $p=\infty$ follows since

$$
\mathbb{E}\,\|Q_s(\widehat{h}_n^{(k)} - K_n^{(k)} * h)\|_\infty \le 2^{s+3}e^2\kappa\|K^{(k)}\|_\infty\sqrt{\frac{\log C_n}{n}}C_n^\ell(1+o(1)). \qquad \square
$$

7.1.3. *Proof of Theorem* 4.2 (upper bound of $\mathbb{L}_p(\mathbb{R})$-risk for trigonometric functions). We only give the proof for the function $f\colon x\mapsto\cos(\ell x)$. Again using the triangle inequality, we aim at bounding $\|\Gamma_f - \mathbb{E}\,\widehat{\Gamma}_{f,n}\|_p$ and $\mathbb{E}\,\|\widehat{\Gamma}_{f,n} - \mathbb{E}\,\widehat{\Gamma}_{f,n}\|_p$, for $2\le p\le\infty$, starting with $\|\Gamma_f - \mathbb{E}\,\widehat{\Gamma}_{f,n}\|_p$ for $2\le p\le\infty$. According to Lemma 2.2, we have

$$
\begin{aligned}
\Gamma_f(y) - \mathbb{E}\,\widehat{\Gamma}_{f,n}(y) = \frac{e^{-\ell^2/2}}{4\pi}\bigg[&e^{i\ell y}\int(1-S_n^\star(t))e^{iyt}e^{-\ell t}h^*(t)\,dt\\
&+e^{-i\ell y}\int(1-S_n^\star(t))e^{iyt}e^{\ell t}h^\star(t)\,dt\bigg],
\end{aligned}
$$

and therefore

$$
(31)\qquad \|\Gamma_f - \mathbb{E}\,\widehat{\Gamma}_{f,n}\|_\infty \le \frac{e^{-(C_n-\ell)^2/2}}{2\pi(C_n-\ell)}(1+o(1)).
$$

We now turn to the control of $\|\Gamma_f - \mathbb{E}\,\widehat{\Gamma}_{f,n}\|_p$ when $2 \le p < \infty$. For $\delta \in \{-1, 1\}$, set

$$\varphi_\delta(t) = (1 - S_n^\star(t))e^{-\ell\delta t}h^*(t).$$

Then we have

$$\|\Gamma_f - \mathbb{E}\,\widehat{\Gamma}_{f,n}\|_p \le \frac{e^{-\ell^2/2}}{4\pi}\big(\|\varphi_1^\star\|_p + \|\varphi_{-1}^\star\|_p\big).$$

Arguing as for (20), we obtain that $\|\varphi_\delta^\star\|_p^p \le 2\pi\|\varphi_\delta\|_1^{p-2}\|\varphi_\delta\|_2^2$, and therefore

$$\|\Gamma_f - \mathbb{E}\,\widehat{\Gamma}_{f,n}\|_p \le \frac{e^{-(C_n-\ell)^2/2}}{(C_n-\ell)^{1-1/p}}(1+o(1)).$$

It remains now to control the variance terms $\mathbb{E}\,\|\widehat{\Gamma}_{f,n} - \mathbb{E}\,\widehat{\Gamma}_{f,n}\|_p$ for $2 \le p < \infty$ and $\mathbb{E}\,\|\widehat{\Gamma}_{f,n} - \mathbb{E}\,\widehat{\Gamma}_{f,n}\|_\infty$. We start with the case $2 \le p < \infty$. The term $\mathbb{E}\,\|\widehat{\Gamma}_{f,n} - \mathbb{E}\,\widehat{\Gamma}_{f,n}\|_\infty$ will be studied by applying Lemma 7.2 and the arguments used for the density or the polynomial case. Write $\widehat{\Gamma}_{f,n} - \mathbb{E}\,\widehat{\Gamma}_{f,n}$ as a sum of independent random variables

$$\widehat{\Gamma}_{f,n}(x) - \mathbb{E}\,\widehat{\Gamma}_{f,n}(x) = \frac{1}{n}\sum_{j=1}^n \mathcal{X}_{n,\ell}(x, Y_j),$$

where the $\mathcal{X}_{n,\ell}(x, Y_j)$ are centered random variables defined by

$$\begin{aligned}\mathcal{X}_{n,\ell}(x, Y_j) = \frac{e^{-\ell^2/2}}{2}e^{i\ell x}\big[S_n(x+i\ell-Y_j) - \mathbb{E}\,S_n(x+i\ell-Y_j)\big]\\ + \frac{e^{-\ell^2/2}}{2}e^{-i\ell x}\big[S_n(x-i\ell-Y_j) - \mathbb{E}\,S_n(x-i\ell-Y_j)\big].\end{aligned}$$

By applying inequality (23) to $\mathbb{E}\,\Big|n^{-1}\sum_{j=1}^n \mathcal{X}_{n,\ell}(x, Y_j)\Big|^p$, we infer that

$$\begin{aligned}\mathbb{E}\left|\frac{1}{n}\sum_{j=1}^n \mathcal{X}_{n,\ell}(x, Y_j)\right|^p \le 2^{p-1}\left[\frac{e\kappa\sqrt{p}}{\sqrt{n}}\mathbb{E}^{\,1/2}|\mathcal{X}_{n,\ell}(x, Y_1)|^2\right]^p\\ + 2^{p-1}\left[\frac{\kappa p}{n^{1-1/p}}\mathbb{E}^{\,1/p}|\mathcal{X}_{n,\ell}(x, Y_1)|^p\right]^p,\end{aligned}$$

and therefore, using inequality (22), the quantity $\mathbb{E}\,\|\widehat{\Gamma}_{f,n} - \mathbb{E}\,\widehat{\Gamma}_{f,n}\|_p$ is bounded by

$$\left\{2^{p-1}\left[\frac{e\kappa\sqrt{p}}{\sqrt{n}}\right]^p \int \mathbb{E}^{\,p/2}|\mathcal{X}_{n,\ell}(x, Y_1)|^2\,dx + 2^{p-1}\left[\frac{\kappa p}{n^{1-1/p}}\right]^p \int \mathbb{E}\,|\mathcal{X}_{n,\ell}(x, Y_1)|^p\,dx\right\}^{1/p}.$$

By using Fubini's Theorem and the fact that $\int h(x+u/c_n)\,dx = 1$, the last term $\int \mathbb{E}\,|\mathcal{X}_{n,\ell}(x, Y_1)|^p\,dx$ is controlled by

$$\int \mathbb{E}\,|\mathcal{X}_{n,\ell}(x, Y_1)|^p\,dx \le 2^{p-1}e^{-p\ell^2/2}C_n^{p-1}\int \big(|S(u+i\ell C_n)|^p + |S(u-i\ell C_n)|^p\big)\,du.$$

Note that for any real number u and any δ in $\{-1;+1\}$

$$|S(u+i\delta\ell C_n)|^2 \le \frac{e^{2\ell C_n}}{\pi^2(u^2+\ell^2 C_n^2)}, \tag{32}$$

and consequently, denoting $\varphi_\ell(y)=(y^2+\ell^2)^{-1}$,

$$\int \mathbb{E}\,|\mathcal{X}_{n,\ell}(x,Y_1)|^p\,dx \le \frac{2^p}{\pi^p}\|\varphi_\ell\|_{p/2}^{p/2}\,e^{-p\ell^2/2}e^{\ell p C_n}. \tag{33}$$

It remains now to bound the term $\int \mathbb{E}^{\,p/2}|\mathcal{X}_{n,\ell}(x,Y_1)|^2\,dx$. First notice that

$$\mathbb{E}^{\,p/2}|\mathcal{X}_{n,\ell}(x,Y_1)|^2 \le e^{-p\ell^2/4}2^{p/2-1}\big(\mathbb{E}^{\,p/2}|S_n(x+i\ell-Y_1)|^2+\mathbb{E}^{\,p/2}|S_n(x-i\ell-Y_1)|^2\big),$$

and secondly, use bound (32) to get that for $\delta\in\{-1,1\}$,

$$\int \mathbb{E}^{\,p/2}|S_n(x+\delta il-Y_1)|^2\,dx \le \frac{e^{\ell p C_n}}{\pi^p}\|h*\varphi_\ell\|_{p/2}^{p/2}.$$

Since $\|h*\varphi_\ell\|_{p/2}\le\|h\|_1\|\varphi_\ell\|_{p/2}=\|\varphi_\ell\|_{p/2}$, the above bound is uniform in $h\in\mathcal{H}$. Consequently for all $2\le p<\infty$, we get the bound

$$\int \mathbb{E}^{\,p/2}|\mathcal{X}_{n,\ell}(x,Y_1)|^2\,dx \le \frac{\|\varphi_\ell\|_{p/2}^{p/2}e^{-p\ell^2/4}2^{p/2}}{\pi^p}e^{\ell p C_n}. \tag{34}$$

Combining (33) and (34) we find that

$$\begin{aligned}\mathbb{E}\,\|\widehat{\Gamma}_{f,n}-\mathbb{E}\,\widehat{\Gamma}_{f,n}\|_p \le \Bigg\{&2^{p-1}\left[\frac{e\kappa\sqrt{p}}{\sqrt{n}}\right]^p\|\varphi_\ell\|_{p/2}^{p/2}\frac{e^{-p\ell^2/4}2^{p/2}}{\pi^p}e^{\ell p C_n}\\ &+2^{p-1}\left[\frac{\kappa p}{n^{1-1/p}}\right]^p\|\varphi_\ell\|_{p/2}^{p/2}\frac{2^p}{\pi^p}e^{-p\ell^2/2}e^{\ell p C_n}\Bigg\}^{1/p}.\end{aligned}$$

Finally by using that $2\le p<\infty$ we conclude that

$$\mathbb{E}\,\|\widehat{\Gamma}_{f,n}-\mathbb{E}\,\widehat{\Gamma}_{f,n}\|_p \le \frac{4e\kappa e^{-\ell^2/4}}{\pi}\|\varphi_\ell\|_{p/2}^{1/2}\left(\frac{\sqrt{p^p}}{\sqrt{n^p}}+\frac{p^p}{n^{p-1}}\right)^{1/p}e^{\ell C_n} \tag{35}$$

$$\le \frac{4(1+\sqrt{2})e\kappa e^{-\ell^2/4}}{\pi}\|\varphi_\ell\|_{p/2}^{1/2}\sqrt{\frac{p}{n}}e^{\ell C_n}(1+o(1)). \tag{36}$$

The result for $2\le p<\infty$ follows by gathering (31), (36) and by taking $C_n=\sqrt{\log n-\frac{1}{2}\log\log n}$.

We now turn to the case $p=\infty$. Note that $\widehat{\Gamma}_{f,n}-\mathbb{E}\,\widehat{\Gamma}_{f,n}$ belongs to $\mathcal{M}_{C_n+\ell,p}$ since its Fourier transform at the point t equals

$$\frac{e^{-t^2/2}}{2n}\sum_{j=1}^n\left[\widetilde{K_n}^{\star}\left(\frac{t+\ell}{C_n}\right)\big(e^{i(t+\ell)Y_j}-h^\star(t+\ell)\big)+\widetilde{K_n}^{\star}\left(\frac{t-\ell}{C_n}\right)\big(e^{i(t-\ell)Y_j}-h^\star(t-\ell)\big)\right],$$

where $\widetilde{K_n}^\star\big((t+\delta\ell)/C_n\big) = S_n^\star\big((t+\delta\ell)/C_n\big)\big/ f_\epsilon^\star(t+\delta\ell)$ (with $\delta \in \{-1;1\}$) is compactly supported on $[-C_n-\ell, C_n+\ell]$. Use the same arguments as for the density or for the polynomials with $\|\varphi_\ell\|_{p/2}^{1/2} \le \|\varphi_\ell\|_1^{1/p}$ (which tends to 1 as $p\to\infty$), apply Lemma 7.2 with $p'=\infty$ and $p=\log C_n$, and then use (36) to get

$$\mathbb{E}\,\|\widehat{\Gamma}_{f,n} - \mathbb{E}\,\widehat{\Gamma}_{f,n}\|_\infty \le \frac{8e^2\kappa e^{-\ell^2/4}}{\pi}\|\varphi_\ell\|_\infty^{1/2}\frac{\sqrt{\log C_n}}{\sqrt{n}}e^{\ell C_n}(1+o(1)). \tag{37}$$

Take $C_n = \sqrt{\log n - \frac{1}{2}\log\log n}$ in (37) to complete the proof. $\square$

7.2. Proofs of lower bounds for the $\mathbb{L}_p(\mathbb{R})$-risk: use of Fano's lemma. The proofs of lower bounds of $\mathbb{L}_p(\mathbb{R})$-risks are based on Fano's lemma [12] (see, e.g., Cover and Thomas [8]) in its new version due to Birgé [3], which we recall here. Consider a metric space (Θ, d) and a set of probability measures $\mathcal{P}$ indexed by Θ: $\mathcal{P} = \{P_\theta;\, \theta\in\Theta\}$. The problem is to give an infimum bound for the minimax risk related to the estimation of θ in Θ from an observation X with law P_θ,

$$R(\Theta) = \inf_{\widehat{\theta}(X)}\sup_{\theta\in\Theta}\mathbb{E}_\theta\big(d(\theta;\widehat{\theta}(X))\big),$$

where the infimum is over all estimators $\widehat{\theta}(X)$ with values in Θ. Consider a finite subset Θ' of Θ of cardinality $|\Theta'|\ge 3$ such that $d(\theta,\theta')\ge\delta>0$ for any pair (θ,θ') of distinct points in Θ', Then we have

$$R(\Theta) \ge \frac{\delta}{2}\inf_{\widehat{T}(X)}\Big(1 - \inf_{\theta\in\Theta'}P_\theta(\widehat{T}(X)=\theta)\Big)$$

when $\widehat{T}(X)$ ranges over all estimators with values in Θ'. Now, we use the following result based on Fano's lemma (see Fano [12] or Birgé [3]).

Lemma 7.3. *Let θ_0 be a fixed point in Θ' and set*

$$\overline{\mathcal{K}} = \frac{1}{|\Theta'|}\sum_{\theta\in\Theta'}\mathcal{K}(P_\theta;P_{\theta_0}),$$

where $\mathcal{K}(P_\theta;P_{\theta_0})$ denotes the Kullback–Leibler divergence between P_θ and P_{θ_0}. There exists an absolute constant α such that if $\widehat{T}(X)$ is an estimator taking values in Θ', we have

$$\inf_{\theta\in\Theta'}P_\theta(\widehat{T}(X)=\theta) \le \alpha \vee \frac{\overline{\mathcal{K}}}{\log(|\Theta'|+1)}.$$

As an immediate consequence, we obtain

$$R(\Theta) \ge \frac{\delta}{2}\bigg(1 - \alpha\vee\frac{\overline{\mathcal{K}}}{\log(|\Theta'|+1)}\bigg).$$

We are going to apply this result to $\mathcal{P} = \{h\,d\lambda;\, h\in\mathcal{H}\}^{\otimes n}$, where $d\lambda$ is the Lebesgue measure on $\mathbb{R}$, and $(\Theta,d) = (\{\Gamma_f; g\in\mathcal{G}_f\}, \|\cdot\|_p)$ when $1\le p\le\infty$ and f is a fixed polynomial or trigonometric function. In other words, we estimate the functional Γ_f and compute an infimum bound for the minimax $\mathbb{L}_p(\mathbb{R})$-risk

$$R_{n,p}(f) = \inf_{T_n} \sup_{g \in \mathcal{G}_f} \mathbb{E}\, \|T_n - \Gamma_f\|_p,$$

where the infimum is taken over all the estimators T_n based on the observations $Y_1, \ldots, Y_n$. Fix an integer $m \geq 3$ and denote by ℓ the degree of the fixed function f when f is a polynomial or a trigonometric function (of the form $\mathcal{C}_{\beta,\ell}$ or $\mathcal{S}_{\beta,\ell}$). The application of Fano's lemma relies on four steps described below.

Step 1: To define a family of probability densities $\{(\varphi_{m,a,\ell})\}_{a\in\mathcal{A}}$ belonging to $\mathcal{H}$ indexed by some finite set $\mathcal{A}$ whose cardinality depends on m. Note that this family will depend on the integer ℓ (implying different regularity conditions on the family for different ℓ).

Step 2: To calculate the Kullback–Leibler distance between points in this family of probabilities by noting that

$$\mathcal{K}(P_{m,a,\ell}^{\otimes n}; P_{m,0,\ell}^{\otimes n}) = n\mathcal{K}(\varphi_{m,a,\ell}; \varphi_{m,0,\ell}),$$

where $P_{m,a,\ell}$ denotes the probability $\varphi_{m,a,\ell}\, d\lambda$.

Step 3: To calculate the minimum distance between two points in the family of parameters $(\Gamma_{f,m,a})_{a\in\mathcal{A}}$ induced by this family of densities, i.e., find δ_m such that $\inf_{a,a'\in\mathcal{A}} \|\Gamma_{f,m,a} - \Gamma_{f,m,a'}\|_p \geq \delta_m$ (for m large enough).

Step 4: To check that there exists a constant $C_1 \leq \alpha$ and a particular point $\varphi_{m,0,\ell}$ in the family such that

$$\frac{n\sum_{a\in\mathcal{A}} \mathcal{K}(\varphi_{m,a,\ell}; \varphi_{m,0,\ell})}{|\mathcal{A}| \log(|\mathcal{A}|+1)} \leq C_1.$$

Then we conclude that

$$R_{n,p}(f) \geq \frac{\alpha}{2}\delta_m.$$

The construction of the family (step 1) will be essentially the same for all lower bounds. Therefore, the first two steps will be checked in a general setting, whereas the last two steps are described in each case.

7.2.1. *Step* 1: *Construction of the family.* Define the functions on $\mathbb{R}$

$$\alpha_\ell(x) = C_\ell \left(\frac{\sin x}{\pi x}\right)^{2\ell+2},$$

where C_ℓ is a normalizing constant such that $\int \alpha_\ell = 1$, $\alpha_{0,\ell} = \alpha_\ell * \alpha_\ell$, and $h_{0,\ell} = \alpha_{0,\ell} * f_\varepsilon$. Note that α_ℓ and hence $\alpha_{0,\ell}$ and $h_{0,\ell}$ are probability densities on $\mathbb{R}$. Moreover, the support of the Fourier transform $\alpha_{0,\ell}^\star$ of $\alpha_{0,\ell}$ is contained in $[-2\ell-2; 2\ell+2]$. In what follows, we fix the kernel K to be the de La Vallée-Poussin kernel V (defined by (3)) or more generally its rescaling $V_\lambda(\cdot) = V(\lambda\,\cdot\,)$ (for some $\lambda \geq 1$) and we put

$$K_{m,j}(x) = \theta_m K(mx - j), \quad 1 \leq j \leq m-1; \qquad K_{m,0} = 0,$$

where m is an integer greater than or equal to 3, and the sequence $(\theta_m)_{m\geq 0}$ is positive and converges to zero as $m \to \infty$.

We denote by $\overline{K}_{m,j}$ the normalizing constant

$$\overline{K}_{m,j} = \int h_{0,\ell}(x) K_{m,j}(x)\, dx \qquad \text{for} \quad 0 \leq j \leq m-1, \tag{38}$$

and consider the following family of functions

$$\varphi_{m,a,\ell}(x) = h_{0,\ell}(x) + \sum_{j=1}^{m-1} a(j) h_{0,\ell}(x)(K_{m,j}(x) - \overline{K}_{m,j}), \tag{39}$$

where a ranges over the set $\mathcal{A}$ of mappings of $\{1;\ldots;m-1\}$ to $\{0;1\}$ with the convention $\varphi_{m,0,\ell} = h_{0,\ell}$. The functions $\varphi_{m,a,\ell}$ are also given by

$$\varphi_{m,a,\ell}(x) = \left(1 - \sum_{j=1}^{m-1} a(j)\overline{K}_{m,j}\right)\alpha_{0,\ell} * f_\varepsilon(x) + \sum_{j=1}^{m-1} a(j) h_{0,\ell}(x) K_{m,j}(x).$$

Note that the constants $\overline{K}_{m,j}$, $1 \leq j \leq m-1$, ensure that $\varphi_{m,a,\ell}$ integrates to one.

Lemma 7.4. *Fix K equal to $V_\lambda(\cdot) = V(\lambda\,\cdot)$ for some parameter $\lambda \geq 1$, then there exists a constant $\mathcal{C}_\lambda$ such that*

$$\sup_{x\in\mathbb{R}} \sum_{j=1}^{m-1} |K(mx-j)| \leq \mathcal{C}_\lambda < +\infty. \tag{40}$$

Moreover we have the bounds

$$\sum_{j=1}^{m-1} |\overline{K}_{m,j}| \leq \theta_m \mathcal{C}_\lambda \tag{41}$$

and

$$|\overline{K}_{m,j}| \leq \frac{\theta_m}{m} \|h_{0,\ell}\|_\infty \|K\|_1. \tag{42}$$

7.2.2. *Proof of Lemma* 7.4. The proof of (40) consists in noting that

$$\sup_{x\in\mathbb{R}} \sum_{j=1}^{m-1} |V_\lambda(mx-j)| = \sup_{k\in\mathbb{Z}} \sup_{x\in[\frac{2k-1}{2m};\frac{2k+1}{2m}]} \sum_{j=1}^{m-1} |V_\lambda(mx-j)|.$$

Now, when x belongs to the interval $[(2k-1)/2m;(2k+1)/2m]$, the quantity $mx-j$ belongs to the interval $[k-j-1/2;k-j+1/2]$. Hence

$$\sup_{x\in\mathbb{R}} \sum_{j=1}^{m-1} |V_\lambda(mx-j)| \leq \sup_{k\in\mathbb{Z}} \Bigg(\sum_{j<k} \frac{2}{\pi\lambda^2(k-j-1/2)^2} + \sum_{j>k} \frac{2}{\pi\lambda^2(k-j+1/2)^2} + \|V\|_\infty \Bigg) \leq \frac{4}{\pi\lambda^2} \sum_{n\in\mathbb{Z}} \frac{1}{(n-1/2)^2} + \|V\|_\infty,$$

which is a finite constant, denoted by $\mathcal{C}_\lambda$. The bound (41) is a direct consequence of (40), and the bound (42) follows from the definition of $\overline{K}_{m,j}$. $\square$

The following lemma states that the set of functions $\{\varphi_{m,a,\ell}\}_{a\in\mathcal{A}}$ is a family of probability densities contained in $\mathcal{H}$ under a suitable assumption on the parameter θ_m related to the support $[-2\lambda; 2\lambda]$ of the Fourier transform $V_\lambda^\star$ of the kernel V_λ.

Condition 2. $m\theta_m e^{2\lambda^2 m^2+4\lambda(\ell+1)m} \underset{m\to\infty}{\longrightarrow} 0.$

Lemma 7.5. *Under Condition* 2 *and for* m *large enough, the family* $\{\varphi_{m,a,\ell}\}_{a\in\mathcal{A}}$ *is contained in* $\mathcal{H}$.

The proof of this lemma is postponed to Section 8.

Let us now compute the Kullback–Leibler distance between two densities, which characterizes the size of the family. We do not need to use the whole family indexed by the set $\mathcal{A}$ (which is large), hence we will restrict our family to the densities parametrized by particular finite subsets $\mathcal{A}_1$ and $\mathcal{A}_2$, which are chosen according to the norm we consider.

7.2.3. *Step* 2: *Calculation of the Kullback–Leibler distance.*

Calculation of the Kullback–Leibler distance for the $\mathbb{L}_\infty(\mathbb{R})$-norm. Following Ibragimov and Hasminskii [15], we consider the subset $\mathcal{A}_1$ of $\mathcal{A}$ consisting of the mappings from $\{1;\dots;m-1\}$ to $\{0;1\}$ taking value 0 everywhere except for one point. The subset $\mathcal{A}_1$ has cardinality $|\mathcal{A}_1| = m$. In this case $\varphi_{m,a,\ell} = h_{0,\ell}(x)(1+K_{m,j}(x)-\overline{K}_{m,j})$ with j being the index such that $a(j)=1$. For the sake of simplicity, we put

$$\varphi_{m,j,\ell}(x) = h_{0,\ell}(x)(1+K_{m,j}(x)-\overline{K}_{m,j}), \qquad \text{for all} \quad 0\le j\le m-1.$$

The resulting family $\{\varphi_{m,a,\ell}\}_{a\in\mathcal{A}_1\cup\{0\}}$ is exactly the family $\{\varphi_{m,j,\ell}\}_{0\le j\le m-1}$. We compute the Kullback–Leibler divergence related to this family,

$$\mathcal{K}(\varphi_{m,j,\ell};\varphi_{m,0,\ell}) = \mathcal{K}(\varphi_{m,j,\ell}; h_{0,\ell}) = \int \log\left(\frac{\varphi_{m,j,\ell}(x)}{h_{0,\ell}(x)}\right)\varphi_{m,j,\ell}(x)\,dx.$$

By using that $\log(1+u)\le u$ for $u\ge 0$, we get

$$\mathcal{K}(\varphi_{m,j,\ell}; h_{0,\ell}) \le \int (K_{m,j}(x)-\overline{K}_{m,j})h_{0,\ell}(x)(1+K_{m,j}(x)-\overline{K}_{m,j})\,dx.$$

Since $\int(K_{m,j}(x)-\overline{K}_{m,j})h_{0,\ell}(x)\,dx = 0$, we have

$$\mathcal{K}(\varphi_{m,j,\ell}; h_{0,\ell}) \le \int (K_{m,j}(x)-\overline{K}_{m,j})^2 h_{0,\ell}(x)\,dx,$$

which gives the bound

$$\mathcal{K}(\varphi_{m,j,\ell}; h_{0,\ell}) \le \int K_{m,j}^2(x)h_{0,\ell}(x)\,dx \le \frac{\theta_m^2}{m}\|h_{0,\ell}\|_\infty\|K\|_2^2. \tag{43}$$

Calculation of the Kullback–Leibler distance for the $\mathbb{L}_p(\mathbb{R})$-norm, $2 \leq p < \infty$. Again following Ibragimov and Hasminskii [15], we select a specific family for the $\mathbb{L}_p(\mathbb{R})$-risk when $2 \leq p < \infty$, and consider the subset $\mathcal{A}_2$ of mappings of $\{1; \dots; m-1\}$ to $\{0; 1\}$ containing the mapping identically equal to zero and such that for all $a \neq a'$ in $\mathcal{A}_2$, we have

$$\sum_{j=1}^{m-1} |a(j) - a'(j)| > \frac{1}{4}(m-1). \tag{44}$$

Therefore the cardinality of this set of mappings satisfies $|\mathcal{A}_2| \geq \exp\{(m-1)/8\}$ (see Ibragimov and Hasminskii [15]). We denote by $\{\varphi_{m,a,\ell}\}_{a \in \mathcal{A}_2}$ the resulting family.

Compute the Kullback–Leibler divergence between $\varphi_{m,a,\ell}$ and $h_{0,\ell}$ by using once again that $\log(1+u) \leq u$. We obtain

$$\begin{aligned} \mathcal{K}(\varphi_{m,a,\ell}; h_{0,\ell}) \leq \sum_{j=1}^{m-1} a(j) \int (K_{m,j}(x) - \overline{K}_{m,j}) h_{0,\ell}(x) \\ \times \left(1 + \sum_{k=1}^{m-1} a(k)(K_{m,k}(x) - \overline{K}_{m,k})\right) dx. \end{aligned}$$

The right-hand side of this expression also equals

$$\sum_{j=1}^{m-1} \sum_{k=1}^{m-1} a(j)a(k) \int (K_{m,j}(x) - \overline{K}_{m,j})(K_{m,k}(x) - \overline{K}_{m,k}) h_{0,\ell}(x)\, dx,$$

and we get the bound

$$\mathcal{K}(\varphi_{m,a,\ell}; h_{0,\ell}) \leq \sum_{j=1}^{m-1} \sum_{k=1}^{m-1} a(j)a(k) \left[\int K_{m,j}(x) K_{m,k}(x) h_{0,\ell}(x)\, dx - \overline{K}_{m,j} \overline{K}_{m,k} \right].$$

Apply inequality (42) to get

$$\begin{aligned} \mathcal{K}(\varphi_{m,a,\ell}; h_{0,\ell}) \leq \theta_m^2 \int \left(\sum_{j=1}^{m-1} |K(mx-j)| \right) \left(\sum_{k=1}^{m-1} |K(mx-k)| \right) h_{0,\ell}(x)\, dx \\ + (m-1)^2 \frac{\theta_m^2}{m^2} \|h_{0,\ell}\|_\infty^2 \|K\|_1^2, \end{aligned}$$

which combined with (40) gives that

$$\mathcal{K}(\varphi_{m,a,\ell}; h_{0,\ell}) \leq \left(\mathcal{C}_\lambda^2 + \|h_{0,\ell}\|_\infty^2 \|V_\lambda\|_1^2 \right) \theta_m^2. \tag{45}$$

7.3. Proofs of lower bounds for $\mathbb{L}_\infty(\mathbb{R})$-risk. In this section, we take $\lambda = 1$ in V_λ, that is K is chosen equal to the de La Vallée-Poussin kernel V defined by (3).

7.3.1. *Proof of Theorem* 5.1 (lower bound of $\mathbb{L}_\infty(\mathbb{R})$-risk for the derivatives). The main purpose of this proof is now to check steps 3 and 4.

Step 3: We need a lower bound for the distance between the parameters in the family that we wish to estimate, that is to say between the derivatives of the densities $\{\varphi^{(\ell)}_{m,j,\ell}\}_{0\leq j\leq m-1}$. This will be done in the following lemma, which will be proved after completing step 4.

Lemma 7.6. *For any integer $\ell \geq 0$ and m large enough, there exists a positive constant C such that for all $j \neq k$ in $\{0;\dots;m-1\}$, we have*

$$\|\varphi^{(\ell)}_{m,j,\ell} - \varphi^{(\ell)}_{m,k,\ell}\|_\infty \geq C m^\ell \theta_m.$$

Step 4: The last step of the proof consists in showing that there exists a constant $C_1 \leq \alpha$ such that

$$\frac{n\sum_{j=1}^{m-1}\mathcal{K}(\varphi_{m,j,\ell}; h_{0,\ell})}{m\log(m+1)} \leq C_1,$$

where $m\log(m+1)$ stands for the quantity $|\mathcal{A}_1|\log(|\mathcal{A}_1|+1)$, by applying inequality (43) which gives the bound on the Kullback–Leibler distance between two points in this family. Then Fano's lemma tells us that there exists a positive constant C such that

$$\inf_{T_n}\sup_{h\in\mathcal{H}}\|h^{(\ell)} - T_n\|_\infty \geq C m^\ell \theta^\star_m(n),$$

where $\theta^\star_m(n)$ is the supremum over all the parameters θ_m satisfying the conditions

$$\|h_{0,\ell}\|_\infty\|V\|_2^2\frac{n\theta_m^2}{m\log(m+1)} \leq \alpha \quad\text{and}\quad m\theta_m e^{2m^2+4(\ell+1)m} \underset{m\to\infty}{\longrightarrow} 0.$$

Choose

$$\theta_m = e^{-3m^2} \quad\text{and}\quad m = \sqrt{\tfrac{1}{6}\log n - \tfrac{1}{12}\log\log n - \tfrac{1}{6}\log\log\log n}.$$

We finally have $\theta_m = e^{-3m^2} = (\log n)^{1/4}(\log\log n)^{1/2}/\sqrt{n}$, which concludes the proof.

Let us now prove Lemma 7.6. The following identity holds for all $0 \leq j \leq m-1$:

$$\varphi^{(\ell)}_{m,j,\ell}(x) = h_{0,\ell}(x)V^{(\ell)}_{m,j}(x) + h^{(\ell)}_{0,\ell}(x)(1+V_{m,j}(x)-\overline{V}_{m,j}) + \sum_{r=1}^{\ell-1}\binom{\ell}{r}h^{(r)}_{0,\ell}(x)V^{(\ell-r)}_{m,j}(x).$$

Therefore, by applying the triangle inequality, we get

$$\begin{aligned}\|\varphi^{(\ell)}_{m,j,\ell} - \varphi^{(\ell)}_{m,k,\ell}\|_\infty \geq{}& \|h_{0,\ell}(V^{(\ell)}_{m,j} - V^{(\ell)}_{m,k})\|_\infty \\ &- \sum_{r=1}^{\ell-1}\binom{\ell}{r}\|h^{(r)}_{0,\ell}\|_\infty\|V^{(\ell-r)}_{m,j} - V^{(\ell-r)}_{m,k}\|_\infty \\ &- \|h^{(\ell)}_{0,\ell}(V_{m,j} - V_{m,k} + \overline{V}_{m,k} - \overline{V}_{m,j})\|_\infty.\end{aligned} \tag{46}$$

The first term in (46) is bounded from below in the following way:

$$\|h_{0,\ell}(V_{m,j}^{(\ell)} - V_{m,k}^{(\ell)})\|_\infty \geq m^\ell \theta_m \inf_{u\in[0;1]} h_{0,\ell}\left(\frac{u+j}{m}\right) \sup_{u\in[0;1]} |V^{(\ell)}(u) - V^{(\ell)}(u+j-k)|.$$

Since for any u in $[0;1]$ and any $0 \leq j \leq m-1$, the real number $(u+j)/m$ belongs to $[0;1]$, we conclude that

$$\|h_{0,\ell}(V_{m,j}^{(\ell)} - V_{m,k}^{(\ell)})\|_\infty \geq m^\ell \theta_m (\inf_{[0;1]} h_{0,\ell}) \inf_{|n|\geq 1} \sup_{u\in[0;1]} |V^{(\ell)}(u) - V^{(\ell)}(u+n)|. \tag{47}$$

By using the definition of $V_{m,j}$ and inequality (42) the remainder terms appearing in (46) are bounded and then

$$\begin{aligned}\|\varphi_{m,j,\ell}^{(\ell)} - \varphi_{m,k,\ell}^{(\ell)}\|_\infty \geq{}& m^\ell \theta_m (\inf_{[0;1]} h_{0,\ell}) \inf_{|n|\geq 1} \sup_{u\in[0;1]} |V^{(\ell)}(u) - V^{(\ell)}(u+n)| \\ &- 2^\ell m^{\ell-1}\theta_m \max_{1\leq r\leq \ell-1} (\|V^{(\ell-r)}\|_\infty \|h_{0,\ell}^{(r)}\|_\infty) \\ &- \|h_{0,\ell}^{(\ell)}\|_\infty \Big(2\theta_m \|V\|_\infty + 2\frac{\theta_m}{m}\|h_{0,\ell}\|_\infty \|V\|_1\Big).\end{aligned}$$

For m large enough and $\ell \geq 1$, we have

$$\|\varphi_{m,j,\ell}^{(\ell)} - \varphi_{m,k,\ell}^{(\ell)}\|_\infty \geq \frac{1}{2} m^\ell \theta_m (\inf_{[0;1]} h_{0,\ell}) \inf_{|n|\geq 1} \sup_{u\in[0;1]} |V^{(\ell)}(u) - V^{(\ell)}(u+n)|.$$

Finally, the uniform distance between two points of this family is bounded from below in the following way:

$$\|\varphi_{m,j,\ell}^{(\ell)} - \varphi_{m,k,\ell}^{(\ell)}\|_\infty \geq C m^\ell \theta_m,$$

where C is a positive constant. When ℓ equals zero, by using the triangle inequality combined with (42) and the lower bound (47), we obtain

$$\begin{aligned}\|\varphi_{m,j,\ell} - \varphi_{m,k,\ell}\|_\infty &\geq \|h_{0,\ell}(V_{m,j} - V_{m,k})\|_\infty - \|h_{0,\ell}\|_\infty |\overline{V}_{m,j} - \overline{V}_{m,k}| \\ &\geq C_1\theta_m - \frac{2}{m}\theta_m \|h_{0,\ell}\|_\infty \|V\|_1.\end{aligned}$$

This concludes the proof of Lemma 7.6 and hence the proof of Theorem 5.1. □

7.3.2. *Proof of Corollary* 5.1 (lower bound of $\mathbb{L}_\infty(\mathbb{R})$-risk for polynomial functions). Using Lemma 2.1, the family $\{\varphi_{m,j,\ell}\}_{0\leq j\leq m-1}$ of densities induces a family of functionals $\{\Gamma_{f,m,j}\}_{0\leq j\leq m-1}$ defined by

$$\Gamma_{f,m,j}(y) = \beta_\ell \varphi_{m,j,\ell}^{(\ell)}(y) + \sum_{k=0}^{\ell-1} Q_{\beta,\ell-k}(y)\varphi_{m,j,\ell}^{(k)}(y), \qquad \text{for} \quad 0 \leq j \leq m-1,$$

where $\beta_\ell \neq 0$ and $Q_{\beta,j}$ is a polynomial function of degree j related to f.

The main purpose of this proof consists again in checking steps 3 and 4.

Step 3: We have to show that the minimum distance between two functionals $\|\Gamma_{f,m,s} - \Gamma_{f,m,t}\|_\infty$ in the family is still bounded from below by a constant times $m^\ell\theta_m$, for m large enough. Then the step 4 and Corollary 5.1 follow by arguing as in the proof of Theorem 5.1.

When $s \neq t$ belong to $\{0, \dots, m-1\}$, using the triangle inequality and Lemma 7.6 we have

$$\|\Gamma_{f,m,s} - \Gamma_{f,m,t}\|_\infty \geq Cm^\ell\theta_m - \sum_{k=0}^{\ell-1} \|Q_{\beta,\ell-k}(\varphi^{(k)}_{m,s,\ell} - \varphi^{(k)}_{m,t,\ell})\|_\infty.$$

When $k = 0$, we immediately deduce the bound

$$\|Q_{\beta,\ell}(\varphi_{m,s,\ell} - \varphi_{m,t,\ell})\|_\infty \leq 2\|Q_{\beta,\ell}h_{0,\ell}\|_\infty \biggl(\theta_m\|V\|_\infty + \frac{\theta_m}{m}\|h_{0,\ell}\|_\infty\|V\|_1\biggr).$$

When $1 \leq k \leq \ell - 1$, we have

$$\begin{aligned}\|Q_{\beta,\ell-k}(\varphi^{(k)}_{m,s,\ell} - \varphi^{(k)}_{m,t,\ell})\|_\infty = \biggl\| &Q_{\beta,\ell-k}h^{(k)}_{0,\ell}(V_{m,s} - V_{m,t} - \overline{V}_{m,s} + \overline{V}_{m,t}) \\ &+ \sum_{j=0}^{k-1}\binom{k}{j} Q_{\beta,\ell-k}h^{(j)}_{0,\ell}(V^{(k-j)}_{m,s} - V^{(k-j)}_{m,t})\biggr\|_\infty,\end{aligned}$$

which is, by (40), bounded for m large enough by

$$\begin{aligned}\|Q_{\beta,\ell-k}(\varphi^{(k)}_{m,s,\ell} - \varphi^{(k)}_{m,t,\ell})\|_\infty &\leq 2\|Q_{\beta,\ell-k}h^{(k)}_{0,\ell}\|_\infty \biggl(\theta_m\|V\|_\infty + \frac{\theta_m}{m}\|h_{0,\ell}\|_\infty\|V\|_1\biggr) \\ &\quad + 2^{k+1}\max_{0\leq j\leq k-1}\|Q_{\beta,\ell-k}h^{(j)}_{0,\ell}\|_\infty\theta_m m^k \max_{1\leq j\leq k}\|V^{(j)}\|_\infty \leq Cm^{\ell-1}\theta_m.\end{aligned}$$

It follows that the minimum distance $\inf_{s\neq t}\|\Gamma_{f,m,s} - \Gamma_{f,m,t}\|_\infty$ is also bounded from below by a constant times $m^\ell\theta_m$ for m large enough. □

7.3.3. *Proof of Theorem* 5.2 (lower bound for $\mathbb{L}_\infty(\mathbb{R})$-risk for trigonometric functions). We only give the proof for the function $f\colon x \mapsto \cos(\ell x)$. Arguing as before, the main point of the proof lies in checking steps 3 and 4.

Step 3: We need a lower bound for the distance between the parameters in the family. Using Lemma 2.2 and the fact that the functions $\{\varphi_{m,j,\ell}\}_{0\leq j\leq m-1}$ admit the analytic continuation on the whole complex plane, we have the following identities for the functionals induced by the family of densities $\{\varphi_{m,j,\ell}\}_{0\leq j\leq m-1}$:

$$\begin{aligned}\Gamma_{f,m,j}(y) &\triangleq \int \cos(\ell x)g_{m,j,\ell}(x)f_\varepsilon(x-y)\,dx \\ &= \frac{e^{-\ell^2/2}}{2}\bigl(e^{i\ell y}\varphi_{m,j,\ell}(y+i\ell) + e^{-i\ell y}\varphi_{m,j,\ell}(y-i\ell)\bigr).\end{aligned}$$

Now, the following result gives the minimum distance between two parameters to be estimated in our family. Its proof is postponed to the end of the current argument.

Lemma 7.7. *For all m large enough, there exists a positive constant C such that for all $j \neq k \in \{0; \dots; m-1\}$, we have*

$$\|\Gamma_{f,m,j} - \Gamma_{f,m,k}\|_\infty \geq C \frac{\theta_m e^{2m\ell}}{m^2}.$$

Step 4: We conclude by using Fano's lemma in the same way as in the proof of Theorem 5.1. More precisely, there exists a positive constant C such that

$$\inf_{T_n} \sup_{g \in \mathcal{G}_f} \|\Gamma_f - T_n\|_\infty \geq C \frac{e^{2m\ell} \theta_m^\star(n)}{m^2},$$

where $\theta_m^\star(n)$ is the supremum over all the parameters θ satisfying the conditions

$$\|h_{0,\ell}\|_\infty \|V\|_2^2 \frac{n\theta_m^2}{m \log(m+1)} \leq \alpha \quad \text{and} \quad m\theta_m e^{2m^2+4(\ell+1)m} \underset{m\to\infty}{\longrightarrow} 0.$$

Choose the parameters

$$\theta_m = \frac{(\log n)^{1/4}\sqrt{\log\log n}}{\sqrt{n}} \quad \text{and} \quad m = \frac{1}{2}\sqrt{\log n} - C, \qquad \text{where} \quad C > \ell + 1,$$

in order to obtain that

$$m\theta_m e^{2m^2+4(\ell+1)m} = O(1)(\log n)^{3/4}\sqrt{\log\log n}\, e^{-2(C-(\ell+1))\sqrt{\log n}} \underset{n\to\infty}{\longrightarrow} 0$$

and that there exists a positive constant C' such that $n\theta_m^2/(m\log(m+1)) \leq C'$, which completes the proof of Theorem 5.2.

Let us now prove Lemma 7.7. By definition of $\Gamma_{f,m,j}$ and by using the triangle inequality, we get

$$\begin{aligned}\|\Gamma_{f,m,j} - \Gamma_{f,m,k}\|_\infty \geq \frac{e^{-\ell^2/2}}{2} \sup_{y\in\mathbb{R}} \Big(& \big|e^{i\ell y} h_{0,\ell}(y+i\ell)\big(V_{m,j}(y+i\ell) - V_{m,k}(y+i\ell)\big) \\ & + e^{-i\ell y} h_{0,\ell}(y-i\ell)\big(V_{m,j}(y-i\ell) - V_{m,k}(y-i\ell)\big)\big| \\ & - |\overline{V}_{m,j} - \overline{V}_{m,k}| |e^{i\ell y} h_{0,\ell}(y+i\ell) + e^{-i\ell y} h_{0,\ell}(y-i\ell)| \Big).\end{aligned}$$

Now, using (42) we have the bound

$$\begin{aligned}&|\overline{V}_{m,j} - \overline{V}_{m,k}| |e^{i\ell y} h_{0,\ell}(y+i\ell) + e^{-i\ell y} h_{0,\ell}(y-i\ell)| \\ &\qquad\qquad \leq \frac{2\theta_m}{m} \|h_{0,\ell}\|_\infty \|V\|_1 \sup_{y\in\mathbb{R}} \big|e^{i\ell y} h_{0,\ell}(y+i\ell) + e^{-i\ell y} h_{0,\ell}(y-i\ell)\big|,\end{aligned}$$

which implies that

$$\begin{aligned}\|\Gamma_{f,m,j} - \Gamma_{f,m,k}\|_\infty \geq \frac{e^{-\ell^2/2}}{2} \sup_{y\in\mathbb{R}} \Big| & e^{i\ell y} h_{0,\ell}(y+i\ell)\big(V_{m,j}(y+i\ell) - V_{m,k}(y+i\ell)\big) \\ & + e^{-i\ell y} h_{0,\ell}(y-i\ell)\big(V_{m,j}(y-i\ell) - V_{m,k}(y-i\ell)\big)\Big| - C\frac{\theta_m}{m}.\end{aligned}$$

We now come to the study of the main term. Straightforward calculations provide that, for $\delta \in \{-1, 1\}$,

$$V_{m,j}(y + \delta i\ell) = \frac{-\theta_m e^{2m\ell} e^{-2\delta imy}(e^{2\delta ij} + r_{m,j,\delta}(y))}{2\pi(my + mi\delta\ell - j)^2} \tag{48}$$

with $|r_{m,j,\delta}(y)| \leq 3e^{-m\ell}$. It follows that

$$\begin{aligned}\|\Gamma_{f,m,j} - \Gamma_{f,m,k}\|_\infty \geq\ & \frac{e^{-\ell^2/2} e^{2m\ell}\theta_m}{4\pi m^2} \\ & \times \sup_{y\in\mathbb{R}} \left| e^{i(\ell-2m)y} h_{0,\ell}(y+i\ell)\left(\frac{e^{2ij} + r_{m,j,1}(y)}{(y+i\ell-j/m)^2} - \frac{e^{2ik} + r_{m,k,1}(y)}{(y+i\ell-k/m)^2}\right)\right. \\ & \left. + e^{-i(\ell-2m)y} h_{0,\ell}(y-i\ell)\left(\frac{e^{-2ij} + r_{m,j,-1}(y)}{(y-i\ell-j/m)^2} - \frac{e^{-2ik} + r_{m,k,-1}(y)}{(y-i\ell-k/m)^2}\right)\right| - C\frac{\theta_m}{m}.\end{aligned}$$

Now consider the particular point $y = j/m$ and bound this quantity from below in the following way:

$$\begin{aligned}\|\Gamma_{f,m,j} - \Gamma_{f,m,k}\|_\infty \geq\ & \frac{e^{-\ell^2/2} e^{2m\ell}\theta_m}{4\pi m^2} \\ & \times \inf_{j\neq k} \left| e^{i(\ell/m-2)j} h_{0,\ell}(j/m+i\ell)\left(\frac{e^{2ij}}{\ell^2} + \frac{e^{2ik}}{(i\ell+(j-k)/m)^2}\right)\right. \\ & \left. + e^{-i(\ell/m-2)j} h_{0,\ell}(j/m-i\ell)\left(\frac{e^{-2ij}}{\ell^2} + \frac{e^{-2ik}}{(-i\ell+(j-k)/m)^2}\right) + O(e^{-m\ell})\right| - C\frac{\theta_m}{m}.\end{aligned}$$

We finally obtain

$$\|\Gamma_{m,j} - \Gamma_{m,k}\|_\infty \geq \frac{e^{-\ell^2/2}|h_{0,\ell}(i\ell)| e^{2m\ell}\theta_m}{2\pi m^2}\left[\inf_{j\neq k}\left|\frac{1-\cos(2(k-j))}{\ell^2}\right| + o(1)\right],$$

which ends up the proof of Lemma 7.7 and hence the proof of Theorem 5.2. $\square$

7.4. Proofs of lower bounds for $\mathbb{L}_p(\mathbb{R})$-risk, $2 \leq p < \infty$. We use the function K equal to $V_\lambda = V(\lambda\,\cdot)$ for some fixed well-chosen $\lambda \geq 1$, where V is the analogue of the de La Vallée-Poussin kernel defined by (3).

7.4.1. *Proof of Theorem* 5.3 (lower bound of $\mathbb{L}_p(\mathbb{R})$-risk, $2 \leq p < \infty$, for the derivatives of h).

Step 3: We compute the minimum distance of $\mathbb{L}_p(\mathbb{R})$-norm between two points in our family of parameters $\{\varphi^{(\ell)}_{m,a,\ell}\}_{a\in\mathcal{A}_2}$.

Lemma 7.8. *For any pair (ℓ, p) in $\mathbb{N}\times]1;+\infty[$ or in $\mathbb{N}^\star \times [1;+\infty[$ and m large enough, there exists a positive constant C such that for all $a \neq a'$ in $\mathcal{A}_2$*

$$\|\varphi^{(\ell)}_{m,a,\ell} - \varphi^{(\ell)}_{m,a',\ell}\|_p \geq C m^\ell \theta_m.$$

This lemma, proved in Section 8, is obtained by a careful generalization of the methods used to prove Lemma 7.6.

Step 4: We apply Fano's lemma and claim that for any pair (ℓ, p) in $\mathbb{N}\times]1;+\infty[$ or in $\mathbb{N}^\star \times [1;+\infty[$, there exists a positive constant C such that

$$\inf_{\widehat{h}_n} \sup_{h\in\mathcal{H}} \|h^{(\ell)} - \widehat{h}_n\|_p \geq C' m^\ell \theta^\star_m(n),$$

where $\theta^\star_m(n)$ is the supremum over all the parameters θ_m satisfying the conditions

$$\frac{n\sum_{a\in\mathcal{A}_2} \mathcal{K}(\varphi_{m,a,\ell}; h_{0,\ell})}{|\mathcal{A}_2|\log(|\mathcal{A}_2|+1)} \leq \alpha \quad \text{and} \quad m\theta_m e^{2\lambda^2 m^2 + 4\lambda(\ell+1)m} \underset{m\to\infty}{\longrightarrow} 0.$$

Now, remember that the cardinality of $\mathcal{A}_2$ satisfies $\log(|\mathcal{A}_2|) \geq (m-1)/8$, and inequality (45) gives a bound on the Kullback–Leibler distance of the family. We will look for the supremum of the parameters θ_m satisfying the conditions

$$16(C_\lambda^2 + \|h_{0,\ell}\|_\infty^2 \|V_\lambda\|_1^2)\frac{\theta_m^2}{m-1} \leq \frac{\alpha}{n} \quad \text{and} \quad m\theta_m e^{2\lambda^2 m^2 + 4\lambda(\ell+1)m} \underset{m\to\infty}{\longrightarrow} 0.$$

Choose

$$m = \sqrt{\frac{\log n}{2(2\lambda^2+1)} - \frac{\log\log n}{4(2\lambda^2+1)}} \qquad \text{and}$$
$$\theta_m = e^{-(2\lambda^2+1)m^2} = (\log n)^{1/4}/\sqrt{n}$$

to obtain that there exists a positive constant C such that

$$\inf_{\widehat{h}_n} \sup_{h\in\Sigma} \|h^{(\ell)} - \widehat{h}_n\|_p \geq C m^\ell \theta_m = C\frac{(\log n)^{(2\ell+1)/4}}{\sqrt{n}}$$

for any pair (ℓ, p) in $\mathbb{N}\times]1;+\infty[$ or in $\mathbb{N}^\star \times [1;+\infty[$, which entails the result of Theorem 5.3. $\square$

7.4.2. *Proof of Corollary* 5.2 (lower bound for $\mathbb{L}_p(\mathbb{R})$-risk, $2 \leq p < \infty$, for polynomial functions). Using Lemma 2.1, the family of densities $(\varphi_{m,a,\ell})_{a\in\mathcal{A}_2}$ induces a family of functionals $\{\Gamma_{f,m,a}\}_{a\in\mathcal{A}}$,

$$\Gamma_{f,m,a}(y) = \beta_\ell \varphi^{(\ell)}_{m,a,\ell}(y) + \sum_{k=0}^{\ell-1} Q_{\beta,\ell-k}(y)\varphi^{(k)}_{m,a,\ell}(y),$$

where $\beta_\ell \neq 0$ and Q_j is a polynomial function of degree j related to f.

We calculate the minimum distance between two functionals $\|\Gamma_{f,m,a} - \Gamma_{f,m,a'}\|_p$ in the family which is still bounded from below by a constant times $m^\ell\theta_m$. Then step 4 and the conclusion follow arguing as in the proof of Theorem 5.3.

Step 3: Fix $a \neq a'$ in $\mathcal{A}_2$. We use the same kernel as in the previous proof, so that the lower bound for $\|\varphi^{(\ell)}_{m,a,\ell} - \varphi^{(\ell)}_{m,a',\ell}\|_p$ obtained in Lemma 7.8 remains valid. Using the triangle inequality yields

$$\|\Gamma_{f,m,a} - \Gamma_{f,m,a'}\|_p \geq Cm^\ell \theta_m - \sum_{k=0}^{\ell-1} \|Q_{\beta,\ell-k}(\varphi^{(k)}_{m,a,\ell} - \varphi^{(k)}_{m,a',\ell})\|_p.$$

When $k = 0$, we have the bound

$$\|Q_{\beta,\ell}(\varphi_{m,a,\ell} - \varphi_{m,a',\ell})\|_p \leq \sum_{j=1}^{m-1} \|Q_{\beta,\ell} h_{0,\ell}\|_\infty \|K_{m,j} - \overline{K}_{m,j}\|_p \leq Cm^{1-1/p}\theta_m,$$

which is negligible with respect to $m^\ell \theta_m$.

When $1 \leq k \leq \ell - 1$, and for m large enough,

$$\begin{aligned}\|Q_{\beta,\ell-k}(\varphi^{(k)}_{m,a,\ell} - \varphi^{(k)}_{m,a',\ell})\|_p \leq \sum_{j=1}^{m-1} \Bigg[& \|Q_{\beta,\ell-k} h^{(k)}_{0,\ell}(K_{m,j} - \overline{K}_{m,j})\|_p \\ & + \sum_{r=0}^{k-1} \binom{k}{r} \|Q_{\beta,\ell-k} h^{(r)}_{0,\ell} K^{(k-r)}_{m,j}\|_p \Bigg] \leq Cm^{1+k-1/p}\theta_m,\end{aligned}$$

which is still negligible with respect to $m^\ell \theta_m$, and the proof is complete. □

8. Proofs of Technical Lemmas

Proof of Lemma 7.5. Combine (39) and Lemma 7.4 to get that for m large enough $\varphi_{m,a,\ell}$ is a positive function which integrates to one and hence $\varphi_{m,a,\ell}$ is a probability density. The main point of the proof is checking that there exists a family of probability densities $\{\alpha_{m,a,\ell}\}_{a \in \mathcal{A}}$ on $\mathbb{R}$ such that for all integer m,

$$\text{for all } a \in \mathcal{A}, \qquad \alpha_{m,a,\ell} * f_\varepsilon = \varphi_{m,a,\ell}.$$

We start with the proof of the existence of a family of real-valued functions $\{\beta_{m,j,\ell}\}_{1 \leq j \leq m-1}$ such that

$$h_{0,\ell} K_{m,j} = \beta_{m,j,\ell} * f_\varepsilon, \qquad 1 \leq j \leq m - 1.$$

We follow the lines of Theorem 2.2.1 in Taupin [20]. Consider the function $a_{m,j,\ell}$ defined for all t in $\mathbb{R}$ by

$$a_{m,j,\ell}(t) = e^{t^2/2}(h_{0,\ell} K_{m,j})^*(t).$$

The Fourier transform of the product $h_{0,\ell} K_{m,j}$ is the convolution product of $h^*_{0,\ell} = \alpha^*_{0,\ell} f^*_\varepsilon$ and

$$K^*_{m,j} \colon\ u \mapsto \frac{\theta_m}{m} e^{-iuj/m} K^*\left(\frac{u}{m}\right),$$

which are all compactly supported. This implies in particular that the function $a_{m,j,\ell}$ is compactly supported, with support contained in the algebraic sum of the preceding ones, that is, contained in $[-2\lambda m - 2(\ell+1); 2\lambda m + 2(\ell+1)]$. In particular, the function $a_{m,j,\ell}$ belongs to $\mathbb{L}_1(\mathbb{R}) \cap \mathbb{L}_2(\mathbb{R})$, and we can define $\beta_{m,j,\ell}$ as its inverse Fourier transform (denoted by $\bar{*}$) defined by

$$\beta_{m,j,\ell}(x) = a^{\bar{*}}_{m,j,\ell}(x) = \frac{1}{2\pi} \int e^{-itx} e^{t^2/2} (h_{0,\ell} K_{m,j})^*(t)\, dt.$$

This leads to the identity $\beta_{m,j,\ell} * f_\varepsilon = h_{0,\ell} K_{m,j}$. By construction $h_{0,\ell} K_{m,j}$ is a real-valued function. Consequently, the imaginary part of $\beta_{m,j,\ell} * f_\varepsilon$, also denoted by $\mathrm{Im}(\beta_{m,j,\ell} * f_\varepsilon)$, equals zero and therefore $f_\varepsilon * \mathrm{Im}(\beta_{m,j,\ell}) \equiv 0$. This implies that $f^*_\varepsilon (\mathrm{Im}(\beta_{m,j,\ell}))^* \equiv 0$ and hence $\mathrm{Im}(\beta_{m,j,\ell}) \equiv 0$. We finally obtain that $\beta_{m,j,\ell}$ is a real-valued function.

Now, the real-valued function $\alpha_{m,a,\ell}$ defined as

$$\alpha_{m,a,\ell} = \Big(1 - \sum_{j=1}^{m-1} a(j) \overline{K}_{m,j}\Big) \alpha_{0,\ell} + \sum_{j=1}^{m-1} a(j) \beta_{m,j,\ell}$$

satisfies $\alpha_{m,a,\ell} * f_\varepsilon = \varphi_{m,a,\ell}$. The last thing to prove is that $\alpha_{m,a,\ell}$ is a probability density on $\mathbb{R}$. The main point lies in proving that $\alpha_{m,a,\ell}$ is positive. This, combined with the fact that $\varphi_{m,a,\ell}$ is a probability density, $\varphi_{m,a,\ell} = \alpha_{m,a,\ell} * f_\varepsilon$, and Fubini's Theorem, will give us that it integrates to one. It remains thus to prove that, for m large enough, $\alpha_{m,a,\ell}$ is a nonnegative function.

The first step is to note that

$$\|\alpha_{m,a,\ell} - \alpha_{0,\ell}\|_\infty = \Big\| \sum_{j=1}^{m-1} a(j) \overline{K}_{m,j} \alpha_{0,\ell} - \sum_{j=1}^{m-1} a(j) \beta_{m,j,\ell} \Big\|_\infty \underset{m\to\infty}{\longrightarrow} 0, \tag{49}$$

and therefore, since $\alpha_{0,\ell}$ is a nonnegative function, for each compact set $\mathcal{C}$ in $\mathbb{R}$, we can find an integer m_0 large enough such that, for all $m \geq m_0$ and for all $x \in \mathcal{C}$, we have $\alpha_{m,a,\ell}(x) \geq 0$. Let us establish this convergence. Using the definition of $\beta_{m,j,\ell}$ and the fact that $a_{m,j,\ell}$ is compactly supported we may write

$$\begin{aligned} \Big\| \sum_{j=1}^{m-1} a(j) \beta_{m,j,\ell} \Big\|_\infty &\leq \frac{1}{2\pi} \int e^{t^2/2} \Big| \Big(h_{0,\ell} \sum_{j=1}^{m-1} K_{m,j} \Big)^*(t) \Big| \mathbf{1}_{|t| \leq 2(\ell+1)+2\lambda m}\, dt \\ &\leq \frac{1}{2\pi} e^{2\lambda^2 m^2 + 4\lambda(\ell+1)m + 2(\ell+1)^2} \int \Big| \Big(h_{0,\ell} \sum_{j=1}^{m-1} K_{m,j} \Big)^*(t) \Big| \mathbf{1}_{|t| \leq 2(\ell+1)+2\lambda m}\, dt. \end{aligned}$$

Now, using Lemma 7.4 we see that

$$\Big| \Big(h_{0,\ell} \sum_{j=1}^{m-1} K_{m,j} \Big)^*(t) \Big| \leq \mathcal{C}_\lambda \theta_m$$

and hence

$$\Big\| \sum_{j=1}^{m-1} a(j)\beta_{m,j,\ell} \Big\|_\infty \le e^{2(\ell+1)^2} \frac{\mathcal{C}_\lambda}{2\pi} 4(\lambda m + \ell + 1)\theta_m e^{2\lambda^2 m^2 + 4(\ell+1)\lambda m}.$$

Apply inequality (41) to get that

$$\Big\| \sum_{j=1}^{m-1} a(j)\overline{K}_{m,j}\alpha_{0,\ell} \Big\|_\infty \le \|\alpha_{0,\ell}\|_\infty \mathcal{C}_\lambda \theta_m,$$

and the uniform convergence of $\alpha_{m,a,\ell}$ to $\alpha_{0,\ell}$ in (49) follows under Condition 2.

The second step deals with the case of large $|x|$. For this write $\beta_{m,j,\ell}$ in the form

$$\begin{aligned}
\beta_{m,j,\ell}(x) &= \frac{1}{2\pi} \iint e^{-itx} e^{t^2/2} h^\star_{0,\ell}(u-t) K^\star_{m,j}(u)\, du\, dt \\
&= \frac{\theta_m}{2\pi m} \iint e^{-itx} e^{t^2/2} e^{iuj/m} e^{-(u-t)^2/2} \alpha^*_{0,\ell}(u-t) K^*\left(\frac{u}{m}\right) du\, dt \\
&= \frac{\theta_m}{2\pi m} \int e^{-ixu+iuj/m} \left(\int e^{v(ix-u)} \alpha^*_{0,\ell}(v)\, dv \right) e^{u^2/2} K^* \left(\frac{u}{m}\right) du.
\end{aligned}$$

The function $\alpha^*_{0,\ell}$ is equal C^2_ℓ times the square of the convolution product of $2\ell+2$ times the function $S^* = \mathbf{1}_{[-1;1]}$. More precisely, $\alpha^*_{0,\ell} = C^2_\ell (S^* * \cdots * S^*)^2$, where the convolution product appears $2\ell+2$ times. This is an even piecewise polynomial function with support $[-2\ell-2; 2\ell+2]$ and 2ℓ times continuously differentiable. The behavior of $\beta_{m,j,\ell}(x)$ is related to the quantity

$$\int e^{v(ix-u)} \alpha^*_{0,\ell}(v)\, dv = I(ix-u) + I(-ix+u),$$

where

$$I(c) = \int_0^{2\ell+2} e^{cv} \alpha^*_{0,\ell}(v)\, dv.$$

Integrating by parts, we get

$$\begin{aligned}
I(c) = \left[\frac{e^{cv}}{c} \alpha^*_{0,\ell}(v) \right]_0^{2\ell+2} &- \left[\frac{e^{cv}}{c^2} (\alpha^*_{0,\ell})'(v) \right]_0^{2\ell+2} + \cdots - \left[\frac{e^{cv}}{c^{2\ell}} (\alpha^*_{0,\ell})^{(2\ell-1)}(v) \right]_0^{2\ell+2} \\
&+ \left[\frac{e^{cv}}{c^{2\ell+1}} (\alpha^*_{0,\ell})^{(2\ell)}(v) \right]_0^{2\ell+2} - \int_0^{2\ell+2} \frac{e^{cv}}{c^{2\ell+1}} (\alpha^*_{0,\ell})^{(2\ell+1)}(v)\, dv.
\end{aligned}$$

Since $\alpha^*_{0,\ell}$ is 2ℓ times continuously differentiable and equals zero outside $[-2\ell-2; 2\ell+2]$, its derivatives up to the order 2ℓ are equal to zero at the point $2\ell+2$. Moreover, $\alpha_{0,\ell}$ is an even function, so that $(\alpha^*_{0,\ell})^{(2k-1)}(0) = 0$ for all $1 \le k \le \ell$. Now adding $I(c)$ with $I(-c)$, all but the final terms vanish and integrating once more by parts we get

$$\begin{aligned}
I(c) + I(-c) = &- \left[\frac{e^{cv} - e^{-cv}}{c^{2\ell+2}} (\alpha^*_{0,\ell})^{(2\ell+1)}(v) \right]_0^{2\ell+2} \\
&+ \int_0^{2\ell+2} \frac{e^{cv} - e^{-cv}}{c^{2\ell+2}} (\alpha^*_{0,\ell})^{(2\ell+2)}(v)\, dv.
\end{aligned}$$

Consequently,

$$\int e^{v(ix-u)}\alpha^*_{0,\ell}(v)\,dv = -\left[\frac{e^{(ix-u)v}-e^{-(ix-u)v}}{(ix-u)^{2\ell+2}}(\alpha^*_{0,\ell})^{(2\ell+1)}(v)\right]_0^{2\ell+2} + \int_0^{2\ell+2}\frac{e^{(ix-u)v}-e^{-(ix-u)v}}{(ix-u)^{2\ell+2}}(\alpha^*_{0,\ell})^{(2\ell+2)}(v)\,dv,$$

and hence is bounded as follows:

$$\left|\int e^{v(ix-u)}\alpha^*_{0,\ell}(v)\,dv\right| \le \gamma_\ell \frac{e^{2(\ell+1)|u|}}{(u^2+x^2)^{\ell+1}}.$$

We use this expression and the fact that the support of $K^\star(\,\cdot\,/m)$ is contained in $[-2\lambda m; 2\lambda m]$ to conclude that

$$\left|\sum_{j=1}^{m-1} a(j)\beta_{m,j,\ell}(x)\right| \le \frac{Cm\theta_m e^{2\lambda^2 m^2+4\lambda(\ell+1)m}}{x^{2\ell+2}}, \tag{50}$$

with the numerator converging to zero under Condition 2. Arguing as in Taupin [20], we obtain that for $|x|$ large enough,

$$\left(1-\sum_{j=1}^{m-1} a(j)\overline{K}_{m,j}\right)\alpha_{0,\ell}(x) \ge \frac{\delta}{x^{2\ell+2}}, \tag{51}$$

where δ is a positive constant. This lower bound is based on the following inequalities: for any $x>0$ setting $x=k\pi+t$, where $k>0$ and $0<t<\pi$, we get

$$\alpha_{0,\ell}(x) \ge C_\ell^2 \int_{|u|\le\pi/2} \sin^{2\ell+2}(t-u)\,\sin^{2\ell+2}(u)\,[u(k\pi+t-u)]^{-2\ell-2}\,du.$$

Since $(k-\frac{1}{2})\pi < k\pi+t-u < (k+1)\pi-\pi/2$, we have

$$\alpha_{0,\ell}(x) \ge C_\ell^2\left[\left(k+\frac{3}{2}\right)\pi\right]^{-2\ell-2}\int_{|u|\le\pi/2}\sin^{2\ell+2}(t-u)\,\sin^{2\ell+2}(u)u^{-2\ell-2}\,du,$$

and consequently $\alpha_{0,\ell}(x) \ge C'[(k+3/2)\pi]^{-2\ell-2}$. We finally use that $x^{2\ell+2} = (k\pi+t)^{2\ell+2} > (k\pi)^{2\ell+2}$ and that there exists a constant C'' such that for k large enough, $[(k+\frac{3}{2})\pi]^{2\ell+2} \le C_1(k\pi)^{2\ell+2} \le C'' x^{2\ell+2}$. Combining (50) and (51), we get that for $|x|$ large enough the function $\alpha_{m,a,\ell}(x)$ is positive. □

Proof of Lemma 7.8. We will consider separately the cases $\ell=0$ (estimation of the density of the observations) and $\ell\ge 1$ (estimation of the derivatives of this density). We fix the kernel K equal to $V_\lambda = V(\lambda\cdot)$, where V is the de la Vallée Poussin kernel (see (3)) and $\lambda\ge 1$ will be chosen in the proof.

First case: $\ell = 0$. Fix two functions a and a' in $\mathcal{A}_2$. By definition, the distance $\|\varphi_{m,a,\ell} - \varphi_{m,a',\ell}\|_p$ is given by

$$\|\varphi_{m,a,\ell} - \varphi_{m,a',\ell}\|_p = \left\| \sum_{j=1}^{m-1} (a(j) - a'(j)) h_{0,\ell}(K_{m,j} - \overline{K}_{m,j}) \right\|_p .$$

Now, we restrict our attention to the interval $[1/(2m); (2m-1)/(2m)]$ and bound the integral $\|\varphi_{m,a,\ell} - \varphi_{m,a',\ell}\|_p^p$ from below by the integral over this subset. Writing

$$\left[\frac{1}{2m}; \frac{2m-1}{2m}\right] = \bigcup_{1\le k\le m-1} \left[\frac{2k-1}{2m}; \frac{2k+1}{2m}\right],$$

we get

$$\|\varphi_{m,a,\ell} - \varphi_{m,a',\ell}\|_p^p \ge \sum_{k=1}^{m-1} \int_{\frac{2k-1}{2m}}^{\frac{2k+1}{2m}} \left| \sum_{j=1}^{m-1} (a(j) - a'(j)) h_{0,\ell}(x)(K_{m,j}(x) - \overline{K}_{m,j}) \right|^p dx.$$

We use the following inequality for real numbers, valid for all $p \ge 1$,

$$|a|^p \ge 2^{-p}|a-b|^p - |b|^p, \tag{52}$$

to obtain that

$$\begin{aligned}\|\varphi_{m,a,\ell} - \varphi_{m,a',\ell}\|_p^p \ge \sum_{k=1}^{m-1} \int_{\frac{2k-1}{2m}}^{\frac{2k+1}{2m}} \Bigg[& 2^{-p}|a(k) - a'(k)|^p h_{0,\ell}^p(x) |K_{m,k}(x) - \overline{K}_{m,k}|^p \\ & - \Bigg| \sum_{\substack{1\le j\le m-1 \\ j\ne k}} \big(a(j) - a'(j)\big) h_{0,\ell}(x) \big(K_{m,j}(x) - \overline{K}_{m,j}\big) \Bigg|^p \Bigg] dx.\end{aligned}$$

Since the functions a and a' take values in $\{0;1\}$, the quantity $|a(k) - a'(k)|^p$ equals $|a(k) - a'(k)|$ and is bounded by 1, for all $1 \le k \le m-1$. Use also the triangle inequality to obtain

$$\begin{aligned}\|\varphi_{m,a,\ell} - \varphi_{m,a',\ell}\|_p^p \ge \sum_{k=1}^{m-1} \int_{\frac{2k-1}{2m}}^{\frac{2k+1}{2m}} \Bigg[& 2^{-p}|a(k) - a'(k)| h_{0,\ell}^p(x) |K_{m,k}(x) - \overline{K}_{m,k}|^p \\ & - \Bigg(\sum_{\substack{1\le j\le m-1 \\ j\ne k}} h_{0,\ell}(x) |K_{m,j}(x) - \overline{K}_{m,j}| \Bigg)^p \Bigg] dx.\end{aligned}$$

The interval $[1/(2m); (2m-1)/(2m)]$ is contained in $[0;1]$ and the density $h_{0,\ell}$ is positive on the compact interval $[0;1]$, so that we have

$$\begin{aligned}&\|\varphi_{m,a,\ell} - \varphi_{m,a',\ell}\|_p^p \\ &\quad \ge 2^{-p} \sum_{k=1}^{m-1} |a(k) - a'(k)| \big(\inf_{[0;1]} h_{0,\ell}^p\big) \int_{\frac{2k-1}{2m}}^{\frac{2k+1}{2m}} |K_{m,k}(x) - \overline{K}_{m,k}|^p \, dx \\ &\quad - \sum_{k=1}^{m-1} \int_{\frac{2k-1}{2m}}^{\frac{2k+1}{2m}} \Bigg(\sum_{\substack{1\le j\le m-1 \\ j\ne k}} h_{0,\ell}(x) |K_{m,j}(x) - \overline{K}_{m,j}| \Bigg)^p dx.\end{aligned} \tag{53}$$

Consider the first term and use inequality (52) to get that

$$\int_{\frac{2k-1}{2m}}^{\frac{2k+1}{2m}} |K_{m,k}(x) - \overline{K}_{m,k}|^p \, dx \geq 2^{-p} \int_{\frac{2k-1}{2m}}^{\frac{2k+1}{2m}} |K_{m,k}(x)|^p \, dx - \frac{|\overline{K}_{m,k}|^p}{m}.$$

Therefore by using (42) and the definition of $K_{m,k}$

$$\int_{\frac{2k-1}{2m}}^{\frac{2k+1}{2m}} |K_{m,k}(x) - \overline{K}_{m,k}|^p \, dx \geq 2^{-p} \frac{\theta_m^p}{m} \int_{-1/2}^{1/2} |K|^p - \frac{\theta_m^p}{m^{p+1}} \|h_{0,\ell}\|_\infty^p \|K\|_1^p.$$

Coming back to (53), we get

$$\begin{aligned}
&\|\varphi_{m,a,\ell} - \varphi_{m,a',\ell}\|_p^p \\
&\quad \geq 2^{-p} \frac{\theta_m^p}{m} \big(\inf_{[0;1]} h_{0,\ell}^p \big) \sum_{k=1}^{m-1} |a(k) - a'(k)| \Big(2^{-p} \int_{-1/2}^{1/2} |K|^p - \frac{\|h_{0,\ell}\|_\infty^p \|K\|_1^p}{m^p} \Big) \\
&\quad - \sum_{k=1}^{m-1} \int_{\frac{2k-1}{2m}}^{\frac{2k+1}{2m}} h_{0,\ell}(x)^p \Big(\sum_{\substack{1 \leq j \leq m-1 \\ j \neq k}} |K_{m,j}(x) - \overline{K}_{m,j}| \Big)^p dx.
\end{aligned}$$

By using (44) we deduce that

$$\begin{aligned}
(54) \qquad &\|\varphi_{m,a,\ell} - \varphi_{m,a',\ell}\|_p^p \\
&\quad \geq \frac{2^{-p}}{8} \theta_m^p \big(\inf_{[0;1]} h_{0,\ell}^p \big) \Big(2^{-p} \int_{-1/2}^{1/2} |K|^p - \frac{\|h_{0,\ell}\|_\infty^p \|K\|_1^p}{m^p} \Big) \\
&\quad - \sum_{k=1}^{m-1} \int_{\frac{2k-1}{2m}}^{\frac{2k+1}{2m}} h_{0,\ell}(x)^p \Big(\sum_{\substack{1 \leq j \leq m-1 \\ j \neq k}} |K_{m,j}(x) - \overline{K}_{m,j}| \Big)^p dx.
\end{aligned}$$

Consider the second term in (54). Apply the triangle inequality, bound (42), and the definition of the kernel $K = V_\lambda$ to get that

$$\sum_{\substack{1 \leq j \leq m-1 \\ j \neq k}} |K_{m,j}(x) - \overline{K}_{m,j}| \leq \frac{\theta_m}{\lambda} \Big(\sum_{\substack{1 \leq j \leq m-1 \\ j \neq k}} \frac{2}{\lambda \pi (mx - j)^2} + \|h_{0,\ell}\|_\infty \|V\|_1 \Big).$$

Arguing as in the proof of Lemma 7.4, when $x \in [(2k-1)/(2m); (2k+1)(2m)]$, we obtain

$$\begin{aligned}
\sum_{\substack{1 \leq j \leq m-1 \\ j \neq k}} \frac{2}{\lambda \pi (mx - j)^2} &\leq \sum_{1 \leq j \leq k-1} \frac{2}{\lambda \pi (k - j - \frac{1}{2})^2} + \sum_{k+1 \leq j \leq m-1} \frac{2}{\lambda \pi (k - j + \frac{1}{2})^2} \\
&\leq \frac{4}{\lambda \pi} \sum_{n \geq 1} \frac{1}{(n - \frac{1}{2})^2}.
\end{aligned}$$

The last quantity is a finite constant to be denoted by C_1/λ. It follows that, when x belongs to the interval $[(2k-1)/(2m);(2k+1)/(2m)]$, we have

$$\sum_{\substack{1\le j\le m-1\\ j\ne k}} |K_{m,j}(x)-\overline{K}_{m,j}| \le \frac{\theta_m}{\lambda}\big(C_1/\lambda+\|h_{0,\ell}\|_\infty\|V\|_1\big).$$

Now, coming back to (54), we obtain

$$\|\varphi_{m,a,\ell}-\varphi_{m,a',\ell}\|_p^p \ge \theta_m^p\bigg[\frac{4^{-p}}{8}\big(\inf_{[0;1]} h_{0,\ell}^p\big)\int_{-1/2}^{1/2}|K|^p \\ -\lambda^{-p}\big(C_1/\lambda+\|h_{0,\ell}\|_\infty\|V\|_1\big)^p\|h_{0,\ell}\|_p^p-\frac{2^{-p}}{8}\big(\inf_{[0;1]} h_{0,\ell}^p\big)\frac{\|h_{0,\ell}\|_\infty^p\|V\|_1^p}{\lambda^p m^p}\bigg].$$

Let us study the right-hand side of this inequality. The third term is negligible with respect to the two others as $m\to\infty$. The last thing to check is that the quantity

$$\frac{4^{-p}}{8}\big(\inf_{[0;1]} h_{0,\ell}^p\big)\int_{-1/2}^{1/2}|K|^p-\lambda^{-p}\big(C_1/\lambda+\|h_{0,\ell}\|_\infty\|V\|_1\big)^p\|h_{0,\ell}\|_p^p \tag{55}$$

is positive. By using that $K=V_\lambda$, we infer that

$$\int_{-1/2}^{1/2}|K|^p=\frac{1}{\lambda}\int_{-\lambda/2}^{\lambda/2}|V|^p,$$

and therefore the quantity (55) becomes

$$\frac{1}{\lambda}\bigg[\frac{4^{-p}}{8}\big(\inf_{[0;1]} h_{0,\ell}^p\big)\int_{-\lambda/2}^{\lambda/2}|V|^p-\frac{1}{\lambda^{p-1}}\bigg(\frac{C_1}{\lambda}+\|h_{0,\ell}\|_\infty\|V\|_1\bigg)^p\|h_{0,\ell}\|_p^p\bigg],$$

which is positive for some λ_0 large enough, since $p>1$. Hence we conclude that there exists a positive constant C such that $\|\varphi_{m,a,\ell}-\varphi_{m,a',\ell}\|_p^p\ge C\theta_m^p$.

Second case: $\ell\ge 1$. By definition $\|\varphi^{(\ell)}_{m,a,\ell}-\varphi^{(\ell)}_{m,a',\ell}\|_p$ satisfies

$$\|\varphi^{(\ell)}_{m,a,\ell}-\varphi^{(\ell)}_{m,a',\ell}\|_p \ge A_1-A_2-A_3 \tag{56}$$

with

$$A_1=\bigg\|\sum_{j=1}^{m-1}\big(a(j)-a'(j)\big)h_{0,\ell}K^{(\ell)}_{m,j}\bigg\|_p,\quad A_2=\bigg\|\sum_{j=1}^{m-1}\big(a(j)-a'(j)\big)h^{(\ell)}_{0,\ell}(K_{m,j}-\overline{K}_{m,j})\bigg\|_p,$$

$$\text{and}\qquad A_3=\bigg\|\sum_{j=1}^{m-1}\big(a(j)-a'(j)\big)\sum_{s=1}^{\ell-1}\binom{\ell}{s}h^{(s)}_{0,\ell}K^{(\ell-s)}_{m,j}\bigg\|_p.$$

The purpose is to bound A_1 from below, and then to bound A_2 and A_3 from above by quantities negligible with respect to A_1. We begin with the first term

A_1, which will be treated in the same way as the case $\ell = 0$. We restrict our attention to the interval $[1/(2m); (2m-1)/(2m)]$, which is the disjoint union of the intervals $[(2k-1)/(2m); (2k+1)/2m]$, when $1 \le k \le m-1$, and use the convexity inequality (52) to write

$$A_1^p \ge m^{\ell p}\theta_m^p \sum_{k=1}^{m-1} \int_{\frac{2k-1}{2m}}^{\frac{2k+1}{2m}} \Big[2^{-p}|a(k)-a'(k)|h_{0,\ell}^p(x)|K^{(\ell)}(mx-k)|^p - \Big| \sum_{\substack{1\le j\le m-1\\ j\ne k}} \big(a(j)-a'(j)\big)h_{0,\ell}(x)K^{(\ell)}(mx-j)\Big|^p \Big] dx.$$

The same calculations as in the case $\ell = 0$ give

$$(57)\qquad A_1^p \ge m^{\ell p}\theta_m^p \Big[\frac{4^{-p}}{8}\big(\inf_{[0;1]} h_{0,\ell}^p\big) \int_{-1/2}^{1/2} |K^{(\ell)}|^p - \sum_{k=1}^{m-1} \int_{\frac{2k-1}{2m}}^{\frac{2k+1}{2m}} h_{0,\ell}^p(x) \Big(\sum_{1\le j\le m-1; j\ne k} |K^{(\ell)}(mx-j)| \Big)^p dx \Big],$$

and the point is to bound the quantity

$$\sum_{1\le j\le m-1; j\ne k} |K^{(\ell)}(mx-j)|$$

from above by a finite constant. Recall the definition of the function $K = V(\lambda\,\cdot)$, where

$$V(u) = \frac{\cos(u) - \cos(2u)}{\pi u^2} = \frac{1}{\pi}\frac{P(\cos(u))}{Q(u)}$$

with P a polynomial function and $Q(u) = u^2$. Now, compute the derivative with respect to u,

$$K^{(\ell)}(u) = \lambda^\ell V^{(\ell)}(\lambda u) = \frac{\lambda^\ell}{\pi} \sum_{k=0}^{\ell} \binom{\ell}{k} (P\circ\cos)^{(\ell-k)}(\lambda u) \Big(\frac{1}{Q}\Big)^{(k)}(\lambda u).$$

Easy calculations give

$$\Big(\frac{1}{Q}\Big)^{(k)}(\lambda u) = \frac{(-1)^k(k+1)!}{\lambda^{2+k}u^{2+k}},$$

and by obvious upper bounds on trigonometric functions, there exists a constant M_ℓ such that, for all integer $0 \le k \le \ell$, $|(P\circ\cos)^{(\ell-k)}(\lambda u)| \le M_\ell$. Therefore the derivative $K^{(\ell)}$ satisfies

$$|K^{(\ell)}(u)| \le \frac{\lambda^{\ell-2}M_\ell}{\pi} \sum_{k=0}^{\ell} \binom{\ell}{k} \frac{(k+1)!}{|u|^{2+k}} \le \begin{cases} \lambda^{\ell-2}v_\ell/|u|^2 & \text{if } |u| \ge 1, \\ \lambda^{\ell-2}v_\ell/|u|^{2+\ell} & \text{if } |u| \le 1, \end{cases}$$

where v_ℓ is a positive constant. For $x \in [(2k-1)/(2m); (2k+1)/(2m)]$, the variable $mx - j$ ranges over the interval $[k-j-1/2; k-j+1/2]$, hence we have

$$\sum_{1\le j\le m-1; j\ne k} |K^{(\ell)}(mx-j)| \le |K^{(\ell)}(mx-k+1)| + |K^{(\ell)}(mx-k-1)| + 2\lambda^{\ell-2} v_\ell \sum_{n\ge 2} \frac{1}{(n-1/2)^2}.$$

Finally,

$$\sum_{1\le j\le m-1; j\ne k} |K^{(\ell)}(mx-j)| \le 2\lambda^{\ell-2} v_\ell (1+2^{2+\ell}) + 2\lambda^{\ell-2} v_\ell \sum_{n\ge 2} \frac{1}{(n-1/2)^2}. \tag{58}$$

The last quantity is a finite constant to be denoted by $\lambda^{\ell-2} C_\ell$. Combining (57) with (58), we get

$$A_1^p \ge m^{\ell p} \theta_m^p \left[\frac{4^{-p}}{8} \big(\inf_{[0;1]} h_{0,\ell}^p \big) \int_{-1/2}^{1/2} |K^{(\ell)}|^p - \lambda^{p(\ell-2)} C_\ell^p \|h_{0,\ell}\|_p^p \right].$$

The next thing to check is that the following quantity is positive:

$$\frac{2^{-p}}{8} \big(\inf_{[0;1]} h_{0,\ell}^p \big) \int_{-1/2}^{1/2} |K^{(\ell)}|^p - \lambda^{p(\ell-2)} C_\ell^p \|h_{0,\ell}\|_p^p > 0. \tag{59}$$

Since $\int_{-1/2}^{1/2} |K^{(\ell)}|^p = \lambda^{\ell p-1} \int_{-\lambda/2}^{\lambda/2} |V^{(\ell)}|^p$, the quantity in (59) becomes

$$\lambda^{\ell p-1} \left(\frac{2^{-p}}{8} \big(\inf_{[0;1]} h_{0,\ell}^p \big) \int_{-\lambda/2}^{\lambda/2} |V^{(\ell)}|^p - \frac{(C'_\ell)^p}{\lambda^{2p-1}} \|h_{0,\ell}\|_p^p \right).$$

For $p > \frac{1}{2}$, the second term converges to zero as $\lambda \to \infty$, while the first one converges to $(2^{-p}/8)(\inf_{[0;1]} h_{0,\ell}^p)\|V^{(\ell)}\|_p^p$. Consequently we obtain that there exists a $\lambda_0 \ge 1$ large enough for which (59) holds. Using this remark, we conclude that there exists a positive constant C_1 such that for m large enough,

$$A_1 = \left\| \sum_{j=1}^{m-1} \big(a(j) - a'(j)\big) h_{0,\ell} K_{m,j}^{(\ell)} \right\|_p \ge C_1 m^\ell \theta_m. \tag{60}$$

Now we estimate the second term A_2 from above. Apply the triangle inequality to infer that

$$A_2 \le \sum_{j=1}^{m-1} \Big(\|h_{0,\ell}^{(\ell)}\|_\infty \|K_{m,j}\|_p + |\overline{K}_{m,j}| \; \|h_{0,\ell}^{(\ell)}\|_p \Big).$$

Again using the definition of $K_{m,j}$, we get $\|K_{m,j}\|_p = \theta_m m^{-1/p} \|K\|_p$, which combined with inequality (42) provides that for $p \ge 1$,

$$A_2 \le (m-1) \left(\frac{\theta_m}{m^{1/p}} \|K\|_p \|h_{0,\ell}^{(\ell)}\|_\infty + \frac{\theta_m}{m} \|h_{0,\ell}\|_\infty \|K\|_1 \|h_{0,\ell}^{(\ell)}\|_p \right), \tag{61}$$

and this quantity is negligible with respect to $m^\ell\theta_m$ when $\ell \geq 1$. Finally, the third term A_3 satisfies

$$A_3 \leq \sum_{j=1}^{m-1} \sum_{k=1}^{\ell-1} \binom{\ell}{k} \|h_{0,\ell}^{(k)}\|_\infty \|K_{m,j}^{(\ell-k)}\|_p$$

with

$$\|K_{m,j}^{(\ell-k)}\|_p \leq m^{\ell-1-1/p}\theta_m \max_{1\leq r\leq \ell-1} \|K^{(r)}\|_p.$$

Therefore

$$A_3 \leq 2^\ell m^{\ell-1/p}\theta_m \max_{1\leq r\leq \ell-1} \|h_{0,\ell}^{(r)}\|_\infty \max_{1\leq r\leq \ell-1} \|K^{(r)}\|_p, \tag{62}$$

which is also negligible with respect to $m^\ell\theta_m$. Combining (56), (60), (61), and (62), we conclude that for any $\ell \geq 1$ there exists a positive constant C such that for m large enough, $\|\varphi_{m,a,\ell}^{(\ell)} - \varphi_{m,a',\ell}^{(\ell)}\|_p \geq Cm^\ell\theta_m$, and the proof is complete. $\square$

References

[1] L. M. Artiles Martinez, *Adaptive minimax estimation in a class of smooth functions*, PhD Thesis, Utrecht University, Budapestlaan 6, 3584 CD Utrecht, The Netherlands, 1991.

[2] L. M. Artiles Martinez and B. Y. Levit, *Adaptive estimation of analytic functions on an interval*, Math. Methods Statist., 12 (2003), 62–94.

[3] L. Birgé, *A new look at an old result: Fano's lemma*, Technical report, Prépublication 632, Laboratoire de Probabilités et Modèles Aléatoires, Universités Paris 6 and Paris 7, CNRS (UMR 7599), 4, Place Jussieu, 75252 Paris Cedex 05, 2001.

[4] C. Butucea, *Deconvolution of supersmooth densities with smooth noise*, Canad. J. Statist., 32 (2004), 181–192.

[5] C. Butucea and A. Tsybakov, *Sharp optimality and some effects of dominating bias in density deconvolution*, Technical report, Laboratoire de Probabilités et Modèles Aléatoires, Universités Paris 6 and Paris 7, CNRS (UMR 7599), 4, Place Jussieu, 75252 Paris Cedex 05, 2004.

[6] E. Cator, *Deconvolution with arbitrarily smooth kernels*, Statist. Probab. Lett., 54 (2001), 205–214.

[7] F. Comte and M. L. Taupin, *Penalized contrast estimator for density deconvolution with mixing variables*, Technical report, Prépublication 2003-30, Laboratoire de mathématiques, CNRS UMR 8628, Université Paris-Sud - Bât. 425, 91405 Orsay Cedex (France), 2003.

[8] T. M. Cover and J. A. Thomas, *Elements of Information Theory*, Wiley, New York, 1991.

[9] K. B. Davis, *Mean square error properties of density estimates*, Ann. Statist., 3 (1975), 1025–1030.

[10] K. B. Davis, *Mean integrated square error properties of density estimates*, Ann. Statist., 5 (1977), 530–535.

[11] J. Fan, *On the optimal rates of convergence for nonparametric deconvolution problems*, Ann. Statist., 19 (1991), 1257–1272.

[12] R. M. Fano, *Class Notes for Transmission of Information*, Course 6.574. MIT, Cambridge, 1952.

[13] R. D. Gill and B. Y. Levit, *Applications of the van Trees inequality: a Bayesian Cramér–Rao bound*, Bernoulli, 1 (1995), 59–79.

[14] G. K. Golubev and B. Y. Levit, *Asymptotically efficient estimation for analytic distributions*, Math. Methods Statist., 5 (1996), 357–368.

[15] I. A. Ibragimov and R. Z. Hasminskii, *Estimation of distribution density*, J. Soviet. Math., 25 (1983), 40–57.

[16] S. M. Nikol'skii, *Approximations of Functions of Several Variables and Embedding Theorems*, Springer, 1975.

[17] M. Pensky and B. Vidakovic, *Adaptive wavelet estimator for nonparametric density deconvolution*, Ann. Statist., 27 (1999), 2033–2053.

[18] I. Pinelis, *Optimum bounds on moment of sum of independent random vectors*, Sibirean Adv. Math., 5 (1995), 141–150.

[19] H. P. Rosenthal, *On the subspaces of* l^p $(p > 2)$, *spanned by sequences of independent random variables*, Israel J. Math., 8 (1970), 273–303.

[20] M. L. Taupin, *Semi-parametric estimation in the nonlinear errors-in-variables model*, PhD thesis, Université Paris-Sud, 91405 Orsay France, 1998.

[21] M. L. Taupin, *Semi-parametric estimation in the nonlinear structural errors-in-variables model*, Ann. Statist., 29 (2001), 66–93.

[*Received October* 2002; *revised April* 2004]