

Estimation of integral functionals of a density and its derivatives

BÉATRICE LAURENT

Université Paris Sud, Bât. 425, 91405 Orsay Cédex, France

We consider the problem of estimating a functional of a density of the type $\int \phi(f, f', \dots, f^{(k)}, \cdot)$. The estimation of $\int \phi(f, \cdot)$ has already been studied by the author: starting from efficient estimators of linear and quadratic functionals of f and its derivatives and using a Taylor expansion of ϕ , we construct estimators which achieve the $n^{-1/2}$ rate whenever f is smooth enough. Moreover, we show that these estimators are efficient. We also obtain the optimal rate of convergence when the $n^{-1/2}$ rate is not achievable and when $k > 0$. Concerning the estimation of quadratic functionals, more precisely of integrated squared density derivatives, Bickel and Ritov have already constructed efficient estimators. Here we propose an alternative construction based on projections, an approach which seems more natural.

Keywords: estimation of a density and its derivatives; projection methods; kernel estimators; Fourier series; semi-parametric Cramér–Rao bound

1. Introduction

Let X_1, \dots, X_n be i.i.d. random variables with common density f defined over a compact set S of \mathbb{R} . Our purpose is to estimate quantities of the type $T(f) = \int \phi(f, f', \dots, f^{(k)}, \cdot)$ where f is assumed to belong to some Sobolev space of index $s > k$. This problem is motivated by statistical applications. For example, let \hat{f}_n be a kernel estimator of the density f : $\hat{f}_n(x) = nh^{-1} \sum_{i=1}^n K((x - X_i)/h)$. If f is a regular function, the optimal value of the parameter h in order to minimize the mean integrated square error depends on integral functionals of the type $\int (f^{(k)})^2$, as shown in Deheuvels and Hominal (1980). The same type of functional occurs in penalization estimation related to smoothness (see Wahba 1990). The knowledge of $\int (f^{(k)})^2$ allows the original problem to be converted into a constrained estimation problem (see Gu 1994). Moreover, the estimate of the Fisher information $\int (f')^2 / f$ will give a value for the Cramér–Rao bound in a translation model.

The estimation of integral functionals was first studied by Levit (1978), who constructed efficient estimators of this kind of functional under regularity properties for the density f which are not optimal. It is also worth mentioning the paper by Ibragimov *et al.* (1987) which deals with differentiable functionals in Gaussian white noise. The problem of estimating $\int (f^{(k)})^2$ has already been studied by Bickel and Ritov (1988) and by Donoho and Nussbaum (1990). The same results are obtained in these papers: if $s > 2k + \frac{1}{4}$, $\int (f^{(k)})^2$

may be estimated at the semi-parametric rate $n^{-1/2}$; if $k < s \leq 2k + \frac{1}{4}$, the rate of convergence has order $n^{-4(s-k)/(1+4s)}$. Moreover, Bickel and Ritov show that their estimator is efficient and that the nonparametric rate of convergence is optimal.

In this paper we generalize the results obtained in Laurent (1996) concerning the problem of estimating $\int \phi(f, \cdot)$. Before stating our results, we shall explain how our methods work: ϕ is assumed to be a smooth function. So, finding the Taylor expansion of ϕ up to terms of second order provides an expression for $T(f) - T(\hat{f})$, where \hat{f} is a nonparametric preliminary estimator of the density f , constructed with a small part of the n -sample. With the remainder of the sample, we construct estimators of the terms, up to the second order, which appear in the Taylor expansion. Some of these terms are linear functionals of f and its derivatives while others are quadratic functionals of the type $\int f^{(j)} f^{(j')} K_{j,j'}(\hat{f})$. So, a crucial point in making our methods work will be the construction of estimators of quantities of the type $\int f^{(j)} f^{(j')} \psi$ where ψ is a fixed smooth function, since we shall work conditionally to \hat{f} . This will be the purpose of Section 2. The estimator of $\int f^{(j)} f^{(j')} \psi$ that we propose is based on the orthogonal projection of f on the Fourier basis. f and ψ are supposed to be periodic functions defined on a compact set. For certain functionals, this assumption of periodicity is necessary to achieve the rate of convergence that we obtain. The main idea of the construction of the estimator is to take a linear combination of several estimators of $\int f^{(j)} f^{(j')} \psi$ in order to reduce the bias. The order of the expansion on the Fourier basis will be determined by the usual trade-off between the bias term and the variance. For the particular problem of estimating $\int (f^{(k)})^2$ we obtain the same results as Bickel and Ritov, and as Donoho and Nussbaum; moreover, the formulation of the estimator is very simple.

Our main result is stated in Section 3, and may be summarized as follows. We can construct an estimator \hat{T}_n of $T(f)$ such that:

(i) if $s > 2k + \frac{1}{4}$, $\sqrt{n}(\hat{T}_n - T(f)) \rightarrow N(0, C(f, \phi))$ where $N(0, \sigma^2)$ denotes the normal distribution and $nE(\hat{T}_n - T(f))^2 \rightarrow C(f, \phi)$, with

$$C(f, \phi) = \int \left[\sum_{j=0}^k (-1)^j (\phi'_j(f, f', \dots, f^{(k)}, \cdot))^{(j)} \right]^2 f - \left[\sum_{j=0}^k \int \phi'_j(f, f', \dots, f^{(k)}, \cdot) f^{(j)} \right]^2,$$

in which $\phi'_j(x_0, \dots, x_k, t) = \partial \phi(x_0, \dots, x_k, t) / \partial x_j$ and $(\phi'_j(f, f', \dots, f^{(k)}, \cdot))^{(j)}$ is the derivative of order j of the function $x \mapsto \phi'_j(f(x), f'(x), \dots, f^{(k)}(x), x)$;

(ii) if $k > 0$ and $k < s \leq 2k + \frac{1}{4}$ then $E(\hat{T}_n - T(f))^2 = O(n^{-8(s-k)/(1+4s)})$.

In case (i) $C(f, \phi)$ is the semi-parametric information bound for the problem of estimating the smooth functional $T(f)$ when f is assumed to belong to a set of regular functions of order s as is shown in the Appendix. Hence our estimator is asymptotically efficient. On the other hand, the rates which appear in (ii) cannot be improved for most functionals $T(f)$, as shown in Birgé and Massart (1995).

When $k = 0$ and $s < \frac{1}{4}$ we do not obtain the optimal rate. Our estimator converges at the rate $n^{-3s/(1+2s)}$. Actually, in this very case the remainder term in the Taylor expansion is precisely of the order $n^{-3s/(1+2s)}$. So, it is necessary to do the Taylor expansion up to the

third order and to estimate $\int f^3\psi$. Kerkyacharian and Picard (1996) gave an estimator of $\int f^3\psi$ with the optimal rate of convergence $n^{-4s/(1+4s)}$.

This paper is organized as follows. In Section 2, we construct an estimator for the quantity $\int f^{(j)}f^{(j')}\psi$. In Section 3, we propose an estimator for $T(f)$. The proofs of the theorems are postponed to Section 4.

2. Estimation of $\int f^{(j)}f^{(j')}\psi$

We will first set out some notation. The expression $A_n \approx B_n$ will mean that $\nu_1 \leq A_n/B_n \leq \nu_2$, where ν_1 and ν_2 are positive constants. We shall denote (j, j') by J and $\int f^{(j)}f^{(j')}\psi$ by T^J_ψ . For any function g defined over $S = [-\pi, \pi]$, we define:

$$\|g\|_\infty = \sup_{x \in S} |g(x)|, \quad \|g\|_\alpha = \sup_{x \neq y \in S} \frac{|g(x) - g(y)|}{|x - y|^\alpha} \quad \text{for } 0 < \alpha \leq 1.$$

Definition 1. Let $s = p + \alpha$, where $p \in \mathbb{N}$ and $\alpha \in]0, 1]$. We shall denote by $F_{s,C}$ the set of densities f with support on $S = [-\pi, \pi]$ satisfying the following conditions:

- (i) $f^{(l)}(\pi) = f^{(l)}(-\pi)$ for all $l \in \{0, \dots, p - 1\}$;
- (ii) $\|f^{(p)}\|_\alpha \leq C$.

When necessary, the function f will be extended by periodicity. $C^k(S)$ will be the space of k -times differentiable functions on S .

Let $(p_i, i \in \mathbb{N})$ be the orthonormal Fourier basis of $\mathbb{L}^2([-\pi, \pi])$:

$$p_0(x) = \frac{1}{\sqrt{2\pi}}, \quad p_{2i}(x) = \frac{\cos ix}{\sqrt{\pi}}, \quad p_{2i-1}(x) = \frac{\sin ix}{\sqrt{\pi}} \quad \text{for } i > 0.$$

We set $S_m f(x) = \sum_{i=0}^m a_i(f) p_i(x)$ with $a_i(f) = \int f p_i$. When no confusion is likely, we shall write a_i instead of $a_i(f)$.

We wish to estimate $T^J_\psi = \int_{-\pi}^\pi f^{(j)}f^{(j')}\psi(x) dx$. We assume that $j \leq j' < s$, $\psi \in C^{j'}(S)$ and that $\psi^{(l)}(\pi) = \psi^{(l)}(-\pi)$ for all $l \in \{0, \dots, j' - 1\}$.

The problem of estimating $\int f^2\psi$ has already been studied in a more general framework, including the multidimensional case in Laurent (1996). Moreover, in order to estimate $\int f^2\psi$, we do not have to suppose that f and ψ are periodic functions since these assumptions are used to carry out integration by parts. The construction of the estimator of $\int_{-\pi}^\pi f^{(j)}f^{(j')}\psi(x) dx$ will be of the same type. We recall that the estimator of $\int f^2\psi$ proposed in Laurent (1996) is

$$\begin{aligned} \hat{T}^{\text{OO}}_\psi &= \frac{2}{n(n-1)} \sum_{i=0}^{m(n)} \sum_{l_1 \neq l_2=1}^n p_i(X_{l_1}) p_{i'}(X_{l_2}) \psi(X_{l_2}) \\ &\quad - \frac{1}{n(n-1)} \sum_{i,i'=0}^{m(n)} \sum_{l_1 \neq l_2=1}^n p_i(X_{l_1}) p_{i'}(X_{l_2}) \int p_i p_{i'} \psi. \end{aligned}$$

Let us write m instead of $m(n)$. The bias of this estimator is equal to $-\int(S_m f - f)^2 \psi$. We shall look for a similar bias for the estimation of T_{ψ}^J . More precisely, we wish the bias to be equal to

$$-\int_{-\pi}^{\pi} ((S_m f)^{(j)} - f^{(j)})(S_m f)^{(j')} - f^{(j')} \psi = \int_{-\pi}^{\pi} (f^{(j)}(S_m f)^{(j')} + f^{(j')}(S_m f)^{(j)}) \psi - \int_{-\pi}^{\pi} (S_m f)^{(j)}(S_m f)^{(j')} \psi - \int_{-\pi}^{\pi} f^{(j)} f^{(j')} \psi.$$

Indeed, by special properties of the Fourier basis, this bias will be equal to $-\int(S_m(f^{(j)} - f^{(j)})(S_m(f^{(j')} - f^{(j')}))\psi$. Moreover, when f is a regular function of order s , the \mathbb{L}^2 -norm of $(S_m(f^{(l)} - f^{(l)}))$ is well controlled if $l \leq s$. Hence, the aim of this construction is to provide an estimator with small bias, which will lead to the optimal rate of convergence in most cases. The problem is to find an estimator with expectation

$$\int_{-\pi}^{\pi} (f^{(j)}(S_m f)^{(j')} + f^{(j')}(S_m f)^{(j)}) \psi - \int_{-\pi}^{\pi} (S_m f)^{(j)}(S_m f)^{(j')} \psi.$$

Since $f^{(l)}(\pi) = f^{(l)}(-\pi)$ and $\psi^{(l)}(\pi) = \psi^{(l)}(-\pi)$ for all $l \in \{0, \dots, j' - 1\}$, repeated integration by parts leads to $\int_{-\pi}^{\pi} f^{(j)}(S_m f)^{(j')} \psi = (-1)^j \int_{-\pi}^{\pi} ((S_m f)^{(j')} \psi)^{(j)} f$. It is therefore easy to see that the estimator

$$\begin{aligned} \hat{T}_{\psi,1}^J &= \frac{1}{n(n-1)} \sum_{i=0}^m \sum_{l_1 \neq l_2=1}^n p_i(X_{l_1})(-1)^j (p_i^{(j)} \psi)^{(j)}(X_{l_2}) \\ &+ \frac{1}{n(n-1)} \sum_{i=0}^m \sum_{l_1 \neq l_2=1}^n p_i(X_{l_1})(-1)^{j'} (p_i^{(j')} \psi)^{(j')}(X_{l_2}) \end{aligned} \tag{2.1}$$

will have expectation $\int_{-\pi}^{\pi} (f^{(j)}(S_m f)^{(j')} + f^{(j')}(S_m f)^{(j)}) \psi$. To obtain the term $-\int_{-\pi}^{\pi} (S_m f)^{(j)}(S_m f)^{(j')} \psi$, we propose the estimator

$$\hat{T}_{\psi,2}^J = -\frac{1}{n(n-1)} \sum_{i,i'=0}^m \sum_{l_1 \neq l_2=1}^n p_i(X_{l_1}) p_{i'}(X_{l_2}) \int_{-\pi}^{\pi} p_i^{(j)} p_{i'}^{(j')} \psi(x) dx. \tag{2.2}$$

Theorem 1. *Let X_1, X_2, \dots, X_n be i.i.d. variables with density f belonging to the set $F_{s,C}$ defined by Definition 1, and let ψ belong to $C^j(S)$ with $j' < s$ and satisfy $\psi^{(l)}(\pi) = \psi^{(l)}(-\pi)$ for all $l \in \{0, \dots, j' - 1\}$. We wish to estimate $T_{\psi}^J = \int_{-\pi}^{\pi} f^{(j)} f^{(j')} \psi$, with $j \leq j'$.*

Define $\hat{T}_{\psi}^J = \hat{T}_{\psi,1}^J + \hat{T}_{\psi,2}^J$, where $\hat{T}_{\psi,1}^J$ and $\hat{T}_{\psi,2}^J$ are given by (2.1) and (2.2) respectively, with $m \approx n^{2/(1+4s)}$. Then:

(i) *If $s > j + j' + \frac{1}{4}$, and $\lambda_1 = \sup_{l=0, \dots, j'} \|\psi^{(l)}\|_{\infty}$, we obtain*

$$E(\hat{T}_{\psi}^J - T_{\psi}^J)^2 \leq \frac{C_1 \lambda_1^2}{n}. \tag{2.3}$$

Assuming that ξ satisfies the same assumptions as ψ , let $\hat{T}_{\xi}^{J_1}$ be defined as \hat{T}_{ψ}^J , replacing

(j, j') by (j_1, j'_1) and ψ by ξ , and $T_{\xi}^{J_1} = \int_{-\pi}^{\pi} f^{(j_1)} f^{(j'_1)} \xi$. Let $s > \sup(j + j' + \frac{1}{4}, j_1 + j'_1 + \frac{1}{4})$, $j'_1 \geq j_1$ and

$$\mu_1 = \sup [\sup_{l=0, \dots, j'} (\|\psi^{(l)}\|_{\infty}, \|\psi^{(l)}\|_{\alpha}); \sup_{l=0, \dots, j'_1} (\|\xi^{(l)}\|_{\infty}, \|\xi^{(l)}\|_{\alpha})];$$

then

$$|nE(\hat{T}_{\psi}^J - T_{\psi}^J)(\hat{T}_{\xi}^{J_1} - T_{\xi}^{J_1}) - \Lambda_{JJ_1}(f, \psi, \xi)| \leq C'_1 \mu_1^2 (m^{-\alpha} + m^{\frac{1}{2} + j + j'_1 j_1 + j'_1 - 2s}), \tag{2.4}$$

where

$$\begin{aligned} \Lambda_{JJ_1}(f, \psi, \xi) = & \int ((-1)^j (f^{(j)} \psi)^{(j)} + (-1)^{j'} (f^{(j')} \psi)^{(j')}) (-1)^{j_1} (f^{(j_1)} \xi)^{(j_1)} \\ & + (-1)^{j'_1} (f^{(j'_1)} \xi)^{(j'_1)} f - 4 \int_{-\pi}^{\pi} f^{(j)} f^{(j')} \psi \int_{-\pi}^{\pi} f^{(j_1)} f^{(j'_1)} \xi. \end{aligned}$$

(ii) If $s \leq j + j' + \frac{1}{4}$ and

$$\lambda_2 = \sup \left(\sup_{l=0, \dots, p-j} \frac{\|\psi^{(l)}\|_{\infty} + \|\psi^{(l)}\|_{\alpha}}{m^{1/6}} \sup_{l=0, \dots, j'} \frac{\|\psi^{(l)}\|_{\infty}}{m^{(j+j'-s+\frac{1}{4}) \wedge l}} \right),$$

then

$$E(\hat{T}_{\psi}^J - T_{\psi}^J)^2 \leq C_2 \lambda_2^2 n^{-(8s+4j+4j')/(1+4s)}. \tag{2.5}$$

C_1, C'_1 and C_2 are absolute constants.

Comments

(1) If $j = j'$, $s < 2j + \frac{1}{4}$ and ψ is either positive on S or negative on S , the rate of convergence is optimal (see Birgé and Massart 1995). Otherwise, we do not know whether the optimal rate is $n^{-(8s+4j+4j')/(1+4s)}$ or not.

(2) For the estimation of $\theta_k = \int (f^{(k)})^2$ the expression for our estimator is very simple:

$$\hat{\theta}_k = \frac{1}{n(n-1)} \sum_{i=0}^m \sum_{l_1 \neq l_2=1}^n q_i(k) p_i(X_{l_1}) p_i(X_{l_2}),$$

where $q_i(k) = \int_{-\pi}^{\pi} (p_i^{(k)})^2 = (-1)^k \int_{-\pi}^{\pi} p_i^{(2k)} p_i$. Hence $q_{2i}(k) = q_{2i-1}(k) = i^{2k}$ for all $i > 0$, $q_0(k) = 0$ if $k \geq 1$, $q_0(0) = 1$. It has the same properties as Bickel and Ritov's estimator.

(3) In the next section, ψ is a random function depending on n . This explains why we need bounds depending explicitly on ψ . In particular, the technical results (2.4) and (2.5) are formulated in this way with a view to proving Theorem 2.

(4) The assumption of periodicity for f and its derivatives is necessary: suppose that we want to estimate $2 \int_{-\pi}^{\pi} ff'$. This is equal to $f^2(\pi) - f^2(-\pi)$; hence, this problem is the same as estimating the density at one point. Farrell (1972) showed that the rate of convergence for the problem of estimating the density at one point never achieves the $n^{-1/2}$ rate.

3. Estimation of integrated functionals of f and its derivatives

3.1. MAIN RESULTS

The purpose of this section is to estimate $T(f) = \int_{-\pi}^{\pi} \phi(f(x), f'(x), \dots, f^{(k)}(x), x) dx$ efficiently when this is possible. Assuming that $f \in F_{s,C}$ defined by Definition 1, and $k < s$, we also give the rates of convergence of the estimators when the $n^{-1/2}$ rate is not achievable. We would like to start with some preliminary estimator \hat{f} of the density f constructed on a small part of the initial sample and do a Taylor expansion of ϕ in a neighbourhood of $(\hat{f}(x), \dots, \hat{f}^{(k)}(x), x)$. In order to give a sense to this expansion, we shall assume the following:

Assumption A1. ϕ belong to $C^{k+3}(\Omega)$ for some open set Ω and for all $0 \leq i \leq k$, $f^{(i)}(S) \subset [a_i, b_i]$ where $a_i, b_i \in \mathbb{R}$ and $\prod_{i=0}^k [a_i, b_i] \times S \subset \Omega$. For instance, if the functional to be estimated is the entropy $\int f \log(f)$ or the Fisher information $\int (f')^2/f$, we assume that f is bounded from below by some positive constant.

Since $\prod_{i=0}^k \hat{f}^{(i)}(S) \times S$ is not almost surely included in Ω when \hat{f} is a standard preliminary estimator of f , such as a kernel or projection-based estimator, we shall have to modify it in order to get a feel for $\phi(\hat{f}(x), \hat{f}'(x), \dots, \hat{f}^{(k)}(x), x)$. Moreover, we will need some periodicity conditions for \hat{f} to carry out integration by parts. In order to control the remainder term of the Taylor expansion, we will also require that the quantities $E(\|\hat{f}^{(l)} - f^{(l)}\|_q^q)$ achieve the optimal rate of convergence. More precisely, for $f \in F_{s,C}$ satisfying Assumption A1, we will need a preliminary estimator \hat{f} verifying the following conditions:

Conditions A2

- (a) $\hat{f} \in C^{2k \vee p}(S)$, $\hat{f}^{(l)}(\pi) = \hat{f}^{(l)}(-\pi)$ for all $l \in \{0, \dots, 2k - 1\}$.
- (b) $\hat{f}^{(i)}(S) \subset [a_i - \epsilon, b_i + \epsilon]$ for all $i \in \{0, \dots, k\}$, where $\epsilon > 0$ satisfies $\prod_{i=0}^k [a_i - \epsilon, b_i + \epsilon] \times S \subset \Omega$.
- (c) $E(\|\hat{f}^{(l)} - f^{(l)}\|_q^q) \leq \gamma_1(q) n_1^{q(l-s)/(1+2s)}$ for all $l \leq p$, and all $2 \leq q < \infty$,
- (d) $E(\|\hat{f}^{(l)}\|_\infty^q) \leq \gamma_2(q)(1 + n_1^{q(l-s)/(1+2s)})$ for all $s' < s$ and all $0 \leq l \leq 2k$, $q \geq 1$,
- (e) $E(\|\hat{f}^{(l)}\|_\alpha^q) \leq \gamma_3(q)(1 + n_1^{q(l+\alpha-s)/(1+2s)})$ for all $s' < s$ and $0 \leq l \leq 2k$, $q \geq 1$,

where $\gamma_1(q), \gamma_2(q), \gamma_3(q)$ are absolute constants independent of $f \in F_{s,C}$.

Such estimators exist. This is the purpose of Lemma 1.

Lemma 1. Suppose $f \in F_{s,C}$ satisfies Assumption A1. Let \tilde{f} be an estimator of f based on projection methods constructed with the n_1 last observations:

$$\tilde{f}(x) = \frac{1}{n_1} \sum_{j=n-n_1+1}^n \sum_{i=0}^{m_1} p_i(X_j) p_i(x),$$

where $m_1 \approx n_1^{1/(1+2s)}$. Let A_n be the event

$$A_n = \{\tilde{f}^{(i)}(S) \subset [a_i - \epsilon, b_i + \epsilon], \forall i \in \{0, \dots, k\}\},$$

where $\epsilon > 0$ satisfies $K_\epsilon = \prod_{i=0}^k [a_i - \epsilon, b_i + \epsilon] \times S \subset \Omega$, and let f_0 a density which satisfies Conditions A2(a) and A2(b). We define

$$\hat{f} = \tilde{f} \mathbf{1}_{A_n} + f_0 \mathbf{1}_{A_n^c}.$$

Then \hat{f} satisfies Conditions A2, as will be shown in the Appendix.

For \hat{f} satisfying Conditions A2, $\phi(\hat{f}(x), \hat{f}'(x), \dots, \hat{f}^{(k)}(x), x)$ is well defined, and it is legitimate to carry out a Taylor expansion of ϕ in a neighbourhood of $(\hat{f}(x), \dots, \hat{f}^{(k)}(x), x)$. We shall use the following notation for partial derivatives:

$$\phi'_j = \frac{\partial \phi}{\partial y_j}(y_0, \dots, y_{k+1}), \quad \phi''_{j'j''} = \frac{\partial^2 \phi}{\partial y_j \partial y_{j'}}(y_0, \dots, y_{k+1})$$

$$\|\phi^{(l)}\|_\infty = \sup_{j_1, \dots, j_l \in \{0, \dots, k\}} \sup_{(y_0, \dots, y_{k+1}) \in K_\epsilon} |\phi^{(l)}_{j_1, \dots, j_l}(y_0, \dots, y_{k+1})|.$$

The expansion of $T(f)$ using Taylor's formula is

$$\begin{aligned} T(f) &= \int \phi(\hat{f}(x), \dots, \hat{f}^{(k)}(x), x) dx + \sum_{j=0}^k \int \phi'_j(\hat{f}(x), \dots, \hat{f}^{(k)}(x), x) (f^{(j)} - \hat{f}^{(j)})(x) \\ &+ \frac{1}{2} \sum_{j=0}^k \sum_{j'=0}^k \int \phi''_{j'j''}(\hat{f}(x), \dots, \hat{f}^{(k)}(x), x) (f^{(j)} - \hat{f}^{(j)}) (f^{(j')} - \hat{f}^{(j')})(x) + \Gamma_n, \end{aligned}$$

where Γ_n is a remainder term which will be proved to be negligible compared to the linear and quadratic terms. It is convenient to write $T(f)$ as follows:

$$T(f) = \int G(\hat{f}) + \sum_{j=0}^k \int H_j(\hat{f}) f^{(j)} + \sum_{j,j'=0}^k \int K_{j'j''}(\hat{f}) f^{(j)} f^{(j')} + \Gamma_n,$$

where

$$G(\hat{f}) = \phi(\hat{f}, \dots, \hat{f}^{(k)}, \cdot) - \sum_{j=0}^k \phi'_j(\hat{f}, \dots, \hat{f}^{(k)}, \cdot) \hat{f}^j + \frac{1}{2} \sum_{j,j'=0}^k \phi''_{j'j''}(\hat{f}, \dots, \hat{f}^{(k)}, \cdot) \hat{f}^{(j)} \hat{f}^{(j')}, \tag{3.1}$$

$$H_j(\hat{f}) = \phi'_j(\hat{f}, \dots, \hat{f}^{(k)}, \cdot) - \sum_{j'=0}^k \phi''_{j'j''}(\hat{f}, \dots, \hat{f}^{(k)}, \cdot) \hat{f}^{(j')}, \tag{3.2}$$

$$K_{j'j''}(\hat{f}) = K_{j''j'}(\hat{f}) = \frac{1}{2} \phi''_{j'j''}(\hat{f}, \dots, \hat{f}^{(k)}, \cdot). \tag{3.3}$$

We have to estimate two types of terms: $\int_{-\pi}^\pi H_j(\hat{f}) f^{(j)}$, which is a linear functional of

f ; and $T^J(K_J(\hat{f})) = \int_{-\pi}^{\pi} K_J(\hat{f}) f^{(j)} f^{(j)}$, which is a quadratic functional of the type we have studied in Section 2.

Before stating the results, let us present some ideas about the estimation of these terms. Since $f^{(l)}(\pi) = f^{(l)}(-\pi)$ for all $l \in \{0, \dots, k-1\}$, and $\hat{f}^{(l)}(\pi) = \hat{f}^{(l)}(-\pi)$ for all $l \in \{0, \dots, 2k-1\}$, assuming that ϕ'_j , and $\phi''_{jj'}$, which are functions of $k+2$ variables are periodic, with respect to the last variable, as well as their partial derivatives up to order $k-1$, we obtain by repeated integration by parts,

$$\int_{-\pi}^{\pi} H_j(\hat{f}) f^{(j)} = \int_{-\pi}^{\pi} (-1)^j [H_j(\hat{f})]^{(j)} f.$$

Setting $n_2 = n - n_1$, we can estimate H_j by $\hat{H}_j = 1/n_2 \{ \sum_{l=1}^{n_2} (-1)^j [H_j(\hat{f})]^{(j)}(X_l) \}$ with the data independent of \hat{f} .

As to $T^J(K_J(\hat{f})) = \int f^{(j)} f^{(j)} K_J(\hat{f})$, its estimation has been studied in the previous section. $K_J(\hat{f})$ is a random function based on the n_1 last observations and $T^J(K_J(\hat{f}))$ has to be estimated with the remainder of the n -sample, which leads to

$$\begin{aligned} \hat{T}^J(K_J(\hat{f})) &= \frac{1}{n_2(n_2-1)} \sum_{i=0}^m \sum_{l_1 \neq l_2=1}^{n_2} p_i(X_{l_1}) (-1)^j (p_i^{(j)} K_J(\hat{f}))^{(j)}(X_{l_2}) \\ &+ \frac{1}{n_2(n_2-1)} \sum_{i=0}^m \sum_{l_1 \neq l_2=1}^{n_2} p_i(X_{l_1}) (-1)^{j'} (p_i^{(j')} K_J(\hat{f}))^{(j')}(X_{l_2}) \\ &- \frac{1}{n_2(n_2-1)} \sum_{i,i'=0}^m \sum_{l_1 \neq l_2=1}^{n_2} p_i(X_{l_1}) p_{i'}(X_{l_2}) \int_{-\pi}^{\pi} p_i^{(j)} p_{i'}^{(j')} K_J(\hat{f})(x) dx. \end{aligned}$$

Theorem 2. Let X_1, X_2, \dots, X_n be i.i.d. random variables with common density f belonging to $F_{s,C}$ defined by Definition 1. $T(f) = \int \phi(f(x), \dots, f^{(k)}(x), x) dx$ is to be estimated with $k < s$; let us assume that Assumption A1 holds and that $\phi'_j, \phi''_{jj'}$, are 2π periodic with respect to the last variable for all $j, j' \in \{0, \dots, k\}$ as well as their partial derivatives up to order $k-1$.

Consider a preliminary estimator \hat{f} of f satisfying Condition A2 and based on the n_1 last observations where $n_1 \approx n/\log(n)$. Let

$$\begin{aligned} \hat{T}_n &= \int_{-\pi}^{\pi} G(\hat{f}) + \sum_{j=0}^k \frac{1}{n_2} \sum_{l=1}^{n_2} (-1)^j [H_j(\hat{f})]^{(j)}(X_l) \\ &+ \sum_{j,j'=0}^k \sum_{i=0}^m \frac{2}{n_2(n_2-1)} \sum_{l_1 \neq l_2=1}^{n_2} (-1)^{j'} p_i(X_{l_1}) [K_J(\hat{f}) p_i^{(j)}]^{(j')}(X_{l_2}) \\ &- \sum_{j,j'=0}^k \sum_{i,i'=0}^m \frac{1}{n_2(n_2-1)} \sum_{l_1 \neq l_2=1}^{n_2} p_i(X_{l_1}) p_{i'}(X_{l_2}) \int_{-\pi}^{\pi} p_i^{(j)} p_{i'}^{(j')} K_J(\hat{f}) dx, \end{aligned}$$

where G, H_j, K_J are defined by (3.1), (3.2) and (3.3), and $m \approx n^{2/(1+4s)}$. The following properties hold:

(i) If $s > 2k + \frac{1}{4}$ then

$$\sqrt{n}(\hat{T}_n - T(f)) \rightarrow N(0, C(f, \phi)) \tag{3.4}$$

$$nE(\hat{T}_n - T(f))^2 \rightarrow C(f, \phi), \tag{3.5}$$

where $C(f, \phi)$ is given by

$$C(f, \phi) = \int_{-\pi}^{\pi} \left[\sum_{j=0}^k (-1)^j (\phi'_j(f, \dots, f^{(k)}, \cdot))^{(j)} \right]^2 f - \left[\sum_{j=0}^k \int_{-\pi}^{\pi} \phi'_j(f, \dots, f^{(k)}, \cdot) f^{(j)} \right]^2.$$

(ii) If $k > 0$ and $k < s \leq 2k + \frac{1}{4}$ then

$$E(\hat{T}_n - T(f))^2 = O(n^{-8(s-k)/(1+4s)}).$$

Comment. When $k > 0$ and $k < s \leq 2k + \frac{1}{4}$, assuming that T is not degenerate, Birgé and Massart (1995) have proved that the rate of convergence that we obtain is optimal; moreover, in the semi-parametric case, the asymptotic variance is optimal, as will be shown in the Appendix.

3.2. FISHER INFORMATION ESTIMATION

As an example, let us give the precise expression of our estimator of $\int_{-\pi}^{\pi} f'^2/f$. Using the Taylor expansion of $\phi(f)$ in a neighbourhood of (\hat{f}, \hat{f}') , we obtain

$$\begin{aligned} \int_{-\pi}^{\pi} \frac{f'^2}{f} &= - \int_{-\pi}^{\pi} \frac{\hat{f}'^2}{\hat{f}^2} f + 2 \int_{-\pi}^{\pi} \frac{\hat{f}'}{\hat{f}} f' + \int_{-\pi}^{\pi} \frac{\hat{f}''}{\hat{f}^3} f^2 + \int_{-\pi}^{\pi} \frac{f'^2}{\hat{f}} - 2 \int_{-\pi}^{\pi} \frac{\hat{f}'}{\hat{f}^2} f f' + \Gamma_n \\ &= \int_{-\pi}^{\pi} \left(\frac{\hat{f}''}{\hat{f}^2} - 2 \frac{\hat{f}'''}{\hat{f}} \right) f + \int_{-\pi}^{\pi} \frac{\hat{f}''}{\hat{f}^3} f^2 + \int_{-\pi}^{\pi} \frac{f'^2}{\hat{f}} - 2 \int_{-\pi}^{\pi} \frac{\hat{f}'}{\hat{f}^2} f f' + \Gamma_n. \end{aligned}$$

We have to suppose that f is bounded from below by a positive constant (Assumption A1); Condition A2(b) ensures that \hat{f} is also bounded from below by a positive constant. The estimator has the following expression:

$$\begin{aligned} \hat{T}_n &= \frac{1}{n_2} \sum_{l=1}^{n_2} \left(\frac{\hat{f}'^2}{\hat{f}^2} - 2 \frac{\hat{f}''}{\hat{f}} \right) (X_l) + \sum_{j,j'=0}^1 \sum_{i=0}^m \frac{2}{n_2(n_2-1)} \\ &\quad \times \sum_{l_1 \neq l_2=1}^{n_2} (-1)^{j'} p_i(X_{l_1}) [K_{j,j'}(\hat{f}) p_i^{(j)}]^{(j')} (X_{l_2}) \\ &\quad - \sum_{j,j'=0}^1 \sum_{i,i'=0}^m \frac{1}{n_2(n_2-1)} \sum_{l_1 \neq l_2=1}^{n_2} p_i(X_{l_1}) p_{i'}(X_{l_2}) \int_{-\pi}^{\pi} p_i^{(j)} p_{i'}^{(j')} K_{j,j'}(\hat{f}), \end{aligned}$$

where $K_{0,0} = \hat{f}'^2/\hat{f}^3$, $K_{0,1} = K_{1,0} = -\hat{f}''/\hat{f}^2$, $K_{1,1} = 1/\hat{f}$. The remainder term Γ_n is

bounded by the product of $\|\phi^{(3)}\|_\infty$ and quantities of the type $E(\|\hat{f}^{(l)} - f^{(l)}\|_q^q)$ for $l = 0, 1$, which are controlled by Condition A2(c). We obtain the following result. If $s > \frac{9}{4}$,

$$\lim_{n \rightarrow \infty} nE \left(\hat{T}_n - \int_{-\pi}^{\pi} \frac{f'{}^2}{f} \right)^2 = \int_{-\pi}^{\pi} \left(\frac{2f''}{f} - \frac{f'{}^2}{f^2} \right)^2 f - \left(\int_{-\pi}^{\pi} \frac{f'{}^2}{f} \right)^2.$$

If $1 < s \leq \frac{9}{4}$,

$$E \left(\hat{T}_n - \int_{-\pi}^{\pi} \frac{f'{}^2}{f} \right)^2 = O(n^{-8(s-1)/(1+4s)}).$$

4. Proofs

4.1. PROOF OF THEOREM 1

We recall that $s = p + \alpha$, where $p \in \mathbb{N}$ and $\alpha \in]0, 1]$. The following lemmas will prove useful.

Lemma 2 (Lorentz’s inequality).

$$\sup \left\{ \sum_{i=m+1}^{\infty} i^{2p}(a_{2i}^2(f) + a_{2i-1}^2(f)), f \in F_{s,C} \right\} \leq \frac{C^2}{m^{2\alpha}} \left(\frac{\pi^{2\alpha+1}}{4^\alpha - 1} \right)$$

$$\sup \left\{ \sum_{i=1}^{\infty} i^{2p+2\beta}(a_{2i}^2(f) + a_{2i-1}^2(f)), f \in F_{s,C} \right\} < \frac{\pi^{2\alpha+1} C^2}{4^{\alpha-\beta} - 1} \quad \text{for all } 0 < \beta < \alpha.$$

This lemma is proved in Bary (1964 vol. 1, pp. 215–216).

Lemma 3.

$$\sup \{ \|f^{(l)}\|_\infty, l = 0, \dots, p, f \in F_{s,C} \} < \infty.$$

See Bickel and Ritov (1988) for the proof.

In order to prove Theorem 1, we use the classical decomposition:

$$E(\hat{T}_\psi^J - T_\psi^J)^2 = (\text{Bias}(\hat{T}_\psi^J))^2 + \text{var}(\hat{T}_\psi^J).$$

In the proofs, we shall write m for $m(n)$ and assume that m is even.

We first examine the bias. We noticed above that

$$|\text{Bias}(\hat{T}_\psi^J)| = \left| \int_{-\pi}^{\pi} ((S_m f)^{(j)} - f^{(j)})((S_m f)^{(j')} - f^{(j')})\psi \right|$$

$$\leq \|\psi\|_\infty \left(\int_{-\pi}^{\pi} ((S_m f)^{(j)} - f^{(j)})^2 \right)^{1/2} \left(\int_{-\pi}^{\pi} ((S_m f)^{(j')} - f^{(j')})^2 \right)^{1/2}.$$

We claim that, since m is even

$$\sum_{i=0}^m p_i(x)p_i^{(j)}(y) = (-1)^j \sum_{i=0}^m p_i^{(j)}(x)p_i(y). \tag{4.1}$$

This equality implies that $S_m(g^{(l)}) = (S_m g)^{(l)}$, for all $g \in C^l(S)$. Hence

$$|\text{Bias}(\hat{T}_\psi^J)|^2 \leq \|\psi\|_\infty^2 \sum_{i=m+1}^\infty a_i^2(f^{(j)}) \sum_{i=m+1}^\infty a_i^2(f^{(j')}) = \|\psi\|_\infty^2 \sum_{i=m+1}^\infty q_i(j)a_i^2 \sum_{i=m+1}^\infty q_i(j')a_i^2,$$

where $q_{2i}(j) = q_{2i-1}(j) = i^{2j}$ for all $i > 0$. Since $0 \leq q_i(j) \leq i^{2j}$, we obtain

$$|\text{Bias}(\hat{T}_\psi^J)|^2 \leq \|\psi\|_\infty^2 \sum_{i=m+1}^\infty i^{2p} a_i^2 i^{2j-2p} \sum_{i=m+1}^\infty i^{2p} a_i^2 i^{2j'-2p}.$$

From Lemma 2 and the fact that $j - p \leq 0$ and $j' - p \leq 0$, we derive

$$|\text{Bias}(\hat{T}_\psi^J)|^2 = O(\|\psi\|_\infty^2 (m^{2j+2j'-4s})).$$

Here, and in the remainder of the proof, the O s are independent of ψ and its derivatives, and of $f \in F_{s,C}$. Since $j \leq j' \leq s$, the bias is always a decreasing function of m .

Turning now to the variance, we have that $\hat{T}_\psi^J = \hat{T}_{\psi,1}^J + \hat{T}_{\psi,2}^J$ as defined by (2.1) and (2.2). We shall first prove (2.3) and (2.5), so we just have to determine the order of magnitude of $\text{var}(\hat{T}_\psi^J)$. We notice that \hat{T}_ψ^J can be written in the following form:

$$\hat{T}_\psi^J = \frac{1}{n(n-1)} \sum_{l_1 \neq l_2=1}^n h_{jj'}(X_{l_1}, X_{l_2})$$

where $h_{jj'}$ is symmetric. More precisely,

$$\hat{T}_{\psi,1}^J = \frac{1}{2n(n-1)} \sum_{l_1 \neq l_2=1}^n (H_1^{jj'} + H_1^{j'j})(X_{l_1}, X_{l_2}) + (H_1^{jj'} + H_1^{j'j})(X_{l_2}, X_{l_1})$$

where $H_1^{jj'}(x, y) = \sum_{i=0}^m p_i(x)(-1)^j (p_i^{(j')}\psi)^{(j)}(y)$. Moreover,

$$\hat{T}_{\psi,2}^J = -\frac{1}{2n(n-1)} \sum_{l_1 \neq l_2=1}^n (H_2^{jj'} + H_2^{j'j})(X_{l_1}, X_{l_2})$$

with $H_2^{jj'}(x, y) = \sum_{i,i'=0}^m p_i(x)p_{i'}(y) \int_{-\pi}^\pi p_i^{(j)} p_{i'}^{(j')} \psi$. We notice that $H_2^{jj'}(x, y) = H_2^{j'j}(y, x)$. $h_{jj'}$ is defined by

$$h_{jj'}(x, y) = \frac{1}{2}[H_1^{jj'} + H_1^{j'j}](x, y) + (H_1^{jj'} + H_1^{j'j})(y, x) - (H_2^{jj'} + H_2^{j'j})(x, y)].$$

\hat{T}_ψ^J is thus a U -statistic; it follows from Hoeffding's results concerning the computation of the variance for U -statistics (see Hoeffding 1948; or Serfling 1980, p. 183) that

$$\text{var}(\hat{T}_\psi^J) = \frac{4(n-2)}{n(n-1)} \xi_1 + \frac{2}{n(n-1)} \xi_2,$$

where $\xi_1 = \text{var}(\bar{h}_{jj'}(X_1))$, $\xi_2 = \text{var}(h_{jj'}(X_1, X_2))$, and $\bar{h}_{jj'}(x) = E(h_{jj'}(x, X_2))$.

We will see that in case (i), when $s > j + j' + \frac{1}{4}$, the main contribution to the variance is made by the term $4\xi_1/n$, which can be viewed as a linear term. It is of order λ_1^2/n . In case (ii), the term $(2/n(n-1))\xi_2$ is dominant, it is an increasing function of m , so the optimal value of m makes the trade-off between this term and the square of the bias. In this case, the rate $1/n$ is not achieved.

We first compute $\bar{h}_{jj'}(X_1)$.

$$\begin{aligned} E(H_1^{jj'}(x, X_2)) &= E\left[\sum_{i=0}^m p_i(x)(-1)^j(p_i^{(j')}\psi)^{(j)}(X_2)\right] \\ &= \sum_{i=0}^m p_i(x) \int p_i^{(j')}\psi f^{(j)} \quad \text{integrating by parts.} \\ &= \sum_{i=0}^m (-1)^{j'} p_i^{(j')}(x) \int \psi f^{(j)} p_i \quad \text{by (4.1).} \\ &= (-1)^{j'} S_m^{(j')}[\psi f^{(j)}](x). \end{aligned}$$

By similar computations, we get

$$\begin{aligned} E(H_1^{jj'}(X_2, x)) &= (-1)^j (S_m f^{(j')}\psi)^{(j)}(x) \\ E(H_2^{jj'}(x, X_2)) &= (-1)^j S_m[(S_m f)^{(j')}\psi]^{(j)}(x). \end{aligned}$$

Denote

$$Y_1^{jj'} = (-1)^{j'} S_m^{(j')} [f^{(j)}\psi](X_1) \tag{4.2}$$

$$Y_2^{jj'} = (-1)^j (S_m f^{(j')}\psi)^{(j)}(X_1) \tag{4.3}$$

$$Y_3^{jj'} = -(-1)^j S_m[(S_m f)^{(j')}\psi]^{(j)}(X_1) \tag{4.4}$$

$$\bar{h}_{jj'}(X_1) = \frac{1}{2} \sum_{l=1}^3 (Y_l^{jj'} + Y_l^{j'j}).$$

It follows that

$$\xi_1 = \text{var}[\bar{h}_{jj'}(X_1)] \leq E[(\bar{h}_{jj'}(X_1))^2] \leq \frac{3}{2} \sum_{l=1}^3 E[(Y_l^{jj'})^2] + E[(Y_l^{j'j})^2].$$

We will now bound from above the quantities $E[(Y_l^{jj'})^2]$ for $l \in \{1, 2, 3\}$.

For $E[(Y_1^{jj'})^2]$ we can write

$$E[(Y_1^{jj'})^2] = \int [S_m^{(j')}[\psi f^{(j)}]]^2 f.$$

If $s > j + j' + \frac{1}{4}$ (case (i)) then $p \geq j + j'$ and f is $j + j'$ times continuously differentiable, so we can write, using the fact that $S_m(g^{(l)}) = (S_m g)^{(l)}$,

$$E[(Y_1^{jj'})^2] = \int S_m^2[(\psi f^{(j)})^{(j')}] f$$

$$\begin{aligned} &\leq \|f\|_\infty \|(\psi f^{(j)})^{(j')}\|_2^2 \quad \text{since } S_m \text{ is a projection} \\ &\leq \|f\|_\infty \left\| \sum_{l=0}^{j'} \binom{j'}{l} \psi^{(l)} f^{(j+j'-l)} \right\|_2^2. \end{aligned}$$

Since by Lemma 3, $\sup \{\|f^{(l)}\|_\infty, l = 0, \dots, p, f \in F_{s,C}\} < \infty$, we obtain

$$E[(Y_1^{j'})^2] \leq C_1 \lambda_2^2,$$

where C_1 is independent of ψ and $f \in F_{s,C}$.

If $s \leq j + j' + \frac{1}{4}$ (case (ii)) then $p \leq j + j'$. If $p = j + j'$, as above, we can write

$$E[(Y_1^{j'})^2] = \int S_m^2 [(\psi f^{(j)})^{(j')}] f = O\left(\sup_{l=0, \dots, j'} \|\psi^{(l)}\|_\infty^2\right).$$

Moreover, by definition of λ_2

$$\sup_{l=0, \dots, j'} \|\psi^{(l)}\|_\infty^2 \leq \lambda_2^2 m^{2(j+j'-s+\frac{1}{4})}.$$

Let us now consider the case where $j + j' > p$. In this case we also have $j + j' \geq s$. Using the fact that $S_m(g^{(l)}) = (S_m g)^{(l)}$, where g is l times continuously differentiable, we obtain:

$$\begin{aligned} E[(Y_1^{j'})^2] &= \int (S_m^{(j+j'-p)} [(\psi f^{(j)})^{(p-j)}])^2 f \\ &\leq \|f\|_\infty \sum_{i=0}^m i^{2j+2j'-2p} b_i^2 \text{ by orthogonality of the } p_i \text{ and their derivatives,} \end{aligned}$$

where b_i denotes the i th Fourier coefficient of the function $(\psi f^{(j)})^{(p-j)}$. We will evaluate the α -norm of this function in order to apply Lemma 2.

$$(\psi f^{(j)})^{(p-j)} = \sum_{l=0}^{p-j} \binom{p-j}{l} \psi^{(l)} f^{(p-l)}.$$

Using the fact that for any functions g, h , $\|gh\|_\alpha \leq \|g\|_\infty \|h\|_\alpha + \|h\|_\infty \|g\|_\alpha$, we obtain

$$\begin{aligned} \|(\psi f^{(j)})^{(p-j)}\|_\alpha &\leq \sum_{l=0}^{p-j} \binom{p-j}{l} (\|\psi^{(l)}\|_\infty \|f^{(p-l)}\|_\alpha + \|\psi^{(l)}\|_\alpha \|f^{(p-l)}\|_\infty) \\ &= O\left(\sup_{l=0, \dots, p-j} (\|\psi^{(l)}\|_\infty + \|\psi^{(l)}\|_\alpha)\right) \quad \text{by Lemma 3.} \end{aligned}$$

By definition of λ_2 ,

$$\sup_{l=0, \dots, p-j} (\|\psi^{(l)}\|_\infty + \|\psi^{(l)}\|_\alpha) \leq \lambda_2 m^{1/6}.$$

From Lemma 2 and since $j + j' \geq s > p + \beta$ for all $0 < \beta < \alpha$, we derive

$$\sum_{i=0}^m i^{2j+2j'-2p} b_i^2 = \sum_{i=0}^m i^{2j+2j'-2p-2\beta} i^{2\beta} b_i^2$$

$$\begin{aligned}
 &= O(m^{2j+2j'-2p-2\beta} \|(\psi f^{(j)})^{(p-j)}\|_\alpha^2) \\
 &= O(\lambda_2^2 m^{1/3} m^{2j+2j'-2s+\frac{1}{2}} m^{-\frac{1}{2}+2(\alpha-\beta)}).
 \end{aligned}$$

Since β is arbitrarily close to α , $m^{1/3} m^{-1/2+2(\alpha-\beta)} \rightarrow 0$. Collecting the above evaluations, we obtain, for $s \leq j + j' + \frac{1}{4}$,

$$E[(Y_1^{jj'})^2] \leq C_2 \lambda_2^2 (m^{2j+2j'-2s+\frac{1}{2}}).$$

Turning now to $E[(Y_2^{jj'})^2]$, we have

$$\begin{aligned}
 E[(Y_2^{jj'})^2] &= \int [(S_m f^{(j')}\psi)^{(j)}]^2 f \leq \|f\|_\infty \| (S_m f^{(j')}\psi)^{(j)} \|_2^2 \\
 &= O[\sup_{l=0,\dots,j'} \|\psi^{(l)}\|_\infty^2 \| (S_m f)^{(j+j'-1)} \|_2^2].
 \end{aligned}$$

If $s > j + j' + \frac{1}{4}$ (case (i)) then $p \geq j + j'$ and since f is p times continuously differentiable, by Lemma 3 we get $E[(Y_2^{jj'})^2] \leq C_1 \lambda_1^2$.

If $s \leq j + j' + \frac{1}{4}$ (case (ii)) we shall consider two subcases. If $j + j' - l \leq p$ then $\|S_m(f^{(j+j'-l)})\|_2^2 \leq 2\pi \sup_{l=0,\dots,p} \|f^{(l)}\|_\infty^2$, it follows that

$$E[(Y_2^{jj'})^2] = O(\sup_{l=0,\dots,j'} \|\psi^{(l)}\|_\infty^2) = O(\lambda_2^2 m^{2j+2j'-2s+\frac{1}{2}}).$$

If however, $j + j' - l \geq p + 1 \geq s$, then

$$\begin{aligned}
 \|\psi^{(l)}\|_\infty^2 \| (S_m f)^{(j+j'-l)} \|_2^2 &\leq \|\psi^{(l)}\|_\infty^2 \sum_{i=0}^m i^{2j+2j'-2l} a_i^2 \\
 &\leq \|\psi^{(l)}\|_\infty^2 \sum_{i=0}^m i^{2j+2j'-2l-2p-2\beta} i^{2p+2\beta} a_i^2 \quad \text{for all } 0 < \beta < \alpha.
 \end{aligned}$$

Since $j + j' - l \geq s > p + \beta$, for all $0 < \beta < \alpha$, and by Lemma 2, we obtain

$$\begin{aligned}
 \|\psi^{(l)}\|_\infty^2 \| (S_m f)^{(j+j'-l)} \|_2^2 &\leq \|\psi^{(l)}\|_\infty^2 m^{2j+2j'-2l-2p-2\beta} \sum_{i=0}^\infty i^{2p+2\beta} a_i^2 \\
 &= O[(\|\psi^{(l)}\|_\infty m^{-l})^2 m^{2j+2j'-2p-2\beta}] \\
 &= O[\lambda_2^2 m^{2j+2j'-2p-2\beta}].
 \end{aligned}$$

The above equality leads to $E[(Y_2^{jj'})^2] \leq C_2 \lambda_2^2 (m^{2j+2j'-2s+\frac{1}{2}})$.

All that is left is $E[(Y_3^{jj'})^2]$,

$$E[(Y_3^{jj'})^2] = \int [S_m(S_m f^{(j')}\psi)^{(j)}]^2 f \leq \|f\|_\infty \| (S_m f^{(j')}\psi)^{(j)} \|_2^2,$$

and the computation of this term has just been done above.

Of course, the above results hold for $E[(Y_l^{j'j})^2]$, $l = 1, 2, 3$. It follows that

$$\xi_1 \leq \begin{cases} C_1 \lambda_1^2 & s > j + j' + \frac{1}{4} \\ C_2 \lambda_2^2 (m^{2j+2j'-2s+\frac{1}{2}}) & s \leq j + j' + \frac{1}{4}. \end{cases} \quad (4.5)$$

We now bound ξ_2 :

$$\begin{aligned} \xi_2 &= \text{var}(h_{jj'}(X_1, X_2)) \leq E(h_{jj'}^2(X_1, X_2)) \\ &\leq 3E[(H_1^{jj'}(X_1, X_2))^2 + (H_1^{j'j}(X_1, X_2))^2 + (H_2^{jj'}(X_1, X_2))^2]. \end{aligned}$$

The inner bracketed terms will be evaluated separately, beginning with $E[(H_1^{jj'}(X_1, X_2))^2]$:

$$\begin{aligned} E[(H_1^{jj'}(X_1, X_2))^2] &= \iint \left[\sum_{i=0}^m p_i(x)(p_i^{(j)}\psi)^{(j)}(y) \right]^2 f(x)f(y) \, dx \, dy \\ &\leq \|f\|_\infty^2 \iint \sum_{i,i'=0}^m p_i p_{i'}(x)(p_i^{(j)}\psi)^{(j)}(y)(p_{i'}^{(j)}\psi)^{(j)}(y) \, dx \, dy \\ &\leq \|f\|_\infty^2 \sum_{i=0}^m \int [(p_i^{(j)}\psi)^{(j)}]^2 \quad \text{for orthonormality} \\ &\leq \|f\|_\infty^2 \sum_{i=0}^m \int \sum_{l,l'=0}^j \binom{j}{l} \binom{j}{l'} p_i^{(j+j'-l)} p_i^{(j+j'-l')} \psi^{(l)} \psi^{(l')}. \end{aligned}$$

Since $\|p_i^{(l)}\|_\infty \leq i^l$ our expression has order

$$O \left[\sum_{l,l'=0}^j \sum_{i=0}^m i^{2j+2j'-l-l'} \| \psi^{(l)} \|_\infty \| \psi^{(l')} \|_\infty \right] = O \left[m^{1+2j+2j'} \left(\sum_{l=0}^j m^{-l} \| \psi^{(l)} \|_\infty \right)^2 \right].$$

Turning to $E[(H_2^{jj'}(X_1, X_2))^2]$, we have that

$$\begin{aligned} E[(H_2^{jj'}(X_1, X_2))^2] &= \sum_{i,i',i_1,i_1'=0}^m \int p_i p_{i_1} f \int p_{i'} p_{i_1'} f \int p_i^{(j)} p_{i_1}^{(j')} \psi \int p_{i_1}^{(j)} p_{i_1'}^{(j')} \psi \\ &= \iint \left(\sum_{i,i'=0}^m \left(\int p_i^{(j)} p_{i'}^{(j')} \psi \right) p_i(x) p_{i'}(y) \right)^2 f(x) f(y) \, dx \, dy \\ &\leq \|f\|_\infty^2 \sum_{i,i'=0}^m \int p_i^{(j)} p_{i'}^{(j')} \psi \int p_{i_1}^{(j)} p_{i_1'}^{(j')} \psi \quad \text{by orthonormality} \\ &\leq \|f\|_\infty^2 \iint \left(\sum_{i=0}^m p_i^{(j)}(z) p_i^{(j)}(t) \right) \left(\sum_{i'=0}^m p_{i'}^{(j')}(z) p_{i'}^{(j')}(t) \right) \psi(z) \psi(t) \, dt \, dz. \end{aligned}$$

By the Cauchy–Schwarz inequality, this expression is bounded by

$$\|f\|_\infty^2 \|\psi\|_\infty^2 \left[\iint \left(\sum_{i=0}^m p_i^{(j)}(z) p_i^{(j)}(t) \right)^2 dt dz \iint \left(\sum_{i'=0}^m p_{i'}^{(j)}(z) p_{i'}^{(j)}(t) \right)^2 dt dz \right]^{1/2}.$$

Moreover, by orthogonality of the $p_i^{(j)}$,

$$\iint \left(\sum_{i=0}^m p_i^{(j)}(z) p_i^{(j)}(t) \right)^2 dt dz \leq \sum_{i=0}^m i^{4j} \leq m^{1+4j}.$$

It follows that $E[(H_2^{jj'}(X_1, X_2))^2] \leq \|f\|_\infty^2 \|\psi\|_\infty^2 m^{1+2j+2j'}$. Hence we obtain

$$\xi_2 \leq \begin{cases} C_1 \lambda_1^2 m^{1+2j+2j'} & s > j + j' + \frac{1}{4} \\ C_2 \lambda_2^2 m^{1+2j+2j'} & s \leq j + j' + \frac{1}{4}. \end{cases} \tag{4.7}$$

We recall that

$$\text{var}(\hat{T}_\psi^J) = \frac{4(n-2)}{n(n-1)} \xi_1 + \frac{2}{n(n-1)} \xi_2.$$

From (4.5) and (4.7) we obtain, possibly enlarging C_1 , that

$$\text{if } s > j + j' + \frac{1}{4} \text{ then } \text{var}(\hat{T}_\psi^J) \leq C_1 \lambda_1^2 \left(\frac{1}{n} + \frac{m^{1+2j+2j'}}{n^2} \right).$$

From (4.6) and (4.8) we obtain, possibly enlarging C_2 , that

$$\text{if } s \leq j + j' + \frac{1}{4} \text{ then } \text{var}(\hat{T}_\psi^J) \leq C_2 \lambda_2^2 \left(\frac{m^{2j+2j'-2s+\frac{1}{2}}}{n} + \frac{m^{1+2j+2j'}}{n^2} \right).$$

We recall that $\text{Bias}^2(\hat{T}_\psi^J) = O(\|\psi\|_\infty^2 m^{2j+2j'-4s})$ and that $m \approx n^{2/(1+4s)}$, which is the value of m minimizing the sum $\text{Bias}^2(\hat{T}_\psi^J) + \text{var}(\hat{T}_\psi^J)$. Hence

$$E(\hat{T}_\psi^J - T_\psi^J)^2 \leq \begin{cases} C_1 \frac{\lambda_1^2}{n} & s > j + j' + \frac{1}{4} \\ C_2 \lambda_2^2 n^{(-8s+4j+4j')/(1+4s)} & s \leq j + j' + \frac{1}{4}. \end{cases}$$

This completes the proof of (2.3) and (2.5). We shall now prove (2.4).

We first notice that when $s > \sup(j + j' + \frac{1}{4}, j_1 + j'_1 + \frac{1}{4})$ then $j + j' \leq p$ and $j_1 + j'_1 \leq p$. We recall that

$$\hat{T}_\psi^J = \frac{1}{n(n-1)} \sum_{l_1 \neq l_2=1}^n h_{jj'\psi}(X_{l_1}, X_{l_2}),$$

where $h_{jj'\psi}$ is symmetric. We now write $h_{jj'\psi}$ instead of $h_{jj'}$ since \hat{T}_ψ^J is an estimator of $\int f^{(j)} f^{(j')}\psi$ while $\hat{T}_\xi^{J_1}$ estimates $\int f^{(j_1)} f^{(j'_1)} \xi$. Since

$$E(\hat{T}_\psi^J - T_\psi^J)(\hat{T}_\xi^{J_1} - T_\xi^{J_1}) = \text{Bias}(\hat{T}_\psi^J) \text{Bias}(\hat{T}_\xi^{J_1}) + \text{cov}(\hat{T}_\psi^J, \hat{T}_\xi^{J_1})$$

we shall use the following lemma:

Lemma 4. Let T and U be U -statistics respectively defined by

$$T = \frac{1}{n(n-1)} \sum_{j \neq k=1}^n a(X_j, X_k), \quad U = \frac{1}{n(n-1)} \sum_{j \neq k=1}^n b(X_j, X_k)$$

where a and b are symmetric. Then

$$\text{cov}(T, U) = \frac{4(n-2)}{n(n-1)} \xi_3 + \frac{2}{n(n-1)} \xi_4$$

where $\xi_3 = \text{cov}[\bar{a}(X_1), \bar{b}(X_1)]$, $\xi_4 = \text{cov}[a(X_1, X_2), b(X_1, X_2)]$, $\bar{a}(x) = E[a(x, X_2)]$ and $\bar{b}(x) = E[b(x, X_2)]$.

This equality is a generalization for the covariance of U -statistics of Hoeffding's formula for the variance. The proof is immediate and will be omitted (for more details, see Laurent 1993). It follows from Lemma 4 that

$$nE(\hat{T}_\psi^J - T_\psi^J)(\hat{T}_\xi^{J_1} - T_\xi^{J_1}) = n\text{Bias}(\hat{T}_\psi^J)\text{Bias}(\hat{T}_\xi^{J_1}) + \frac{4(n-2)}{n-1} \xi_3 + \frac{2}{n-1} \xi_4,$$

where $\xi_3 = \text{cov}[\bar{h}_{jj'\psi}(X_1), \bar{h}_{j_1, j_1 \xi}(X_1)]$, $\xi_4 = \text{cov}[h_{jj'\psi}(X_1, X_2), h_{j_1, j_1 \xi}(X_1, X_2)]$ and $\bar{h}_{jj'\psi}(x) = E[h_{jj'\psi}(x, X_2)]$. Using the results obtained in the first part of the proof, we obtain

$$n|\text{Bias}(\hat{T}_\psi^J)\text{Bias}(\hat{T}_\xi^{J_1})| \leq C'_1 \|\psi\|_\infty \|\xi\|_\infty n m^{j+j'+j_1+j_1-4s} \leq C'_1 \mu_1^2 m^{\frac{1}{2}+j+j'+j_1+j_1-2s}$$

since $n \approx m^{\frac{1}{2}+2s}$. We want to prove that $(2/(n-1))\xi_4$ is bounded by a similar quantity.

$$\xi_4 \leq (E(h_{jj'\psi}^2(X_1, X_2)))^{1/2} (E(h_{j_1, j_1 \xi}^2(X_1, X_2)))^{1/2}.$$

From (4.7), which is actually an upper bound for $E(h_{jj'\psi}^2(X_1, X_2))$, we derive, possibly enlarging C'_1 .

$$\frac{2}{n-1} \xi_4 \leq C'_1 \frac{\mu_1^2}{n} m^{1+j+j'+j_1+j_1} = C'_1 \mu_1^2 m^{\frac{1}{2}+j+j'+j_1+j_1-2s}.$$

It follows that the asymptotic covariance $\Lambda_{J_1}(f, \psi)$ can only come from the term $(4(n-2)/(n-1))\xi_3$. Since, $|\xi_3| \leq C'_1 \mu_1^2$ (this result follows from the computation of ξ_1), to show (2.4), we just have to prove that $|4\xi_3 - \Lambda_{J_1}(f, \psi, \xi)| \leq C'_1 \mu_1^2 m^{-\alpha}$.

We recall that

$$\bar{h}_{jj'\psi}(X_1) = \frac{1}{2} \sum_{l=1}^3 (Y_l^{jj'} + Y_l^{j'j})$$

where $Y_l^{jj'}$, $l = 1, 2, 3$, are defined by (4.2), (4.3) and (4.4). We have to keep in mind that $Y_l^{jj'}$ depends on ψ while $Y_l^{j_1 j_1 \xi}$ depends on ξ .

$$4 \text{cov}(\bar{h}_{jj'\psi}(X_1), \bar{h}_{j_1, j_1 \xi}(X_1)) = \sum_{l, l'=1}^3 \text{cov}((Y_l^{jj'} + Y_l^{j'j}), (Y_{l'}^{j_1 j_1 \xi} + Y_{l'}^{j_1 j_1 \xi})).$$

We claim that for all $l, l' \in \{1, 2, 3\}^2$

$$|\text{cov}((Y_l^{jj'} + Y_l^{j'j}), (Y_{l'}^{jj} + Y_{l'}^{jj})) - \epsilon_{ll'} \Lambda_{JJ_1}(f, \psi, \xi)| \leq C'_1 \mu_1^2 m^\alpha \tag{4.9}$$

where

$$\epsilon_{ll'} = -1 \text{ if } (l, l') = (1, 3), (2, 3), (3, 1), (3, 2)$$

$$\epsilon_{ll'} = 1 \text{ if } (l, l') = (1, 1), (2, 2), (3, 3), (1, 2), (2, 1).$$

Of course, this concludes the proof of (2.4). Since, the proof employs standard arguments, we content ourself with sketching it (a complete proof for the case $l = 1$, and $l' = 3$, is given in Laurent 1993). To prove (4.9) in the case $l = 1$, and $l' = 3$ for example, we have bound from above by $C'_1 \mu_1^2 m^{-\alpha}$ quantities of the type

$$\left| \int S_m((f^{(j')} \psi)^{(j)}) S_m((S_m f^{(jj)} \xi)^{(j_1)}) f - \int (f^{(j')} \psi)^{(j)} (f^{(jj)} \xi)^{(j_1)} f \right|.$$

Introducing auxiliary terms, this quantity is bounded by

$$C'_1 \mu_1 [\|S_m((S_m f^{(jj)} \xi)^{(j_1)}) - S_m((f^{(jj)} \xi)^{(j_1)})\|_2 + \|S_m((f^{(jj)} \xi)^{(j_1)}) - (f^{(jj)} \xi)^{(j_1)}\|_2].$$

We bound this term from above by $C'_1 \mu_1^2 m^{-\alpha}$ using the fact that for any function g , $\|S_m(g) - g\|_2 = O(\|g\|_\alpha m^{-\alpha})$ by Lemma 2. This concludes the proof of Theorem 1.

4.2. PROOF OF THEOREM 2

Before proving properties (i) and (ii), we first control the remainder term Γ_n :

$$|\Gamma_n| \leq \frac{1}{6} \|\phi^{(3)}\|_\infty \sum_{j, j', j''=0}^k \int |f^{(j)} - \hat{f}^{(j)}| |f^{(j')} - \hat{f}^{(j')}| |f^{(j'')} - \hat{f}^{(j'')}|(x) dx.$$

Hence,

$$E(\Gamma_n^2) = O\left(\sup_{jj'j''} E\left[\left(\int |\hat{f}^{(j)} - f^{(j)}| |\hat{f}^{(j')} - f^{(j')}| |\hat{f}^{(j'')} - f^{(j'')}|\right)^2\right]\right).$$

Using the generalized Hölder inequality and Cauchy–Schwarz, we get

$$E(\Gamma_n^2) = O\left(\sup_{j=0, \dots, k} E(\|\hat{f}^{(j)} - f^{(j)}\|_6^6)\right).$$

Using Condition A2(c), which ensures that the quantities $E(\|\hat{f}^{(l)} - f^{(l)}\|_q^q)$ achieve the optimal rate of convergence, we see that $E(\Gamma_n^2) = O(n_1^{-6(s-k)/(1+2s)})$ where $n_1 \approx n/(\log(n))$. Thus:

$$\text{if } s > 2k + \frac{1}{4} \text{ then } E(\Gamma_n^2) = o\left(\frac{1}{n}\right); \tag{4.10}$$

$$\text{if } \frac{1}{4} < s \leq 2k + \frac{1}{4} \text{ then } E(\Gamma_n^2) = o(n^{-8(s-k)/(1+4s)}). \tag{4.11}$$

This proves that in both cases the remainder term is negligible. We are now in a position to prove properties (i) and (ii).

Taking case (ii) first, we first look at the order of magnitude of $E(\hat{T}_n - T(f))^2$ in order to obtain the rate of convergence in the nonparametric case, that is when $s \leq 2k + \frac{1}{4}$. $\hat{T}_n - T(f)$ is composed of a sum of a linear term, a quadratic term, and the remainder term Γ_n :

$$\begin{aligned} \hat{T}_n - T(f) &= \sum_{j=0}^k \frac{1}{n_2} \sum_{l=1}^{n_2} \left((-1)^j (H_j(\hat{f}))^{(j)}(X_l) - \int_{-\pi}^{\pi} H_j(\hat{f}) f^{(j)} \right) \\ &+ \sum_{j,j'=0}^k (\hat{T}^J(K_J(\hat{f})) - T^J(K_J(\hat{f}))) + \Gamma_n. \end{aligned}$$

The expectation of the square of the linear term is

$$O \left[\frac{1}{n_2} \sup_{j \in \{0, \dots, k\}} E(\|H_j(\hat{f})\|_{\infty}^2) \right].$$

As to the evaluation of the quadratic term, we shall apply the results of Section 2. In Section 2, we studied the properties of estimators of $\int_{-\pi}^{\pi} f^{(j)} f^{(j')} \psi$. Here, ψ is the random function $K_J(\hat{f})$ and $T^J(K_J(\hat{f}))$ is estimated independently of \hat{f} . Therefore, working conditionally on \hat{f} , we can apply the results of Section 2. $\hat{T}^J(K_J(\hat{f}))$ will now be denoted by \hat{T}^J for short.

We have to determine the order of magnitude of $E(\|(K_J(\hat{f}))^{(l)}\|_{\infty}^2)$ and $E(\|K_J(\hat{f})\|_{\alpha}^2)$ for $l \in \{1, \dots, k\}$ in order to apply Theorem 1. We recall that $K_J(\hat{f}) = \frac{1}{2} \phi_{j'j}''(\hat{f}, \dots, \hat{f}^{(k)}, \cdot)$; $(K_J(\hat{f}))^{(l)}$ is the derivative of order l of the function $x \mapsto \frac{1}{2} \phi_{j'j}''(\hat{f}(x), \dots, \hat{f}^{(k)}(x), x)$. It involves a sum of term of type $\sum_{i=1}^l \phi_{j'j''j_1, \dots, j_i}^{(2+i)}(\hat{f}, \dots, \hat{f}^{(k)}, \cdot) \hat{f}^{(j_1+n_1)} \times \dots \times \hat{f}^{(j_i+n_i)} \mathbf{1}_{n_1+\dots+n_i=l}$ where the indices j_1, \dots, j_i belong to $\{0, 1, \dots, k\}$. Since $\phi \in C^{k+3}(K_c)$, $\|\phi^{(2+i)}\|_{\infty} < \infty$ for all $i \in \{1, \dots, k\}$ and we have to determine the order of magnitude of $E(\|\hat{f}^{(j_1+n_1)} \times \dots \times \hat{f}^{(j_i+n_i)}\|_{\infty}^2)$ and $E(\|\hat{f}^{(j_1+n_1)} \times \dots \times \hat{f}^{(j_i+n_i)}\|_{\alpha}^2)$, where $n_1 + \dots + n_i = l$. Put $h = n_1^{-1/(1+2s)}$ for short. By Condition A2(d) of Hölder's inequality, it follows that

$$E(\|\hat{f}^{(j_1+n_1)} \times \dots \times \hat{f}^{(j_i+n_i)}\|_{\infty}^2) = O[(h^{2(s'-j_1-n_1)} + 1) \times \dots \times (h^{2(s'-j_i-n_i)} + 1)]$$

for all $s' < s$.

Let $k < s' < s$, since $\{\sum_{v=1}^i n_v = l$ and $j_v \leq k$, we obtain

$$E(\|\hat{f}^{(j_1+n_1)} \times \dots \times \hat{f}^{(j_i+n_i)}\|_{\infty}^2) = O(h^{2(s'-k-l)} + 1). \tag{4.12}$$

As to the evaluation of the α -norm, $E(\|\hat{f}^{(j_1+n_1)} \times \dots \times \hat{f}^{(j_i+n_i)}\|_{\alpha}^2)$, we use the equality

$$\begin{aligned} &(\hat{f}^{(j_1+n_1)} \times \dots \times \hat{f}^{(j_i+n_i)})(x) - (\hat{f}^{(j_1+n_1)} \times \dots \times \hat{f}^{(j_i+n_i)})(y) \\ &= \sum_{i_0=0}^{i-1} \left(\left[\prod_{i_1=1}^{i-i_0} \hat{f}^{(j_{i_1+n_{i_1}})}(x) \prod_{i_1=i-i_0+1}^i \hat{f}^{(j_{i_1+n_{i_1}})}(y) \right] \right. \\ &\quad \left. - \left[\prod_{i_1=1}^{i-i_0-1} \hat{f}^{(j_{i_1+n_{i_1}})}(x) \prod_{i_1=i-i_0}^i \hat{f}^{(j_{i_1+n_{i_1}})}(y) \right] \right). \end{aligned}$$

This leads to

$$\|\hat{f}^{(j_1+n_1)} \times \dots \times \hat{f}^{(j_i+n_i)}\|_\alpha \leq \sum_{i_0=1}^i \left(\prod_{i_1 \neq i_0} \|\hat{f}^{(j_{i_1}+n_{i_1})}\|_\infty \right) \|\hat{f}^{(j_{i_0}+n_{i_0})}\|_\alpha.$$

It follows from (4.12) that

$$\mathbb{E} \left(\prod_{i_1 \neq i_0} \|\hat{f}^{(j_{i_1}+n_{i_1})}\|_\infty^2 \mathbf{1}_{\{\sum_{i_1 \neq i_0} n_{i_1} = l - n_{i_0}\}} \right) = O(h^{2(s'-k-l+n_{i_0})} + 1).$$

Moreover, by Condition A2(e) we obtain

$$\begin{aligned} \mathbb{E}(\|\hat{f}^{(j_1+n_1)} \times \dots \times \hat{f}^{(j_i+n_i)}\|_\alpha^2) &= O \left[\sum_{i_0=1}^i (h^{2(s'-k-l+n_{i_0})} + 1)(h^{2(s'-k-\alpha-n_{i_0})} + 1) \right] \\ &= O(h^{2(s'-k-\alpha-l)} + 1). \end{aligned}$$

So,

$$\begin{aligned} \mathbb{E}(\|(K_J(\hat{f}))^{(l)}\|_\infty^2) &= O(1 + h^{2(s'-k-l)}), \\ \mathbb{E}(\|(K_J(\hat{f}))^{(l)}\|_\alpha^2) &= O(1 + h^{2(s'-k-l-\alpha)}). \end{aligned}$$

By similar arguments, we can see that for all $j \in \{0, \dots, k\}$,

$$\mathbb{E}(\|H_j(\hat{f})\|_\infty^{(j)}) = O(h^{2s'-4k} + 1).$$

Hence, the expectation of the square of the linear term has order $O((h^{2s'-4k} + 1)/n_2) = O(n^{-8(s-k)/(1+4s)} + n^{-1})$, since s' is arbitrary close to s and $n/n_2 \rightarrow 1$.

We can now apply the results of Theorem 1 to obtain the order of magnitude of $\mathbb{E}(\hat{T}^J - T^J)^2$. Since $s \leq 2k + \frac{1}{4}$ we have to consider two cases corresponding to cases (i) and (ii) of Theorem 1. First, if $j + j' + \frac{1}{4} < s$, then

$$\sup_{l=0, \dots, j'} \mathbb{E}(\|(K_J(\hat{f}))^{(l)}\|_\infty^2) = O(h^{2(s'-k-j')}) = O(h^{2s'-4k}) \text{ since } j' \leq k.$$

Hence, by Theorem 1,

$$\mathbb{E}(\hat{T}^J - T^J)^2 \leq C_1 \frac{h^{2s'-4k}}{n_2} \leq C_1 n^{-8(s-k)/(1+4s)}.$$

Second, if $j + j' + \frac{1}{4} \geq s$ we have to evaluate the order of magnitude of $\mathbb{E}(\lambda_2^2)$ for all j, j' . We want to prove that $\mathbb{E}(\lambda_2^2) = O(h^{2j-2k})$. To do this, we shall prove that

$$\sup_{l=0, \dots, p-j} \mathbb{E} \left[\frac{\|(K_J(\hat{f}))^{(l)}\|_\infty^2 + \|(K_J(\hat{f}))^{(l)}\|_\alpha^2}{m^{1/3}} \right] = O(h^{2j-2k}); \quad (4.13)$$

$$\sup_{l=0, \dots, j'} \mathbb{E} \left[\frac{\|(K_J(\hat{f}))^{(l)}\|_\infty^2}{m^{(2j+2j'-2s+\frac{1}{2}) \wedge 2l}} \right] = O(h^{2j-2k}). \quad (4.14)$$

To prove (4.13), we observe that, for all $l = 0, \dots, p - j$,

$$\begin{aligned} \mathbb{E} \left[\frac{\| (K_J(\hat{f}))^{(l)} \|_\infty^2 + \| (K_J(\hat{f}))^{(l)} \|_\alpha^2}{m^{1/3}} \right] &= O \left(\frac{h^{2(s'-k-p+j-\alpha)}}{m^{1/3}} \right) \\ &= O(h^{2j-2k})(h^{2s'-2p-2\alpha} m^{-1/3}). \end{aligned}$$

Since s' is arbitrarily close to $s = p + \alpha$ we get $\lim_{m \rightarrow \infty} h^{2s'-2p-2\alpha} m^{-1/3} = 0$. Thus (4.13) is proved.

To prove (4.14) we first assume that $2j + 2j' - 2s + \frac{1}{2} \leq l$. Then we have to show that, for all $l \leq j'$,

$$\mathbb{E} \left(\frac{\| (K_J(\hat{f}))^{(l)} \|_\infty^2}{m^{(2j+2j'-2s+\frac{1}{2})}} \right) = O(h^{2j-2k}).$$

From the above evaluations of $\mathbb{E}(\| (K_J(\hat{f}))^{(l)} \|_\infty^2)$, we derive

$$\begin{aligned} \mathbb{E} \left(\frac{\| (K_J(\hat{f}))^{(l)} \|_\infty^2}{m^{(2j+2j'-2s+\frac{1}{2})}} \right) &= O \left(\frac{1 + h^{2s'-2k-2l}}{m^{(2j+2j'-2s+\frac{1}{2})}} \right) \\ &= O(h^{2j-2k})(h^{2k-2j} m^{-2j-2j'+2s-\frac{1}{2}} + h^{2s'-2j-2l} m^{-2j-2j'+2s-\frac{1}{2}}). \end{aligned}$$

Since $j \leq k$, $s \leq j + j' + \frac{1}{4}$, $h \rightarrow 0$ and $m \rightarrow +\infty$, we obtain $h^{2k-2j} m^{-2j-2j'+2s-\frac{1}{2}} \rightarrow 0$. Moreover, for all $l \leq j'$,

$$h^{2s'-2j-2l} m^{-2j-2j'+2s-\frac{1}{2}} \leq h^{2s'-2j-2j'} m^{-2j-2j'+2s-\frac{1}{2}} \leq (mh)^{2s-2j-2j'} h^{2s'-2s} m^{-1/2}.$$

Using the fact that $mh \rightarrow +\infty$ and that $s \leq j + j' + \frac{1}{4}$, the above expression is bounded by $(mh)^{1/2} h^{2s'-2s} m^{-1/2} = h^{\frac{1}{2}+2s'-2s}$ which converges towards zero since s' is arbitrarily close to s .

To conclude the proof of (4.14) we have to consider the case of $l > 2j + 2j' - 2s + \frac{1}{4}$, and show that $\mathbb{E}(\| (K_J(\hat{f}))^{(l)} \|_\infty^2 / m^{2l}) = O(h^{2j-2k})$.

$$\begin{aligned} \mathbb{E} \left(\frac{\| (K_J(\hat{f}))^{(l)} \|_\infty^2}{m^{2l}} \right) &= O((h^{2s'-2k-2l} + 1)m^{-2l}) \\ &= O(h^{2j-2k})(h^{2s'-2j}(mh)^{-2l} + h^{2k-2j} m^{-2l}) \\ &= O(h^{2j-2k}) \text{ since } mh \rightarrow +\infty \text{ and } h \rightarrow 0. \end{aligned}$$

Expression (4.14) is thus proved.

Finally, if $s \leq 2k + \frac{1}{4}$ by Theorem 1 we obtain

$$\mathbb{E} \left(\left(\sum_{j,j'=0}^k \hat{T}^J - T^J \right)^2 \right) = O \left(\sum_{j,j'=0}^k h^{2j-2k} n^{(-8s+4j+4j')/(1+4s)} + n^{-8(s-k)/(1+4s)} \right)$$

$$\begin{aligned}
 &= O\left(\sup_{j=0,\dots,k} (hn^{2/(1+4s)})^2 j h^{-2k} n^{(-8s+4k)/(1+4s)} + n^{-8(s-k)/(1+4s)}\right) \\
 &= O(n^{-8(s-k)/(1+4s)}) \text{ since } hn^{2/(1+4s)} = mh \rightarrow +\infty.
 \end{aligned}$$

Collecting the above evaluations and the computation of the remainder and linear terms, we get in case (ii),

$$E(\hat{T}_n - T(f))^2 = O(n^{-8(s-k)/(1+4s)}).$$

Turning to case (i), let $s > 2k + \frac{1}{4}$. We are now going to prove asymptotic efficiency. We define

$$R = \sqrt{n} \left[\hat{T}_n - T(f) - \frac{1}{n_2} \sum_{l=1}^{n_2} \sum_{j=0}^k (-1)^j (\phi'_j(f, \dots, f^{(k)}, \cdot))^{(j)}(X_l) + \int \phi'_j(f, \dots, f^{(k)}, \cdot) f^{(j)} \right].$$

Of course, to ensure that both (3.4) and (3.5) hold, it is enough to show that $E(R^2) \rightarrow 0$. Introducing auxiliary terms, we notice that $R = R_1 + R_2$, where

$$\begin{aligned}
 R_1 &= \sqrt{n} \left[\hat{T}_n - T(f) \right. \\
 &\quad \left. - \frac{1}{n_2} \sum_{l=1}^{n_2} \left(\sum_{j=0}^k (-1)^j (\phi'_j(\hat{f}, \dots, \hat{f}^{(k)}, \cdot))^{(j)}(X_l) - \int_{-\pi}^{\pi} \phi'_j(\hat{f}, \dots, \hat{f}^{(k)}, \cdot) f^{(j)} \right) \right]
 \end{aligned}$$

and

$$\begin{aligned}
 R_2 &= -\sqrt{n} \sum_{j=0}^k \left[\frac{1}{n_2} \sum_{l=1}^{n_2} (-1)^j (\phi'_j(f, \dots, f^{(k)}, \cdot) - \phi'_j(\hat{f}, \dots, \hat{f}^{(k)}, \cdot))^{(j)}(X_l) \right] \\
 &\quad + \sqrt{n} \sum_{j=0}^k \left[\int_{-\pi}^{\pi} \phi'_j(f, \dots, f^{(k)}, \cdot) f^{(j)} - \int_{-\pi}^{\pi} \phi'_j(\hat{f}, \dots, \hat{f}^{(k)}, \cdot) f^{(j)} \right].
 \end{aligned}$$

We want to prove that both $E(R_1^2)$ and $E(R_2^2) \rightarrow 0$. Plugging the expressions for \hat{T}_n and $T(f)$ into R_1 we obtain

$$\begin{aligned}
 R_1 &= \sqrt{n} \sum_{j,j'=0}^k \left[\frac{-2}{n_2} \sum_{l=1}^{n_2} (-1)^j (K_J(\hat{f}) \hat{f}^{(j')})^{(j)}(X_l) + 2 \int_{-\pi}^{\pi} K_J(\hat{f}) \hat{f}^{(j')} f^{(j)} \right] \\
 &\quad + \sqrt{n} \left[\sum_{j,j'=0}^k (\hat{T}^J - T^J) + \Gamma_n \right].
 \end{aligned}$$

Because of (4.10), we just have to prove that the expectation of the square of $\sqrt{n} \sum_{j,j'=0}^k (\hat{L}^J - L^J + \hat{T}^J - T^J)$ converges to zero, where $L^J = 2 \int K_J(\hat{f}) \hat{f}^{(j')} f^{(j)}$ and $\hat{L}^J = (2/n_2) \sum_{l=1}^{n_2} (-1)^j (K_J(\hat{f}) (\hat{f}^{(j')})^{(j)}(X_l))$. We will first evaluate $nE(\sum_{j,j'=0}^k \hat{L}^J - L^J)^2$ and $nE(\sum_{j,j'=0}^k \hat{T}^J - T^J)^2$, and then we will show that the sum of these terms is asymptotically the opposite of the covariance term $2nE(\sum_{j,j'=0}^k \hat{L}^J - L^J)(\sum_{j,j'=0}^k \hat{T}^J - T^J)$.

First, then, we have that

$$nE[(\hat{L}^J - L^J)(\hat{L}^{J_1} - L^{J_1})|\hat{f}] = 4 \frac{n}{n_2} \int_{-\pi}^{\pi} (-1)^{j+j_1} (K_J(\hat{f})\hat{f}^{(j)})^{(j)} (K_{J_1}(\hat{f})\hat{f}^{(j_1)})^{(j_1)} f \\ - 4 \frac{n}{n_2} \int_{-\pi}^{\pi} K_J(\hat{f})\hat{f}^{(j)} f^{(j)} \int_{-\pi}^{\pi} K_{J_1}(\hat{f})\hat{f}^{(j_1)} f^{(j_1)}.$$

The expectation of this expression converges towards the same expression with f instead of \hat{f} . So,

$$\lim_{n \rightarrow \infty} nE \left[\left(\sum_{j,j'=0}^k \hat{L}^J - L^J \right)^2 \right] = \sum_{j,j',j_1,j'_1=0}^k \Lambda_{JJ_1}(f, K_J(f), K_{J_1}(f))$$

where $\Lambda_{JJ_1}(f, \psi, \xi)$ is defined in Theorem 1.

Next, since $s > 2k + \frac{1}{4}$, for all $j, j' \in \{0, \dots, k\}$, $s > j + j' + \frac{1}{4}$; therefore, by Theorem 1,

$$|nE[(\hat{T}^J - T^J)(\hat{T}^{J_1} - T^{J_1})|\hat{f}] - \Lambda_{JJ_1}(\hat{f}, K_J(\hat{f}), K_{J_1}(\hat{f}))| \leq C'_1 \mu_1^2 (m^{-\alpha} + m^{\frac{1}{2}+j+j'+j_1+j'_1-2s}),$$

where

$$\mu_1 = \sup_{l=0, \dots, j'} [\sup_{l=0, \dots, j'} (\|K_J(\hat{f})\|_{\infty}^{(l)} + \|(K_J(\hat{f}))^{(l)}\|_a); \sup_{l=0, \dots, j'_1} (\|K_{J_1}(\hat{f})\|_{\infty}^{(l)} + \|(K_{J_1}(\hat{f}))^{(l)}\|_a)].$$

As already proved, $E(\mu_1^2) = O(h^{2s'-4k-2\alpha} + 1)$. Since s' is arbitrarily close to $s = p + \alpha$ and $s > 2k + \frac{1}{4}$, we obtain $p \geq 2k$ and $\lim_{n \rightarrow \infty} E(C'_1 \mu_1^2 (m^{-\alpha} + m^{\frac{1}{2}+j+j'+j_1+j'_1-2s})) = 0$. So,

$$\lim_{n \rightarrow \infty} nE(\hat{T}^J - T^J)(\hat{T}^{J_1} - T^{J_1}) = \lim_{n \rightarrow \infty} E(\Lambda_{JJ_1}(\hat{f}, K_J(\hat{f}), K_{J_1}(\hat{f}))) \\ = \Lambda_{JJ_1}(f, K_J(f), K_{J_1}(f)).$$

This implies that

$$\lim_{n \rightarrow \infty} nE \left[\left(\sum_{j,j'=0}^k \hat{T}^J - T^J \right)^2 \right] = \sum_{j,j',j_1,j'_1=0}^k \Lambda_{JJ_1}(f, K_J(f), K_{J_1}(f)).$$

We now show that

$$\lim_{n \rightarrow \infty} 2nE \left[\left(\sum_{j,j'=0}^k \hat{L}^J - L^J \right) \left(\sum_{j,j'=0}^k \hat{T}^J - T^J \right) \right] = -2 \sum_{j,j',j_1,j'_1=0}^k \Lambda_{JJ_1}(f, K_J(f), K_{J_1}(f)).$$

We begin by noticing that

$$E[(\hat{T}^J - T^J)(\hat{L}^{J_1} - L^{J_1})|\hat{f}] = E(\hat{T}^J \hat{L}^{J_1}|\hat{f}) - E(\hat{T}^J|\hat{f})L^{J_1}.$$

We recall that $\hat{T}^J = \hat{T}_0^{jj'} + \hat{T}_0^{jj} + \hat{T}_2^J$, where \hat{T}_2^J is defined by (2.2) with $K_J(\hat{f})$ in place of ψ . Moreover,

$$\hat{T}_0^{j'j} = \frac{1}{n_2(n_2 - 1)} \sum_{i=0}^m \sum_{l_1 \neq l_2=1}^{n_2} (-1)^j p_i(X_{l_1})(K_J(\hat{f})p_i^{(j')})^{(j)}(X_{l_2}).$$

We will do the computations for $\hat{T}_0^{j'j}$ instead of \hat{T}^J . Integrating by parts, we obtain

$$\begin{aligned} \mathbb{E}(\hat{T}_0^{j'j} \hat{L}^{J_1} | \hat{f}) - \mathbb{E}(\hat{T}_0^{j'j} | \hat{f}) L^{J_1} &= -\frac{2}{n_2} \left(\int_{-\pi}^{\pi} (-1)^{j'+j_1} S_m[(K_{J_1}(\hat{f})\hat{f}^{(j_1)})^{(j)} f](K_J(\hat{f})f^{(j)})^{(j')} \right) \\ &\quad - \frac{2}{n_2} \left(\int_{-\pi}^{\pi} (-1)^{j+j_1} (K_J(\hat{f})S_m f^{(j')})^{(j)} (K_{J_1}(\hat{f})\hat{f}^{(j_1)})^{(j)} f \right) \\ &\quad + \frac{4}{n_2} \int_{-\pi}^{\pi} K_J(\hat{f})S_m f^{(j')} f^{(j)} \int_{-\pi}^{\pi} K_{J_1}(\hat{f})\hat{f}^{(j_1)} f^{(j_1)}. \end{aligned}$$

So

$$\lim_{n \rightarrow \infty} 2n \sum_{j'j_1j_1=0}^k [\mathbb{E}(\hat{T}_0^{j'j} \hat{L}^{J_1}) - \mathbb{E}(\hat{T}_0^{j'j} L^{J_1})] = -2 \sum_{j',j_1,j_1=0}^k \Lambda_{J_1}(f, K_J(f), K_{J_1}(f)).$$

Of course, the same result holds with $\hat{T}_0^{j'j}$ replacing $\hat{T}_0^{j'j}$. By similar computations for $\mathbb{E}(\hat{T}_0^{j'j} \hat{L}^{J_1} | \hat{f}) - \mathbb{E}(\hat{T}_0^{j'j} | \hat{f}) L^{J_1}$, the above results holds for \hat{T}^J instead of $\hat{T}_0^{j'j}$. Thus, finally,

$$\lim_{n \rightarrow \infty} n \mathbb{E} \left[\left(\sum_{j'j=0}^k \hat{T}^J - T^J + \hat{L}^J - L^J \right)^2 \right] = 0.$$

It remains to prove that $R_2 \xrightarrow{\mathbb{L}^2} 0$.

$$\begin{aligned} \mathbb{E}(R_2^2) &= \frac{n}{n_2} \mathbb{E} \left[\int_{-\pi}^{\pi} \left(\sum_{j=0}^k (-1)^j (\phi'_j(\hat{f}, \dots, \hat{f}^{(k)}, \cdot))^{(j)} - \sum_{j=0}^k (-1)^j (\phi'_j(f, \dots, f^{(k)}, \cdot))^{(j)} \right)^2 f \right] \\ &\quad - \frac{n}{n_2} \mathbb{E} \left[\left(\sum_{j=0}^k \int_{-\pi}^{\pi} \phi'_j(\hat{f}, \dots, \hat{f}^{(k)}, \cdot) f^{(j)} - \int_{-\pi}^{\pi} \phi'_j(f, \dots, f^{(k)}, \cdot) f^{(j)} \right)^2 \right] \end{aligned}$$

Using the fact that ϕ and its derivatives are bounded over K_ϵ and that \hat{f} and its p first derivatives converge in \mathbb{L}^{-1} -norm towards f and its derivatives respectively, we show as above that

$$\mathbb{E} \left[\int_{-\pi}^{\pi} ((\phi'_j(\hat{f}, \dots, \hat{f}^{(k)}, \cdot))^{(j)})^2 f \right] \rightarrow \int_{-\pi}^{\pi} ((\phi'_j(f, \dots, f^{(k)}, \cdot))^{(j)})^2 f.$$

By similar arguments for all the terms which appear in the expression for $\mathbb{E}(R_2^2)$ and since $n/n_2 \rightarrow 1$, we obtain $R_2 \xrightarrow{\mathbb{L}^2} 0$. This concludes the proof of Theorem 2.

Appendix

Proof of Lemma 1

The density f is assumed to satisfy Assumption A1. \tilde{f} is an estimator of the density f based on projection methods:

$$\tilde{f}(x) = \frac{1}{n_1} \sum_{j=n-n_1}^n \sum_{i=0}^{m_1} p_i(X_j) p_i(x), \quad \text{where } m_1 \approx n_1^{1/(1+2s)}.$$

We define the event $A_n = \{\tilde{f}^{(l)}(S) \subset [a_i - \epsilon, b_i + \epsilon] \text{ for all } i \in \{0, \dots, k\}\}$, where $\epsilon > 0$ is such that $K_\epsilon = \prod_{i=0}^k [a_i - \epsilon, b_i + \epsilon] \times S \subset \Omega$, and we set $\hat{f} = \tilde{f} \mathbf{1}_{A_n} + f_0 \mathbf{1}_{A_n^c}$ where f_0 satisfies Condition A2(a) and (b).

It is clear that \hat{f} also satisfies Condition A2(a) and (b); in particular, this implies that $\|\hat{f}^{(l)}\|_\infty$ is bounded for $l = 0, \dots, k$. Conditions A2(c)–(e) will also be satisfied because they are satisfied for \tilde{f} and because $P(A_n^c)$ is of smaller order than the rates we wish to obtain.

We will first prove Condition A2(c) with \tilde{f} instead of \hat{f} . Let us first evaluate the bias term $\|\tilde{f}^{(l)} - f^{(l)}\|_q^q$, where $\tilde{f}^{(l)} - E(\tilde{f}^{(l)}) = (S_{m_1} f)^{(l)} = S_{m_1}(f^{(l)})$, for m_1 even.

Let \mathcal{P}_{m_1} be the set of trigonometric polynomials with degree not bigger than m_1 . It follows from M. Riesz's theorem that for $1 < q < +\infty$ and for $f \in \mathbb{L}^q(S)$, $\|S_m(f)\|_q \leq C_q \|f\|_q$ (see Bary 1964, vol. 2, p. 137). Therefore for $P \in \mathcal{P}_{m_1}$,

$$\begin{aligned} \|S_{m_1}(f^{(l)}) - f^{(l)}\|_q &= \|S_{m_1}(f^{(l)} - P) - (f^{(l)} - P)\|_q \\ &\leq C_q \|f^{(l)} - P\|_q \leq (2\pi)^{1/q} C_q \|f^{(l)} - P\|_\infty, \end{aligned}$$

hence,

$$\|S_{m_1}(f^{(l)} - f^{(l)})\|_q \leq (2\pi)^{1/q} C_q \inf_{P \in \mathcal{P}_{m_1}} \|f^{(l)} - P\|_\infty \leq \frac{C'_q}{m_1^{s-1}}$$

(see Zygmund 1968, vol. 1, p. 117). This gives

$$\|\tilde{f}^{(l)} - f^{(l)}\|_q^q = O(n_1^{q(l-s)/(1+2s)}).$$

Here, and in the remainder of the proof, the O s do not depend on $f \in F_{s,C}$ but depend on q . We will show that $E(\|\tilde{f}^{(l)} - \tilde{f}^{(l)}\|_q^q)$ is of the same order. Denote

$$Y_j(x) = \frac{\sum_{i=0}^{m_1} p_i(X_j) p_i^{(l)}(x) - \sum_{i=0}^{m_1} a_i(f) p_i^{(l)}(x)}{m_1^{1+l}}.$$

Since $\|p_i^{(l)}\|_\infty \leq i^l$ and $|a_i(f)| \leq \int f^2 \int p_i^2 \leq \|f\|_\infty$, we obtain $|Y_j(x)| \leq 2(1 + \|f\|_\infty)$ for m_1 large enough. Moreover, by orthonormality of the p_i we obtain $\sum_{j=n-n_1}^n E(Y_j^2(x)) \leq$

$\lambda(n_1/m_1)$, where $\lambda \in \mathbb{R}$ is independent of x . The variables $Y_j(x)$ are independent and centred. It follows from Rosenthal's (1970) inequality that

$$E\left(\left|\sum_{j=n-n_1}^n Y_j(x)\right|^q\right) \leq A_q \left(\frac{n_1}{m_1}\right)^{q/2},$$

where A_q is independent of x , which leads to

$$E\left(\left\|\sum_{j=n-n_1}^n Y_j\right\|_q^q\right) \leq 2\pi A_q \left(\frac{n_1}{m_1}\right)^{q/2}.$$

It follows that

$$E(\|\tilde{f}^{(l)} - \bar{f}^{(l)}\|_q^q) = O(n_1^{q(l-s)/(1+2s)}).$$

It remains to show that the same result holds when \tilde{f} is replaced by \hat{f} . To do this, we will bound $P(A_n^c)$ from above. Since $f \in F_{s,C}$, $\|S_{m_1}(f^{(l)}) - f^{(l)}\|_\infty = O(\log(m_1)/m_1^s)$ for all $l \leq p$ (see Zygmund 1968, vol. 1, p. 64). Hence for n large enough, independently of $f \in F_{s,C}$, $\|\tilde{f}^{(l)} - f^{(l)}\|_\infty \leq \epsilon/2$ for all $l \in \{0, \dots, p\}$. If $\|\tilde{f}^{(l)} - \bar{f}^{(l)}\|_\infty$ is also bounded by $\epsilon/2$ for $l = 0, \dots, p$, then $\tilde{f} = \hat{f}$. So,

$$P(A_n^c) \leq \sum_{l=0}^p P\left(\|\tilde{f}^{(l)} - \bar{f}^{(l)}\|_\infty \geq \frac{\epsilon}{2}\right).$$

Denote by D_{m_1} Dirichlet's kernel:

$$D_{m_1}(t) = \frac{\sin(m_1 + \frac{1}{2})t}{2\pi \sin \frac{t}{2}}.$$

Then

$$\tilde{f}^{(l)}(x) = \frac{1}{n_1} \sum_{j=n-n_1}^n D_{m_1}^{(l)}(x - X_j),$$

where $\tilde{f}^{(l)}$ is the convolution product $f * D_{m_1}^{(l)}$. We have that $\|D_{m_1}^{(l)}\|_\infty \leq m_1^{1+l}$ and $E((D_{m_1}^{(l)}(x - X_j))^2) \leq m_1^{1+2l}$. It follows from Bernstein's inequality that

$$P\left(\left|\tilde{f}^{(l)} - \bar{f}^{(l)}\right|(x) \geq \frac{\epsilon}{2}\right) \leq 2 \exp\left\{-\frac{1}{2} \left[\frac{n_1 \epsilon^2 / 4}{m_1^{1+2l} + \frac{1}{6} \epsilon m_1^{1+l}}\right]\right\}.$$

Let x_0, \dots, x_N be $N + 1$ points from $[-\pi, \pi]$ such that $x_0 = -\pi$, $x_N = \pi$, $x_i \leq x_{i+1}$, $|x_{i+1} - x_i| = \delta$. Then

$$\left|\tilde{f}^{(l)}(x) - \bar{f}^{(l)}(x) - (\tilde{f}^{(l)}(x_i) - \bar{f}^{(l)}(x_i))\right| \leq (\|\tilde{f}^{(l+1)}\|_\infty + \|\bar{f}^{(l+1)}\|_\infty)|x - x_i| \leq 2m_1^{2+l}|x - x_i|.$$

Let δ be such that $2\delta m_1^{2+l} = \epsilon/4$. Then

$$P\left(\sup_{|x-x_i|\leq\delta} |\tilde{f}^{(l)} - \bar{f}^{(l)}|(x) > \frac{\epsilon}{2}\right) \leq P\left(|\tilde{f}^{(l)} - \bar{f}^{(l)}|(x_i) > \frac{\epsilon}{4}\right) \leq 2 \exp\left\{-\frac{1}{2} \left[\frac{n_1 \epsilon^2 / 16}{m_1^{1+2l} + \frac{1}{12} \epsilon m_1^{1+l}}\right]\right\}.$$

Finally

$$P\left(\|\tilde{f}^{(l)} - \bar{f}^{(l)}\|_\infty \geq \frac{\epsilon}{2}\right) \leq 2\left(\frac{2\pi}{\delta}\right) \exp\left\{-\frac{1}{2} \left(\frac{n_1 \epsilon^2 / 16}{m_1^{1+2l} + \frac{1}{12} \epsilon m_1^{1+l}}\right)\right\},$$

which we can also write as

$$P\left(\|\tilde{f}^{(l)} - \bar{f}^{(l)}\|_\infty \geq \frac{\epsilon}{2}\right) \leq K_0 m_1^{2+l} \exp\left\{-\left(\frac{K_1 n_1 \epsilon^2}{m_1^{1+2l} + \frac{1}{12} \epsilon m_1^{1+l}}\right)\right\}, \tag{5.1}$$

where K_0, K_1 are positive absolute constants. Since $m_1 = n_1^{1/(1+2s)}$, for n_1 large enough, and for $l \leq p < s$, $P(A_n^c) \leq \sum_{l=0}^p P(\|\tilde{f}^{(l)} - \bar{f}^{(l)}\|_\infty \geq \epsilon/2) \leq K_0 n_1^{-qs/(1+2s)}$. We can now write

$$E(\|\hat{f}^{(l)} - f^{(l)}\|_q^q) \leq E(\|\tilde{f}^{(l)} - f^{(l)}\|_q^q) + \|f^{(l)} - f_0^{(l)}\|_q^q K_0 n_1^{-qs/(1+2s)} \leq \gamma_1 n_1^{-q(s-l)/(1+2s)}.$$

This concludes the proof of Condition A2(c).

Let us now control the order of magnitude of $E(\|\hat{f}^{(l)}\|_\infty^q)$ for all $l \leq 2k$ in order to prove Condition A2(d). From (5.1) we deduce

$$P\left(\frac{\|\tilde{f}^{(l)} - \bar{f}^{(l)}\|_\infty^q}{m_1^{q(l-s')}} \geq t\right) \leq K_0 m_1^{2+l} \exp[-K_1 n_1 t^{2/q} m_1^{-1-2s'}].$$

Since the above inequality holds for $t \geq t_0 > 0$, we obtain, for n_1 large enough,

$$E\left(\frac{\|\tilde{f}^{(l)} - \bar{f}^{(l)}\|_\infty^q}{m_1^{q(l-s')}}\right) \leq t_0 + \int_{t_0}^{+\infty} K_0 m_1^{2+l} \exp[-K_1 n_1^{2(s-s')/(1+2s)} t^{2/q}] dt \leq 2t_0 \quad \text{for all } s' < s.$$

To conclude the proof of Condition A2(d), it remains to prove that

$$\|\bar{f}^{(l)}\|_\infty = O(1 + m_1^{l-s'}) \quad \text{for all } s' < s, \quad \text{for all } 0 \leq l \leq 2k.$$

We can see that for all $l \leq p$, $\|\bar{f}^{(l)}\|_\infty$ is bounded by some constant independent of $f \in F_{s,C}$. This follows from Lemma 3 and the fact that for all $l \leq p$, $\|\bar{f}^{(l)} - f^{(l)}\|_\infty = O(\log m_1)/m_1^\alpha$. We shall now prove that $\|\bar{f}^{(p+1)}\|_\infty = O(m_1^{1-\alpha'})$ for all $0 < \alpha' < \alpha$. Since f is extended by periodicity, we have that

$$\bar{f}^{(p+1)}(x) = \int_{-\pi}^{\pi} D'_{m_1}(t)[f^{(p)}(x-t) - f^{(p)}(x)] dt$$

$$= \int_{-\pi}^{\pi} \left[\frac{(m_1 + \frac{1}{2}) \cos(m_1 + \frac{1}{2})t}{2\pi \sin \frac{t}{2}} - \frac{\sin(m_1 + \frac{1}{2})t \cos \frac{t}{2}}{4\pi \left(\sin \frac{t}{2}\right)^2} \right] [f^{(p)}(x-t) - f^{(p)}(x)] dt.$$

Moreover, for all x and for $t \in [-\pi, \pi]$, we have $|f^{(p)}(x-t) - f^{(p)}(x)| \leq 2C|t|^\alpha$ since $f \in F_{s,C}$ and f is extended by periodicity. So,

$$\begin{aligned} |\bar{f}^{(p+1)}(x)| &\leq \left(m_1 + \frac{1}{2}\right) \left| \int_{-\pi}^{\pi} \left[\frac{\cos(m_1 + \frac{1}{2})t}{2\pi \left(\sin \frac{t}{2}\right)} \right] [f^{(p)}(x-t) - f^{(p)}(x)] dt \right| \\ &\quad + 2Cm_1^{-\alpha} \int_{-m_1\pi}^{m_1\pi} \left| \frac{\sin\left(1 + \frac{1}{2m_1}\right)y}{\left(\sin \frac{y}{2m_1}\right)^2} \right| |y|^\alpha \frac{dy}{m_1}. \end{aligned}$$

The second term of the sum is bounded by

$$C'm_1^{1-\alpha} \int_{-m_1\pi}^{m_1\pi} |y|^{\alpha-2} \left| \sin\left(1 + \frac{1}{2m_1}\right)y \right| dy = \begin{cases} O(m_1^{1-\alpha}) & \text{if } 0 < \alpha < 1 \\ O(\log m_1) & \text{if } \alpha = 1. \end{cases}$$

As to the first term, we observe that

$$\int_{-\pi}^{\pi} \left[\frac{(\cos(m_1 + \frac{1}{2})t)}{2\pi \sin \frac{t}{2}} \right] (f^{(p)}(x-t) - f^{(p)}(x)) dt = S_{m_1}(\check{f}^{(p)})(x) - \check{f}^{(p)}(x),$$

where $S_m(\check{f}^{(p)})$ denotes the conjugate Fourier series of $f^{(p)}$ and $\check{f}^{(p)}$ the conjugate function of $f^{(p)}$ (see Bary 1964, vol. 2, pp. 51–53). Since $f^{(p)}$ is an α -Hölderian function, if $0 < \alpha < 1$, $\check{f}^{(p)}$ is also an α -Hölderian function, as proved in Bary (1964, vol. 2, p. 99). Therefore,

$$\|S_{m_1}(\check{f}^{(p)}) - \check{f}^{(p)}\|_\infty = O(m_1^{-\alpha'})$$

for all α' such that $0 < \alpha' < \alpha \leq 1$. This leads to

$$\|\bar{f}^{(p+1)}\|_\infty = O(m_1^{1-\alpha'}) \quad \text{for all } \alpha' < \alpha.$$

From Bernstein's inequality concerning trigonometric polynomials (see Zygmund 1968, vol. 1, p. 118), it follows that for $l \geq 1$ $\|\bar{f}^{(p+l)}\|_\infty \leq (2m_1)^{l-1} \|\bar{f}^{(p+1)}\|_\infty$. Moreover, since $\|\bar{f}^{(p+1)}\|_\infty = O(m_1^{1-\alpha'})$ for all $\alpha' < \alpha$, we obtain $\|\bar{f}^{(p+l)}\|_\infty = O(m_1^{l-\alpha'})$ for all $l \geq 1$, and $\alpha' < \alpha$. This completes the proof of Condition A2(d). So, for all $s' < s$,

$$E(\|\hat{f}^{(l)}\|_a^q) = O(m_1^{q(l-s')} + 1). \tag{5.2}$$

Finally for all $s' < s$, we have to show that

$$E(\|\hat{f}^{(l)}\|_a^q) \leq \gamma_3(1 + n_1^{q(l+\alpha-s)/(1+2s)}).$$

We use the inequality

$$\frac{|\hat{f}^{(l)}(x) - \hat{f}^{(l)}(y)|}{|x - y|^\alpha} \leq \|\hat{f}^{(l+1)}\|_\infty |x - y|^{1-\alpha} \quad \text{for all } x, y \in S,$$

and we consider two cases, first, if $l + 1 < s$, then $E(\|\hat{f}^{(l+1)}\|_\infty^q) = O(1)$ by (5.2) and we also have $E(\|\hat{f}^{(l)}\|_\alpha^q) = O(1)$. Second, if $l + 1 \geq s = p + \alpha$, then $l \geq p$; there are two subcases to deal with. If $|x - y| \leq 1/m_1$, then

$$\frac{|\hat{f}^{(l)}(x) - \hat{f}^{(l)}(y)|}{|x - y|^\alpha} \leq \|\hat{f}^{(l+1)}\|_\infty m_1^{\alpha-1}.$$

If however $|x - y| \geq 1/m_1$, then, for all $0 < \alpha' < \alpha$,

$$\frac{|\hat{f}^{(l)}(x) - \hat{f}^{(l)}(y)|}{|x - y|^\alpha} = \frac{|\hat{f}^{(l)}(x) - \hat{f}^{(l)}(y)|}{|x - y|^{\alpha-\alpha'}} |x - y|^{\alpha'} \leq 2\|\hat{f}^{(l)}\|_\infty (2\pi)^{\alpha'} m_1^{\alpha-\alpha'}.$$

Finally,

$$E(\|\hat{f}^{(l)}\|_\alpha^q) = O[E(\|\hat{f}^{(l+1)}\|_\infty^q) m_1^{q(\alpha-1)} + E(\|\hat{f}^{(l)}\|_\infty^q) m_1^{q(\alpha-\alpha')}].$$

From (5.2) we derive that, for all $s' < s$,

$$E(\|\hat{f}^{(l)}\|_\alpha^q) = O(m_1^{q(l+1-s')} m_1^{q(\alpha-1)} + (1 + m_1^{q(l-s')}) m_1^{q(\alpha-\alpha')}) = O(1 + m_1^{q(l+\alpha-s')}).$$

This completes the proof of Lemma 1.

Semi-parametric information bound

We refer to Koshevnik and Levit (1976) and Levit (1978). We suppose that f belongs to $F_{s,C}$, with $s = p + \alpha$. We want to estimate $T(f) = \int \phi(f(x), \dots, f^{(k)}(x), x) dx$ with $k \leq p$. Let ξ be a bounded function, infinitely differentiable, such that $\int f \xi = 0$. We define

$$f_t = f(1 + t\xi).$$

$f_t \in F_{s,C}$ is a density for t small enough. Now

$$\begin{aligned} T(f_t) - T(f) &= \int \sum_{j=0}^k \frac{\partial \phi}{\partial x_j}(f, \dots, f^{(k)}, \cdot)(f_t^{(j)} - f^{(j)}) + o(t) \\ &= \int \sum_{j=0}^k (-1)^j \left(\frac{\partial \phi}{\partial x_j}(f, \dots, f^{(k)}, \cdot) \right)^{(j)} (f_t - f) + o(t) \text{ by integration by parts} \\ &= \int \sum_{j=0}^k \left[(-1)^j \left(\frac{\partial \phi}{\partial x_j}(f, \dots, f^{(k)}, \cdot) \right)^{(j)} - \int \left(\frac{\partial \phi}{\partial x_j}(f, \dots, f^{(k)}, \cdot) f^{(j)} \right) \right] \\ &\quad \times (f_t - f) + o(t) \end{aligned}$$

since $\int f_t = \int f = 1$. Moreover, $\|\sqrt{f_t} - \sqrt{f} - At\|_2 \rightarrow 0$ as $n \rightarrow \infty$, where $At = t\xi\sqrt{f}/2$.

Now, $T(f_t) - T(f) = Bt + o(t)$, where $Bt = E_f(Gt\xi/2)$ and

$$G = 2 \left[\sum_{j=0}^k (-1)^j \left(\frac{\partial \phi}{\partial x_j}(f, \dots, f^{(k)}, \cdot) \right)^{(j)} - \sum_{j=0}^k \int \frac{\partial \phi}{\partial x_j}(f, \dots, f^{(k)}, \cdot) f^{(j)} \right].$$

so the function G satisfies $E_f(G) = 0$. Since the functions ξ are dense in the set of functions $g \in \mathbb{L}^2([-\pi, \pi])$ such that $E_f(g) = 0$, there exists a sequence ξ_n such that

$$\lim_{n \rightarrow \infty} \int_{-\pi}^{\pi} \left(\frac{\xi_n}{2} - G \right)^2 f(x) dx = 0.$$

According to Theorems 1 and 2 of Koshevnik and Levit (1976),

$$\inf_{\epsilon} \liminf_{n \rightarrow \infty} \sup_{f_1 \in F_{s,c}, \|f_1 - f\|_2 \leq \epsilon} n E(\hat{T}_n - T(f))^2 \geq \frac{1}{4} E_f(G^2).$$

Finally,

$$\frac{1}{4} E_f(G^2) = \int \left[\sum_{j=0}^k (-1)^j \left(\frac{\partial \phi}{\partial x_j}(f, \dots, f^{(k)}, \cdot) \right)^{(j)} \right]^2 f d\mu - \left(\int \sum_{j=0}^k \frac{\partial \phi}{\partial x_j}(f, \dots, f^{(k)}, \cdot) f^{(j)} d\mu \right)^2.$$

For the problem of estimating $\int (f^{(k)})^2$, we get for the analogue of the Cramér–Rao bound:

$$4 \left[\int (f^{(2k)})^2 f - \left(\int (f^{(k)})^2 \right)^2 \right].$$

Acknowledgements

The author wishes to express his thanks to Pascal Massart and to Lucien Birgé for their helpful suggestions. Thanks are also due to the referee and editors for their constructive comments.

References

- Bary, N.K. (1964). *A Treatise on Trigonometric Series*. Oxford: Pergamon Press.
- Bickel, P.J. and Ritov, Y. (1988) Estimating integrated squared density derivatives: sharp best order of convergence estimates. *Sankhya Ser. A*, **50**, 381–393.
- Birgé, L. and Massart, P. (1995) Estimation of integral functionals of a density. *Ann. Statist.*, **23**(1), 11–29.
- Deheuvels, P. and Hominal, P. (1980) Estimation automatique de la densité. *Rev. Statist. Appl.*, **28**, 25–55.
- Donoho, D. and Nussbaum, M. (1990) Minimax quadratic estimation of a quadratic functional. *J. Complexity*, **6**, 290–323.

- Farell, R.H. (1972) On the best obtainable asymptotic rates of convergence in estimation of a density function at a point. *Ann. Math. Statist.*, **43**, 170–180.
- Gu, C. (1994) Model indexing and model selection in nonparametric function estimation. Technical report, Purdue University.
- Hoeffding, W. (1948) A class of statistics with asymptotically normal distribution. *Ann. Math. Statist.*, **19**, 293–325.
- Ibragimov, I.A. and Has'minskii, R.Z. (1991) Asymptotic normal families of distribution and effective estimation. *Ann. Statist.*, **19**, 1681–1724.
- Ibragimov, I.A., Nemirovskii, A.S. and Has'minskii, R.Z. (1987) Some problems on nonparametric estimation in Gaussian white noise. *Theory Probab. Appl.*, **31**, 391–406.
- Kerkycharian, G. and Picard, D. (1996) Estimating nonquadratic functionals of a density using Haar wavelets. *Ann. Statist.*, **24**, 485–507.
- Koshevnik, Y.A. and Levit, B.Y. (1976) On a non parametric analogue of the information matrix. *Theory Probab. Appl.*, **21**, 738–753.
- Laurent, B. (1996) Efficient estimation of integral functionals of a density. *Ann. Statist.*, **24**, 659–681.
- Laurent, B. (1993) Estimation de fonctionnelles intégrales non linéaires d'une densité et de ses dérivées. Ph.D. thesis, Université Paris Sud.
- Levit, B.Y. (1978) Asymptotically efficient estimation of nonlinear functionals. *Problems Inform. Transmission*, **14**, 65–72.
- Rosenthal, H.P. (1970) On the subspace of L_p , $p > 2$, spanned by sequences of independent variables. *Israel J. Math.*, **8**, 273–303.
- Serfling, R.J. (1980) *Approximation Theorems of Mathematical Statistics*. New York: Wiley.
- Wahba, G. (1990) *Spline Models for Observational Data* CBMS-NSF Regional Conference Series. Philadelphia: SIAM.
- Zygmund, A. (1968) *Trigonometric Series*. Cambridge: Cambridge University Press.

Received May 1994 and Revised January 1996