The Annals of Statistics
1996, Vol. 24, No. 2, 659–681

EFFICIENT ESTIMATION OF INTEGRAL FUNCTIONALS OF A DENSITY

BY BÉATRICE LAURENT

Université Paris Sud

We consider the problem of estimating a functional of a density of the type $\int\phi(f,\cdot)$. Starting from efficient estimators of linear and quadratic functionals of f and using a Taylor expansion of ϕ, we build estimators that achieve the $n^{-1/2}$ rate whenever f is smooth enough. Moreover, we show that these estimators are efficient. Concerning the estimation of quadratic functionals (more precisely, of integrated squared density) Bickel and Ritov have already built efficient estimators. We propose here an alternative construction based on projections, which seems more natural.

1. Introduction. Let $X_1,\dots,X_n$ be i.i.d. with common density f with respect to some measure μ. When μ is the Lebesgue measure on the real line, Bickel and Ritov (1988) have studied the problem of estimating $\int(f^{(k)})^2$, where f is supposed to belong to a nonparametric set of densities Θ_s included in some compact set of smooth functions of order s. They built an efficient estimator if $s > 2k + \frac{1}{4}$. If $s \le 2k + \frac{1}{4}$, they showed that the best order of convergence is $n^{(-4(s-k))/(1+4s)}$. It is quite a remarkable result in the sense that the critical regularity $2k + \frac{1}{4}$ is completely unusual. In fact, one could think at first glance that this critical regularity should be $2k + \frac{1}{2}$ [see Hall and Marron (1987), where some statistical motivations for studying these functionals are also provided].

This problem has also been treated by Donoho and Nussbaum (1990) for the white noise model. It is also worth mentioning the paper by Ibragimov, Nemirovskii and Khas'minskii (1986), which deals with differentiable functionals in the same framework.

Bickel and Ritov's estimator is a quite intricate expression based on kernel estimators of the density. In this paper, we propose an alternative and somehow simpler method of estimation based on orthogonal projections. This method will allow us to treat the more general problem of estimating $\psi f^2\psi\, d\mu$ when f belongs to some ellipsoid $\mathscr{E} = \{\sum_{i\in D} a_i p_i; \sum_{i\in D}|a_i^2/c_i^2| \le 1\}$, where $(p_i)_{i\in D}$ is an orthonormal basis of $\mathbb{L}^2(d\mu)$. This generalization is crucial to achieve the aim of this paper, which is to construct efficient estimators of functionals of the type $T(f) = \int\phi(f(x), x)\, d\mu(x)$, $f\in\mathscr{E}$, when it is possible. This problem was first studied by Levit (1978), who built efficient estimators

Received October 1994.
AMS 1991 *subject classifications*. 62G06, 62G07, 62G20.
Key words and phrases. Estimation of density, projection methods, kernel estimators, Fourier series, semiparametric Cramér–Rao bound.

of this kind of functionals, under regularity properties for the density that are more restrictive than our conditions.

A typical motivating example of such functionals is the Shannon entropy $\int f \log(f)$. Dudewicz and Van der Meulen (1981) showed how estimators of the entropy may be used to test uniformity of a n-sample with density f concentrated on the interval $[0, 1]$. Moreover, Vasicek (1976) proposed a test of normality that is also based on estimators of the entropy.

Before stating our results, let us explain how the two above-mentioned problems are connected. ϕ is assumed to be a smooth function. So expanding ϕ up to the second order with Taylor's formula provides an expansion of $T(f) - T(\hat{f})$, 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 build 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 $\int f^2 K(\hat{f})$, which have been studied in the first section. Our main result is proved in Section 3 and may be summarized as follows. When f belongs to some Hölder's space with index s over $\mathbb{R}^d$, we can build an estimator $\hat{T}_n$ of $T(f)$ such that if $s > d/4$, $\sqrt{n}(\hat{T}_n - T(f)) \to \mathscr{N}(0, C(f, \phi))$, where $\mathscr{N}(0, \sigma^2)$ denotes the normal distribution and $nE(\hat{T}_n - T(f))^2 \to C(f, \phi)$, where

$$C(f, \phi) = \int (\phi'_1(f, \cdot))^2 f - \left(\int \phi'_1(f, \cdot) f \right)^2$$

and $\phi'_1(u, v) = \partial\phi/\partial u(u, v)$. Moreover, $C(f, \phi)$ is the semiparametric information bound for the problem of estimating $T(f)$ as will be shown in the Appendix; hence our estimator is asymptotically efficient.

When $s < d/4$ we do not know what the optimal rate is, except that it is smaller than $n^{-3s/(d+2s)}$. Actually, in this very case the remainder term in the Taylor expansion is precisely of order $n^{-3s/(d+2s)}$. Birgé and Massart (1995) have proved that the rate cannot be smaller than $n^{-4s/(d+4s)}$, so it would be necessary to do the Taylor expansion up to the third order and to estimate $\int f^3$. This was made in the case where $d = 1$ by Kerkyacharian and Picard (1993). They propose an estimator of $\int f^3$ based on wavelet methods which achieve the optimal rate of convergence when $s < \frac{1}{4}$.

2. Estimation of $\int f^2 \psi$. Suppose $X_1, \dots, X_n$ are i.i.d. random variables with common density $f \in \mathbb{L}^2(d\mu)$. In the following, all integrals will be taken with respect to μ. Let $(p_i)_{i \in D}$ be an orthonormal basis of $\mathbb{L}^2(d\mu)$, where D is a countable set. Let $a_i = \int f p_i$ and consider the ellipsoid $\mathscr{E} = \{\sum_{i \in D} a_i p_i; \sum_{i \in D} |a_i^2 / c_i^2| \le 1\}$.

Our purpose is to estimate $\theta = \int f^2 \psi$ when $f \in \mathscr{E}$. Let us first look at the case $\psi = 1$. Since $\int f^2 = \sum_{i \in D} a_i^2$, a natural idea would be to estimate this integral by $\tilde{\theta} = \sum_{i \in M} \hat{a}_i^2$, where $\hat{a}_i$ is the empirical estimator of a_i: $\hat{a}_i = (1/n)\sum_{j=1}^n p_i(X_j)$ and M is a subset of D. Therefore $\tilde{\theta} = (1/n^2)\sum_{i \in M} \sum_{j,l=1}^n p_i(X_j) p_i(X_l)$. The computation of the bias of this estimator shows that it can be reduced by removing the diagonal terms, that is, the terms of the type $(1/n^2)p_i^2(X_l)$. In this way, we get an estimator of $\int f^2$ with bias $-\int(S_M f - f)^2$, where $S_M f$ denotes $\sum_{i\in M} a_i p_i$.

We intend to build an estimator with a similar bias for the estimation of $\int f^2\psi$. More precisely, we wish the bias to be equal to

$$-\int (S_M f - f)^2 \psi = 2\int (S_M f) f\psi - \int (S_M f)^2 \psi - \int f^2\psi.$$

Hence, the problem is to find an estimator with expectation $2\int(S_M f)f\psi - \int(S_M f)^2\psi$. For the part $\int(S_M f)f\psi$, we propose the following estimator:

$$\hat{\theta}_1 = \frac{1}{n(n-1)} \sum_{i\in M} \sum_{j\neq k=1}^{n} p_i(X_j)(p_i\psi)(X_k). \tag{2.1}$$

To get the term $\int(S_M f)^2\psi$, we propose the estimator

$$\hat{\theta}_2 = \frac{1}{n(n-1)} \sum_{i,i'\in M} \sum_{j\neq k=1}^{n} p_i(X_j)p_{i'}(X_k)\int p_i p_{i'}\psi(x)\,dx. \tag{2.2}$$

This explains the expression of the estimator $\hat{\theta}$ proposed in the following theorem.

THEOREM 1. *Let $X_1,\dots,X_n$ be i.i.d. random variables with common density f belonging to some Hilbert space $\mathbb{L}^2(d\mu)$. Let $(p_i)_{i\in D}$ be an orthonormal basis of $\mathbb{L}^2(d\mu)$. We assume that f is uniformly bounded and belongs to the ellipsoid $\mathscr{E} = \{\sum_{i\in D} a_i p_i; \sum_{i\in D}|a_i^2/c_i^2| \le 1\}$. Suppose that the following condition holds: We can find a subset M_n of D such that*

$$\left(\sup_{i\notin M_n} |c_i|^2\right)^2 \approx \frac{|M_n|}{n^2},$$

where $|M_n|$ denotes the cardinality of M_n and

$$\forall\, g \in \mathbb{L}^2(d\mu), \qquad \int \left(S_{M_n} g - g\right)^2 d\mu \to 0 \quad \text{as } n\to\infty.$$

$\theta = \int f^2\psi$ is to be estimated, where ψ is a bounded function, and let

$$\hat{\theta} = \frac{2}{n(n-1)} \sum_{i\in M_n} \sum_{j\neq k=1}^{n} p_i(X_j)(p_i\psi)(X_k) - \frac{1}{n(n-1)} \sum_{i,i'\in M_n} \sum_{j\neq k=1}^{n} p_i(X_j)p_{i'}(X_k)\int p_i p_{i'}\psi(x)\,dx.$$

$\hat{\theta}$ has the following asymptotic properties to estimate $\int f^2\psi$:

(i) *If $|M_n|/n \to 0$ as $n\to\infty$, then*

$$\sqrt{n}\,(\hat{\theta} - \theta) \to \mathscr{N}(0, \Lambda(f,\psi)), \tag{2.3}$$

$$\begin{aligned} &|nE(\hat{\theta}-\theta)^2 - \Lambda(f,\psi)| \\ &\quad \le \gamma_1\left[\frac{|M_n|}{n} + \|S_{M_n} f - f\|_2 + \|S_{M_n}(f\psi) - f\psi\|_2\right], \end{aligned} \tag{2.4}$$

where

$$\Lambda(f,\psi) = 4\left[\int f^3\psi^2 - \left(\int f^2\psi\right)^2\right].$$

(ii) *Otherwise*

$$E(\hat{\theta}-\theta)^2 \le \gamma_2 \frac{|M_n|}{n^2},$$

where γ_1 *and* γ_2 *depend only on* $\|f\|_\infty$ *and* $\|\psi\|_\infty$. *Moreover, they are both increasing functions of* $\|f\|_\infty$ *and* $\|\psi\|_\infty$.

The notation $A_n \approx B_n$ used in Theorem 1 means that $\lambda_1 \le A_n/B_n \le \lambda_2$, where λ_1 and λ_2 are positive constants.

COMMENTS. Of course, it follows from (2.4) that $\lim_{n\to\infty} nE(\hat{\theta}-\theta)^2 = 4[\int f^3\psi^2 - (\int f^2\psi)^2]$. In the next section, ψ will be a random function depending on n, which is why we need some bound which depends explicitly on ψ.

We shall prove in the Appendix that the asymptotic variance is optimal and our estimator is therefore efficient.

EXAMPLE 1. Let $X_1,\dots,X_n$ i.i.d. d-dimensional random variables with density f belonging to $\mathbb{L}^2(d\mu)$, where μ is the Lebesgue measure over $\mathbb{R}^d$. f is supposed to belong to the ellipsoid $\mathscr{E} = \{\sum_{i\in\mathbb{Z}^d} a_i p_i;\ \sum_{(i_1,\dots,i_d)\in\mathbb{Z}^d}(|i_1|^{2s_1} + \cdots + |i_d|^{2s_d})|a_{i_1,\dots,i_d}|^2 \le 1\}$, where $\forall\ j\in\{1,\dots,d\}$, $s_j > 0$. Let $s\in\mathbb{R}$ be defined by $d/s = \sum_{j=1}^d 1/s_j$. Let $M_n = \{(i_1,\dots,i_d)\in\mathbb{Z}^d,\ |i_j| \le m^{s/ds_j}\}$. The cardinality of M_n equals $2m$. Moreover, $\sup_{i\notin M_n}|c_i|^2 \approx (m^{-2s/d})$. Let $m \approx n^{2/(1+4s/d)}$. Then $(\sup_{i\notin M_n}|c_i|^2)^2/|M_n| \approx 1/n^2$. Hence, as soon as $s > d/4$, the estimator $\hat{\theta}$ defined in Theorem 1 has the properties defined by (2.3) and (2.4) to estimate $\theta = \int f^2\psi$.

When $s \le d/4$, we get $E(\hat{\theta}-\theta)^2 \le \lambda n^{-8s/(d+4s)}$ for some $\lambda\in\mathbb{R}$. In the framework of this example, when f belongs to $\mathbb{L}^2([0,1]^d)$ and when $(p_i)_{i\in\mathbb{Z}^d}$ is the Fourier orthonormal basis of $\mathbb{L}^2([0,1]^d)$, we shall prove (see proof of Corollary 1) that the condition $f\in\mathscr{E}$ generalizes a condition of the type: f belongs to some Hölder space of any index greater than s. In this particular framework and when ψ is either always positive or always negative, Birgé and Massart (1995) have proved lower bounds for the rates of convergence which agree with the rates of Theorem 1. Our result is therefore optimal.

EXAMPLE 2. Suppose that $f\in\mathbb{L}^2([0,1]^d)$. We shall give some example to show how wavelets fit in our framework. Denote

$$\mathscr{E} = \left\{\sum_{j\ge 0}\sum_{\lambda\in\Lambda_j}\alpha(\lambda)\tilde{\psi}_\lambda,\ \sum_{j\ge 0}\sum_{\lambda\in\Lambda_j}2^{2js}|\alpha(\lambda)|^2 \le 1\right\},$$

where $\tilde{\psi}$ is a wavelet and

$$\Lambda_j = \left\{ \left(\frac{k_1}{2^j} + \frac{\varepsilon_1}{2^{j+1}}, \dots, \frac{k_d}{2^j} + \frac{\varepsilon_d}{2^{j+1}} \right), 0 \le k_i \le 2^j - 1; \right.$$
$$\left. (\varepsilon_1, \dots, \varepsilon_d) \in \{0,1\}^d \backslash (0, \dots, 0) \right\}.$$

If f belongs to some Hölder space of any index greater than s and if $\tilde{\psi}$ has regularity $r > s$, then $f \in \mathscr{E}$, where the equality $f = \sum_{j \ge 0} \sum_{\lambda \in \Lambda_j} \alpha(\lambda) \tilde{\psi}_\lambda$ holds in $\mathbb{L}^2$ [see Meyer (1990), page 108].

In order to apply Theorem 1, we set

$$M_n = \left\{ \lambda \in \Lambda_j, j \le j_0, 2^{j_0} = n^{2/(d+4s)} \right\}$$

and for $\lambda \in \Lambda_j$, $c_\lambda = 2^{-js}$. We have

$$|M_n| \approx n^{2d/(d+4s)}, \qquad \sup_{\lambda \notin M_n} |c_\lambda^2| \approx 2^{-2j_0 s} = n^{-4s/(d+4s)}$$

and the results of Theorem 1 hold.

3. Estimation of $\int \phi(f)$.

3.1. *Main results*. The purpose of this section is to estimate $T(f) = \int \phi(f(x), x)\, dx$ efficiently when it is possible. As in the previous section, we assume that f belongs to the ellipsoid $\mathscr{E} = \{\sum_{i \in D} a_i p_i; \sum_{i \in D} |a_i^2 / c_i^2| \le 1\}$.

We would like to start with some preliminary estimator $\hat{f}$ of the density f built on a small part of the initial sample and do a Taylor expansion of ϕ in a neighborhood of $(\hat{f}(x), x)$. In order to give a sense to this expansion we shall assume the following:

A1. The function $u \to \phi(u, x)$ belongs to $C^3(\Omega)$, where $C^p(\Omega)$ denotes the class of p times continuously differentiable functions over Ω. $\forall\, x$, $a \le f(x) \le b$, where $a, b \in \mathbb{R}$, with $[a, b] \subset \Omega$.

A2. We can find a preliminary estimator $\hat{f}$ of f constructed with $n_1 \approx n/\log(n)$ data, such that $\forall\, x, a - \varepsilon \le \hat{f}(x) \le b + \varepsilon$ with $[a - \varepsilon, b + \varepsilon] \subset \Omega$. Moreover, $\forall\, 2 \le q < +\infty$, $\forall\, l \in \mathbb{N}^*$, $E_f \|\hat{f} - f\|_q^l \le C(q, l) n_1^{-l\alpha}$ for some $\alpha > \frac{1}{6}$ and for some constant $C(q, l)$ independent of f belonging to the ellipsoid $\mathscr{E}$.

Denote $K_\varepsilon = [a - \varepsilon, b + \varepsilon]$. We shall give an example of such an estimator in the case where f is a density defined over a compact set S of $\mathbb{R}^d$ satisfiying some regularity assumptions.

Assuming that A1 and A2 are verified, it is now legitimate to make a Taylor expansion of ϕ in a neighborhood of $(\hat{f}(x), x)$. We shall use the following notation for partial derivatives:

$$\phi_1' = \frac{\partial \phi}{\partial u}(u, v), \qquad \phi_1'' = \frac{\partial^2 \phi}{\partial u^2}(u, v), \qquad \|\phi_1^{(3)}\|_\infty = \sup_{x, u \in K_\varepsilon} \left| \frac{\partial^3 \phi}{\partial u^3}(u, x) \right|.$$

Then

$$T(f) = \int \phi(\hat{f}(x), x)\, dx + \int \phi_1'(\hat{f}(x), x)(f - \hat{f})(x)\, dx$$
$$+ \frac{1}{2} \int \phi_1''(\hat{f}(x), x)(f - \hat{f})^2(x)\, dx + \Gamma_n,$$

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

$$T(f) = \int G(\hat{f}, \cdot) + \int H(\hat{f}, \cdot) f + \int K(\hat{f}, \cdot) f^2 + \Gamma_n,$$

where

$$G(\hat{f}, \cdot) = \phi(\hat{f}, \cdot) - \phi_1'(\hat{f}, \cdot)\hat{f} + \tfrac{1}{2}\phi_1''(\hat{f}, \cdot)\hat{f}^2, \tag{3.1}$$

$$H(\hat{f}, \cdot) = \phi_1'(\hat{f}, \cdot) - \hat{f}\phi_1''(\hat{f}, \cdot), \tag{3.2}$$

$$K(\hat{f}, \cdot) = \tfrac{1}{2}\phi_1''(\hat{f}, \cdot). \tag{3.3}$$

We have to estimate two types of functionals:

1. $\int H(\hat{f}, \cdot) f$, which is a linear functional of f;
2. $\int K(\hat{f}, \cdot) f^2$, which is a quadratic functional of f of the type studied in Section 2.

If $\hat{f}$ is based on the n_1 last observations, then $\int H(\hat{f}, \cdot) f$ and $\int K(\hat{f}, \cdot) f^2$ are estimated with the n_2 first data, where $n_2 = n - n_1$. The following theorem gives the expression of the estimator $\hat{T}_n$ of $T(f)$ and its properties.

THEOREM 2. *Let $X_1, X_2, \ldots, X_n$ be i.i.d. random variables with common density f belonging to some Hilbert space $\mathbb{L}^2(d\mu)$. Let $(p_i)_{i \in D}$ be an orthonormal basis of $\mathbb{L}^2(d\mu)$, $a_i = \int f p_i$ and suppose that f belongs to the ellipsoid*

$$\mathscr{E} = \left\{ \sum_{i \in D} a_i p_i;\ \sum_{i \in D} \left| \frac{a_i^2}{c_i^2} \right| \le 1 \right\}.$$

$T(f) = \int \phi(f, \cdot)$ is to be estimated. We assume that the hypotheses A1 *and* A2

hold and that $\|\phi_1'\|_\infty$, $\|\phi_1''\|_\infty$ *and* $\|\phi_1^{(3)}\|_\infty$ *are finite. Suppose that we can find a subset* M_n *of* D *such that* $(\sup_{i\notin M_n}|c_i^2|)^2 \approx |M_n|/n^2$ *and such that* $\forall\, g \in \mathbb{L}^2(d\mu)$, $\|S_{M_n} g - g\|_2 \to 0$ *as* $n\to\infty$. *Let*

$$\hat{T}_n = \int G(\hat{f},\cdot) + \frac{1}{n_2}\sum_{j=1}^{n_2} H(\hat{f},\cdot)(X_j)$$
$$+ \frac{2}{n_2(n_2-1)}\sum_{i\in M_n}\sum_{j\neq k=1}^{n} p_i(X_j)\big(K(\hat{f},\cdot)p_i\big)(X_k)$$
$$- \frac{1}{n_2(n_2-1)}\sum_{i,i'\in M_n}\sum_{j\neq k=1}^{n} p_i(X_j)p_{i'}(X_k)\int p_i p_{i'} K(\hat{f},\cdot),$$

where G, H *and* K *are defined by* (3.1), (3.2), *and* (3.3). *The following properties hold*:

$$\text{(3.4)}\qquad \textit{if}\quad \frac{|M_n|}{n}\to 0,\quad \textit{then}\quad \sqrt{n}\big(\hat{T}_n - T(f)\big)\to \mathcal{N}(0, C(f,\phi)),$$

$$\text{(3.5)}\qquad \lim_{n\to\infty} nE\big(\hat{T}_n - T(f)\big)^2 = C(f,\phi),$$

where $C(f,\phi)$ *is defined by* $C(f,\phi) = \int(\phi_1'(f,\cdot))^2 f - (\int\phi_1'(f,\cdot)f)^2$.

COMMENTS.

1. The asymptotic constant $C(f,\phi)$ appearing in Theorem 2 is optimal. This follows from the general theory of efficient estimation as explained in the Appendix.
2. We shall give a corollary of Theorem 2 with an explicit construction of the preliminary estimator $\hat{f}$ in the particular case where f is a density defined over a compact set of $\mathbb{R}^d$, for example, $S = [0,1]^d$, and satisfying regularity conditions. More precisely, let $r = (r_1,\dots,r_d)\in\mathbb{N}^d$ and $\alpha = (\alpha_1,\dots,\alpha_d) \in]0,1]^d$. Let D_j be the derivation operator with respect to the jth variable. We shall denote by $F_{r,\alpha,C}$ the set of densities f defined over $[0,1]^d$ such that

$$D_j^{r_j} f \text{ exists } \forall\, j\in\{0,\dots,d\} \text{ and } D_j^l f \text{ is periodic for } l = 0,\dots,r_j,$$
$$\sup_{\substack{x,y\in S\\ x_j\neq y_j}} \frac{|D_j^{r_j} f(x) - D_j^{r_j} f(y)|}{|x_j - y_j|^{\alpha_j}} \le C.$$

We define $s_j = r_j + \alpha_j$ and s by $d/s = \sum_{j=1}^d 1/s_j$. We will show in Section 4 that if f belongs to $F_{r,\alpha,C}$, then f belongs to some ellipsoid

$$\mathcal{E} = \left\{\sum_{i\in\mathbb{Z}^d} a_i p_i;\ \sum_{(i_1,\dots,i_d)\in\mathbb{Z}^d}\big(|i_1|^{2s_1'} + \cdots + |i_d|^{2s_d'}\big)|a_{i_1,\dots,i_d}|^2 \le \gamma\right\}$$

for all $(s_1',\dots,s_d')$ such that $\forall\, j$, $0 < s_j' < s_j = r_j + \alpha_j$ for some $\gamma > 0$. In this case, the preliminary estimator $\hat{f}$ is defined as follows. Let $\tilde{f}$ be a kernel estimator of f based on the n_1 last observations where $n_1 \approx n/\log n$. Let $f_0 \in F_{p,\alpha,C}$ be such that $f_0(S) \in [a-\varepsilon, b+\varepsilon]$, let $A_n = \{\tilde{f}(S) \subset [a-\varepsilon, b+\varepsilon]\}$ and

$$\hat{f}(x) = \tilde{f}(x)\mathbb{1}_{A_n} + f_0(x)\mathbb{1}_{A_n^c}. \tag{3.6}$$

Theorem 2, together with rates of convergence results by Ibragimov and Khas'minskii ensuring A2 when $s > d/4$ [see Ibragimov and Khas'minskii (1980) and (1981)], implies the following corollary:

COROLLARY 1. *Let $X_1, X_2, \dots, X_n$ be i.i.d. d-dimensional random variables with density f belonging to the set $F_{r,\alpha,C}$, $r \in \mathbb{N}^d$, $\alpha \in]0,1]^d$. Let $s_j = r_j + \alpha_j$ and $d/s = \sum_{j=1}^d 1/s_j$. Suppose that $s > d/4$. Let $(s_1', \dots, s_d')$ be any element of $\mathbb{R}^d$ such that $\forall\ j$, $0 < s_j' < s_j$ and such that s' defined by $d/s' = \sum_{j=1}^d 1/s_j'$ satisfies $s' > d/4$. Let $(p_i)_{i\in\mathbb{Z}^d}$ be the orthonormal Fourier basis of $\mathbb{L}^2([0,1]^d)$ and let $\hat{f}$ be the estimator of f defined by* (3.6). *Let $M_n' = \{(i_1,\dots,i_d) \in \mathbb{Z}^d;\ \forall\ j, |i_j| \le m^{s'/ds_j'}\}$ and $m \approx n^{2d/(d+4s')}$. $\hat{T}_n$ is defined as in Theorem 2 with M_n' instead of M_n. Then*

$$\sqrt{n}\left(\hat{T}_n - T(f)\right) \to \mathcal{N}(0, C(f,\phi))$$

and

$$\lim_{n\to\infty} nE\left(\hat{T}_n - T(f)\right)^2 = C(f,\phi).$$

COMMENT. We do not need the periodicity condition for the partial derivatives of f when $d = 1$ as will be shown in the next section.

3.2. *Entropy estimation.* As an example, let us give the precise shape of our estimator of $\int f\log(f)$. Condition A1 means in this case that $\forall\ x, 0 \le \varepsilon \le f(x) \le b$, that is, that f is bounded from below by some positive constant. Since f is a density, this condition implies that f is defined over a compact set. Hence, this estimator will not be suitable to test the normality, but we may use it to test the uniformity of a density defined over $[0,1]$ [see Dudewicz and Van der Meulen (1981)].

Using the Taylor expansion as before, we get

$$\int f\log(f) = -\frac{1}{2}\int\hat{f} + \int\log(\hat{f})f + \frac{1}{2}\int\frac{f^2}{\hat{f}} + \Gamma_n.$$

The general estimator proposed in Theorem 2 has the expression

$$\hat{T}_n = -\frac{1}{2}\int\hat{f} + \frac{1}{n_2}\sum_{l=1}^{n_2}\log\hat{f}(X_l) + \frac{1}{n_2(n_2-1)}\sum_{l_1\ne l_2=1}^{n_2}\sum_{i\in M_n} p_i(X_{l_1})\left(\frac{p_i}{\hat{f}}\right)(X_{l_2})$$
$$-\frac{1}{2n_2(n_2-1)}\sum_{l_1\ne l_2=1}^{n_2}\sum_{i,i'\in M_n} p_i(X_{l_1})p_{i'}(X_{l_2})\int\left(\frac{p_i p_{i'}}{\hat{f}}\right)(x)\,dx.$$

If $f \in F_{p,\alpha,C}$ and $s > d/4$, then $\hat{T}_n$ has the properties

$$\sqrt{n}\left(\hat{T}_n - T(f)\right) \to \mathscr{N}\left(0, \int \log^2(f) f - \left(\int f \log(f)\right)^2\right),$$

$$\lim_{n\to\infty} nE\left(\hat{T}_n - \int f \log(f)\right)^2 = \int \log^2(f) f - \left(\int f \log(f)\right)^2.$$

4. Proofs.

4.1. *Proof of Theorem* 1. We start from the usual decomposition:

$$E\left(\hat{\theta} - \int f^2\psi\right)^2 = \mathrm{Bias}^2(\hat{\theta}) + \mathrm{Var}(\hat{\theta}),$$

where $\mathrm{Bias}(\hat{\theta}) = E(\hat{\theta}) - \theta$. We shall write M instead of M_n for short. We recall that for any function $g \in \mathbb{L}^2(d\mu)$, $S_M g = \sum_{i\in M}(\int g p_i) p_i$. $\hat{\theta}$ has been constructed in such a way that $\mathrm{Bias}(\hat{\theta}) = -\int (S_M f - f)^2 \psi$. Hence

$$|\mathrm{Bias}(\hat{\theta})| \le \|\psi\|_\infty \int (S_M f - f)^2 = \|\psi\|_\infty \sum_{i \notin M} |a_i|^2 \le \|\psi\|_\infty \sup_{i\notin M} |c_i|^2 \quad \text{since } f \in \mathscr{E}.$$

Let us now evaluate the variance of $\hat{\theta}$. We will denote by m the cardinality of M. Let A and B be the $m \times 1$ vectors with ith components $a_i = \int f p i$ and $b_i = \int f\psi p_i$ for each i in the set M. Let Q and R be the $m \times 1$ vectors of centered functions with ith components $q_i(x) = p_i(x) - a_i$ and $r_i(x) = p_i(x)\psi(x) - b_i$. Let C the $m \times m$ matrix of constants $c_{ii'} = \int p_i p_{i'} \psi$. We shall denote by U_n the process defined by $U_n h = (1/n(n-1))\sum_{j\neq k=1}^n h(X_j, X_k)$ and we denote by P_n the empirical measure $P_n f = (1/n)\sum_{j=1}^n f(X_j)$. Then the U-statistic $\hat{\theta}$ has Hoeffding's decomposition [see Hoeffding (1948) or Serfling (1980)]

$$\hat{\theta} = U_n K + P_n L + 2A'B - A'CA,$$

where

$$K(x, y) = 2Q'(x)R(y) - Q'(x)CQ(y),$$
$$L(x) = 2A'R(x) + 2B'Q(x) - 2A'CQ(x).$$

(In fact, K and L depend on M, hence on n.) $\mathrm{Var}(\hat{\theta}) = \mathrm{Var}(U_n K) + \mathrm{Var}(P_n L) + 2\,\mathrm{Cov}(U_n K, P_n L)$. We will first bound $\mathrm{Var}(U_n K)$. Note that $U_n K$ is centered. Moreover,

$$E\left[(U_n K)^2\right] = E\left(\frac{1}{(n(n-1))^2} \sum_{j\neq k=1}^{n} \sum_{j'\neq k'=1}^{n} K(X_j, X_k) K(X_{j'}, X_{k'})\right)$$
$$= \frac{1}{(n(n-1))} E\left(K^2(X_1, X_2) + K(X_1, X_2)K(X_2, X_1)\right)$$

since all the other terms are equal to zero because Q and R are centered. By the Cauchy–Schwarz inequality, we get

$$\operatorname{Var}(U_n K) \le \frac{2}{n(n-1)} E\big(K^2(X_1, X_2)\big).$$

Moreover, using the fact that $2|E(XY)| \le E(X^2) + E(Y^2)$, we obtain

$$E\big(K^2(X_1, X_2)\big) \le 2\Big[E\Big(\big(2Q'(X_1)R(X_2)\big)^2\Big) + E\Big(\big(Q'(X_1)CQ(X_2)\big)^2\Big)\Big].$$

We will majorize these two terms:

$$\begin{aligned}
&E\Big(\big(2Q'(X_1)R(X_2)\big)^2\Big)\\
&\quad= 4\sum_{i,i'\in M}\Big(\int p_i p_{i'} f(x)\,d\mu(x) - a_i a_{i'}\Big)\Big(\int p_i p_{i'}\psi^2 f(y)\,d\mu(y) - b_i b_{i'}\Big)\\
&\quad= 4\iint\Big(\sum_{i\in M} p_i(x)p_i(y)\Big)^2 f(x)f(y)\psi^2(y)\,d\mu(x)\,d\mu(y)\\
&\qquad -4\int (S_M f)^2\psi^2 f - 4\int \big(S_M(f\psi)\big)^2 f + 4\Big(\sum_{i\in M} a_i b_i\Big)^2\\
&\quad\le 4\|f\|_\infty^2\|\psi\|_\infty^2\iint\Big(\sum_{i\in M} p_i(x)p_i(y)\Big)^2 d\mu(x)\,d\mu(y) + 4\Big(\sum_{i\in M} a_i^2\sum_{i\in M} b_i^2\Big)\\
&\quad\le 4\|f\|_\infty^2\|\psi\|_\infty^2\iint\sum_{i,i'\in M}(p_i p_{i'})(x)(p_i p_{i'})(y)\,d\mu(x)\,d\mu(y)\\
&\qquad +4\Big(\int f^2\int (f\psi)^2\Big).
\end{aligned}$$

By orthogonality of the p_i's the first term is equal to $4\|f\|_\infty^2\|\psi\|_\infty^2 m$. Moreover, since $\int f = 1$, the second term is bounded by $4\|f\|_\infty^2\|\psi\|_\infty^2$. It follows that

$$E\Big(\big(2Q'(X_1)R(X_2)\big)^2\Big) \le 4\|f\|_\infty^2\|\psi\|_\infty^2(m+1).$$

We will now control the second term appearing in $E(K^2(X_1, X_2))$:

$$\begin{aligned}
&E\Big(\big(Q'(X_1)CQ(X_2)\big)^2\Big)\\
&\quad= \sum_{i,i'\in M}\ \sum_{i_1,i_1'\in M}\Big(\int p_i p_{i_1} f - a_i a_{i_1}\Big)\Big(\int p_{i'} p_{i_1'} f - a_{i'} a_{i_1'}\Big)\int p_i p_{i'}\psi\int p_{i_1} p_{i_1'}\psi\\
&\quad= \iint\Big(\sum_{i,i'\in M} p_i(x)p_{i'}(y)\int p_i p_{i'}\psi\Big)^2 f(x)f(y)\,d\mu(x)\,d\mu(y)\\
&\qquad - 2\int \big(S_M[(S_M f)\psi]\big)^2 f + \Big(\int (S_M f)^2\psi\Big)^2
\end{aligned}$$

$$\leq \|f\|_\infty^2 \iint \left(\sum_{i, i' \in M} p_i(x) p_{i'}(y) \int p_i p_{i'} \psi \right)^2 d\mu(x)\, d\mu(y)$$

$$+ \|\psi\|_\infty^2 \left(\int (S_M f)^2 \right)^2$$

$$\leq \|f\|_\infty^2 \sum_{i, i' \in M} \sum_{i_1, i_1' \in M} \left[\iint p_i(x) p_{i'}(y) p_{i_1}(x) p_{i_1'}(y)\, d\mu(x)\, d\mu(y) \right.$$

$$\left. \times \int p_i p_{i'} \psi \int p_{i_1} p_{i_1'} \psi \right] + \|\psi\|_\infty^2 \left(\int f^2 \right)^2$$

since S_M is a projection. Using the orthogonality of the p_i's, the only terms that are not equal to zero in the first term correspond to $i = i_1$ and $i' = i_1'$. So the above term is bounded by

$$\|f\|_\infty^2 \sum_{i, i' \in M} \left(\int p_i p_{i'} \psi \right)^2 + \|\psi\|_\infty^2 \|f\|_\infty^2,$$

which is equal to

$$\|f\|_\infty^2 \iint \left(\sum_{i \in M} p_i(x) p_i(y) \right)^2 \psi(x) \psi(y)\, d\mu(x)\, d\mu(y) + \|\psi\|_\infty^2 \|f\|_\infty^2,$$

so we get

$$E\Big(\big(Q'(X_1) C Q(X_2) \big)^2 \Big) \leq \|\psi\|_\infty^2 \|f\|_\infty^2 (m+1).$$

It follows that

$$\operatorname{Var}(U_n K) \leq \frac{20}{n(n-1)} \|f\|_\infty^2 \|\psi\|_\infty^2 (m+1).$$

Let us now compute the variance of the term $P_n L$:

$$\operatorname{Var}(P_n L) = \frac{1}{n} \operatorname{Var}(L(X_1)),$$

where $L(X_1)$ can also be written in the form

$$L(X_1) = 2(S_M f)\psi(X_1) + 2S_M(f\psi)(X_1) - 2S_M[(S_M f)\psi](X_1)$$

$$- 4 \sum_{i \in M} a_i b_i + 2 \int (S_M f)^2 \psi.$$

Hence,

$$\operatorname{Var}(L(X_1))$$

$$= 4 \operatorname{Var}\big[(S_M f)\psi(X_1) + S_M(f\psi)(X_1) - S_M[(S_M f)\psi](X_1) \big]$$

$$\leq 4E\Big[\big((S_M f)\psi(X_1) + S_M(f\psi)(X_1) - S_M[(S_M f)\psi](X_1) \big)^2 \Big]$$

$$\leq 12 E\Big[\big((S_M f)\psi(X_1) \big)^2 + \big(S_M(f\psi)(X_1) \big)^2 + \big(S_M[(S_M f)\psi](X_1) \big)^2 \Big].$$

We will now control these quantities, using repeatedly the fact that $\|S_M f\|_2 \le \|f\|_2$ since S_M is a projection, and that $\|f\|_2^2 \le \|f\|_\infty$ since $\int f = 1$:

$$E\Big[\big((S_M f)\psi(X_1)\big)^2\Big] = \int (S_M f)^2 \psi^2 f \le \|\psi\|_\infty^2 \|f\|_\infty \|f\|_2^2 \le \|\psi\|_\infty^2 \|f\|_\infty^2,$$

$$E\Big[\big(S_M(f\psi)(X_1)\big)^2\Big] = \int S_M^2(f\psi) f \le \|f\|_\infty \|f\psi\|_2^2 \le \|\psi\|_\infty^2 \|f\|_\infty^2,$$

$$E\Big[\big(S_M(S_M f\psi)(X_1)\big)^2\Big] = \int S_M^2[(S_M f)\psi] f \le \|f\|_\infty \|(S_M f)\psi\|_2^2 \le \|f\|_\infty^2 \|\psi\|_\infty^2.$$

Collecting the preceding evaluations, we obtain $\mathrm{Var}(L(X_1)) \le 36\|f\|_\infty^2 \|\psi\|_\infty^2$. At last,

$$\begin{aligned}\mathrm{Cov}(U_n K, P_n L) &= E(U_n K P_n L) = E\left[\frac{1}{n(n-1)} \sum_{j \ne k = 1}^{n} K(X_j, X_k) \sum_{i=1}^{n} L(X_l)\right] \\ &= \frac{1}{n} E\big(K(X_1, X_2)(L(X_1) + L(X_2))\big) \quad \text{(since } L \text{ is centered)} \\ &= 0 \quad \text{(since } Q \text{ and } R \text{ are centered)}.\end{aligned}$$

Finally, for n large enough and for some $\gamma \in \mathbb{R}$,

$$\mathrm{Var}(\hat{\theta}) \le \gamma \|f\|_\infty^2 \|\psi\|_\infty^2 \left(\frac{m}{n^2} + \frac{1}{n}\right).$$

We recall that $\mathrm{Bias}^2(\hat{\theta}) \le \|\psi\|_\infty^2 (\sup_{i \notin M} |c_i|^2)^2$ and that, by assumption, $(\sup_{i \notin M} |c_i|^2)^2 / m \approx 1/n^2$. Hence we get the following. If $m/n \to 0$, then $E(\hat{\theta} - \theta)^2 = O(1/n)$, else $E(\hat{\theta} - \theta)^2 \le \gamma_2 (m/n^2)$, where γ_2 depends only on $\|f\|_\infty$ and $\|\psi\|_\infty$.

Let us now more precisely look at the semiparametric case, that is, the case where $E(\hat{\theta} - \theta)^2 = O(1/n)$ and show that $\sqrt{n}(\hat{\theta} - \theta) \to \mathscr{N}(0, \Lambda(f, \psi))$ in distribution. We will prove (2.4) at the same time. We recall that $\hat{\theta} = U_n K + P_n L + 2A'B - A'CA$. Hence,

$$\sqrt{n}(\hat{\theta} - \theta) = \sqrt{n}(U_n K) + \sqrt{n}(P_n L) + \sqrt{n}(2A'B - A'CA - \theta).$$

Since $\mathrm{Var}(\sqrt{n}\, U_n K) \le (20/(n-1)) \|f\|_\infty^2 \|\psi\|_\infty^2 (m+1)$, if $m/n \to 0$, then $\sqrt{n}\, U_n K$ converges to 0 in probability as $n \to \infty$.

The empirical process term $P_n L$ will make the main contribution to the central limit theorem. We saw above that

$$\mathrm{Var}(L(X_1)) = 4\,\mathrm{Var}\big[(S_M f)\psi(X_1) + S_M(f\psi)(X_1) - S_M[(S_M f)\psi](X_1)\big].$$

Denote $Y_1 = (S_M f)\psi(X_1)$, $Y_2 = S_M(f\psi)(X_1)$ and $Y_3 = -S_M[(S_M f)\psi](X_1)$. Then $\mathrm{Var}(L(X_1)) = \sum_{i,j=1}^{3} \mathrm{Cov}(Y_i, Y_j)$. In fact, $\forall\ i, j \in \{1, 2, 3\}^2$, we claim that

$$\left|\mathrm{Cov}(Y_i, Y_j) - \varepsilon_{ij}\left[\int f^3\psi^2 - \left(\int f^2\psi\right)^2\right]\right| \le \gamma\big[\|S_M f - f\|_2 + \|S_M(f\psi) - f\psi\|_2\big], \tag{4.1}$$

where $\varepsilon_{ij} = -1$ if $(i,j) = (1,3)$, $(2,3)$, $(3,1)$ or $(3,2)$ and $\varepsilon_{ij} = 1$ otherwise, and where γ depends only on $\|f\|_\infty$ and $\|\psi\|_\infty$ and is an increasing function of these two quantities. We give a complete proof for $i = j = 3$ since the computations are similar in the other cases:

$$\operatorname{Var}(Y_3) = \int S_M^2[(S_M f)\psi] f - \left(\int S_M[(S_M f)\psi] f\right)^2.$$

The computation will be done in two steps. We first bound the quantity $|\int S_M^2[(S_M f)\psi] f - \int f^3\psi^2|$. This expression is bounded by

$$\begin{aligned}
&\int |S_M^2[(S_M f)\psi] f - S_M^2(f\psi) f| + \int |S_M^2(f\psi) f - f^3\psi^2| \\
&\quad \le \|f\|_\infty \big\|S_M[(S_M f)\psi] + S_M(f\psi)\big\|_2 \big\|S_M[(S_M f)\psi] - S_M(f\psi)\big\|_2 \\
&\qquad + \|f\|_\infty \|S_M(f\psi) + f\psi\|_2 \|S_M(f\psi) - f\psi\|_2
\end{aligned}$$

by the Cauchy–Schwarz inequality.

Using repeatedly the fact that since S_M is a projection, $\|S_M g\|_2 \le \|g\|_2$, the sum is bounded by

$$\begin{aligned}
&\|f\|_\infty \|(S_M f)\psi + f\psi\|_2 \|(S_M f)\psi - f\psi\|_2 + 2\|f\|_\infty \|f\psi\|_2 \|S_M(f\psi) - f\psi\|_2 \\
&\quad \le 2\|f\|_\infty \|\psi\|_\infty^2 \|f\|_2 \|S_M f - f\|_2 + 2\|f\|_\infty \|\psi\|_\infty \|f\|_2 \|S_M(f\psi) - f\psi\|_2 \\
&\quad \le 2\|f\|_\infty^{3/2} \|\psi\|_\infty \big(\|\psi\|_\infty \|S_M f - f\|_2 + \|S_M(f\psi) - f\psi\|_2\big).
\end{aligned}$$

The second step consists in bounding the quantity $|(\int S_M[(S_M f)\psi] f)^2 - (\int f^2\psi)^2|$. This term is equal to $|(\int (S_M[(S_M f)\psi] + f\psi) f)(\int (S_M[(S_M f)\psi] - f\psi) f)|$, which by the Cauchy–Schwarz inequality is bounded by

$$\begin{aligned}
&\|f\|_2 \big\|S_M[(S_M f)\psi] + f\psi\big\|_2 \|f\|_2 \|S_M[(S_M f)\psi] - f\psi\|_2 \\
&\quad \le \|f\|_2^2 \big(\|(S_M f)\psi\|_2 + \|f\psi\|_2\big) \\
&\qquad \times \big(\big\|S_M[(S_M f)\psi] - S_M(f\psi)\big\|_2 + \|S_M(f\psi) - f\psi\|_2\big) \\
&\quad \le 2\|f\|_2^2 \|\psi\|_\infty \|f\|_2 \big(\|(S_M f)\psi - f\psi\|_2 + \|S_M(f\psi) - f\psi\|_2\big) \\
&\quad \le 2\|f\|_\infty^{3/2} \|\psi\|_\infty \big(\|\psi\|_\infty \|S_M f - f\|_2 + \|S_M(f\psi) - f\psi\|_2\big).
\end{aligned}$$

Collecting the preceding inequalities we finally get (4.1) for $i = j = 3$.

Since, by assumption, $\forall\, g \in \mathbb{L}^2(d\mu)$, $\|S_M g - g\|_2 \to 0$ as $n \to \infty$, a consequence of (4.1) is that

$$\lim_{n\to\infty} \operatorname{Var}(L(X_1)) = 4\left[\int f^3\psi^2 - \left(\int f^2\psi\right)^2\right] = \Lambda(f,\psi).$$

We shall prove now that $\sqrt{n}\, P_n L \to \mathscr{N}(0, \Lambda(f,\psi))$ in distribution. Since $\sqrt{n}\,(P_n(2f\psi) - 2\int f^2\psi) \to \mathscr{N}(0, \Lambda(f,\psi))$ in distribution, it is enough to show that the expectation of the square of

$$R = \sqrt{n}\left[P_n L - \left(P_n(2f\psi) - 2\int f^2\psi\right)\right]$$

converges to 0:

$$\begin{aligned} E(R^2) &= \mathrm{Var}(R) \\ &= n\,\mathrm{Var}(P_n L) + n\,\mathrm{Var}(P_n(2f\psi)) - 2n\,\mathrm{Cov}(P_n L, P_n(2f\psi)). \end{aligned}$$

By the preceding computations, $n\,\mathrm{Var}(P_n L) \to \Lambda(f, \psi)$. Moreover, the same result holds for $n\,\mathrm{Var}(P_n(2f\psi))$. Hence, it remains to prove that $\lim_{n\to\infty} n\,\mathrm{Cov}(P_n L, P_n(2f\psi)) = \Lambda(f, \psi)$:

$$\begin{aligned} &n\,\mathrm{Cov}(P_n L, P_n(2f\psi)) \\ &\quad = E(L(X_1)2f\psi(X_1)) \quad \text{(since } L \text{ is centered).} \\ &\quad = 4\int (S_M f) f^2\psi^2 + 4\int S_M(f\psi) f^2\psi - 4\int S_M((S_M f)\psi) f^2\psi \\ &\qquad - 8\sum_{i=0}^{m} a_i b_i \int f^2\psi + 4\int (S_M f)^2 \psi \int f^2\psi. \end{aligned}$$

This converges as $n \to \infty$ to $\Lambda(f, \psi) = 4\int f^3\psi^2 - 4(\int f^2\psi)^2$. It follows that

$$\sqrt{n}\,P_n L \to \mathscr{N}(0, \Lambda(f, \psi)) \quad \textit{in distribution}.$$

At last, let us prove that $\sqrt{n}\,(2A'B - A'CA - \theta) \to 0$. This quantity equals

$$\sqrt{n}\left(2\sum_{i\in M} a_i b_i - 2\sum_{i, i'\in M} a_i a_{i'} \int p_i p_{i'} \psi - \int f^2\psi\right).$$

This can also be written in the form

$$\begin{aligned} &\sqrt{n}\left[2\int (S_M f) f\psi - \int (S_M f)^2 \psi - \int f^2\psi\right] \\ &\quad = \sqrt{n}\left[\int (S_M f)[f\psi - (S_M f)\psi] + \int f\psi[S_M f - f]\right] \\ &\quad \le \sqrt{n}\,[\|S_M f\|_2 \|f\psi - (S_M f)\psi\|_2 + \|f\psi\|_2 \|S_M f - f\|_2] \\ &\quad \le 2\sqrt{n}\,\|f\|_2 \|\psi\|_\infty \|S_M f - f\|_2 \\ &\quad \le 2\sqrt{n}\,\|f\|_2 \|\psi\|_\infty \left(\sup_{i\notin M} c_i^2\right)^{1/2} \\ &\quad \approx 2\|f\|_2 \|\psi\|_\infty \sqrt{\frac{m}{n}}, \end{aligned}$$

which tends to zero as $n \to \infty$ since $m/n \to 0$.

Collecting the foregoing evaluations, we get (2.3). It remains to achieve the proof of (2.4). Note that

$$nE(\hat{\theta}-\theta)^2 = n\,\mathrm{Bias}^2(\hat{\theta}) + n\,\mathrm{Var}(\hat{\theta})$$
$$= n\,\mathrm{Bias}^2(\hat{\theta}) + n\,\mathrm{Var}(U_n K) + n\,\mathrm{Var}(P_n L).$$

Moreover, we proved that

$$n\,\mathrm{Bias}^2(\hat{\theta}) \le \lambda\|\psi\|_\infty \frac{m}{n} \quad \text{for some } \lambda \in \mathbb{R},$$
$$n\,\mathrm{Var}(U_n K) \le \mu\|f\|_\infty^2\|\psi\|_\infty^2 \frac{m}{n} \quad \text{for some } \mu \in \mathbb{R}.$$

At last, it follows from (4.1) that

$$|n\,\mathrm{Var}(P_n L) - \Lambda(f,\psi)| \le \gamma\left[\|S_M f - f\|_2 + \|S_M(f\psi) - f\psi\|_2\right],$$

where λ is an increasing function on $\|f\|_\infty$ and $\|\psi\|_\infty$. We finally get (2.4) and this completes the proof of Theorem 1. □

4.2. *Proof of Theorem* 2. We will first control the remainder term Γ_n:

$$|\Gamma_n| \le \tfrac{1}{6}\|\phi_1^{(3)}\|_\infty \int |f-\hat{f}|^3.$$

$\|\phi_1^{(3)}\|_\infty$ is finite, so $E(\Gamma_n^2) = O(E[(\int|f-\hat{f}|^3)^2]) = O(E[\|\hat{f}-f\|_3^6])$. Since $\hat{f}$ is assumed to satisfy condition A2, this quantity has order $O(n_1^{-6\alpha})$, where $n_1 \approx n/\log(n)$ and $\alpha > \frac{1}{6}$, so

$$E(\Gamma_n^2) = o\left(\frac{1}{n}\right).$$

This proves that the remainder term is negligible.

We are now going to prove the asymptotic efficiency. Let

$$R = \sqrt{n}\left[\hat{T}_n - T(f) - \frac{1}{n_2}\sum_{l=1}^{n_2}\phi_1'(f,\cdot)(X_l) - \int \phi_1'(f,\cdot)f\right].$$

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

$$R_1 = \sqrt{n}\left[\hat{T}_n - T(f) - \frac{1}{n_2}\sum_{l=1}^{n_2}\phi_1'(\hat{f},\cdot)(X_l) - \int \phi_1'(\hat{f},\cdot)f\right],$$
$$R_2 = \sqrt{n}\left[\frac{1}{n_2}\sum_{l=1}^{n_2}\left(\phi_1'(\hat{f},\cdot)(X_l) - \int \phi_1'(\hat{f},\cdot)f\right)\right]$$
$$-\sqrt{n}\left[\frac{1}{n_2}\sum_{l=1}^{n_2}\left(\phi_1'(f,\cdot)(X_l) - \int \phi_1'(f,\cdot)f\right)\right].$$

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

$$R_1 = \sqrt{n}\left[\hat{L}' - L' + \hat{Q}' - Q' + \Gamma_n\right],$$

where

$$L' = -\int \hat{f}\phi_1''(\hat{f},\cdot)f; \qquad \hat{L}' = -\frac{1}{n_2}\sum_{j=1}^{n_2} \hat{f}\phi_1''(\hat{f},\cdot)(X_j); \qquad Q' = \tfrac{1}{2}\int \phi_1''(\hat{f},\cdot)f^2$$

and $\hat{Q}'$ is the corresponding estimator. Since $E(\Gamma_n^2) = o(1/n)$, we just have to control the expectation of the square of $\sqrt{n}\,[\hat{L}' - L' + \hat{Q}' - Q']$.

Computation of $\lim_{n\to\infty} nE(\hat{L}' - L')^2$. Note that

$$nE\left[(\hat{L}' - L')^2 | \hat{f}\right] = \frac{n}{n_2}\left[\int\left(\hat{f}\phi_1''(\hat{f},\cdot)\right)^2 f - \left(\int \hat{f}\phi_1''(\hat{f},\cdot)f\right)^2\right],$$

and $n/n_2 \to 1$ as $n \to \infty$. Moreover, we will show that the expectation of the preceding expression converges toward the same expression with f instead of $\hat{f}$. We shall only give the proof for the term $\int(\hat{f}\phi_1''(\hat{f},\cdot))^2 f$. Observe that

$$\left|E\left(\int\left(\hat{f}\phi_1''(\hat{f},\cdot)\right)^2 f - \int f^3(\phi_1'')^2(f,\cdot)\right)\right|$$

$$\leq \|f\|_\infty E\left(\int |\hat{f}\phi_1''(\hat{f},\cdot) + f\phi_1''(f,\cdot)|\,|\hat{f}\phi_1''(\hat{f},\cdot) - f\phi_1''(f,\cdot)|\right).$$

We recall that $a \leq f(x) \leq b$ and $a - \varepsilon \leq \hat{f}(x) \leq b + \varepsilon$ for some $\varepsilon > 0$, so the foregoing difference is bounded by

$$\|f\|_\infty\left[2(\sup(|a|,|b|) + \varepsilon)\|\phi_1''\|_\infty\right] E\left(|\hat{f}\phi_1''(\hat{f},\cdot) - f\phi_1''(f,\cdot)|\right).$$

At last,

$$E\left(|\hat{f}\phi_1''(\hat{f},\cdot) - f\phi_1''(f,\cdot)|\right) \leq E\left(|\hat{f}\phi_1''(\hat{f},\cdot) - \hat{f}\phi_1''(f,\cdot)|\right) + E\left(|\hat{f}\phi_1''(f,\cdot) - f\phi_1''(f,\cdot)|\right).$$

Since $E(\|\hat{f} - f\|_1) \to 0$ and since $\|\phi_1^{(3)}\|_\infty$ and $\|\phi_1''\|_\infty$ are finite, each of these terms converges to zero and we get

$$\lim_{n\to\infty} E(\hat{L}' - L')^2 = \int(\phi_1''(f,\cdot))^2 f^3 - \left(\int \phi_1''(f,\cdot)f^2\right)^2.$$

Computation of $\lim_{n\to\infty} nE(\hat{Q}' - Q')^2$. It follows from (2.4) that if $|M_n|/n \to 0$, then

$$\left|nE\left[(\hat{Q}' - Q')^2|\hat{f}\right] - \left[\int\left(\phi_1''(\hat{f},\cdot)\right)^2 f^3 - \left(\int \phi_1''(\hat{f},\cdot)f^2\right)^2\right]\right|$$

$$\leq \gamma_1(\|f\|_\infty, \|\psi\|_\infty)\left[\frac{|M|}{n} + \|S_M f - f\|_2 + \left\|S_M\left(f\phi_1''(\hat{f},\cdot)\right) - f\phi_1''(\hat{f},\cdot)\right\|_2\right].$$

Hence,

$$\left| nE\left[(\hat{Q}-Q)^2\right] - E\left[\int\left(\phi_1''(\hat{f},\cdot)\right)^2 f^3 - \left(\int \phi_1''(\hat{f},\cdot)f^2\right)^2\right]\right|$$
$$\leq \gamma_1(\|f\|_\infty, \|\psi\|_\infty)\left[\frac{|M|}{n} + \|S_M f - f\|_2 + E\left(\left\|S_M\left(f\phi_1''(\hat{f},\cdot)\right) - f\phi_1''(\hat{f},\cdot)\right\|_2\right)\right].$$

We have to prove that $E(\|S_M(f\phi_1''(\hat{f},\cdot)) - f\phi_1''(\hat{f},\cdot)\|_2) \to 0$ as $n \to \infty$. Since, by similar arguments as before, the expectation of the term $\int(\phi_1''(\hat{f},\cdot))^2 f^3 - (\int\phi_1''(\hat{f},\cdot)f^2)^2$ converges as $n \to \infty$ to $\int(\phi_1''(f,\cdot))^2 f^3 - (\int\phi_1''(f,\cdot)f^2)^2$, this will imply that $nE(\hat{Q}-Q)^2$ converges to the same limit. Note that

$$\begin{aligned}
&E\left(\left\|S_M\left(f\phi_1''(\hat{f},\cdot)\right) - f\phi_1''(\hat{f},\cdot)\right\|_2\right)\\
&\quad\leq E\left(\left\|S_M\left(f\phi_1''(\hat{f},\cdot)\right) - S_M\left(f\phi_1''(f,\cdot)\right)\right\|_2\right)\\
&\qquad + \left\|S_M\left(f\phi_1''(f,\cdot)\right) - f\phi_1''(f,\cdot)\right\|_2 + E\left(\|f\phi_1''(f,\cdot) - f\phi_1''(\hat{f},\cdot)\|_2\right)\\
&\quad\leq 2E\left(\|f\phi_1''(f,\cdot) - f\phi_1''(\hat{f},\cdot)\|_2\right) + \|S_M\left(f\phi_1''(f,\cdot)\right) - f\phi_1''(f,\cdot)\|_2
\end{aligned}$$

since S_M is a projection.

The second term converges to zero since $f\phi_1''(f,\cdot) \in \mathbb{L}^2(d\mu)$ and since $\forall g \in \mathbb{L}^2(d\mu)$, $\|S_M g - g\|_2 \to 0$. Moreover, by similar arguments as in the computation of $nE(\hat{L}' - L')^2$, the first term converges also to zero.

At last, we will see that the sum of the two terms we have just calculated is compensated by the covariance term.

Computation of $\lim_{n\to\infty} 2nE(\hat{L}' - L')(\hat{Q}' - Q')$. Since conditionally to $\hat{f}$, $\hat{L}'$ is an unbiased estimator of L', we get

$$E\left[(\hat{L}' - L')(\hat{Q}' - Q')|\hat{f}\right] = E\left(\hat{L}'\hat{Q}'|\hat{f}\right) - L'E\left(\hat{Q}'\,|\,\hat{f}\right).$$

We recall that $\hat{Q}' = 2\hat{Q}_1 - \hat{Q}_2$, where

$$\hat{Q}_1 = \frac{1}{n_2(n_2-1)}\sum_{i\in M}\sum_{j\neq k=1}^{n} p_i(X_j)\frac{p_i\phi_1''(\hat{f},\cdot)}{2}(X_k),$$

$$\hat{Q}_2 = \frac{1}{n_2(n_2-1)}\sum_{i,i'\in M}\sum_{j\neq k=1}^{n} p_i(X_j)p_{i'}(X_k)\int p_i p_{i'}\frac{\phi_1''(\hat{f},\cdot)}{2}(x)\,dx.$$

We shall only develop the calculations for $\hat{Q}_1$:

$$\begin{aligned}
E\left(\hat{L}'\hat{Q}_1|\hat{f}\right) = &-\frac{1}{n_2}\sum_{i\in M}\int \hat{f}\phi_1''(\hat{f},\cdot)fp_i\int\phi_1''(\hat{f},\cdot)\frac{fp_i}{2}\\
&-\frac{1}{n_2}\sum_{i\in M}a_i\int\hat{f}\left(\phi_1''(\hat{f},\cdot)\right)^2\frac{fp_i}{2} - \left(1-\frac{2}{n_2}\right)L'E\left(\hat{Q}_1|\hat{f}\right).
\end{aligned}$$

So

$$E\left(\hat{L}'\hat{Q}_1|\hat{f}\right) - L'E\left(\hat{Q}_1|\hat{f}\right) = -\frac{1}{2n_2}\int S_M\left[f\phi_1''(\hat{f},\cdot)\right]\hat{f}\phi_1''(\hat{f},\cdot)f$$
$$-\frac{1}{2n_2}\int (S_M f)\hat{f}\left(\phi_1''(\hat{f},\cdot)\right)^2 + \frac{2}{n_2}L'E\left(\hat{Q}_1|\hat{f}\right).$$

Let us show that

$$E\left(S_M\left[f\phi_1''(\hat{f},\cdot)\right]\hat{f}\phi_1''(\hat{f},\cdot)f\right) \to \int f^3(\phi_1''(f,\cdot))^2.$$

Of course, analogous results are valid for the other terms. We shall first show that the expectation of

$$\int S_M\left[f\phi_1''(\hat{f},\cdot)\right]\left[\hat{f}\phi_1''(\hat{f},\cdot)f - f^2\phi_1''(f,\cdot)\right] \to 0.$$

This integral is bounded by

$$\left\|S_M\left(f\phi_1''(\hat{f},\cdot)\right)\right\|_2 \|\hat{f}\phi_1''(\hat{f},\cdot) - f\phi_1''(f,\cdot)\|_2\|f\|_\infty.$$

Moreover,

$$\left\|S_M\left(f\phi_1''(\hat{f},\cdot)\right)\right\|_2 \le \|f(\phi_1'')(\hat{f},\cdot)\|_2 \le \|f\|_2\|\phi_1''\|_\infty \le \|f\|_\infty^{1/2}\|\phi_1''\|_\infty$$

and as was already shown, the expectation of the term $\|\hat{f}\phi_1''(\hat{f},\cdot) - f\phi_1''(f,\cdot)\|_2$ converges to zero. It remains to show that the expectation of

$$\int\left(S_M\left[f\phi_1''(\hat{f},\cdot)\right] - f\phi_1''(f,\cdot)\right)f^2\phi_1''(f,\cdot)$$

converges toward zero. This expression is bounded by

$$\|\phi_1''f\|_\infty\|f\|_2\|S_M\left[f\phi_1''(\hat{f},\cdot)\right] - f\phi_1''(f,\cdot)\|_2.$$

So, we just have to prove that

$$E\left(\|S_M\left[f\phi_1''(\hat{f},\cdot)\right] - f\phi_1''(f,\cdot)\|_2\right) \to 0.$$

This quantity is bounded by

$$E\left(\|S_M\left[f\phi_1''(\hat{f},\cdot)\right] - S_M[f\phi_1''(f,\cdot)]\|_2\right) + \|S_M[f\phi_1''(f,\cdot)] - f\phi_1''(f,\cdot)\|_2$$
$$\le E\left(\|f\phi_1''(\hat{f},\cdot) - f\phi_1''(f,\cdot)\|_2\right) + \|S_M(f\phi_1''(f,\cdot)) - f\phi_1''(f,\cdot)\|_2.$$

Since each of these terms converges to zero, we get

$$\lim_{n\to\infty} nE(\hat{L}' - L')\left(\hat{Q}_1 - Q_1\right) = -\int(\phi_1''(f,\cdot))^2 f^3 + \left(\int\phi_1''(f,\cdot)f^2\right)^2.$$

The same result holds with $\hat{Q}$ instead of $\hat{Q}_1$, so $\lim_{n\to\infty} E(R_1^2) = 0$.

It remains to prove that $E(R_2^2) \to 0$:

$$E(R_2^2) = \frac{n}{n_2} E\left[\int \left(\phi_1'(\hat{f},\cdot) - \phi_1'(f,\cdot)\right)^2 f\right] - \frac{n}{n_2} E\left[\int \phi_1'(\hat{f},\cdot) f - \int \phi_1'(f,\cdot) f\right]^2.$$

Using as before the fact that $E(\|\hat{f} - f\|_1) \to 0$ and that $\|\phi_1'\|_\infty$ is bounded, it is easy to see that $E(R_2^2) \to 0$. This achieves the proof of Theorem 2. □

4.3. *Proof of Corollary* 1. Let us first show that condition A2 is satisfied by the preliminary estimator $\hat{f}$ defined by (3.6) as soon as $s > d/4$. This estimator is based on the kernel estimator $\tilde{f}$ studied by Ibragimov and Khas'minskii, who showed the following properties for $\tilde{f}$:

$$\sup_{f \in F_{r,\alpha,C}} E_f\left(\|\tilde{f} - f\|_q^l\right) \le A_1(q,l) n_1^{-ls/(d+2s)} \qquad \forall\, 2 \le q < +\infty, \forall\, l \in \mathbb{N}^*,$$

$$\sup_{f \in F_{r,\alpha,C}} E_f\left(\|\tilde{f} - f\|_\infty^l\right) \le A_2(l) \left(\frac{n_1}{\log(n_1)}\right)^{-ls/(d+2s)} \qquad \forall\, l \in \mathbb{N}^*.$$

We recall that $\hat{f} = \tilde{f}\mathbb{1}_{A_n} + f_0 \mathbb{1}_{A_n^c}$, where $A_n = \{\tilde{f} \subset [a-\varepsilon, b+\varepsilon]\}$:

$$E\left(\|\hat{f} - f\|_q^l\right) \le E\left(\|\tilde{f} - f\|_q^l\right) + \|f_0 - f\|_q^l P(A_n^c).$$

Hence, to prove condition A2, it is enough to show that $P(A_n^c)$ is small enough. We notice that $A_n^c \subset \{\|\tilde{f} - f\|_\infty \le \varepsilon\}$. Since

$$E_f \|\tilde{f} - f\|_\infty^{l'} \le A_2(l') \left(\frac{n_1}{\log(n_1)}\right)^{-l's/(d+2s)} \qquad \forall\, l' \in \mathbb{N}^*,$$

we get

$$P(A_n^c) \le P\left(\|\tilde{f} - f\|_\infty \ge \varepsilon\right) \le \frac{1}{\varepsilon^{l'}} E_f\left(\|\tilde{f} - f\|_\infty^{l'}\right),$$

so

$$P(A_n^c) \le A_3(l') \left(\frac{n_1}{\log(n_1)}\right)^{-l's/(d+2s)} \qquad \forall\, l' \in \mathbb{N}^*.$$

For n_1 large enough,

$$P(A_n^c) \le A_4(l) n_1^{-ls/(d+2s)}.$$

It follows that

$$E_f\left(\|\hat{f} - f\|_q^l\right) \le C(q,l) n_1^{-ls/(d+2s)}.$$

This proves condition A2 since when $s > d/4$, $s/(d+2s) > \frac{1}{6}$.

Let us now prove the following lemma.

LEMMA 1. *Let $f \in F_{r,\alpha,C}$ with $r \in \mathbb{N}^d$ and $\alpha \in]0,1]^d$. Then f belongs to the ellipsoid*

$$\mathscr{E} = \left\{ \sum_{i \in \mathbb{Z}^d} a_i p_i; \sum_{(i_1,\dots,i_d)\in\mathbb{Z}^d} \left(|i_1|^{2s'_1} + \cdots + |i_d|^{2s'_d}\right)|a_{i_1,\dots,i_d}|^2 \le \gamma \right\}$$

for all $(s'_1,\dots,s'_d)$ such that $\forall\, j$, $0 < s'_j < s_j = r_j + \alpha_j$ for some γ independent of $f \in F_{r,\alpha,C}$ and where $(p_i)_{i\in\mathbb{Z}^d}$ denotes the Fourier orthonormal basis of $\mathbb{L}^2([0,1]^d)$.

PROOF. We have to show that $\forall\ j \in \{1,\dots,d\}$ the quantity $\sum_{(i_1,\dots,i_d)\in\mathbb{Z}^d}(|i_j|^{2s'_j})\,|a_{i_1,\dots,i_d}|^2$ is bounded by some constant γ_j independent of $f \in F_{r,\alpha,C}$. Since $D_j^l f$ is periodic for $l = 0,\dots,r_j$, it is easy to see by integration by parts that

$$\left|\int (D_j^{r_j} f) p_i\right| = (2\pi)^{r_j}|i_j|^{r_j}\left|\int f p_i\right| = (2\pi)^{r_j}|i_j|^{r_j}|a_i|.$$

Using this remark, it is enough to show that if f satisfies the condition

$$|f(x_1,\dots,x_j+h,\dots,x_d) - f(x_s,\dots,x_d)| \le C|h|^{2\alpha_j},$$

then

$$\forall\, j \in \{1,\dots,d\}, \qquad \sum_{(i_1,\dots,i_d)\in\mathbb{Z}^d} \left(|i_j|^{2\beta_j}\right)|a_{i_1 \cdots i_d}|^2 \le \gamma_j \qquad \forall\, 0 < \beta_j < \alpha_j.$$

The Fourier expansion of $f(x_1,\dots,x_j+h,\dots,x_d) - f(x_1,\dots,x_j-h,\dots,x_d)$ is

$$\sum_{(i_1,\dots,i_d)\in\mathbb{Z}^d} a_{i_1 \cdots i_d} \exp(2i\pi(i_1x_1 + \cdots + i_dx_d))2i\sin(2\pi i_j h).$$

Hence

$$\begin{aligned}\int &\left(f(x_1,\dots,x_j+h,\dots,x_d) - f(x_1,\dots,x_j-h,\dots,x_d)\right)^2 \\ &= 4\sum_{(i_1,\dots,i_d)\in\mathbb{Z}^d} |a_{i_1,\dots,i_d}|^2 \sin^2(2\pi i_j h) \\ &\le C^2|h|^{2\alpha_j}.\end{aligned}$$

We shall denote by $(i_1,\dots,\hat{i}_j,\dots,i_d)$ the element $(i_1,\dots,i_{j-1},i_{j+1},\dots,i_d)$ of $\mathbb{Z}^{d-1}$. Let $q \in \mathbb{N}^*$ and $h = 1/8q$. Then, $\forall\ i_j \in \{q,\dots,2q-1\}$, $\sin^2(i_j h) \ge \frac{1}{2}$ and

$$\sum_{i_j=q}^{2_q-1} \sum_{(i_1,\dots,\hat{i}_j,\dots,i_d)\in\mathbb{Z}^{d-1}} |a_{i_1,\dots,i_d}| \le \frac{C^2}{2}\frac{1}{(8q)^{2\alpha_j}}.$$

By analogous arguments,

$$\sum_{i_j=-2q+1}^{-q} \sum_{(i_1,\dots,\hat{i}_j,\dots,i_d)\in\mathbb{Z}^{d-1}} |a_{i_1,\dots,i_d}|^2 \le \frac{C^2}{2}\frac{1}{(8q)^{2\alpha_j}}.$$

Let $0 < \beta_j < \alpha_j$,

$$\sum_{i_j=q}^{2q-1} \sum_{(i_1,\dots,\hat{i}_j,\dots,i_d)\in\mathbb{Z}^{d-1}} |i_j|^{2\beta_j}|a_{i_1,\dots,i_d}|^2 \le \frac{C^2}{2}\frac{(2q)^{2\beta_j}}{(8q)^{2\alpha_j}}.$$

Now, let $q = 2^l$,

$$\sum_{(i_1,\dots,i_d)\in\mathbb{Z}^d} |a_{i_1,\dots,i_d}|^2 \le C^2 \frac{4^{\beta_j}}{8^{2\alpha_j}} \sum_{l=0}^{+\infty} \left(\frac{1}{4^{\alpha_j-\beta_j}}\right)^l.$$

This proves the lemma.

For $(s'_1,\dots,s'_j)$ well chosen as prescribed in Corollary 1, we can apply the results of Theorem 2. □

COMMENT. Of course, if $d = 1$, we do not need to assume that f is periodic. In fact, we can use the same proof as before (except that we do not integrate by parts) to show that, for some $\alpha > \frac{1}{4}$, f belongs to the ellipsoid

$$\mathscr{E} = \left\{ \sum a_i p_i, \sum_{i\in\mathbb{Z}} |i|^{2s}|a_i|^2 \le \gamma \right\} \quad \text{for some constant } \gamma.$$

APPENDIX

Semiparametric information bound. In order to determine the semiparametric Cramér–Rao bound for the problem of estimating $\int\phi(f,\cdot)$ (for $f \in \mathscr{E}$), we are going to apply the results of Ibragimov and Khas'minskii (1991). Similar results can be found in Koshevnik and Levit (1976).

We first have to determine the Fréchet derivative of the functional $T(f) = \int\phi(f,\cdot)\,d\mu$ at a point f_0 belonging to our nonparametric set of densities. Since

$$T(f) - T(f_0) = \int \phi'_1(f_0,\cdot)(f-f_0) + O\left(\int (f-f_0)^2\right),$$

the Fréchet derivative of T at the point f_0 is

$$T'(f_0)\cdot h = \langle \phi'_1(f_0,\cdot), h\rangle,$$

where $\langle\cdot,\cdot\rangle$ denotes the scalar product in $\mathbb{L}^2(d\mu)$. Denote $H(f_0) = \{h \in \mathbb{L}^2(d\mu), \int h\sqrt{f_0}\,d\mu = 0\}$, and $A_n(g) = (1/\sqrt{n})(\sqrt{f_0})g$. $P_{f_0}^{(n)}$ denotes the joint distribution of $(X_1,\dots,X_n)$ under f_0. Since $X_1,\dots,X_n$ are i.i.d., the family $\{P_f^{(n)}, f\in\mathscr{E}\}$ is locally asymptotically normal at all points $f_0 \in \mathscr{E}$ in the direction $H(f_0)$ with normalizing factor $A_n(f_0)$. Under this condition, Ibragimov and Khas'minskii have shown that if we set $K_n = B_n T'(f_0) A_n P_{H(f_0)}$ [where $B_n(h) = \sqrt{n}\,h$], then, if $K_n \to K$ weakly and if $K(h) = \langle g, h\rangle$, then $\forall\ \hat{T}_n$ (estimator of $T(f)$), $\forall\ \{\mathscr{V}(f_0)\}$ family of vicinities of f_0,

$$\inf_{\{\mathscr{V}(f_0)\}} \liminf_{n\to\infty} \sup_{f\in\mathscr{V}(f_0)} nE\big(\hat{T}_n - T(f_0)\big)^2 \ge \|g\|^2_{\mathbb{L}^2(d\mu)}.$$

In our setting, this gives

$$K_n(h) = K(h) = T'(f_0)\cdot\left(\sqrt{f_0}\left(h - \sqrt{f_0}\int h\sqrt{f_0}\right)\right),$$

$$K(h) = \int \phi_1'(f_0,\cdot)\sqrt{f_0}\,h - \int \phi_1'(f_0,\cdot) f_0 \int h\sqrt{f_0}\,,$$

so $K(h) = \langle g, h\rangle$, where

$$g = \phi_1'(f_0,\cdot)\sqrt{f_0} - \left[\int \phi_1'(f_0,\cdot) f_0\right]\sqrt{f_0}\,.$$

Hence, the semiparametric Cramér–Rao bound in our problem equals $\|g\|^2_{\mathbb{L}^2(d\mu)}$, which is also equal to

$$\int (\phi_1'(f_0,\cdot))^2 f_0 - \left(\int \phi_1'(f_0,\cdot) f_0\right)^2.$$

For the problem of estimating $\int f^2\psi$, we get for the analogue of the Cramér–Rao bound

$$4\left[\int f^3\psi^2 - \left(\int f^2\psi\right)^2\right].$$

Acknowledgments. The author wishes to express his thanks to Pascal Massart and to Lucien Birgé for their helpful suggestions and for a meticulous reading of the manuscript. Thanks are also due an Associate Editor and referee for their constructive suggestions.

REFERENCES

BICKEL, P. J. and RITOV, Y. (1988). Estimating integrated squared density derivatives: sharp best order of convergence estimates. *Sankhyā Ser. A* **50** 381–393.

BIRGÉ, L. and MASSART, P. (1995). Estimation of integral functionals of a density. *Ann. Statist.* **23** 11—29.

DONOHO, D. and NUSSBAUM, M. (1990). Minimax quadratic estimation of a quadratic functional. *J. Complexity* **6** 290–323.

DUDEWICZ, E. J. and VAN DER MEULEN, E. C. (1981). Entropy-based tests of uniformity. *J. Amer. Statist. Assoc.* **76** 967–974.

HALL, P. and MARRON, J. S. (1987). Estimation of integrated squared density derivatives. *Statist. Probab. Lett.* **6** 109–115.

HOEFFDING, W. (1948). A class of statistics with asymptotically normal distribution. *Ann. Math. Statist.* **19** 293–325.

IBRAGIMOV, I. A. and KHAS'MINSKII, R. Z. (1980). On the estimation of distribution density. Investigation in the mathematical statistics. *Zap. Naučn Sem. Leningrad. Otdel. Mat. Inst. Steklov* **98** 61–85.

IBRAGIMOV, I. A. and KHAS'MINSKII, R. Z. (1981). On the nonparametric density estimates. Investigation in the mathematical statistics. *Zap. Naučn. Sem. Leningrad. Otdel. Mat. Inst. Steklov* **108** 73–89.

IBRAGIMOV, I. A. and KHAS'MINSKII, R. Z. (1991). Asymptotically normal families of distributions and efficient estimation. *Ann. Statist.* **19** 1681–1724.

IBRAGIMOV, I. A., NEMIROVSKII, A. S. and KHAS'MINSKII, R. Z. (1986). Some problems on nonparametric estimation in Gaussian white noise. *Theory Probab. Appl.* **31** 391–406.

KERKYACHARIAN, G. and PICARD, D. (1993). Estimation of the integral of the cube of the density. Technical report, Univ. Paris VII.

KOSHEVNIK, YU. A. and LEVIT, B. YA. (1976). On a non parametric analogue of the information matrix. *Theory Probab. Appl.* **21** 738–753.

LEVIT, B. YA. (1978). Asymptotically efficient estimation of nonlinear functionals. *Problems Inform. Transmission* **14** 204–209.

MEYER, Y. (1990). *Ondelettes et Opérateurs*. Hermann, Paris.

SERFLING, R. J. (1980). *Approximation Theorems of Mathematical Statistics*. Wiley, New York.

VASICEK, O. (1976). A test of normality based on sample entropy. *J. Roy. Statist. Soc. Ser. B* **38** 54–59.

UNIVERSITÉ PARIS SUD
MATHEMATIQUE BAT. 425
91405 ORSAY CEDEX
FRANCE