



HAL
open science

Efficient Estimation of Nonlinear Conditional Functionals of a Density

Sébastien da Veiga, Fabrice Gamboa

► **To cite this version:**

Sébastien da Veiga, Fabrice Gamboa. Efficient Estimation of Nonlinear Conditional Functionals of a Density. 2008. hal-00266110v1

HAL Id: hal-00266110

<https://hal.science/hal-00266110v1>

Preprint submitted on 24 Mar 2008 (v1), last revised 13 Mar 2012 (v2)

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

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

EFFICIENT ESTIMATION OF NONLINEAR CONDITIONAL FUNCTIONALS OF A DENSITY

BY SÉBASTIEN DA VEIGA

IFP-Lyon, France and Institut de Mathématiques, Toulouse, France
AND

BY FABRICE GAMBOA

Institut de Mathématiques, Toulouse, France

In this paper we address the problem of estimating functional integrals of conditional moments of the form $\mathbb{E}(\psi(\mathbb{E}(\varphi(Y)|X)))$ where (X, Y) is a random vector with joint density f and ψ and φ are functions that are differentiable enough. This problem is motivated by the asymptotically efficient estimation of Sobol sensitivity indices.

1. Introduction. In many applied fields, physicists and engineers are faced with the problem of estimating some sensitivity indices. These indices quantify the impact of some input variables on an output. The general situation may be formalized as follows.

The output Y is a nonlinear regression of input variables $\boldsymbol{\tau} = (\tau_1, \dots, \tau_l)$ ($l \geq 1$ is generally large). This means that Y and $\boldsymbol{\tau}$ satisfy the input-output relationship

$$(1) \quad Y = \Phi(\boldsymbol{\tau})$$

where Φ is a known nonlinear function. Usually, Φ is complicated and has not a closed form, but it may be computed through a computer code ([13]). The estimation of some sensitivity indices occurs for example in chemistry ([2], [15]), space science ([1]), etc. In general, the input $\boldsymbol{\tau}$ is modelled by a random vector, so that Y is also a random variable. A common way to quantify the impact of input variables is to use the so-called Sobol sensitivity indices ([14]). Assuming that all the random variables are square integrable, the Sobol index for the input τ_j ($j = 1, \dots, l$) is

$$(2) \quad \Sigma_j = \frac{\text{Var}(\mathbb{E}(Y|\tau_j))}{\text{Var}(Y)}.$$

AMS 2000 subject classifications: Primary 62G20, ; secondary 62G06, 62G07, 62P30

Keywords and phrases: Density Estimation, semiparametric Cramér-Rao bound, global sensitivity analysis

Observing an i.i.d. sample $(Y_1, \boldsymbol{\tau}^{(1)}), \dots, (Y_n, \boldsymbol{\tau}^{(n)})$ (with $Y_i = \Phi(\boldsymbol{\tau}^{(i)})$, $i = 1, \dots, n$), an interesting statistical problem is then to estimate Σ_j ($j = 1, \dots, l$). Many past and recent works deal with this problem, see [2], [14], [12], [13] or [3]. In these papers the approach is generally empirical in the sense that no proof of convergence for the studied estimators is given (except in [3]).

Obviously, (2) may be rewritten as

$$\Sigma_j = \frac{\mathbb{E}(\mathbb{E}(Y|\tau_j)^2) - \mathbb{E}(Y)^2}{\text{Var}(Y)}.$$

Thus, in order to estimate Σ_j , the hard part is $\mathbb{E}(\mathbb{E}(Y|\tau_j)^2)$. In this paper we will provide an asymptotically efficient estimate for this kind of quantity. More precisely we will tackle the problem of asymptotically efficient estimation of some general nonlinear functional. This functional is built on a density of a pair. Our approach follows the works of Laurent ([8], [9]) where the author was interested in the estimation of less complicated functionals. We also refer to [10] and [7] for general results on nonlinear functionals estimation. Let us specify the functionals we are interested in. Let $(Y_1, X_1), \dots, (Y_n, X_n)$ be a sample of i.i.d. random vectors of \mathbb{R}^2 having a *regular* density f (see Section 2 for the precise frame). We will study the estimation of the nonlinear functional

$$\begin{aligned} T(f) &= \mathbb{E}\left(\psi(\mathbb{E}(\varphi(Y)|X))\right) \\ &= \iint \psi\left(\frac{\int \varphi(y)f(x,y)dy}{\int f(x,y)dy}\right) f(x,y)dx dy \end{aligned}$$

where ψ and φ are regular functions. Hence, the Sobol indices are the particular case obtained with $\psi(\xi) = \xi^2$ and $\varphi(\xi) = \xi$.

The method developed in order to obtain an asymptotically efficient estimate for $T(f)$ is the same as the one developed by Laurent ([8]). Roughly speaking, it involves a preliminary estimate \hat{f} of f built on a small part of the sample. This preliminary estimate is used in a Taylor expansion of $T(f)$ up to the second order in a neighbourhood of \hat{f} . This expansion allows to remove the bias that occurs when using a direct plug-in method. Hence, the bias correction involves a quadratic functional of f . Due to the form of T , this quadratic functional of f may be written as

$$\theta(f) = \iiint \eta(x, y_1, y_2) f(x, y_1) f(x, y_2) dx dy_1 dy_2.$$

This kind of functional does not fall in the frame treated in [8] and have not been studied to the best of our knowledge. We study this problem in

Section 3 where we build an asymptotically efficient estimate for θ . In order to show that our estimates of $T(f)$ and $\theta(f)$ are asymptotically efficient, we calculate the semiparametric Cramér-Rao lower bound using the general frame of Ibragimov and Khas'minskii ([6]).

The paper is organized as follows. In the next section, we precisely give the frame and our assumptions. We then begin to build an estimate for $T(f)$. This first path leads to the problem of estimating $\theta(f)$, which is completely tackled in Section 3. The second path for estimating $T(f)$ is performed in Section 4 where the main theorem (Theorem 3) gives the asymptotic behavior of our estimate. In Section 5 we focus on the particular case where we wish to estimate Sobol indices, which was our first motivation. We also study some numerical examples showing the good behavior of the estimate. Section 6 is a conclusion, we discuss here some possible extensions of the technique. All proofs are postponed in Section 7.

2. Model frame and method. Let $a < b$ and $c < d$, $\mathbb{L}^2(dx dy)$ will denote the set of square integrable functions on $[a, b] \times [c, d]$. Further, $\mathbb{L}^2(dx)$ (resp. $\mathbb{L}^2(dy)$) will denote the set of square integrable functions on $[a, b]$ (resp. $[c, d]$). For sake of simplicity, we work in the whole paper with the Lebesgue measure as reference measure. Nevertheless, most of the results presented can be obtained for a general reference measure on $[a, b] \times [c, d]$. Let $(\alpha_{i_\alpha}(x))_{i_\alpha \in D_1}$ (resp. $(\beta_{i_\beta}(y))_{i_\beta \in D_2}$) be a countable orthonormal basis of $\mathbb{L}^2(dx)$ (resp. of $\mathbb{L}^2(dy)$). We set $p_i(x, y) = \alpha_{i_\alpha}(x)\beta_{i_\beta}(y)$ with $i = (i_\alpha, i_\beta) \in D := D_1 \times D_2$. Obviously $(p_i(x, y))_{i \in D}$ is a countable orthonormal (tensor) basis of $\mathbb{L}^2(dx dy)$. We will also use the following subset of $\mathbb{L}^2(dx dy)$:

$$\mathcal{E} = \left\{ \sum_{i \in D} e_i p_i : (e_i)_{i \in D} \text{ is a sequence with } \sum_{i \in D} \left| \frac{e_i}{c_i} \right|^2 \leq 1 \right\},$$

here $(c_i)_{i \in D}$ is a given fixed positive sequence.

Let (X, Y) having a bounded joint density f on $[a, b] \times [c, d]$ from which we have a sample $(X_i, Y_i)_{i=1, \dots, n}$. We will also assume that f lies in the ellipsoid \mathcal{E} . Recall that we wish to estimate a conditional functional

$$\mathbb{E}(\psi(\mathbb{E}(\varphi(Y)|X)))$$

where φ is a measurable bounded function with $\chi_1 \leq \varphi \leq \chi_2$ and $\psi \in C^3([\chi_1, \chi_2])$ the set of thrice continuously differentiable functions on $[\chi_1, \chi_2]$. This last quantity can be expressed in terms of an integral depending on the joint density f :

$$(3) \quad T(f) = \iint \psi \left(\frac{\int \varphi(y) f(x, y) dy}{\int f(x, y) dy} \right) f(x, y) dx dy.$$

To build our estimator, we propose here to follow the work of Laurent ([8]) on the asymptotically efficient estimation of integral functionals of a density. Denoting by $(Z_i)_{i=1,\dots,n}$ a sample of size n of a d -dimensional random vector Z with density $s \in \mathbb{L}^2(\mathbb{R}^d)$, Laurent gives in her article an asymptotically efficient estimator of

$$\tilde{T}(s) = \int_{\mathbb{R}^d} \phi(s(\mathbf{x}), \mathbf{x}) d\mathbf{x}$$

where ϕ is a given regular function. Under specific assumptions, her key idea is to use first a preliminary estimator $\hat{s}(\mathbf{x})$ computed with a small part of the sample $(Z_i)_{i=1,\dots,n}$, for example the n_1 last observations ($n_1 < n$). In a second step ϕ is expanded in a neighborhood of $(\hat{s}(\mathbf{x}), \mathbf{x})$ using a Taylor expansion :

$$\begin{aligned} \tilde{T}(s) &= \int \phi(\hat{s}(\mathbf{x}), \mathbf{x}) d\mathbf{x} + \int \phi'_1(\hat{s}(\mathbf{x}), \mathbf{x})(s - \hat{s})(\mathbf{x}) d\mathbf{x} \\ &\quad + \frac{1}{2} \int \phi''_1(\hat{s}(\mathbf{x}), \mathbf{x})(s - \hat{s})^2(\mathbf{x}) d\mathbf{x} + \Gamma_n \end{aligned}$$

where $\phi'_1 = \frac{\partial \phi}{\partial u}(u, v)$, $\phi''_1 = \frac{\partial^2 \phi}{\partial u^2}(u, v)$ and Γ_n is a remainder term that is shown to be negligible if \hat{s} is close enough to s . This expansion can be rewritten as

$$(4) \quad \tilde{T}(s) = \int \tilde{G}(\hat{s}, \cdot) + \int \tilde{H}(\hat{s}, \cdot) s + \int \tilde{K}(\hat{s}, \cdot) s^2 + \Gamma_n$$

where

$$\begin{aligned} \tilde{G}(\hat{s}, \cdot) &= \phi(\hat{s}, \cdot) - \phi'_1(\hat{s}, \cdot) \hat{s} + \frac{1}{2} \phi''_1(\hat{s}, \cdot) \hat{s}^2, \\ \tilde{H}(\hat{s}, \cdot) &= \phi'_1(\hat{s}, \cdot) - \hat{s} \phi''_1(\hat{s}, \cdot), \\ \tilde{K}(\hat{s}, \cdot) &= \frac{1}{2} \phi''_1(\hat{s}, \cdot). \end{aligned}$$

Further the linear functional $\int \tilde{H}(\hat{s}, \cdot) s$ is classically estimated by the empirical mean $\frac{1}{n_2} \sum_{j=1}^{n_2} \tilde{H}(\hat{s}, \cdot)(Z_j)$ where $n_2 = n - n_1$ denotes the number of remaining observations. Concerning the quadratic functional $\int \tilde{K}(\hat{s}, \cdot) s^2$, the construction of an asymptotically efficient estimator is more difficult and relies on a preliminary work concerning the asymptotically efficient estimation of quadratic functionals of the type $\int \eta(\mathbf{x}) s(\mathbf{x})^2 d\mathbf{x}$. The reader is

referred to [8] for more details. Now if we envisage to apply the method of Laurent to our problem, we would have to set

$$\phi(f(x, y), (x, y)) = \psi \left(\frac{\int \varphi(y) f(x, y) dy}{\int f(x, y) dy} \right).$$

Here it is not possible to write a direct Taylor expansion of the function $u \rightarrow \phi(u, v)$. Indeed, some integral operators appear in the function ψ . Hence, we propose in our work an extension of the work of Laurent to very general functionals.

Let us come back to the definition (3) of the quantity we are interested in. We write it

$$T(f) = \iint \psi \left(\frac{\int \varphi(y) f(x, y) dy}{\int f(x, y) dy} \right) f(x, y) dx dy = \iint \psi(m(x)) f(x, y) dx dy$$

where $m(x) = \int \varphi(y) f(x, y) dy / \int f(x, y) dy$ is the conditional expectation of $\varphi(Y)$ given $(X = x)$. As before, we suggest as a first step to consider a preliminary estimator \hat{f} of f , and to expand $T(f)$ in a neighborhood of \hat{f} . To achieve this goal we first define $F : [0, 1] \rightarrow \mathbb{R}$:

$$F(u) = T(uf + (1 - u)\hat{f}) \quad (u \in [0, 1]).$$

The Taylor expansion of F between 0 and 1 up to the third order is

$$(5) \quad F(1) = F(0) + F'(0) + \frac{1}{2}F''(0) + \frac{1}{6}F'''(\xi)(1 - \xi)^3$$

for some $\xi \in]0, 1[$. Here, we have

$$F(1) = T(f)$$

and

$$\begin{aligned} F(0) = T(\hat{f}) &= \iint \psi \left(\frac{\int \varphi(y) \hat{f}(x, y) dy}{\int \hat{f}(x, y) dy} \right) \hat{f}(x, y) dx dy \\ &= \iint \psi(\hat{m}(x)) \hat{f}(x, y) dx dy \end{aligned}$$

where $\hat{m}(x) = \int \varphi(y) \hat{f}(x, y) dy / \int \hat{f}(x, y) dy$. Straightforward calculations also

give higher-order derivatives of F :

$$\begin{aligned}
F'(0) &= \iint \left([\varphi(y) - \hat{m}(x)] \dot{\psi}(\hat{m}(x)) + \psi(\hat{m}(x)) \right) \\
&\quad \left(f(x, y) - \hat{f}(x, y) \right) dx dy \\
F''(0) &= \iiint \frac{\ddot{\psi}(\hat{m}(x))}{\left(\int \hat{f}(x, y) dy \right)} (\hat{m}(x) - \varphi(y)) (\hat{m}(x) - \varphi(z)) \\
&\quad \left(f(x, y) - \hat{f}(x, y) \right) \left(f(x, z) - \hat{f}(x, z) \right) dx dy dz \\
F'''(\xi) &= \iiint \frac{\left(\int \hat{f}(x, y) dy \right)^2}{\left(\int \xi f(x, y) + (1 - \xi) \hat{f}(x, y) dy \right)^5} \\
&\quad [(\hat{m}(x) - \varphi(y)) (\hat{m}(x) - \varphi(z)) (\hat{m}(x) - \varphi(t)) \\
&\quad \left(\int \hat{f}(x, y) dy \right) \ddot{\psi}(\hat{r}(\xi, x)) - 3(\hat{m}(x) - \varphi(y)) (\hat{m}(x) - \varphi(z)) \\
&\quad \left(\int [\xi f(x, y) + (1 - \xi) \hat{f}(x, y)] dy \right) \ddot{\psi}(\hat{r}(\xi, x))] \\
&\quad \left(f(x, y) - \hat{f}(x, y) \right) \left(f(x, z) - \hat{f}(x, z) \right) \\
&\quad \left(f(x, t) - \hat{f}(x, t) \right) dx dy dz dt
\end{aligned}$$

where $\hat{r}(\xi, x) = \frac{\int \varphi(y) [\xi f(x, y) + (1 - \xi) \hat{f}(x, y)] dy}{\int [\xi f(x, y) + (1 - \xi) \hat{f}(x, y)] dy}$ and $\dot{\psi}$, $\ddot{\psi}$ and $\ddot{\psi}$ denote the three first derivative of ψ .

Plugging these expressions into (5) leads to the following expansion for $T(f)$:

$$\begin{aligned}
(6) \quad T(f) &= \iint H(\hat{f}, x, y) f(x, y) dx dy \\
&\quad + \iiint K(\hat{f}, x, y, z) f(x, y) f(x, z) dx dy dz + \Gamma_n
\end{aligned}$$

where

$$\begin{aligned}
H(\hat{f}, x, y) &= [\varphi(y) - \hat{m}(x)] \dot{\psi}(\hat{m}(x)) + \psi(\hat{m}(x)), \\
K(\hat{f}, x, y, z) &= \frac{1}{2} \frac{\ddot{\psi}(\hat{m}(x))}{\left(\int \hat{f}(x, y) dy \right)} (\hat{m}(x) - \varphi(y)) (\hat{m}(x) - \varphi(z)), \\
\Gamma_n &= \frac{1}{6} F'''(\xi) (1 - \xi)^3
\end{aligned}$$

for some $\xi \in]0, 1[$. Let us now compare the expansions (4) and (6). Notice that the first term is still a linear functional of the density f . We will estimate it with

$$\frac{1}{n_2} \sum_{j=1}^{n_2} H(\hat{f}, X_j, Y_j).$$

The second one is however different from that in (4). Indeed we have to deal here with a complicated crossed term:

$$(7) \quad \iiint \eta(x, y_1, y_2) f(x, y_1) f(x, y_2) dx dy_1 dy_2$$

where $\eta : \mathbb{R}^3 \rightarrow \mathbb{R}$ is a bounded function verifying $\eta(x, y_1, y_2) = \eta(x, y_2, y_1)$ for all $(x, y_1, y_2) \in \mathbb{R}^3$.

In the next section we focus on the asymptotically efficient estimation of such crossed quadratic functionals. In Section 4, these results are then used to propose an asymptotically efficient estimator for $T(f)$.

3. Efficient estimation of quadratic functionals. In this section, our aim is to build an asymptotically efficient estimate for

$$\theta = \iiint \eta(x, y_1, y_2) f(x, y_1) f(x, y_2) dx dy_1 dy_2.$$

We denote $a_i = \int f p_i$ the scalar product of f with p_i . We will first build a projection estimator achieving a bias equal to

$$- \iiint [S_M f(x, y_1) - f(x, y_1)] [S_M f(x, y_2) - f(x, y_2)] \eta(x, y_1, y_2) dx dy_1 dy_2$$

where $S_M f = \sum_{i \in M} a_i p_i$ and M is a subset of D . Thus, the bias would only be due to projection. Developing the previous expression leads to a goal bias equal to

$$(8) \quad \begin{aligned} & 2 \iiint S_M f(x, y_1) f(x, y_2) \eta(x, y_1, y_2) dx dy_1 dy_2 \\ & - \iiint S_M f(x, y_1) S_M f(x, y_2) \eta(x, y_1, y_2) dx dy_1 dy_2 \\ & - \iiint f(x, y_1) f(x, y_2) \eta(x, y_1, y_2) dx dy_1 dy_2. \end{aligned}$$

Consider now the estimator $\hat{\theta}_n$ defined by

$$\begin{aligned} \hat{\theta}_n &= \frac{2}{n(n-1)} \sum_{i \in M} \sum_{j \neq k=1}^n p_i(X_j, Y_j) \int p_i(X_k, u) \eta(X_k, u, Y_k) du \\ &\quad - \frac{1}{n(n-1)} \sum_{i, i' \in M} \sum_{j \neq k=1}^n p_i(X_j, Y_j) p_{i'}(X_k, Y_k) \\ (9) \quad &\int p_i(x, y_1) p_{i'}(x, y_2) \eta(x, y_1, y_2) dx dy_1 dy_2. \end{aligned}$$

This estimator achieves the desired bias :

LEMMA 1 (Bias of $\hat{\theta}_n$). *The estimator $\hat{\theta}_n$ defined in (9) estimates θ with bias equal to*

$$- \iiint [S_M f(x, y_1) - f(x, y_1)] [S_M f(x, y_2) - f(x, y_2)] \eta(x, y_1, y_2) dx dy_1 dy_2.$$

Since we will carry out an asymptotic analysis, we will work with a sequence $(M_n)_{n \geq 1}$ of subsets of D . We will need an extra assumption concerning this sequence:

- A1. For all $n \geq 1$, we can find a subset $M_n \subset D$ such that $(\sup_{i \notin M_n} |c_i|^2)^2 \approx \frac{|M_n|}{n^2}$ ($A_n \approx B_n$ means $\lambda_1 \leq A_n/B_n \leq \lambda_2$ for some positive constants λ_1 and λ_2). Furthermore, $\forall t \in \mathbb{L}^2(dxdy)$, $\int (S_{M_n} t - t)^2 dx dy \rightarrow 0$ when $n \rightarrow \infty$.

The following theorem gives the most important properties of our estimate $\hat{\theta}_n$:

THEOREM 1. *Assume A1 hold. Then $\hat{\theta}_n$ has the following properties:*

- (i) *If $|M_n|/n \rightarrow 0$ when $n \rightarrow \infty$, then*

$$(10) \quad \sqrt{n} (\hat{\theta}_n - \theta) \rightarrow \mathcal{N}(0, \Lambda(f, \eta)),$$

$$(11) \quad \left| \mathbb{E} (\hat{\theta}_n - \theta)^2 - \Lambda(f, \eta) \right| \leq \gamma_1 \left[\frac{|M_n|}{n} + \|S_{M_n} f - f\|_2 + \|S_{M_n} g - g\|_2 \right],$$

where $g(x, y) := \int f(x, u) \eta(x, y, u) du$ and

$$\Lambda(f, \eta) = 4 \left[\iint g(x, y)^2 f(x, y) dx dy - \left(\iint g(x, y) f(x, y) dx dy \right)^2 \right].$$

(ii) *Otherwise*

$$\mathbb{E} \left(\hat{\theta}_n - \theta \right)^2 \leq \gamma_2 \frac{|M_n|}{n},$$

where γ_1 and γ_2 are constants depending only on $\|f\|_\infty$, $\|\eta\|_\infty$ and Δ_Y (with $\Delta_Y = d - c$). Moreover, these constants are increasing functions of these quantities.

REMARK 1. Since in our main result (to be given in the next section) η will depend on n through the preliminary estimator \hat{f} , we need in (11) a bound that depends explicitly on n . Note however that (11) implies

$$\lim_{n \rightarrow \infty} n \mathbb{E} \left(\hat{\theta}_n - \theta \right)^2 = \Lambda(f, \eta).$$

The asymptotic properties of $\hat{\theta}_n$ are of particular importance, in the sense that they are optimal as stated in the following theorem.

THEOREM 2 (Semiparametric Cramér-Rao bound). *Consider the estimation of*

$$\theta = \theta(f) = \iiint \eta(x, y_1, y_2) f(x, y_1) f(x, y_2) dx dy_1 dy_2.$$

Let $f_0 \in \mathcal{E}$. Then, for all estimator $\hat{\theta}_n$ of $\theta(f)$ and every family $\mathcal{V}(f_0)$ of vicinities of f_0 , we have

$$\inf_{\{\mathcal{V}(f_0)\}} \liminf_{n \rightarrow \infty} \sup_{f \in \mathcal{V}(f_0)} n \mathbb{E} (\hat{\theta}_n - \theta(f_0))^2 \geq \Lambda(f_0, \eta).$$

REMARK 2. In other words, the optimal asymptotic variance for the estimation of θ is $\Lambda(f_0, \eta)$. As our estimator defined in (9) achieves this variance, it is therefore asymptotically efficient.

4. Main Theorem. In this section we come back to our main problem of the asymptotically efficient estimation of

$$T(f) = \iint \psi \left(\frac{\int \varphi(y) f(x, y) dy}{\int f(x, y) dy} \right) f(x, y) dx dy.$$

Recall that we have derived in (6) an expansion for $T(f)$:

$$\begin{aligned} T(f) &= \iint H(\hat{f}, x, y) f(x, y) dx dy \\ &\quad + \iiint K(\hat{f}, x, y, z) f(x, y) f(x, z) dx dy dz + \Gamma_n \end{aligned}$$

where

$$\begin{aligned} H(\hat{f}, x, y) &= [\varphi(y) - \hat{m}(x)]\dot{\psi}(\hat{m}(x)) + \psi(\hat{m}(x)), \\ K(\hat{f}, x, y, z) &= \frac{1}{2} \frac{\ddot{\psi}(\hat{m}(x))}{(\int \hat{f}(x, y) dy)} (\hat{m}(x) - \varphi(y)) (\hat{m}(x) - \varphi(z)), \\ \Gamma_n &= \frac{1}{6} F'''(\xi)(1 - \xi)^3 \end{aligned}$$

for some $\xi \in]0, 1[$ and a preliminary estimator \hat{f} of f .

The key idea is to use here the previous results on the estimation of crossed quadratic functionals. Indeed we have provided an asymptotically efficient estimator for the second term of this expansion, conditionally on \hat{f} . A natural and straightforward estimator for $T(f)$ is then

$$\begin{aligned} \hat{T}_n &= \frac{1}{n_2} \sum_{j=1}^{n_2} H(\hat{f}, X_j, Y_j) \\ &+ \frac{2}{n_2(n_2 - 1)} \sum_{i \in M} \sum_{j \neq k=1}^{n_2} p_i(X_j, Y_j) \int p_i(X_k, u) K(\hat{f}, X_k, u, Y_k) du \\ &- \frac{1}{n_2(n_2 - 1)} \sum_{i, i' \in M} \sum_{j \neq k=1}^{n_2} p_i(X_j, Y_j) p_{i'}(X_k, Y_k) \\ &\int p_i(x, y_1) p_{i'}(x, y_2) K(\hat{f}, x, y_1, y_2) dx dy_1 dy_2. \end{aligned}$$

In the above expression, one can note that the remainder Γ_n does not appear : we will see in the proof of the following theorem that it is negligible comparing to the two first terms.

In order to study the asymptotic properties of \hat{T}_n , some assumptions are required concerning the behavior of the joint density f and its preliminary estimator \hat{f} :

- A2. $\text{supp } f \subset [a, b] \times [c, d]$ and $\forall (x, y) \in \text{supp } f$, $0 < \alpha \leq f(x, y) \leq \beta$ with $\alpha, \beta \in \mathbb{R}$
- A3. One can find an estimator \hat{f} of f built with $n_1 \approx n / \log(n)$ observations, such that

$$\forall (x, y) \in \text{supp } f, 0 < \alpha - \epsilon \leq \hat{f}(x, y) \leq \beta + \epsilon.$$

Moreover,

$$\forall 2 \leq q < +\infty, \forall l \in \mathbb{N}^*, \mathbb{E}_f \| \hat{f} - f \|_q^l \leq C(q, l) n_1^{-l\lambda}$$

for some $\lambda > 1/6$ and some constant $C(q, l)$ not depending on f belonging to the ellipsoid \mathcal{E} .

Here $\text{supp } f$ denotes the set where f is different from 0. Assumption A2 is restrictive in the sense that only densities with compact support can be considered, excluding for example a Gaussian joint distribution.

Assumption A3 imposes to the estimator \hat{f} a convergence fast enough towards f . We will use this result to control the remainder term Γ_n .

We can now state the main theorem of the paper. It investigates the asymptotic properties of \hat{T}_n under assumptions A1, A2 and A3.

THEOREM 3. *Assume that A1, A2 and A3 hold. Then \hat{T}_n has the following properties if $\frac{|M_n|}{n} \rightarrow 0$:*

$$(12) \quad \sqrt{n} \left(\hat{T}_n - T(f) \right) \rightarrow \mathcal{N} \left(0, C(f) \right),$$

$$(13) \quad \lim_{n \rightarrow \infty} n \mathbb{E} \left(\hat{T}_n - T(f) \right)^2 = C(f),$$

where $C(f) = \mathbb{E} \left(\text{Var}(\varphi(Y)|X) \left[\psi(\mathbb{E}(Y|X)) \right]^2 \right) + \text{Var} \left(\psi(\mathbb{E}(\varphi(Y)|X)) \right)$.

We can also compute as in the previous section the semiparametric Cramér-Rao bound for this problem.

THEOREM 4 (Semiparametric Cramér-Rao bound). *Consider the estimation of*

$$T(f) = \iint \psi \left(\frac{\int \varphi(y) f(x, y) dy}{\int f(x, y) dy} \right) f(x, y) dx dy = \mathbb{E} \left(\psi(\mathbb{E}(\varphi(Y)|X)) \right)$$

for a random vector (X, Y) with joint density $f \in \mathcal{E}$. Let $f_0 \in \mathcal{E}$ be a density verifying the assumptions of Theorem 3. Then, for all estimator \hat{T}_n of $T(f)$ and every family $\mathcal{V}(f_0)$ of vicinities of f_0 , we have

$$\inf_{\{\mathcal{V}(f_0)\}} \liminf_{n \rightarrow \infty} \sup_{f \in \mathcal{V}(f_0)} n \mathbb{E}(\hat{T}_n - T(f_0))^2 \geq C(f_0).$$

COROLLARY 1. *\hat{T}_n is asymptotically efficient.*

5. Application to the estimation of sensitivity indices. Now that we have built an asymptotically efficient estimate for $T(f)$, we can apply it to the particular case we were initially interested in: the estimation of Sobol sensitivity indices. Let us then come back to model (1) :

$$Y = \Phi(\boldsymbol{\tau})$$

where we wish to estimate (2):

$$\Sigma_j = \frac{\text{Var}(\mathbb{E}(Y|\tau_j))}{\text{Var}(Y)} = \frac{\mathbb{E}(\mathbb{E}(Y|\tau_j)^2) - \mathbb{E}(Y)^2}{\text{Var}(Y)} \quad j = 1, \dots, l.$$

To do so, we have an i.i.d. sample $(Y_1, \boldsymbol{\tau}^{(1)}), \dots, (Y_n, \boldsymbol{\tau}^{(n)})$. We will only give here the procedure for the estimation of Σ_1 since it will be the same for the other sensitivity indices. Denoting $X := \tau_1$, this problem is equivalent to estimating $\mathbb{E}(\mathbb{E}(Y|X)^2)$ with an i.i.d. sample $(Y_1, X_1), \dots, (Y_n, X_n)$ with joint density f . Note that we only take into account the sample $(X_i)_{i=1, \dots, n} := (\tau_1^{(i)})_{i=1, \dots, n}$ to estimate Σ_j (and thus discard the observations of the other covariates τ_j for $j \neq 1$): we will discuss and justify this method later.

We can hence apply the estimate we developed previously by letting $\psi(\xi) = \xi^2$ and $\varphi(\xi) = \xi$:

$$\begin{aligned} T(f) &= \mathbb{E}(\mathbb{E}(Y|X)^2) \\ &= \iint \left(\frac{\int y f(x, y) dy}{\int f(x, y) dy} \right)^2 f(x, y) dx dy. \end{aligned}$$

The Taylor expansion in this case becomes

$$\begin{aligned} T(f) &= \iint H(\hat{f}, x, y) f(x, y) dx dy \\ &\quad + \iiint K(\hat{f}, x, y, z) f(x, y) f(x, z) dx dy dz + \Gamma_n \end{aligned}$$

where

$$\begin{aligned} H(\hat{f}, x, y) &= 2y\hat{m}(x) - \hat{m}(x)^2, \\ K(\hat{f}, x, y, z) &= \frac{1}{\left(\int \hat{f}(x, y) dy\right)} (\hat{m}(x) - y)(\hat{m}(x) - z) \end{aligned}$$

and the corresponding estimator is

$$\begin{aligned} \widehat{T}_n &= \frac{1}{n_2} \sum_{j=1}^{n_2} H(\widehat{f}, X_j, Y_j) \\ &+ \frac{2}{n_2(n_2 - 1)} \sum_{i \in M} \sum_{j \neq k=1}^{n_2} p_i(X_j, Y_j) \int p_i(X_k, u) K(\widehat{f}, X_k, u, Y_k) du \\ &- \frac{1}{n_2(n_2 - 1)} \sum_{i, i' \in M} \sum_{j \neq k=1}^{n_2} p_i(X_j, Y_j) p_{i'}(X_k, Y_k) \\ &\int p_i(x, y_1) p_{i'}(x, y_2) K(\widehat{f}, x, y_1, y_2) dx dy_1 dy_2. \end{aligned}$$

for some preliminary estimator \widehat{f} of f , an orthonormal basis $(p_i)_{i \in D}$ of $\mathbb{L}^2(dx dy)$ and a subset $M \subset D$ verifying the hypotheses of Theorem 3.

We propose now to investigate the practical behavior of this estimator on two analytical models. In both models, the preliminary estimator \widehat{f} will be a kernel density estimator with bounded support (MATLAB Statistics Toolbox routine 'ksdensity' for example) built on $n_1 = \lceil \log(n)/n \rceil$ observations. Moreover, we choose the Legendre polynomials on $[a, b]$ and $[c, d]$ to build the orthonormal basis $(p_i)_{i \in D}$ and we will take $|M| = \sqrt{n}$. Finally, the integrals in \widehat{T}_n are computed with an adaptive Simpson quadrature (MATLAB routine 'quad').

The first model we investigate is

$$(14) \quad Y = \tau_1 + \tau_2^4$$

where three configurations are considered (τ_1 and τ_2 being independent):

- (a) $\tau_j \sim \mathcal{U}(0, 1)$, $j = 1, 2$;
- (b) $\tau_j \sim \mathcal{U}(0, 3)$, $j = 1, 2$;
- (c) $\tau_j \sim \mathcal{U}(0, 5)$, $j = 1, 2$.

The results obtained with $n = 100$ and $n = 10000$ are given in Table 1.

The asymptotically efficient estimator \widehat{T}_n gives a very accurate approximation of sensitivity indices when $n = 10000$. But surprisingly, it also gives a reasonably accurate estimate when n only equals 100, whereas it has been built to achieve the best *asymptotic* rate of convergence. It is then interesting to compare it with other estimators, more precisely two nonparametric estimators that have been specifically built to give an accurate approximation of sensitivity indices when n is not large: that of [13] and that of [3].

Inputs	$\mathbb{E}(\mathbb{E}(Y \tau_j)^2)$	\widehat{T}_n $n = 100$	\widehat{T}_n $n = 10000$
Configuration (a)			
τ_1	0.5733	0.5894	0.5729
τ_2	0.5611	0.5468	0.5611
Configuration (b)			
τ_1	314.04	305.98	318.27
τ_2	779.85	814.04	787.82
Configuration (c)			
τ_1	16258	18414	16897
τ_2	44034	47667	44473

TABLE 1

Conditional moments for analytical model (14).

We have decided to make such a comparison on the following model :

$$(15) \quad Y = 0.2 \exp(\tau_1 - 3) + 2.2|\tau_2| + 1.3\tau_2^6 - 2\tau_2^2 - 0.5\tau_2^4 - 0.5\tau_1^4 + 2.5\tau_1^2 + 0.7\tau_1^3 + \frac{3}{(8\tau_1 - 2)^2 + (5\tau_2 - 3)^2 + 1} + \sin(5\tau_1) \cos(3\tau_1^2)$$

where τ_1 and τ_2 are independent and uniformly distributed on $[-1, 1]$. This nonlinear function is interesting since it presents a peak and valleys. We estimate the sensitivity indices with a sample of size $n = 100$, the results are given in Table 2.

	True value	Oakley-O'Hagan	Local polynomials	\widehat{T}_n
$\text{Var}(\mathbb{E}(Y X^1))$	1.0932	1.0539	1.0643	1.1701
$\text{Var}(\mathbb{E}(Y X^2))$	0.0729	0.1121	0.0527	0.0939

TABLE 2

Comparison between efficient estimation and nonparametric estimates on analytical model (15).

Globally, the best estimates are given by the local polynomials technique. However, the accuracy of the asymptotically efficient estimator \widehat{T}_n is comparable to that of the nonparametric ones. These results confirm that \widehat{T}_n is a valuable estimator even with a rather complex model and a small sample size (recall that here $n = 100$).

Let us now end this section by discussing the voluntary choice to discard the observations $(Y_i, \boldsymbol{\tau}_{\sim 1}^{(i)})_{i=1, \dots, n}$ (where $\boldsymbol{\tau}_{\sim 1}$ denotes the vector of all the random variables τ_j with $j \neq 1$) when estimating $\mathbb{E}(\mathbb{E}(Y|\tau_1)^2)$. The fundamental question is then the following: does observing other covariates

τ_j for $j \neq 1$ helps to better estimate $\mathbb{E}(\mathbb{E}(Y|\tau_1)^2)$? To answer it, we will place ourselves again in the semiparametric setting: we will compute the semiparametric Cramér-Rao bound associated to the problem of estimating $\mathbb{E}(\mathbb{E}(Y|X)^2)$ from an i.i.d. sample $(X_i, \mathbf{Z}_i, Y_i)_{i=1, \dots, n}$ with joint density $f(x, \mathbf{z}, y) \in \mathbb{L}^2(\Omega)$ and Ω an hypercube of $\mathbb{L}^2(\mathbb{R}^{d+1})$. Obviously, X stands here for τ_1 and \mathbf{Z} for $\tau_{\sim 1}$. The lower bound is given in the following theorem:

THEOREM 5 (Semiparametric Cramér-Rao bound). *Consider the estimation of*

$$T(f) = \iint \left(\frac{\int y f(x, \mathbf{z}, y) d\mathbf{z} dy}{\int f(x, \mathbf{z}, y) d\mathbf{z} dy} \right)^2 f(x, \mathbf{z}, y) dx d\mathbf{z} dy = \mathbb{E}(\mathbb{E}(Y|X)^2)$$

for a random vector (X, \mathbf{Z}, Y) with joint density $f \in \mathcal{E}$. Let $f_0 \in \mathcal{E}$ be a density verifying the assumptions of Theorem 3. Then, for all estimator \hat{T}_n of $T(f)$ and every family $\mathcal{V}(f_0)$ of vicinities of f_0 , we have

$$\inf_{\{\mathcal{V}(f_0)\}} \liminf_{n \rightarrow \infty} \sup_{f \in \mathcal{V}(f_0)} n \mathbb{E}(\hat{T}_n - T(f_0))^2 \geq C(f_0)$$

In other words, the optimal asymptotic variance for the estimation of $T(f)$ is $C(f_0)$.

The previous theorem perfectly answers our question. Indeed, the Cramér-Rao bound with a vector (X, \mathbf{Z}, Y) is the same as with a pair (X, Y) . Note that the estimator \hat{T}_n we built in the previous section asymptotically reaches this bound: it is therefore the best estimate for $T(f)$ (in the asymptotic efficiency sense). Hence, taking into account the observations of an extra-covariates vector \mathbf{Z} does not asymptotically improve the estimation of $\mathbb{E}(\mathbb{E}(Y|X)^2)$ (may \mathbf{Z} be independent of X or not). This result thus justifies the fact to work only with a sample $(Y_i, X_i)_{i=1, \dots, n}$ as we did in our examples. Finally, let us precise we shall qualify this last statement, since it only relies on asymptotic considerations. In practice, using observations of \mathbf{Z} may improve the estimation. However, we are not currently able to give a theoretical result for the finite sample size case.

6. Discussion and conclusions. In this paper, we have shown that it is possible to build an asymptotically efficient estimate for nonlinear conditional functionals. The estimators are both practically computable and have optimal asymptotic properties. Hence, they could be used by practioners to estimate for example the Sobol indices. Interested readers can obtain the Matlab code for computing the estimator of Sobol indices by asking the authors.

The methodology developed here will be extended to other problems in forthcoming work. Firstly, conditional functionals such as Cramér Von Mises may be used to perform goodness-of-fit tests for conditional distributions (see [4]). Secondly, the estimation of conditional covariance matrices in the Sliced Inverse Regression setting ([11]) will also be treated. Finally, a very attractive extension is the construction of an adaptive procedure to calibrate the size of M_n as done in [9] for the \mathbb{L}^2 norm. Nevertheless, this problem seems to be a very difficult one since it would involve treating complicated inequalities on U-statistics such as presented in [5].

7. Proofs of Theorems.

7.1. *Proof of Lemma 1.* Let $\hat{\theta}_n = \hat{\theta}_n^1 - \hat{\theta}_n^2$ where

$$\hat{\theta}_n^1 = \frac{2}{n(n-1)} \sum_{i \in M} \sum_{j \neq k=1}^n p_i(X_j, Y_j) \int p_i(X_k, u) \eta(X_k, u, Y_k) du$$

and

$$\begin{aligned} \hat{\theta}_n^2 &= \frac{1}{n(n-1)} \sum_{i, i' \in M} \sum_{j \neq k=1}^n p_i(X_j, Y_j) p_{i'}(X_k, Y_k) \\ &\quad \int p_i(x, y_1) p_{i'}(x, y_2) \eta(x, y_1, y_2) dx dy_1 dy_2. \end{aligned}$$

Let us first compute $\mathbb{E}(\hat{\theta}_n^1)$:

$$\begin{aligned} \mathbb{E}(\hat{\theta}_n^1) &= 2 \sum_{i \in M} \iint p_i(x, y) f(x, y) dx dy \iiint p_i(x, y) \eta(x, u, y) f(x, y) dx dy du \\ &= 2 \sum_{i \in M} a_i \iiint p_i(x, y) \eta(x, u, y) f(x, y) dx dy du \\ &= 2 \iiint \left(\sum_{i \in M} a_i p_i(x, y) \right) \eta(x, u, y) f(x, y) dx dy du \\ &= 2 \iiint S_M f(x, y) \eta(x, u, y) f(x, y) dx dy du. \end{aligned}$$

Furthermore,

$$\begin{aligned}
\mathbb{E}(\hat{\theta}_n^2) &= \sum_{i,i' \in M} \iint p_i(x,y)f(x,y)dx dy \iint p_{i'}(x,y)f(x,y)dx dy \\
&\quad \int p_i(x,y_1)p_{i'}(x,y_2)\eta(x,y_1,y_2)dx dy_1 dy_2 \\
&= \sum_{i,i' \in M} a_i a_{i'} \int p_i(x,y_1)p_{i'}(x,y_2)\eta(x,y_1,y_2)dx dy_1 dy_2 \\
&= \int \left(\sum_{i \in M} a_i p_i(x,y_1) \right) \left(\sum_{i' \in M} a_{i'} p_{i'}(x,y_2) \right) \eta(x,y_1,y_2) dx dy_1 dy_2 \\
&= \int S_M f(x,y_1) S_M f(x,y_2) \eta(x,y_1,y_2) dx dy_1 dy_2.
\end{aligned}$$

Finally, $\mathbb{E}(\hat{\theta}_n) - \theta = \mathbb{E}(\hat{\theta}_n^1) - \mathbb{E}(\hat{\theta}_n^2) - \theta$ and we get the desired bias with (8).

7.2. *Proof of Theorem 1.* We will write M instead of M_n for readability and denote $m = |M|$. We want to bound the precision of $\hat{\theta}_n$. We first write

$$\mathbb{E} \left(\hat{\theta}_n - \iiint \eta(x,y_1,y_2) f(x,y_1) f(x,y_2) dx dy_1 dy_2 \right)^2 = \text{Bias}^2(\hat{\theta}_n) + \text{Var}(\hat{\theta}_n).$$

The first term of this decomposition can be easily bounded, since $\hat{\theta}_n$ has been built to achieve a bias equal to

$$\begin{aligned}
\text{Bias}(\hat{\theta}_n) &= - \iiint [S_M f(x,y_1) - f(x,y_1)] [S_M f(x,y_2) - f(x,y_2)] \\
&\quad \eta(x,y_1,y_2) dx dy_1 dy_2.
\end{aligned}$$

We then get the following lemma :

LEMMA 2 (Bound for Bias($\hat{\theta}_n$)). *Assuming the hypotheses of Theorem 1 hold, we have*

$$|\text{Bias}(\hat{\theta}_n)| \leq \Delta_Y \|\eta\|_\infty \sup_{i \notin M} |c_i|^2.$$

PROOF.

$$\begin{aligned}
|\text{Bias}(\hat{\theta}_n)| &\leq \|\eta\|_\infty \int \left(\int |S_M f(x, y_1) - f(x, y_1)| dy_1 \right) \\
&\quad \left(\int |S_M f(x, y_2) - f(x, y_2)| dy_2 \right) dx \\
&\leq \|\eta\|_\infty \int \left(\int |S_M f(x, y) - f(x, y)| dy \right)^2 dx \\
&\leq \Delta_Y \|\eta\|_\infty \iint (S_M f(x, y) - f(x, y))^2 dx dy \\
&\leq \Delta_Y \|\eta\|_\infty \sum_{i \notin M} |a_i|^2 \leq \Delta_Y \|\eta\|_\infty \sup_{i \notin M} |c_i|^2.
\end{aligned}$$

Indeed, $f \in \mathcal{E}$ and the last inequality follows from Hölder inequality. \square

Bounding the variance of $\hat{\theta}_n$ is however less straightforward. Let A and B be the $m \times 1$ vectors with components

$$\begin{aligned}
a_i &:= \iint f(x, y) p_i(x, y) dx dy \quad i = 1, \dots, m \\
b_i &:= \iiint p_i(x, y_1) f(x, y_2) \eta(x, y_1, y_2) dx dy_1 dy_2 \\
&= \iint g(x, y) p_i(x, y) dx dy \quad i = 1, \dots, m
\end{aligned}$$

where $g(x, y) = \int f(x, u) \eta(x, y, u) du$ for each $i \in M$. a_i et b_i are the components of f and g onto the i th component of the basis. Let Q and R be the $m \times 1$ vectors of the centered functions $q_i(x, y) = p_i(x, y) - a_i$ and $r_i(x, y) = \int p_i(x, u) \eta(x, u, y) du - b_i$ for $i = 1, \dots, m$. Let C be the $m \times m$ matrix of constants $c_{i'i'} = \iiint p_i(x, y_1) p_{i'}(x, y_2) \eta(x, y_1, y_2) dx dy_1 dy_2$ for $i, i' = 1, \dots, m$. Take care that here $c_{i'i'}$ is double subscript unlike in the (c_i) sequence appearing in the definition of the ellipsoid \mathcal{E} . We denote by U_n the process $U_n h = \frac{1}{n(n-1)} \sum_{j \neq k=1}^n h(X_j, Y_j, X_k, Y_k)$ and by P_n the empirical measure $P_n f = \frac{1}{n} \sum_{j=1}^n f(X_j, Y_j)$. With the previous notation, $\hat{\theta}_n$ has the following Hoeffding's decomposition (see chapter 11 of [16]):

$$(16) \quad \hat{\theta}_n = U_n K + P_n L + 2^t AB - {}^t ACA$$

where

$$\begin{aligned} K(x_1, y_1, x_2, y_2) &= 2^t Q(x_1, y_1)R(x_2, y_2) - {}^t Q(x_1, y_1)CQ(x_2, y_2), \\ L(x_1, y_1) &= 2^t AR(x_1, y_1) + 2^t BQ(x_1, y_1) - 2^t ACQ(x_1, y_1). \end{aligned}$$

Then $\text{Var}(\hat{\theta}_n) = \text{Var}(U_n K) + \text{Var}(P_n L) + 2 \text{Cov}(U_n K, P_n L)$. We have to get bounds for each of these terms : they are given in the three following lemmas.

LEMMA 3 (Bound for $\text{Var}(U_n K)$). *Assuming the hypotheses of Theorem 1 hold, we have*

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

PROOF. Since $U_n K$ is centered, $\text{Var}(U_n K)$ equals

$$\begin{aligned} &\mathbb{E} \left(\frac{1}{(n(n-1))^2} \sum_{j \neq k=1}^n \sum_{j' \neq k'=1}^n K(X_j, Y_j, X_k, Y_k) K(X_{j'}, Y_{j'}, X_{k'}, Y_{k'}) \right) \\ &= \frac{1}{n(n-1)} \mathbb{E}(K^2(X_1, Y_1, X_2, Y_2) + K(X_1, Y_1, X_2, Y_2)K(X_2, Y_2, X_1, Y_1)). \end{aligned}$$

By the Cauchy-Schwarz inequality,

$$\text{Var}(U_n K) \leq \frac{2}{n(n-1)} \mathbb{E}(K^2(X_1, Y_1, X_2, Y_2)).$$

Moreover, the inequality $2|\mathbb{E}(XY)| \leq \mathbb{E}(X^2) + \mathbb{E}(Y^2)$ leads to

$$\begin{aligned} \mathbb{E}(K^2(X_1, Y_1, X_2, Y_2)) &\leq 2 \left[\mathbb{E} \left((2Q'(X_1, Y_1)R(X_2, Y_2))^2 \right) \right. \\ &\quad \left. + \mathbb{E} \left((Q'(X_1, Y_1)CQ(X_2, Y_2))^2 \right) \right]. \end{aligned}$$

We have to bound these two terms. The first one is

$$\mathbb{E} \left((2Q'(X_1, Y_1)R(X_2, Y_2))^2 \right) = 4(W_1 - W_2 - W_3 + W_4)$$

where

$$\begin{aligned}
W_1 &= \iiint \sum_{i,i'} p_i(x,y) p_{i'}(x,y) p_i(x',u) p_{i'}(x',v) \eta(x',u,y') \eta(x',v,y') \\
&\quad f(x,y) f(x',y') du dv dx dy dx' dy' \\
W_2 &= \iint \sum_{i,i'} b_i b_{i'} p_i(x,y) p_{i'}(x,y) f(x,y) dx dy \\
W_3 &= \iiint \sum_{i,i'} a_i a_{i'} p_i(x,u) p_{i'}(x,v) \eta(x,u,y) \eta(x,v,y) f(x,y) dx dy \\
W_4 &= \sum_{i,i'} a_i a_{i'} b_i b_{i'}.
\end{aligned}$$

Straightforward manipulations show that $W_2 \geq 0$ and $W_3 \geq 0$. This implies that

$$\mathbb{E} \left((2Q'(X_1, Y_1)R(X_2, Y_2))^2 \right) \leq 4(W_1 + W_4).$$

On the one hand,

$$\begin{aligned}
W_1 &= \iiint \sum_{i,i'} p_i(x,y) p_{i'}(x,y) \int p_i(x',u) \eta(x',u,y') du \\
&\quad \int p_{i'}(x',v) \eta(x',v,y') dv f(x,y) f(x',y') dx dy dx' dy' \\
&\leq \iiint \left(\sum_i p_i(x,y) \int p_i(x',u) \eta(x',u,y') du \right)^2 \\
&\quad f(x,y) f(x',y') dx dy dx' dy' \\
&\leq \|f\|_\infty^2 \iiint \left(\sum_i p_i(x,y) \int p_i(x',u) \eta(x',u,y') du \right)^2 dx dy dx' dy' \\
&\leq \|f\|_\infty^2 \iiint \sum_{i,i'} p_i(x,y) p_{i'}(x,y) \int p_i(x',u) \eta(x',u,y') du \\
&\quad \int p_{i'}(x',v) \eta(x',v,y') dv dx dy dx' dy' \\
&\leq \|f\|_\infty^2 \sum_{i,i'} \iint p_i(x,y) p_{i'}(x,y) dx dy \iint \left(\int p_i(x',u) \eta(x',u,y') du \right) \\
&\quad \left(\int p_{i'}(x',v) \eta(x',v,y') dv \right) dx' dy' \\
&\leq \|f\|_\infty^2 \sum_i \iint \left(\int p_i(x',u) \eta(x',u,y') du \right)^2 dx' dy'
\end{aligned}$$

since the p_i are orthonormal. Moreover,

$$\begin{aligned} \left(\int p_i(x', u) \eta(x', u, y') du \right)^2 &\leq \left(\int p_i(x', u)^2 du \right) \left(\int \eta(x', u, y')^2 du \right) \\ &\leq \|\eta\|_\infty^2 \Delta_Y \int p_i(x', u)^2 du, \end{aligned}$$

and then

$$\begin{aligned} \iint \left(\int p_i(x', u) \eta(x', u, y') du \right)^2 dx' dy' &\leq \|\eta\|_\infty^2 \Delta_Y^2 \iint p_i(x', u)^2 du dx' \\ &\quad \|\eta\|_\infty^2 \Delta_Y^2. \end{aligned}$$

Finally,

$$W_1 \leq \|\eta\|_\infty^2 \|f\|_\infty^2 \Delta_Y^2 m.$$

On the other hand,

$$W_4 = \left(\sum_i a_i b_i \right)^2 \leq \sum_i a_i^2 \sum_i b_i^2 \leq \|f\|_2^2 \|g\|_2^2 \leq \|f\|_\infty \|g\|_2^2.$$

By the Cauchy-Scharzw inequality we have $\|g\|_2^2 \leq \|\eta\|_\infty^2 \|f\|_\infty \Delta_Y^2$ and then

$$W_4 \leq \|\eta\|_\infty^2 \|f\|_\infty^2 \Delta_Y^2$$

which leads to

$$\mathbb{E} \left((2Q'(X_1, Y_1)R(X_2, Y_2))^2 \right) \leq 4\|\eta\|_\infty^2 \|f\|_\infty^2 \Delta_Y^2 (m+1).$$

Let us bound now the second term $\mathbb{E} \left((Q'(X_1, Y_1)CQ(X_2, Y_2))^2 \right) = W_5 - 2W_6 + W_7$ where

$$\begin{aligned} W_5 &= \iiint \sum_{i, i'} \sum_{i_1, i'_1} c_{ii'} c_{i_1 i'_1} p_i(x, y) p_{i_1}(x, y) p_{i'}(x', y') p_{i'_1}(x', y') \\ &\quad f(x, y) f(x', y') dx dy dx' dy' \\ W_6 &= \sum_{i, i'} \sum_{i_1, i'_1} \iint c_{ii'} c_{i_1 i'_1} a_i a_{i_1} p_{i'}(x, y) p_{i'_1}(x, y) f(x, y) dx dy \\ W_7 &= \sum_{i, i'} \sum_{i_1, i'_1} c_{ii'} c_{i_1 i'_1} a_i a_{i_1} a_{i'} a_{i'_1}. \end{aligned}$$

Following the previous manipulations, we show that $W_6 \geq 0$. Thus,

$$\mathbb{E} \left((Q'(X_1, Y_1)CQ(X_2, Y_2))^2 \right) \leq W_5 + W_7.$$

First, observe that

$$\begin{aligned}
W_5 &= \iiint\!\!\!\int \left(\sum_{i,i'} c_{ii'} p_i(x,y) p_{i'}(x',y') \right)^2 f(x,y) f(x',y') dx dy dx' dy' \\
&\leq \|f\|_\infty^2 \iiint\!\!\!\int \left(\sum_{i,i'} c_{ii'} p_i(x,y) p_{i'}(x',y') \right)^2 dx dy dx' dy' \\
&\leq \|f\|_\infty^2 \sum_{i,i'} \sum_{i_1,i'_1} c_{ii'} c_{i_1 i'_1} \iiint\!\!\!\int p_i(x,y) p_{i_1}(x,y) \\
&\quad p_{i'}(x',y') p_{i'_1}(x',y') dx dy dx' dy' \\
&\leq \|f\|_\infty^2 \sum_{i,i'} c_{ii'}^2
\end{aligned}$$

since the p_i are orthonormal. Besides,

$$\begin{aligned}
\sum_{i,i'} c_{ii'}^2 &= \iint \sum_{i_\alpha, i'_\alpha} \alpha_{i_\alpha}(x) \alpha_{i'_\alpha}(x) \alpha_{i_\alpha}(x') \alpha_{i'_\alpha}(x') \\
&\quad \sum_{i_\beta, i'_\beta} \left(\iint \beta_{i_\beta}(y_1) \beta_{i'_\beta}(y_2) \eta(x, y_1, y_2) dy_1 dy_2 \right) \\
&\quad \left(\iint \beta_{i_\beta}(y_1) \beta_{i'_\beta}(y_2) \eta(x', y_1, y_2) dy_1 dy_2 \right) dx dx' \\
&= \iint \left(\sum_{i_\alpha} \alpha_{i_\alpha}(x) \alpha_{i_\alpha}(x') \right)^2 \\
&\quad \sum_{i_\beta, i'_\beta} \left(\iint \beta_{i_\beta}(y_1) \beta_{i'_\beta}(y_2) \eta(x, y_1, y_2) dy_1 dy_2 \right) \\
&\quad \left(\iint \beta_{i_\beta}(y_1) \beta_{i'_\beta}(y_2) \eta(x', y_1, y_2) dy_1 dy_2 \right) dx dx'.
\end{aligned}$$

But

$$\begin{aligned}
& \sum_{i_\beta, i'_\beta} \left(\iint \beta_{i_\beta}(y_1) \beta_{i'_\beta}(y_2) \eta(x, y_1, y_2) dy_1 dy_2 \right) \\
& \left(\iint \beta_{i_\beta}(y_1) \beta_{i'_\beta}(y_2) \eta(x', y_1, y_2) dy_1 dy_2 \right) \\
= & \sum_{i_\beta, i'_\beta} \iiint \beta_{i_\beta}(y_1) \beta_{i'_\beta}(y_2) \eta(x, y_1, y_2) \beta_{i_\beta}(y'_1) \beta_{i'_\beta}(y'_2) \\
& \eta(x', y'_1, y'_2) dy_1 dy_2 dy'_1 dy'_2 \\
= & \iint \sum_{i_\beta} \left(\int \beta_{i_\beta}(y_1) \eta(x, y_1, y_2) dy_1 \right) \beta_{i_\beta}(y'_1) \\
& \sum_{i'_\beta} \left(\int \beta_{i'_\beta}(y_2) \eta(x', y'_1, y'_2) dy'_2 \right) \beta_{i'_\beta}(y_2) dy'_1 dy_2 \\
= & \iint \eta(x, y'_1, y_2) \eta(x', y'_1, y_2) dy'_1 dy_2 \\
\leq & \Delta_Y^2 \|\eta\|_\infty^2
\end{aligned}$$

using the fact that (β_i) is an orthonormal basis. We then get

$$\begin{aligned}
\sum_{i, i'} c_{ii'}^2 & \leq \Delta_Y^2 \|\eta\|_\infty^2 \iint \left(\sum_{i_\alpha} \alpha_{i_\alpha}(x) \alpha_{i_\alpha}(x') \right)^2 dx dx' \\
& \leq \Delta_Y^2 \|\eta\|_\infty^2 \iint \sum_{i_\alpha, i'_\alpha} \alpha_{i_\alpha}(x) \alpha_{i'_\alpha}(x) \alpha_{i_\alpha}(x') \alpha_{i'_\alpha}(x') dx dx' \\
& \leq \Delta_Y^2 \|\eta\|_\infty^2 \sum_{i_\alpha, i'_\alpha} \left(\int \alpha_{i_\alpha}(x) \alpha_{i'_\alpha}(x) dx \right)^2 \\
& \leq \Delta_Y^2 \|\eta\|_\infty^2 \sum_{i_\alpha} \left(\int \alpha_{i_\alpha}(x)^2 dx \right)^2 \\
& \leq \Delta_Y^2 \|\eta\|_\infty^2 m
\end{aligned}$$

since the α_i are orthonormal. Finally,

$$W_5 \leq \|\eta\|_\infty^2 \|f\|_\infty^2 \Delta_Y^2 m.$$

Besides,

$$W_7 = \left(\sum_{i, i'} c_{ii'} a_i a_{i'} \right)^2$$

with

$$\begin{aligned} \left| \sum_{i,i'} c_{ii'} a_i a_{i'} \right| &\leq \|\eta\|_\infty \iiint |S_M f(x, y_1) S_M f(x, y_2)| dx dy_1 dy_2 \\ &\leq \|\eta\|_\infty \iint \left(\int |S_M f(x, y_1) S_M f(x, y_2)| dx \right) dy_1 dy_2. \end{aligned}$$

By using the Cauchy-Schwarz inequality twice, we get

$$\begin{aligned} \left(\sum_{i,i'} c_{ii'} a_i a_{i'} \right)^2 &\leq \Delta_Y^2 \|\eta\|_\infty^2 \iint \left(\int |S_M f(x, y_1) S_M f(x, y_2)| dx \right)^2 dy_1 dy_2 \\ &\leq \Delta_Y^2 \|\eta\|_\infty^2 \iint \left(\int S_M f(u, y_1)^2 du \right) \\ &\quad \left(\int S_M f(v, y_2)^2 dv \right) dy_1 dy_2 \\ &\leq \Delta_Y^2 \|\eta\|_\infty^2 \iiint \int S_M f(u, y_1)^2 S_M f(v, y_2)^2 du dv dy_1 dy_2 \\ &\leq \Delta_Y^2 \|\eta\|_\infty^2 \left(\iint S_M f(x, y)^2 dx dy \right)^2 \\ &\leq \Delta_Y^2 \|\eta\|_\infty^2 \|f\|_\infty^2. \end{aligned}$$

Finally,

$$\mathbb{E} \left((Q'(X_1, Y_1) C Q(X_2, Y_2))^2 \right) \leq \|\eta\|_\infty^2 \|f\|_\infty^2 \Delta_Y^2 (m+1).$$

Collecting this inequalities, we obtain

$$\text{Var}(U_n K) \leq \frac{20}{n(n-1)} \|\eta\|_\infty^2 \|f\|_\infty^2 \Delta_Y^2 (m+1)$$

which concludes the proof of Lemma 3. \square

Let us now deal with the second term of the Hoeffding's decomposition of $\hat{\theta}_n$:

LEMMA 4 (Bound for $\text{Var}(P_n L)$). *Assuming the hypotheses of Theorem 1 hold, we have*

$$\text{Var}(P_n L) \leq \frac{36}{n} \Delta_Y^2 \|f\|_\infty^2 \|\eta\|_\infty^2.$$

PROOF. First note that

$$\text{Var}(P_n L) = \frac{1}{n} \text{Var}(L(X_1, Y_1)).$$

We can write $L(X_1, Y_1)$ as

$$\begin{aligned} L(X_1, Y_1) &= 2A'R(X_1, Y_1) + 2B'Q(X_1, Y_1) - 2A'CQ(X_1, Y_1) \\ &= 2 \sum_i a_i \left(\int p_i(X_1, u) \eta(X_1, u, Y_1) du - b_i \right) \\ &\quad + 2 \sum_i b_i (p_i(X_1, Y_1) - a_i) - 2 \sum_{i, i'} c_{ii'} a_{i'} (p_i(X_1, Y_1) - a_i) \\ &= 2 \int \sum_i a_i p_i(X_1, u) \eta(X_1, u, Y_1) du + 2 \sum_i b_i p_i(X_1, Y_1) \\ &\quad - 2 \sum_{i, i'} c_{ii'} a_{i'} p_i(X_1, Y_1) - 4A'B + 2A'CA \\ &= 2 \int S_M f(X_1, u) \eta(X_1, u, Y_1) du + 2S_M g(X_1, Y_1) \\ &\quad - 2 \sum_{i, i'} c_{ii'} a_{i'} p_i(X_1, Y_1) - 4A'B + 2A'CA. \end{aligned}$$

Let $h(x, y) = \int S_M f(x, u) \eta(x, u, y) du$, we have

$$\begin{aligned} S_M h(z, t) &= \sum_i \left(\iint h(x, y) p_i(x, y) dx dy \right) p_i(z, t) \\ &= \sum_i \left(\iiint S_M f(x, u) \eta(x, u, y) p_i(x, y) du dx dy \right) p_i(z, t) \\ &= \sum_{i, i'} \left(\iiint a_{i'} p_{i'}(x, u) \eta(x, u, y) p_i(x, y) du dx dy \right) p_i(z, t) \\ &= \sum_{i, i'} c_{ii'} a_{i'} p_i(z, t) \end{aligned}$$

and we can write

$$L(X_1, Y_1) = 2h(X_1, Y_1) + 2S_M g(X_1, Y_1) - 2S_M h(X_1, Y_1) - 4A'B + 2A'CA.$$

Thus,

$$\begin{aligned} \text{Var}(L(X_1, Y_1)) &= 4\text{Var}[h(X_1, Y_1) + S_M g(X_1, Y_1) - S_M h(X_1, Y_1)] \\ &\leq 4\mathbb{E}[(h(X_1, Y_1) + S_M g(X_1, Y_1) - S_M h(X_1, Y_1))^2] \\ &\leq 12\mathbb{E}[(h(X_1, Y_1))^2 + (S_M g(X_1, Y_1))^2 + (S_M h(X_1, Y_1))^2]. \end{aligned}$$

Each of these three terms has to be bounded :

$$\begin{aligned}
\mathbb{E}((h(X_1, Y_1))^2) &= \iint \left(\int S_M f(x, u) \eta(x, u, y) du \right)^2 f(x, y) dx dy \\
&\leq \Delta_Y \iiint S_M f(x, u)^2 \eta(x, u, y)^2 f(x, y) dx dy du \\
&\leq \Delta_Y^2 \|f\|_\infty \|\eta\|_\infty^2 \iint S_M f(x, u)^2 dx du \\
&\leq \Delta_Y^2 \|f\|_\infty \|\eta\|_\infty^2 \|S_M f\|_2^2 \\
&\leq \Delta_Y^2 \|f\|_\infty \|\eta\|_\infty^2 \|f\|_2^2 \\
&\leq \Delta_Y^2 \|f\|_\infty^2 \|\eta\|_\infty^2
\end{aligned}$$

$$\mathbb{E}((S_M g(X_1, Y_1))^2) \leq \|f\|_\infty \|S_M g\|_2^2 \leq \|f\|_\infty \|g\|_2^2 \leq \Delta_Y^2 \|f\|_\infty^2 \|\eta\|_\infty^2$$

$$\mathbb{E}((S_M h(X_1, Y_1))^2) \leq \|f\|_\infty \|S_M h\|_2^2 \leq \|f\|_\infty \|h\|_2^2 \leq \Delta_Y^2 \|f\|_\infty^2 \|\eta\|_\infty^2$$

from previous calculations. Finally,

$$\text{Var}(L(X_1, Y_1)) \leq 36 \Delta_Y^2 \|f\|_\infty^2 \|\eta\|_\infty^2.$$

□

The last term of the Hoeffding's decomposition can also be controled :

LEMMA 5 (Computation of $\text{Cov}(U_n K, P_n L)$). *Assuming the hypotheses of Theorem 1 hold, we have*

$$\text{Cov}(U_n K, P_n L) = 0.$$

PROOF. Since $U_n K$ et $P_n L$ are centered, we have

$$\begin{aligned}
\text{Cov}(U_n K, P_n L) &= \mathbb{E}(U_n K P_n L) \\
&= \mathbb{E} \left[\frac{1}{n^2(n-1)} \sum_{j \neq k=1}^n K(X_j, Y_j, X_k, Y_k) \sum_{i=1}^n L(X_i, Y_i) \right] \\
&= \frac{1}{n} \mathbb{E}(K(X_1, Y_1, X_2, Y_2)(L(X_1, Y_1) + L(X_2, Y_2))) \\
&= 0
\end{aligned}$$

since K , L , Q and R are centered. □

The four previous lemmas give the expected result on the precision of $\hat{\theta}_n$:

LEMMA 6 (Bound for $\mathbb{E}(\hat{\theta}_n - \theta)^2$). *Assuming the hypotheses of Theorem 1 hold, we have :*

- If $m/n \rightarrow 0$,

$$\mathbb{E}(\hat{\theta}_n - \theta)^2 = O\left(\frac{1}{n}\right),$$

- Otherwise,

$$\mathbb{E}(\hat{\theta}_n - \theta)^2 \leq \gamma_2(m/n^2)$$

where γ_2 only depends on $\|f\|_\infty$, $\|\eta\|_\infty$ and Δ_Y .

PROOF. Lemmas 3, 4 and 5 imply

$$\text{Var}(\hat{\theta}_n) \leq \frac{20}{n(n-1)} \Delta_Y^2 \|f\|_\infty^2 \|\eta\|_\infty^2 (m+1) + \frac{36}{n} \Delta_Y^2 \|f\|_\infty^2 \|\eta\|_\infty^2.$$

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

$$\text{Var}(\hat{\theta}_n) \leq \gamma \Delta_Y^2 \|f\|_\infty^2 \|\eta\|_\infty^2 \left(\frac{m}{n^2} + \frac{1}{n}\right).$$

Lemma 2 gives

$$\text{Bias}^2(\hat{\theta}_n) \leq \Delta_Y^2 \|\eta\|_\infty^2 \left(\sup_{i \notin M} |c_i|^2\right)^2$$

and by assumption $(\sup_{i \notin M} |c_i|^2)^2 \approx m/n^2$. If $m/n \rightarrow 0$, then $\mathbb{E}(\hat{\theta}_n - \theta)^2 = O(\frac{1}{n})$. Otherwise $\mathbb{E}(\hat{\theta}_n - \theta)^2 \leq \gamma_2(m/n^2)$ where γ_2 only depends on $\|f\|_\infty$, $\|\eta\|_\infty$ and Δ_Y . \square

The lemma we just proved gives the result of Theorem 1 when m/n does not converge to 0. Let us now study more precisely the semiparametric case, that is when $\mathbb{E}(\hat{\theta}_n - \theta)^2 = O(\frac{1}{n})$, to prove the asymptotic normality (10) and the bound in (11). We have

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

We will study the asymptotic behavior of each of these three terms. The first one is easily treated :

LEMMA 7 (Asymptotics for $\sqrt{n}(U_n K)$). *Assuming the hypotheses of Theorem 1 hold, we have*

$$\sqrt{n}U_n K \rightarrow 0$$

in probability when $n \rightarrow \infty$ if $m/n \rightarrow 0$.

PROOF. Since $\text{Var}(\sqrt{n}U_nK) \leq \frac{20}{(n-1)} \|\eta\|_\infty^2 \|f\|_\infty^2 \Delta_Y^2(m+1)$, $\sqrt{n}U_nK$ converges to 0 in probability when $n \rightarrow \infty$ if $m/n \rightarrow 0$. \square

The random variable P_nL will be the most important term for the central limit theorem. Before studying its asymptotic normality, we need the following lemma concerning the asymptotic variance of $\sqrt{n}(P_nL)$:

LEMMA 8 (Asymptotic variance of $\sqrt{n}(P_nL)$). *Assuming the hypotheses of Theorem 1 hold, we have*

$$n \text{Var}(P_nL) \rightarrow \Lambda(f, \eta)$$

where

$$\Lambda(f, \eta) = 4 \left[\iint g(x, y)^2 f(x, y) dx dy - \left(\iint g(x, y) f(x, y) dx dy \right)^2 \right].$$

PROOF. We proved in Lemma 4 that

$$\begin{aligned} \text{Var}(L(X_1, Y_1)) &= 4\text{Var}[h(X_1, Y_1) + S_M g(X_1, Y_1) - S_M h(X_1, Y_1)] \\ &= 4\text{Var}[A_1 + A_2 + A_3] \\ &= 4 \sum_{i,j=1}^3 \text{Cov}(A_i, A_j). \end{aligned}$$

We will show that $\forall i, j \in \{1, 2, 3\}^2$, we have

$$\begin{aligned} &\left| \text{Cov}(A_i, A_j) - \epsilon_{ij} \left[\iint g(x, y)^2 f(x, y) dx dy - \left(\iint g(x, y) f(x, y) dx dy \right)^2 \right] \right| \\ (17) \quad &\leq \gamma [\|S_M f - f\|_2 + \|S_M g - g\|_2] \end{aligned}$$

where $\epsilon_{ij} = -1$ if $i = 3$ or $j = 3$ and $i \neq j$ and $\epsilon_{ij} = 1$ otherwise, and where γ depends only on $\|f\|_\infty$, $\|\eta\|_\infty$ and Δ_Y .

We shall give the details only for the case $i = j = 3$ since the calculations are similar for the other configurations. We have

$$\text{Var}(A_3) = \iint S_M^2[h(x, y)]f(x, y) dx dy - \left(\iint S_M[h(x, y)]f(x, y) dx dy \right)^2$$

We first study the quantity

$$\left| \iint S_M^2[h(x, y)]f(x, y) dx dy - \iint g(x, y)^2 f(x, y) dx dy \right|.$$

It is bounded by prout prout prout prout prout prout prout prout prout prout

$$\begin{aligned} & \iint \left| S_M^2[h(x, y)]f(x, y) - S_M^2[g(x, y)]f(x, y) \right| dx dy \\ & + \iint \left| S_M^2[g(x, y)]f(x, y) - g(x, y)^2 f(x, y) \right| dx dy \\ & \leq \|f\|_\infty \|S_M h + S_M g\|_2 \|S_M h - S_M g\|_2 + \|f\|_\infty \|S_M g + g\|_2 \|S_M g - g\|_2. \end{aligned}$$

Using the fact that S_M is a projection, this sum is bounded by

$$\begin{aligned} & \|f\|_\infty \|h + g\|_2 \|h - g\|_2 + 2\|f\|_\infty \|g\|_2 \|S_M g - g\|_2 \\ & \leq \|f\|_\infty (\|h\|_2 + \|g\|_2) \|h - g\|_2 + 2\|f\|_\infty \|g\|_2 \|S_M g - g\|_2. \end{aligned}$$

We saw previously that $\|g\|_2 \leq \Delta_Y \|f\|_\infty^{1/2} \|\eta\|_\infty$ and $\|h\|_2 \leq \Delta_Y \|f\|_\infty^{1/2} \|\eta\|_\infty$. The sum is then bounded by

$$2\Delta_Y \|f\|_\infty^{3/2} \|\eta\|_\infty \|h - g\|_2 + 2\Delta_Y \|f\|_\infty^{3/2} \|\eta\|_\infty \|S_M g - g\|_2$$

We now have to deal with $\|h - g\|_2$:

$$\begin{aligned} \|h - g\|_2^2 &= \iint \left(\int (S_M f(x, u) - f(x, u)) \eta(x, u, y) du \right)^2 dx dy \\ &\leq \iint \left(\int (S_M f(x, u) - f(x, u))^2 du \right) \left(\int \eta(x, u, y)^2 du \right) dx dy \\ &\leq \Delta_Y^2 \|\eta\|_\infty^2 \|S_M f - f\|_2^2. \end{aligned}$$

Finally, the sum is bounded by

$$2\Delta_Y \|f\|_\infty^{3/2} \|\eta\|_\infty (\Delta_Y \|\eta\|_\infty \|S_M f - f\|_2 + \|S_M g - g\|_2).$$

Let us now study the second quantity

$$\left| \left(\iint S_M[h(x, y)]f(x, y) dx dy \right)^2 - \left(\iint g(x, y)f(x, y) dx dy \right)^2 \right|.$$

It is equal to

$$\begin{aligned} & \left| \left(\iint (S_M[h(x, y)] + g(x, y))f(x, y) dx dy \right) \right. \\ & \left. \left(\iint (S_M[h(x, y)] - g(x, y))f(x, y) dx dy \right) \right|. \end{aligned}$$

By using the Cauchy-Schwarz inequality, it is bounded by

$$\begin{aligned}
& \|f\|_2 \|S_M h + g\|_2 \|f\|_2 \|S_M h - g\|_2 \\
& \leq \|f\|_2^2 (\|h\|_2 + \|g\|_2) (\|S_M h - S_M g\|_2 + \|S_M g - g\|_2) \\
& \leq 2\Delta_Y \|f\|_\infty^{3/2} \|\eta\|_\infty (\|h - g\|_2 + \|S_M g - g\|_2) \\
& \leq 2\Delta_Y \|f\|_\infty^{3/2} \|\eta\|_\infty (\Delta_Y \|\eta\|_\infty \|S_M f - f\|_2 + \|S_M g - g\|_2)
\end{aligned}$$

by using the previous calculations. Collecting the two inequalities gives (17) for $i = j = 3$.

Finally, since by assumption $\forall t \in \mathbb{L}^2(d\mu)$, $\|S_M t - t\|_2 \rightarrow 0$ when $n \rightarrow \infty$, a direct consequence of (17) is that

$$\begin{aligned}
& \lim_{n \rightarrow \infty} \text{Var}(L(X_1, Y_1)) \\
& = 4 \left[\iint g(x, y)^2 f(x, y) dx dy - \left(\iint g(x, y) f(x, y) dx dy \right)^2 \right] \\
& = \Lambda(f, \eta).
\end{aligned}$$

We then conclude by noting that $\text{Var}(\sqrt{n}(P_n L)) = \text{Var}(L(X_1, Y_1))$. \square

We can now study the convergence of $\sqrt{n}(P_n L)$, which is given in the following lemma:

LEMMA 9 (Asymptotic normality of $\sqrt{n}(P_n L)$). *Assuming the hypotheses of Theorem 1 hold, we have*

$$\sqrt{n} P_n L \xrightarrow{\mathcal{L}} \mathcal{N}(0, \Lambda(f, \eta)).$$

PROOF. We first note that

$$\sqrt{n} \left(P_n(2g) - 2 \iint g(x, y) f(x, y) dx dy \right) \rightarrow \mathcal{N}(0, \Lambda(f, \eta))$$

where $g(x, y) = \int \eta(x, y, u) f(x, u) du$.

It is then sufficient to show that the expectation of the square of

$$R = \sqrt{n} \left[P_n L - \left(P_n(2g) - 2 \iint g(x, y) f(x, y) dx dy \right) \right]$$

converges to 0. We have

$$\begin{aligned}
\mathbb{E}(R^2) &= \text{Var}(R) \\
&= n \text{Var}(P_n L) + n \text{Var}(P_n(2g)) - 2n \text{Cov}(P_n L, P_n(2g))
\end{aligned}$$

We know that $n\text{Var}(P_n(2g)) \rightarrow \Lambda(f, \eta)$ and Lemma 8 shows that $n\text{Var}(P_n L) \rightarrow \Lambda(f, \eta)$. Then, we just have to prove that

$$\lim_{n \rightarrow \infty} n\text{Cov}(P_n L, P_n(2g)) = \Lambda(f, \eta).$$

We have

$$n\text{Cov}(P_n L, P_n(2g)) = \mathbb{E}(2L(X_1, Y_1)g(X_1, Y_1))$$

because L is centered. Since

$$L(X_1, Y_1) = 2h(X_1, Y_1) + 2S_M g(X_1, Y_1) - 2S_M h(X_1, Y_1) - 4A'B + 2A'CA,$$

we get

$$\begin{aligned} n\text{Cov}(P_n L, P_n(2g)) &= 4 \iint h(x, y)g(x, y)f(x, y)dx dy \\ &+ 4 \iint S_M g(x, y)g(x, y)f(x, y)dx dy \\ &- 4 \iint S_M h(x, y)g(x, y)f(x, y)dx dy - 8 \sum_i a_i b_i \iint g(x, y)f(x, y)dx dy \\ &+ 4A'CA \iint g(x, y)f(x, y)dx dy \end{aligned}$$

which converges to $4 \left[\iint g(x, y)^2 f(x, y)dx dy - \left(\iint g(x, y)f(x, y)dudxdy \right)^2 \right]$ which is equal to $\Lambda(f, \eta)$. We finally deduce that

$$\sqrt{n}P_n L \rightarrow \mathcal{N}(0, \Lambda(f, \eta))$$

in distribution. □

In order to prove the asymptotic normality of $\hat{\theta}_n$, the last step is to control the remainder term in the Hoeffding's decomposition:

LEMMA 10 (Asymptotics for $\sqrt{n}(2A'B - A'CA)$). *Assuming the hypotheses of Theorem 1 hold, we have*

$$\sqrt{n}(2A'B - A'CA - \theta) \rightarrow 0.$$

PROOF. $\sqrt{n}(2A'B - A'CA - \theta) \rightarrow 0$ is equal to

$$\begin{aligned} &\sqrt{n} \left[2 \iint g(x, y)S_M f(x, y)dx dy \right. \\ &- \iiint S_M f(x, y_1)S_M f(x, y_2)\eta(x, y_1, y_2)dx dy_1 dy_2 \\ &\left. - \iiint f(x, y_1)f(x, y_2)\eta(x, y_1, y_2)dx dy_1 dy_2 \right]. \end{aligned}$$

By replacing g we get

$$\begin{aligned} & \sqrt{n} \left[2 \iiint S_M f(x, y_1) f(x, y_2) \eta(x, y_1, y_2) dx dy_1 dy_2 \right. \\ & - \iiint S_M f(x, y_1) S_M f(x, y_2) \eta(x, y_1, y_2) dx dy_1 dy_2 \\ & \left. - \iiint f(x, y_1) f(x, y_2) \eta(x, y_1, y_2) dx dy_1 dy_2 \right] \end{aligned}$$

With integral manipulation, we show it is also equal to

$$\begin{aligned} & \sqrt{n} \left[\iiint S_M f(x, y_1) (f(x, y_2) - S_M f(x, y_2)) \eta(x, y_1, y_2) dx dy_1 dy_2 \right. \\ & \left. - \iiint f(x, y_2) (S_M f(x, y_1) - f(x, y_1)) \eta(x, y_1, y_2) dx dy_1 dy_2 \right] \\ & \leq \sqrt{n} \Delta_Y \|\eta\|_\infty (\|S_M f\|_2 \|S_M f - f\|_2 + \|f\|_2 \|S_M f - f\|_2) \\ & \leq 2\sqrt{n} \Delta_Y \|f\|_2 \|\eta\|_\infty \|S_M f - f\|_2 \\ & \leq 2\sqrt{n} \Delta_Y \|f\|_2 \|\eta\|_\infty \left(\sup_{i \notin M} |c_i|^2 \right)^{1/2} \\ & \approx 2\Delta_Y \|f\|_2 \|\eta\|_\infty \sqrt{\frac{m}{n}}, \end{aligned}$$

which converges to 0 when $n \rightarrow \infty$ since $m/n \rightarrow 0$. \square

Collecting now the results of Lemmas 7, 8 and 10 we get (10) since

$$\sqrt{n} (\hat{\theta}_n - \theta) \rightarrow \mathcal{N}(0, \Lambda(f, \eta))$$

in distribution. We finally have to prove (11). Remark that

$$\begin{aligned} n\mathbb{E} (\hat{\theta}_n - \theta)^2 &= n\text{Bias}^2(\hat{\theta}_n) + n\text{Var}(\hat{\theta}_n) \\ &= n\text{Bias}^2(\hat{\theta}_n) + n\text{Var}(U_n K) + n\text{Var}(P_n L) \end{aligned}$$

We previously proved that

$$\begin{aligned} n\text{Bias}^2(\hat{\theta}_n) &\leq \lambda \Delta_Y^2 \|\eta\|_\infty^2 \frac{m}{n} \text{ for some } \lambda \in \mathbb{R}, \\ n\text{Var}(U_n K) &\leq \mu \Delta_Y^2 \|f\|_\infty^2 \|\eta\|_\infty^2 \frac{m}{n} \text{ for some } \mu \in \mathbb{R}. \end{aligned}$$

Moreover, (17) imply

$$|n\text{Var}(P_n L) - \Lambda(f, \eta)| \leq \gamma [\|S_M f - f\|_2 + \|S_M g - g\|_2],$$

where γ is a increasing function of $\|f\|_\infty, \|\eta\|_\infty$ and Δ_Y . We then deduce (11) which ends the proof of Theorem 1.

7.3. *Proof of Theorem 2.* To prove the inequality we will use the work of [6] (see also chapter 25 of [16]) on efficient estimation. The first step is the computation of the Fréchet derivative of $\theta(f)$ at a point f_0 . Straightforward calculations show that

$$\begin{aligned} \theta(f) - \theta(f_0) &= \iint \left[2 \int \psi(x, y, z) f_0(x, z) dz \right] (f(x, y) - f_0(x, y)) dx dy \\ &\quad + O \left(\iint (f(x, y) - f_0(x, y))^2 dx dy \right) \end{aligned}$$

from which we deduce that the Fréchet derivative of $\theta(f)$ at f_0 is

$$\theta'(f_0) \cdot u = \left\langle 2 \int \psi(x, y, z) f_0(x, z) dz, u \right\rangle \quad (u \in \mathbb{L}^2(dxdy)),$$

where $\langle \cdot, \cdot \rangle$ is the scalar product in $\mathbb{L}^2(dxdy)$. We can now use the results of [6]. Denote $H(f_0) = H(f_0) = \left\{ u \in \mathbb{L}^2(dxdy), \iint u(x, y) \sqrt{f_0(x, y)} dx dy = 0 \right\}$ the set of functions in $\mathbb{L}^2(dxdy)$ orthogonal to $\sqrt{f_0}$, $\text{Proj}_{H(f_0)}$ the projection on $H(f_0)$, $A_n(t) = (\sqrt{f_0})t/\sqrt{n}$ and $P_{f_0}^{(n)}$ the joint distribution of (X_1, \dots, X_n) under f_0 . Since here X_1, \dots, X_n are i.i.d., $\{P_f^{(n)}, f \in \mathcal{E}\}$ is locally asymptotically normal at all points $f_0 \in \mathcal{E}$ in the direction $H(f_0)$ with normalizing factor $A_n(f_0)$. Ibragimov and Khas'minskii result say that under these conditions, denoting $K_n = B_n \theta'(f_0) A_n \text{Proj}_{H(f_0)}$ with $B_n(u) = \sqrt{nu}$, if $K_n \rightarrow K$ weakly and if $K(u) = \langle t, u \rangle$, then for every estimator $\hat{\theta}_n$ of $\theta(f)$ and every family $\mathcal{V}(f_0)$ of vicinities of f_0 , we have

$$\inf_{\{\mathcal{V}(f_0)\}} \liminf_{n \rightarrow \infty} \sup_{f \in \mathcal{V}(f_0)} n \mathbb{E}(\hat{\theta}_n - \theta(f_0))^2 \geq \|t\|_{\mathbb{L}^2(dxdy)}^2.$$

Here,

$$K_n(u) = \sqrt{n} \theta'(f_0) \cdot \frac{1}{\sqrt{n}} \sqrt{f_0} \text{Proj}_{H(f_0)}(u) = \theta'(f_0) \cdot \left(\sqrt{f_0} \left(u - \sqrt{f_0} \int u \sqrt{f_0} \right) \right)$$

does not depend on n and

$$\begin{aligned} K(u) &= \iint \left[2 \int \psi(x, y, z) f_0(x, z) dz \right] \sqrt{f_0(x, y)} \\ &\quad \left(u(x, y) - \sqrt{f_0(x, y)} \int u \sqrt{f_0} \right) dx dy \\ &= \iint \left[2 \int \psi(x, y, z) f_0(x, z) dz \right] \sqrt{f_0(x, y)} u(x, y) dx dy \\ &\quad - \iint \left[2 \int \psi(x, y, z) f_0(x, z) dz \right] f_0(x, y) dx dy \int u \sqrt{f_0} \\ &= \langle t, u \rangle \end{aligned}$$

where

$$t(x, y) = \left[2 \int \psi(x, y, z) f_0(x, z) dz \right] \sqrt{f_0(x, y)} - \left(\iint \left[2 \int \psi(x, y, z) f_0(x, z) dz \right] f_0(x, y) dx dy \right) \sqrt{f_0(x, y)}.$$

The semiparametric Cramér-Rao bound for our problem is $\|t\|_{\mathbb{L}^2(dx dy)}^2$:

$$\begin{aligned} \|t\|_{\mathbb{L}^2(dx dy)}^2 &= 4 \iint \left[\int \psi(x, y, z) f_0(x, z) dz \right]^2 f_0(x, y) dx dy \\ &\quad - 4 \left(\iint \left[\int \psi(x, y, z) f_0(x, z) dz \right] f_0(x, y) dx dy \right)^2 \\ &= 4 \iint g_0(x, y)^2 f_0(x, y) dx dy - 4 \left(\iint g_0(x, y) f_0(x, y) \right)^2 \end{aligned}$$

where $g_0(x, y) = \int \psi(x, y, z) f_0(x, z) dz$. Finally, we recognize the expression of $\Lambda(f_0, \psi)$ given in Theorem 1.

7.4. *Proof of Theorem 3.* We will first control the remainder term Γ_n :

$$\Gamma_n = \frac{1}{6} F'''(\xi) (1 - \xi)^3.$$

Let us recall that

$$\begin{aligned} F'''(\xi) &= \iiint \frac{\left(\int \hat{f}(x, y) dy \right)^2}{\left(\int \xi f(x, y) + (1 - \xi) \hat{f}(x, y) dy \right)^5} \\ &\quad [(\hat{m}(x) - \varphi(y))(\hat{m}(x) - \varphi(z))(\hat{m}(x) - \varphi(t)) \\ &\quad \left(\int \hat{f}(x, y) dy \right) \ddot{\psi}(\hat{r}(\xi, x)) - 3(\hat{m}(x) - \varphi(y))(\hat{m}(x) - \varphi(z)) \\ &\quad \left(\int [\xi f(x, y) + (1 - \xi) \hat{f}(x, y)] dy \right) \ddot{\psi}(\hat{r}(\xi, x))] \\ &\quad (f(x, y) - \hat{f}(x, y)) (f(x, z) - \hat{f}(x, z)) \\ &\quad (f(x, t) - \hat{f}(x, t)) dx dy dz dt \end{aligned}$$

Assumptions A2 and A3 ensure that the first part of the integrand is bounded by a constant μ :

$$\begin{aligned}\Gamma_n &\leq \frac{1}{6}\mu \iiint\limits_t |f(x, y) - \hat{f}(x, y)| |f(x, z) - \hat{f}(x, z)| \\ &\quad |f(x, t) - \hat{f}(x, t)| dx dy dz dt \\ &\leq \frac{1}{6}\mu \int \left(\int |f(x, y) - \hat{f}(x, y)| dy \right)^3 dx \\ &\leq \frac{1}{6}\mu \Delta_Y^2 \iint |f(x, y) - \hat{f}(x, y)|^3 dx dy\end{aligned}$$

by the Hölder inequality. Then $\mathbb{E}(\Gamma_n^2) = O(\mathbb{E}[(f|f-\hat{f}|^3)^2]) = O(\mathbb{E}[\|f-\hat{f}\|_3^6])$. Since \hat{f} verifies assumption A2, this quantity has order $O(n_1^{-6\lambda})$. If we further assume that $n_1 \approx n/\log(n)$ and $\lambda > 1/6$, we get $E(\Gamma_n^2) = o(\frac{1}{n})$, which proves that the remainder term Γ_n is negligible. We will now show that $\sqrt{n}(\hat{T}_n - T(f))$ and $Z_n = \frac{1}{n_2} \sum_{j=1}^{n_2} H(f, X_j, Y_j) - \iint H(f, x, y) f(x, y) dx dy$ have the same asymptotic behavior. The idea is that we can easily get a central limit theorem for Z_n with asymptotic variance

$$C(f) = \iint H(f, x, y)^2 f(x, y) dx dy - \left(\iint H(f, x, y) f(x, y) dx dy \right)^2,$$

which imply both (12) and (13) (we will show at the end of the proof that $C(f)$ can be expressed such as in the theorem). In order to show that $\sqrt{n}(\hat{T}_n - T(f))$ and Z_n have the same asymptotic behavior, we will prove that

$$R = \sqrt{n} \left[\hat{T}_n - T(f) - \left(\frac{1}{n_2} \sum_{j=1}^{n_2} H(f, X_j, Y_j) - \iint H(f, x, y) f(x, y) dx dy \right) \right]$$

has a second-order moment converging to 0. Let us note that $R = R_1 + R_2$ where

$$\begin{aligned}R_1 &= \sqrt{n} \left[\hat{T}_n - T(f) \right. \\ &\quad \left. - \left(\frac{1}{n_2} \sum_{j=1}^{n_2} H(\hat{f}, X_j, Y_j) - \iint H(\hat{f}, x, y) f(x, y) dx dy \right) \right], \\ R_2 &= \sqrt{n} \left[\frac{1}{n_2} \sum_{j=1}^{n_2} \left(H(\hat{f}, X_j, Y_j) - \iint H(\hat{f}, x, y) f(x, y) dx dy \right) \right] \\ &\quad - \sqrt{n} \left[\frac{1}{n_2} \sum_{j=1}^{n_2} \left(H(f, X_j, Y_j) - \iint H(f, x, y) f(x, y) dx dy \right) \right].\end{aligned}$$

We propose to show that both $\mathbb{E}(R_1^2)$ and $\mathbb{E}(R_2^2)$ converge to 0. We can write R_1 as follows :

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

where

$$\begin{aligned} Q' &= \iiint K(\hat{f}, x, y, z) f(x, y) f(x, z), \\ K(\hat{f}, x, y, z) &= \frac{1}{2} \frac{\ddot{\psi}(\hat{m}(x))}{\left(\int \hat{f}(x, y) dy \right)} (\hat{m}(x) - \varphi(y)) (\hat{m}(x) - \varphi(z)) \end{aligned}$$

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

LEMMA 11 (Asymptotics for $\sqrt{n} \left(\hat{Q}' - Q' \right)$). *Assuming the hypotheses of Theorem 3 hold, we have*

$$\lim_{n \rightarrow \infty} n \mathbb{E} \left(\hat{Q}' - Q' \right)^2 = 0.$$

PROOF. The bound given in (11) states that if $|M_n|/n \rightarrow 0$ we have

$$\begin{aligned} & \left| n \mathbb{E} \left[\left(\hat{Q}' - Q' \right)^2 \mid \hat{f} \right] \right. \\ & \quad \left. - 4 \left[\iint \hat{g}(x, y)^2 f(x, y) dx dy - \left(\iint \hat{g}(x, y) f(x, y) dx dy \right)^2 \right] \right| \\ & \leq \gamma_1 (\|f\|_\infty, \|\psi\|_\infty, \Delta_Y) \left[\frac{|M_n|}{n} + \|S_M f - f\|_2 + \|S_M \hat{g} - \hat{g}\|_2 \right] \end{aligned}$$

where $\hat{g}(x, y) = \int K(\hat{f}, x, y, z) f(x, z) dz$. By deconditioning, we get

$$\begin{aligned} & \left| n \mathbb{E} \left[\left(\hat{Q}' - Q' \right)^2 \right] \right. \\ & \quad \left. - 4 \mathbb{E} \left[\iint \hat{g}(x, y)^2 f(x, y) dx dy - \left(\iint \hat{g}(x, y) f(x, y) dx dy \right)^2 \right] \right| \\ & \leq \gamma_1 (\|f\|_\infty, \|\psi\|_\infty, \Delta_Y) \left[\frac{|M_n|}{n} + \|S_M f - f\|_2 + \mathbb{E} (\|S_M \hat{g} - \hat{g}\|_2) \right]. \end{aligned}$$

Note that

$$\begin{aligned} \mathbb{E} (\|S_M \hat{g} - \hat{g}\|_2) &\leq \mathbb{E} (\|S_M \hat{g} - S_M g\|_2) + \mathbb{E} (\|S_M g - g\|_2) \\ &\leq \mathbb{E} (\|\hat{g} - g\|_2) + \mathbb{E} (\|S_M g - g\|_2) \end{aligned}$$

where $g(x, y) = \int K(f, x, y, z)f(x, z)dz$. The second term converges to 0 since $g \in \mathbb{L}^2(dxdy)$ and $\forall t \in \mathbb{L}^2(dxdy)$, $\int (S_M t - t)^2 d\mu \rightarrow 0$. Moreover

$$\begin{aligned}
\|\hat{g} - g\|_2^2 &= \iint [\hat{g}(x, y) - g(x, y)]^2 f(x, y) dx dy \\
&= \iint \left[\int \left(K(\hat{f}, x, y, z) - K(f, x, y, z) \right) f(x, z) dz \right]^2 f(x, y) dx dy \\
&\leq \iint \left[\int \left(K(\hat{f}, x, y, z) - K(f, x, y, z) \right)^2 dz \right] \\
&\quad \left[\int f(x, z)^2 dz \right] f(x, y) dx dy \\
&\leq \Delta_Y^2 \|f\|_\infty^3 \iiint \left(K(\hat{f}, x, y, z) - K(f, x, y, z) \right)^2 dx dz \\
&\leq \delta \Delta_Y^3 \|f\|_\infty^3 \iint (f(x, y) - \hat{f}(x, y))^2 dx dy
\end{aligned}$$

for some constant δ by applying the mean value theorem to $K(f, x, y, z) - K(\hat{f}, x, y, z)$. Of course, the bound δ is obtained here by considering assumptions A1, A2 and A3. Since $\mathbb{E}(\|f - \hat{f}\|_2) \rightarrow 0$, we get $\mathbb{E}(\|\hat{g} - g\|_2) \rightarrow 0$. Let us now show that the expectation of

$$\iint \hat{g}(x, y)^2 f(x, y) dx dy - \left(\iint \hat{g}(x, y) f(x, y) dx dy \right)^2$$

converges to 0. We will only develop the proof for the first term :

$$\begin{aligned}
&\left| \iint \hat{g}(x, y)^2 f(x, y) dx dy - \iint g(x, y)^2 f(x, y) dx dy \right| \\
&\leq \iint \left| \hat{g}(x, y)^2 - g(x, y)^2 \right| f(x, y) dx dy \\
&\leq \lambda \iint (\hat{g}(x, y) - g(x, y))^2 dx dy \\
&\leq \lambda \|\hat{g} - g\|_2^2
\end{aligned}$$

for some constant λ . By taking the expectation of both sides, we see it is enough to show that $\mathbb{E}(\|\hat{g} - g\|_2^2) \rightarrow 0$, which is done exactly as above. Besides, we can verify that

$$\begin{aligned}
g(x, y) &= \int K(f, x, y, z)f(x, z)dz \\
&= \frac{1}{2} \frac{\ddot{\psi}(m(x))}{\left(\int f(x, y) dy \right)} (m(x) - \varphi(y)) \\
&\quad \left(m(x) \int f(x, z) dz - \int \varphi(z) f(x, z) dz \right) \\
&= 0,
\end{aligned}$$

which proves that the expectation of $\iint \hat{g}(x, y)^2 f(x, y) dx dy$ converges to 0. Similar considerations show that the expectation of the second term $\left(\iint \hat{g}(x, y) f(x, y) dx dy\right)^2$ also converges to 0. We finally have

$$\lim_{n \rightarrow \infty} n \mathbb{E} \left(\hat{Q}' - Q' \right)^2 = 0.$$

□

Lemma 11 imply that $\mathbb{E}(R_1^2) \rightarrow 0$. We will now prove that $\mathbb{E}(R_2^2) \rightarrow 0$:

$$\begin{aligned} \mathbb{E}(R_2^2) &= \frac{n}{n_2} \mathbb{E} \left[\iint \left(H(f, x, y) - H(\hat{f}, x, y) \right)^2 f(x, y) dx dy \right] \\ &\quad - \frac{n}{n_2} \mathbb{E} \left[\iint H(f, x, y) f(x, y) dx dy - \iint H(\hat{f}, x, y) f(x, y) dx dy \right]^2. \end{aligned}$$

The same arguments as before (mean value theorem and assumptions A2 and A3) show that $\mathbb{E}(R_2^2) \rightarrow 0$. At last, we can give another expression for the asymptotic variance :

$$C(f) = \iint H(f, x, y)^2 f(x, y) dx dy - \left(\iint H(f, x, y) f(x, y) dx dy \right)^2.$$

We will prove that

$$C(f) = \mathbb{E} \left(\text{Var}(\varphi(Y)|X) \left[\dot{\psi}(\mathbb{E}(Y|X)) \right]^2 \right) + \text{Var}(\psi(\mathbb{E}(\varphi(Y)|X))).$$

Remark that

$$\begin{aligned} \iint H(f, x, y) f(x, y) dx dy &= \iint \left([\varphi(y) - m(x)] \dot{\psi}(m(x)) + \psi(m(x)) \right) \\ &\quad f(x, y) dx dy \\ &= \iint m(x) \dot{\psi}(m(x)) f(x, y) dx dy \\ &\quad - \iint m(x) \dot{\psi}(m(x)) f(x, y) dx dy \\ &\quad + \iint \psi(m(x)) f(x, y) dx dy \\ (18) \qquad \qquad \qquad &= \mathbb{E}(\psi(\mathbb{E}(\varphi(Y)|X))). \end{aligned}$$

Moreover,

$$\begin{aligned} H(f, x, y)^2 &= [\varphi(y) - m(x)]^2 \dot{\psi}(m(x))^2 + \psi(m(x))^2 \\ &\quad + 2 [\varphi(y) - m(x)] \dot{\psi}(m(x)) \psi(m(x)) \\ &= \varphi(y)^2 \dot{\psi}(m(x))^2 + m(x)^2 \dot{\psi}(m(x))^2 - 2\varphi(y)m(x) \dot{\psi}(m(x))^2 \\ &\quad + \psi(m(x))^2 + 2 [\varphi(y) - m(x)] \dot{\psi}(m(x)) \psi(m(x)). \end{aligned}$$

We can then rewrite $\iint H(f, x, y)^2 f(x, y) dx dy$ as:

$$\begin{aligned}
& \iint \varphi(y)^2 \dot{\psi}(m(x))^2 f(x, y) dx dy + \iint m(x)^2 \dot{\psi}(m(x))^2 f(x, y) dx dy \\
& - 2 \iint \varphi(y) m(x) \dot{\psi}(m(x))^2 f(x, y) dx dy + \iint \psi(m(x))^2 f(x, y) dx dy \\
& + 2 \iint \varphi(y) \dot{\psi}(m(x)) \psi(m(x)) f(x, y) dx dy \\
& - 2 \iint m(x) \dot{\psi}(m(x)) \psi(m(x)) f(x, y) dx dy \\
= & \iint v(x) \dot{\psi}(m(x))^2 f(x, y) dx dy - \iint m(x)^2 \dot{\psi}(m(x)) f(x, y) dx dy \\
& + \iint \psi(m(x))^2 f(x, y) dx dy \\
= & \iint \left([v(x) - m(x)^2] \dot{\psi}(m(x))^2 + \psi(m(x))^2 \right) f(x, y) dx dy \\
= & \mathbb{E} \left([v(X) - m(X)^2] \dot{\psi}(m(X))^2 \right) + \mathbb{E} \left(\psi(m(X))^2 \right) \\
= & \mathbb{E} \left(\left[\mathbb{E}(\varphi(Y)^2 | X) - \mathbb{E}(\varphi(Y) | X)^2 \right] \left[\dot{\psi}(\mathbb{E}(\varphi(Y) | X)) \right]^2 \right) \\
& + \mathbb{E} \left(\psi(\mathbb{E}(\varphi(Y) | X))^2 \right) \\
= & \mathbb{E} \left(\text{Var}(\varphi(Y) | X) \left[\dot{\psi}(\mathbb{E}(\varphi(Y) | X)) \right]^2 \right) + \mathbb{E} \left(\psi(\mathbb{E}(\varphi(Y) | X))^2 \right)
\end{aligned}$$

where we have set $v(x) = \int \varphi(y)^2 f(x, y) dy / \int f(x, y) dy$. This result and (18) give the desired form for $C(f)$ which ends the proof of Theorem 3.

7.5. *Proof of Theorem 4.* We follow the proof of Theorem 2. Assumptions A2 and A3 imply that

$$\begin{aligned}
T(f) - T(f_0) &= \iint \left([\varphi(y) - m_0(x)] \dot{\psi}(m_0(x)) + \psi(m_0(x)) \right) \\
& \quad \left(f(x, y) - f_0(x, y) \right) dx dy + O \left(\int (f - f_0)^2 \right)
\end{aligned}$$

where $m_0(x) = \int \varphi(y) f_0(x, y) dy / \int f_0(x, y) dy$. This result shows that the Fréchet derivative of $T(f)$ at f_0 is $T'(f_0) \cdot h = \langle H(f_0, \cdot), h \rangle$ where

$$H(f_0, x, y) = \left([\varphi(y) - m_0(x)] \dot{\psi}(m_0(x)) + \psi(m_0(x)) \right).$$

We then deduce that

$$\begin{aligned}
K(h) &= T'(f_0) \cdot \left(\sqrt{f_0} \left(h - \sqrt{f_0} \int h \sqrt{f_0} \right) \right) \\
&= \int H(f_0, \cdot) \sqrt{f_0} h - \int H(f_0, \cdot) \sqrt{f_0} \int h \sqrt{f_0} \\
&= \langle t, h \rangle
\end{aligned}$$

with

$$t = H(f_0, \cdot) \sqrt{f_0} - \left(\int H(f_0, \cdot) f_0 \right) \sqrt{f_0}.$$

The semiparametric Cramér-Rao bound for this problem is thus

$$\|t\|_{\mathbb{L}^2(dx dy)}^2 = \int H(f_0, \cdot)^2 f_0 - \left(\int H(f_0, \cdot) f_0 \right)^2 = C(f_0)$$

where we recognize the expression of $C(f_0)$ in Theorem 4.

7.6. Proof of Theorem 5. We follow the proof of Theorem 4. We write $T(f) - T(f_0)$ as

$$\begin{aligned}
&\iint \frac{1}{\left(\int f_0(x, \mathbf{z}, y) d\mathbf{z} dy \right)^2} \left[2y \int y f_0(x, \mathbf{z}, y) d\mathbf{z} dy \int f_0(x, \mathbf{z}, y) d\mathbf{z} dy \right. \\
&\quad \left. - \left(\int y f_0(x, \mathbf{z}, y) d\mathbf{z} dy \right)^2 \right] (f(x, \mathbf{z}, y) - f_0(x, \mathbf{z}, y)) dx d\mathbf{z} dy \\
&+ O \left(\iiint (f(x, \mathbf{z}, y) - f_0(x, \mathbf{z}, y))^2 dx d\mathbf{z} dy \right).
\end{aligned}$$

The Fréchet derivative of $T(f)$ at f_0 is then given by

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

where $H(f_0, x, \mathbf{z}, y)$ equals

$$\frac{2y \int y f_0(x, \mathbf{z}, y) d\mathbf{z} dy \int f_0(x, \mathbf{z}, y) d\mathbf{z} dy - \left(\int y f_0(x, \mathbf{z}, y) d\mathbf{z} dy \right)^2}{\left(\int f_0(x, \mathbf{z}, y) d\mathbf{z} dy \right)^2}.$$

Note that $H(f_0, x, \mathbf{z}, y)$ does not depend on \mathbf{z} . Using the identities

$$\begin{aligned}
\int f_0(x, \mathbf{z}, y) d\mathbf{z} dy &= \int f(x, y) dy, \\
\int y f_0(x, \mathbf{z}, y) d\mathbf{z} dy &= \int y f(x, y) dy,
\end{aligned}$$

where $f(x, y)$ is the density of (X, Y) , H can be written as:

$$H(f_0, x, \mathbf{z}, y) = \frac{2y \int y f_0(x, y) dy \int f_0(x, y) dy - (\int y f_0(x, y) dy)^2}{(\int f_0(x, y) dy)^2}.$$

We hence remark that $H(f_0, x, \mathbf{z}, y)$ equals $H(f_0, x, y)$, the function appearing in the derivative of $T(f)$ when we only consider the pair (X, Y) to estimate $T(f)$. Thus, the semiparametric Cramér-Rao bound equals the bound $C(f_0)$ computed in Theorem 4.

Acknowledgements. Many thanks are due to A. Antoniadis, B. Laurent and F. Wahl for helpful discussion.

References.

- [1] CARRASCO, N., BANASZKIEWICZ, M., THISSEN, R., DUTUIT, O., AND PERNOT, P. (2007). Uncertainty analysis of bimolecular reactions in Titan ionosphere chemistry model. *Planetary and Space Science* 55, 141–157.
- [2] CUKIER, R. I., FORTUIN, C. M., SHULER, K. E., PETSCHKE, A. G., AND SCHAIBLY, J. H. (1973). Study of the sensitivity of coupled reaction systems to uncertainties in rate coefficients. I Theory. *The Journal of Chemical Physics* 59, 3873–3878.
- [3] DA VEIGA, S., WAHL, F., AND GAMBOA, F. (2006). Local polynomial estimation for sensitivity analysis on models with correlated inputs. *submitted to Technometrics*.
- [4] FERRIGNO, S. AND DUCHARME, G. R. (2005). Un test d'adéquation global pour la fonction de répartition conditionnelle. *Comptes rendus. Mathématique* 341, 313–316.
- [5] HOUDRÉ, C. AND REYNAUD, P. (2002). Stochastic inequalities and applications. In *Euroconference on Stochastic inequalities and applications*. Birkhauser.
- [6] IBRAGIMOV, I. A. AND KHAS'MINSKII, R. Z. (1991). Asymptotically normal families of distributions and efficient estimation. *The Annals of Statistics* 19, 1681–1724.
- [7] KERKYACHARIAN, G. AND PICARD, D. (1996). Estimating nonquadratic functionals of a density using haar wavelets. *The Annals of Statistics* 24, 485–507.
- [8] LAURENT, B. (1996). Efficient estimation of integral functionals of a density. *The Annals of Statistics* 24, 659–681.
- [9] LAURENT, B. (2005). Adaptive estimation of a quadratic functional of a density by model selection. *ESAIM: Probability and Statistics* 9, 1–19.
- [10] LEVIT, B. Y. (1978). Asymptotically efficient estimation of nonlinear functionals. *Problems Inform. Transmission* 14, 204–209.
- [11] LI, K. C. (1991). Sliced inverse regression for dimension reduction. *Journal of the American Statistical Association* 86, 316–327.
- [12] MCKAY, M. D. (1995). Evaluating prediction uncertainty. Tech. Rep. NUREG/CR-6311, U.S. Nuclear Regulatory Commission and Los Alamos National Laboratory.
- [13] OAKLEY, J. E. AND O'HAGAN, A. (2004). Probabilistic sensitivity analysis of complex models : a bayesian approach. *Journal of the Royal Statistical Society Series B* 66, 751–769.
- [14] SOBOL', I. M. (1993). Sensitivity estimates for nonlinear mathematical models. *MMCE* 1, 407–414.
- [15] TURANYI, T. (1990). Sensitivity analysis of complex kinetic systems. *Journal of Mathematical Chemistry* 5, 203–248.

- [16] VAN DER VAART, A. W. (1998). *Asymptotic Statistics*. Cambridge: Cambridge University Press.

IFP-LYON
BP3
69390 VERNAISON
FRANCE
[HTTP://WWW.LSP.UPS-TLSE.FR/Fp/DAVEIGA](http://www.lsp.ups-tlse.fr/Fp/DAVEIGA)
E-MAIL: sebastien.da-veiga@ifp.fr

INSTITUT DE MATHÉMATIQUES
UNIVERSITÉ PAUL SABATIER
31062 TOULOUSE CEDEX 9
FRANCE
[HTTP://WWW.LSP.UPS-TLSE.FR/Fp/GAMBOA](http://www.lsp.ups-tlse.fr/Fp/GAMBOA)
E-MAIL: gamboa@cict.fr