

# Cross-validation method for bivariate measure with certain mixture

**Rachid Sabre**

LE2I UMR6306, CNRS, ENSAM, Univ. Bourgogne Franche-Comté, Agrosup, Département Sciences Pour l'Ingénieur et Procédés, 26, Bd Docteur Petitjean 21000 Dijon France

E-mail: [r.sabre@agrosupdijon.fr](mailto:r.sabre@agrosupdijon.fr)

**Abstract.** We consider a pair of random variables  $(X, Y)$  whose probability measure is the sum of an absolutely continuous measure, a discrete measure and a finite number of absolutely continuous measures on several lines. An asymptotically unbiased and consistent estimate of the density of the continuous part is given in [13]. In this work, we focus on the choice of these parameters so that this estimate will be optimal and the rate of convergence will be better. we as well as its rate of convergence. To achieve this we use the cross-validation technics.

## 1. Introduction

It has been studied in Sabre [13] the estimation of the bivariate density of the continuous part of a certain mixture. More precisely, he considered a pair of the random variables  $(X, Y)$  whose probability measure is the sum of an absolutely continuous measure with respect to Lebesgue measure, a discrete measure and a finite number of absolutely continuous measures on several lines:

$$d\mu = f(x, y)dxdy + \sum_{m=1}^q a'_m \delta_{(\omega_{1m}, \omega_{2m})} + \sum_{k=1}^{q'} \varphi_k(u_1) \delta_{(u_1, a_k u_1 + b_k)}, \quad (1)$$

where  $\delta$  is a Dirac measure,  $f$  and  $\varphi_i$  are nonnegative uniformly continuous integrable functions. The motivation for the choice of such model is illustrated through a concrete example in Sabre [14]. It concerns the study of structural fissuration of the agricultural soil. On a homogenous soil, measure of the resistance variable  $X$  and the humidity variable  $Y$  are taken on several locations at a depth of 30 cm. The measurement values are distributed according to a continuous law, except in certain locations where the experimentalist finds small galleries where measurement values of resistance and humidity decrease (the presence of jumps). When the measures are made in places where the passage of tractors is frequent, the variable  $Y$  becomes linear with respect to the variable  $X$  and their measures follow a new distribution noted  $\varphi_i$  (the presence of some measures continuous on the lines determined by frequent passage of tractors).

An asymptotically unbiased and consistent estimate of the density of the continuous part "f" is given in Sabre [13]. The estimate was based on the work of Sabre [12]. This estimate uses some kernels with smoothing parameters. Our work is devoted to determining how to choose the smoothing parameters for the best rate of convergence of bias and variance. To do this we will use the method of cross validation allowing us to choose the smoothing parameters of the kernel density estimators from data. This method was introduced by Rudemo [10] and



Bowman [2], for choosing the smoothing parameter to minimize the mean integrated square where the probability measure is absolutely continuous. Hardle and Marron [7] developed the same technic for regression function estimation. Hall and Marron[6], [9]have shown, under reasonable assumptions that the parameter given by the cross validation procedure is optimal. Cross-validation method was used in several works including [1], [3] and [15]...

We briefly indicate the organization of this paper: the second section recalls the estimate and its convergence given in Sabre [13]. The third section provides criteria for selecting parameters by using cross validation method and estimation of integrated square error. In the forth section we give the asymptotic optimality theorems. The last section is reserved to numerical results and simulation.

## 2. Kernel estimate of the density function and amplitudes of jumps:

As Sabre[13], consider a pair of random variables  $(X, Y)$  whose probability measure,  $\mu$ , is defined by (1). The numbers  $q$  and  $q'$  are assumed nonnegative integers and known.  $f$  is the density of the continuous variable which is assumed to be a nonnegative uniformly continuous function. The real positive number  $a'_j$  is the amplitude of the *jump* at  $(\omega_{1j}, \omega_{2j})$  assumed unknown. The densities  $\varphi_i$  are nonnegative uniformly continuous functions assumed unknown. The coefficients of the lines  $a_i, b_i$  are real numbers assumed unknown.  $\delta$  is the Dirac measure. Suppose that the *jump points* don't belong to the *jump lines* (ie.  $w_{2j} \neq a_i w_{1j} + b_i$  for all  $i$  and  $j$ ).

Sabre [13] has assumed, that any *jump point*  $(\omega_{1j}, \omega_{2j})$  can be localized in a small block  $[\alpha_{1j}, \beta_{1j}] \times [\alpha_{2j}, \beta_{2j}]$ . Thus, the estimate of the density  $f(x, y)$  proposed is different according to the position of  $(x, y)$ :

$$\hat{h}_H(x, y) = \begin{cases} f_n(x, y) & \text{if } (x, y) \notin \mathcal{A} \\ g_n(x, y) & \text{if } (x, y) \in \mathcal{A} \end{cases}.$$

$$\begin{aligned} \text{with } f_n(x, y) &= \sum_{i=1}^n \frac{1}{nh_n^2} K\left(\frac{x-x_i}{h_n}, \frac{y-y_i}{h_n}\right) \\ g_n(x, y) &= \int S_n(x-u_1)R_n(y-u_2)f_n(u_1, u_2)du_1du_2 \end{aligned} \quad (2)$$

$$\text{where } \mathcal{A} = \cup_{j=1}^q ([\alpha_{1j}, \beta_{1j}] \times \mathbb{R}) \cup (\mathbb{R} \times [\alpha_{2j}, \beta_{2j}]) \cup \mathcal{B}$$

$$\text{with } \mathcal{B} = \left\{ (x, y) \in \mathbb{R}^2 \text{ such as } \exists i \in \{1, \dots, q'\} : y = a_i x + b_i \right\}.$$

The kernel  $K$  is defined by  $K(u, v) = K_1(u)K_2(v)$  with  $K_1$  and  $K_2$  two continuous, even, decreasing bounded function such that:

$\int \|y^2\| K_i(y)dy < \infty \quad i = 1, 2$ . The smoothing parameter  $h_n$ , converges to zero and  $nh_n^2$  converges to the infinite, where

$$S_n(z) = \frac{W_n^{(2)}(z) - \frac{M_n^{(2)}}{M_n^{(1)}} W_n^{(1)}(z)}{1 - \frac{M_n^{(2)}}{M_n^{(1)}}} \text{ and } R_n(t) = \frac{W_n^{(3)}(t) - \frac{L_n^{(1)}}{L_n^{(1)}} W_n^{(4)}(t)}{1 - \frac{L_n^{(2)}}{L_n^{(1)}}} \quad (3)$$

The windows functions are defined as follows:

$$\begin{aligned} W_n^{(1)}(t) &= M_n^{(1)} W^{(1)}(tM_n^{(1)}); \quad W_n^{(2)}(t) = M_n^{(2)} W^{(2)}(tM_n^{(2)}); \\ W_n^{(3)}(t) &= L_n^{(1)} W^{(3)}(tL_n^{(1)}) \text{ and } W_n^{(4)}(t) = L_n^{(2)} W^{(4)}(tL_n^{(2)}) \end{aligned}$$

where  $M_n^{(1)}, M_n^{(2)}, L_n^{(1)}$  and  $L_n^{(2)}$  are nonnegative real sequences satisfying:

$$M_n^{(r)} \rightarrow +\infty; \quad L_n^{(r)} \rightarrow +\infty; \quad M_n^{(r)} h_n \rightarrow 0; \quad L_n^{(r)} h_n \rightarrow 0; \quad \frac{M_n^{(2)}}{M_n^{(1)}} \rightarrow 0 \text{ and } \frac{L_n^{(2)}}{L_n^{(1)}} \rightarrow 0$$

The function  $W^{(i)}$  is a nonnegative, even, integrable function vanishing outside the interval  $[-1, 1]$  such that  $\int_{-1}^1 W^{(i)}(x)dx = 1$ ,  $i = 1, 2, 3, 4$ . Moreover  $W^{(i)}$  satisfying the following equalities:

$$W^{(2)}(M_n^{(2)}\theta) - W^{(1)}(M_n^{(1)}\theta) = 0 \quad \forall \theta \in \left] \frac{-1}{M_n^{(1)}}, \frac{1}{M_n^{(1)}} \right[. \quad (4)$$

$$W^{(4)}(L_n^{(2)}\theta) - W^{(3)}(L_n^{(1)}\theta) = 0 \quad \forall \theta \in \left] \frac{-1}{L_n^{(1)}}, \frac{1}{L_n^{(1)}} \right[. \quad (5)$$

We assume that  $\frac{1}{h_n}K_1\left(\frac{1}{h_n M_n^{(1)}}\right)$  and  $\frac{1}{h_n}K_2\left(\frac{1}{h_n L_n^{(1)}}\right)$  converge to zero, for example  $K_1(x) = \frac{1}{(2\pi)^{\frac{1}{2}}} \exp(-\frac{x^2}{2})$ . The estimator  $\hat{h}$  depends to the smoothing parameters  $M_n^{(1)}, M_n^{(2)}, L_n^{(1)}, L_n^{(2)}, \frac{1}{h_n}$ . In the following we denote  $\hat{h}_H$  instead  $\hat{h}$ . Let  $H$  the vector of smoothing parameters :  $H = (M_n^{(1)}, M_n^{(2)}, L_n^{(1)}, L_n^{(2)}, \frac{1}{h_n})$ . The goal of the following sections is to gives the best vector  $H$  for having the optimal quadratical convergence of the estimator. To simplify without lost the generality we consider in the following of this paper that:

$$d\mu = f(x, y)dx dy + \sum_{m=1}^q a'_m \delta_{(\omega_{1m}, \omega_{2m})} \text{ and } \mathcal{A} = \cup_{j=1}^q ([\alpha_{1j}, \beta_{1j}] \times \mathbb{R}) \cup (\mathbb{R} \times [\alpha_{2j}, \beta_{2j}]).$$

### 3. Estimation of integrated square error

In this section we use the method of cross validation to select smoothing parameters to have the optimal rate of convergence. First, we study the estimation of integrated square error. In deed, our gaol is to give the vector  $H$ , minimizing the integrated square error ( $ISE = \int (\hat{h}_H(x, y) - f(x, y))^2 w(x, y) dx dy$ ) where  $w(x, y)$  is a nonnegative bounded weight function.

$$ISE = Q_1(\hat{h}_H) - 2Q_2(\hat{h}_H, f) + \int f^2(x, y)w(x, y)dx dy,$$

where  $Q_1(\hat{h}_H) = \int \hat{h}_H^2(x, y)w(x, y)dx dy$  and  $Q_2(\hat{h}_H, f) = \int \hat{h}_H(x, y)f(x, y)w(x, y)dx dy$ . Since the last term of the expression of  $ISE$  is independent of  $H$ , to minimize  $ISE$  is equivalent to minimize  $Q(\hat{h}_H, f) = Q_1(\hat{h}_H) - 2Q_2(\hat{h}_H, f)$ . According to the cross-validation method used in [9], we propose the following estimator of  $Q_2(\hat{h}, f)$  :

$$\hat{Q}_2(\hat{h}, f) = \frac{1}{n} \sum_{j=1}^n \hat{h}_{-j}(x_j, y_j) w(x_j, y_j) \text{ where } \hat{h}_{-j}(x, y) = \begin{cases} \hat{f}_{-j}(x, y) & \text{if } (x, y) \notin \mathcal{A} \\ \hat{g}_{-j}(x, y) & \text{if } (x, y) \in \mathcal{A} \end{cases}$$

$$\text{with } \hat{f}_{-j}(x, y) = \frac{1}{(n-1)h_n^2} \sum_{i \neq j}^n K\left(\frac{x-x_i}{h_n}, \frac{y-y_i}{h_n}\right) \text{ and } \hat{g}_{-j}(x, y) = \int S_n(x-u)R_n(y-v)\hat{f}_{-j}(u, v)dudv$$

where  $S_n$  and  $R_n$  are defined in (3).

**Theorem 3.1** *We have*

$$\bullet \text{ i) } EQ_2(\hat{h}, f) - EQ_2(\hat{h}, f) \rightarrow 0.$$

$$\bullet \text{ ii) } \text{The variance of } \hat{Q}_2(\hat{h}, f) \text{ converges to zero:}$$

$$Var(\hat{Q}_2(\hat{h}, f)) = O\left(\frac{1}{nh_n^4} \left[ \sup \left\{ K_1\left(\frac{1}{h_n M_n^{(1)}}\right) K_2\left(\frac{1}{h_n L_n^{(1)}}\right); K_1\left(\frac{1}{h_n M_n^{(1)}}\right); K_2\left(\frac{1}{h_n L_n^{(1)}}\right) \right\}^2 \right]\right).$$

#### 4. Asymptotic optimality theorem

The least squares method used above consists in choosing  $H$  which minimizes  $LS(H)$  where

$$LS(H) = \int \hat{h}^2(x, y) w(x, y) dx dy - \frac{2}{n} \sum_{j=1}^n \hat{h}_{j-}(x_j, y_j) w(x_j, y_j).$$

Before starting the asymptotic optimal theorems we define the following distances:

- Average square error:  $d_A(\hat{h}_H, f) = n^{-1} \sum_{j=1}^n [\hat{h}_H(x_j, y_j) - f(x_j, y_j)]^2 f(x_j, y_j)^{-1} w(x_j, y_j).$
- Integrated square error:  $d_I(\hat{h}_H, f) = \int [\hat{h}_H(x, y) - f(x, y)]^2 w(x, y) dx dy = ISE.$
- Mean integrated square error:  $d_M(\hat{h}_H, f) = E \int [\hat{h}_H(x, y) - f(x, y)]^2 w(x, y) dx dy = MISE.$

In the following we give some notations that we will use afterwards:

$$\begin{aligned} R &= \frac{1}{n} \sum_{j=1}^n f(x_j, y_j) w(x_j, y_j) + E[f(x_j, y_j) w(x_j, y_j)] \\ S &= \frac{2}{n} \sum_{j=1}^n f(x_j, y_j) w(x_j, y_j) [1 - \log f(x_j, y_j)] - R \\ T &= - \int f^2(x, y) w(x, y) dx dy - 2R \\ \mathcal{P} &= LS(H) - d_I(\hat{h}_H, f) - T \end{aligned}$$

$$B_H(x, y) = \begin{cases} \int \frac{1}{h_n^2} K\left(\frac{x-u}{h_n}, \frac{y-v}{h_n}\right) f(u, v) du dv - f(x, y) & \text{if } (x, y) \notin \mathcal{A} \\ \int S_n(x, u_1) R_n(y, u_2) \left( \int \frac{1}{h_n^2} K\left(\frac{u_1-v_1}{h_n}, \frac{u_2-v_2}{h_n}\right) dv_1 dv_2 \right) du_1 du_2 - f(x, y) & \text{if } (x, y) \in \mathcal{A} \end{cases}$$

As in [9], we assume the following assumptions. These assumptions are verified by the examples given in [6] and [7].

- a)  $\lim_{n \rightarrow \infty} \sup_{H \in \Lambda_n} \left| \frac{d_A(\hat{h}_H, f) - d_M(\hat{h}_H, f)}{d_M(\hat{h}_H, f)} \right| = 0 \quad a.s.,$
- b)  $\lim_{n \rightarrow \infty} \sup_{H \in \Lambda_n} \left| \frac{d_I(\hat{h}_H, f) - d_M(\hat{h}_H, f)}{d_M(\hat{h}_H, f)} \right| = 0 \quad a.s.,$  where  $\Lambda_n$  is a finite subset of  $\mathbb{R}^5$  whose cardinality grows algebraically fast, with  $card(\Lambda_n) \leq n^C$ ,  $C$  is some nonnegative constant, and every vector  $H = (H_1, H_2, H_3, H_4, H_5) \in \Lambda_n$  verifies the fact that:  
 $C^{-1} n^\delta \leq H_k \leq C n^{1-\delta}$ ,  $k = 1, 2, \dots, 5$ ;  $\delta$  is a constant satisfying  $0 < \delta < 1$ ,  
 $0 < H_2 < n^{-\beta} H_1$  and  $0 < H_4 < n^{-\beta'} H_3$  where  $\beta$  and  $\beta'$  are some nonnegative constants.
- c)  $w(x, y) \leq C \quad \forall (x, y) \in \mathbb{R}^2$

- d)  $f(x, y) \leq C, \quad \forall (x, y) \in \mathcal{S}$ , where  $\mathcal{S}$  denotes the support of the weight function  $w$ .
- e)  $B_H(x, y) \leq Cn^{-\delta}, \quad \forall H \in \Lambda_n, \quad \forall (x, y) \in \mathcal{S}$ .
- f) For all  $i \in \{1, 2, \dots, n > Q\}$

$$\lim_{n \rightarrow \infty} \sup_{\lambda \in \Lambda_n} \left| \frac{\int \text{Var} [K_{H,i}(x, y)] w(x, y) dx dy}{C' H_k} - 1 \right| = 0, \quad k = 1, 2, \dots, 5$$

$$\text{where } \widehat{K}_{H,i}(x, y) = \begin{cases} \frac{1}{h_n^2} K\left(\frac{x-x_i}{h_n}, \frac{y-y_i}{h_n}\right) & \text{if } (x, y) \notin \mathcal{A} \\ \int S_n(x, u_1) R_n(y, u_2) \frac{1}{h_n^2} K\left(\frac{u_1-x_i}{h_n}, \frac{u_2-y_i}{h_n}\right) du_1 du_2 & \text{if } (x, y) \in \mathcal{A} \end{cases}$$

- g)  $\int \dots \int K_H(x_{i_1}, x_{j_1}) \dots K_H(x_{i_k}, x_{j_k}) dx_1 dx_2 \dots dx_m \leq C_k (\inf(H_k))^{k-\frac{m}{2}}$ , where  $i_1, j_1, \dots, i_k, j_k = 1 \dots m$  subject to  $i_1 \neq j_1, \dots, i_k \neq j_k$ .

**Theorem 4.1** Under the assumptions a)- g)

$$\lim_{n \rightarrow \infty} \sup_{H \in \Lambda_n} \left| \frac{LS(H) - d_I(\hat{h}_H, f) - T}{d_M(\hat{h}_H, f)} \right| = 0 \quad a.s.,$$

The immediate consequence of this theorem is the following corollary

**Corollary 4.1** Under the assumptions a)- g), if  $\hat{h}_H$  is the minimizer of  $LS(H)$  over  $\Lambda_n$ , then

$$\lim_{n \rightarrow \infty} \frac{d_I(\hat{h}_H, f)}{\inf_{\lambda \in \Lambda_n} d_I(\hat{h}_H, f)} = 1 \quad a.s.$$

## 5. Proofs

### 5.1. Proof of the theorem 3.1

We have  $E[\widehat{Q}_2(\hat{h}_H, f)] = F_1 + F_2$  where

$$\begin{aligned} F_1 &= \frac{1}{n(n-1)h_n^2} E \sum_{j=1}^n \sum_{i \neq j}^n \left[ K\left(\frac{x_j - x_i}{h_n}, \frac{y_j - y_i}{h_n}\right) w(x_j, y_j) \mathbf{1}_{\mathcal{A}^c}(x_j, y_j) \right] \\ F_2 &= \frac{1}{n(n-1)h} \sum_{j=1}^n \sum_{i \neq j}^n \rho(i, j) \end{aligned}$$

where

$$\rho(i, j) = E \int \mathbf{1}_{\mathcal{A}}(x_j, y_j) S_n(x_j - t_1) R_n(y_j - t_2) K\left(\frac{t_1 - x_i}{h_n^2}, \frac{t_2 - y_i}{h_n}\right) w(x_j, y_j) dt_1 dt_2$$

and  $\mathcal{A}^c$  is the complementary of  $\mathcal{A}$ .

$$\begin{aligned} F_1 &= \frac{1}{(n-1)h_n^2} \sum_{i \neq j}^n \left[ \int K\left(\frac{u-x}{h_n}, \frac{v-y}{h_n}\right) w(u, v) \mathbf{1}_{\mathcal{A}^c}(u, v) f(u, v) dP_{(x_i, y_i)}(x, y) dudv \right. \\ &\quad \left. + \int \sum_{m=1}^q \mathbf{1}_{\mathcal{A}^c}(\lambda_{1m}, \lambda_{2m}) a_m K\left(\frac{\lambda_{1m}-x}{h_n}, \frac{\lambda_{2m}-y}{h_n}\right) w(\lambda_{1m}, \lambda_{2m}) dP_{(x_i, y_i)}(x, y) \right] \end{aligned}$$

As  $(\lambda_{1m}, \lambda_{2m}) \notin \mathcal{A}^c$ , the second term is null.

$$\begin{aligned} F_1 &= \frac{1}{h_n^2} \left[ \int K\left(\frac{u-x}{h_n}, \frac{v-y}{h_n}\right) w(u, v) \mathbf{1}_{\mathcal{A}^c}(u, v) f(u, v) f(x, y) dy dx dudv \right. \\ &\quad \left. + \int \sum_{m=1}^q \mathbf{1}_{\mathcal{A}^c}(u, v) a_m K\left(\frac{u-\lambda_{1m}}{h_n}, \frac{v-\lambda_{2m}}{h_n}\right) w(u, v) f(u, v) dudv \right]. \end{aligned}$$

The second term of the last equality tends to  $\sum_{m=1}^q a_m \mathbf{1}_{\mathcal{A}^c}(\lambda_{1m}, \lambda_{2m}) f(\lambda_{1m}, \lambda_{2m}) = 0$ , because  $(\lambda_{1m}, \lambda_{2m}) \notin \mathcal{A}^c$ . For  $F_2$ , we use the fact that  $dP_{(x_j, y_j)}(u, v) = f(u, v) du dv + \sum_{m=1}^q a_m \delta_{(\lambda_{1m}, \lambda_{2m})}$  for all  $(x_j, y_j)$  and interchanging the integrals, we have

$$\begin{aligned} F_2 &= \frac{1}{(n-1)h_n^2} \sum_{i \neq j}^n \left\{ \int \left[ \int \left( \int K \left( \frac{t_1 - x}{h_n}, \frac{t_2 - y}{h_n} \right) S_n(t_1 - u) R_n(t_2 - v) \mathbf{1}_{\mathcal{A}}(u, v) f(u, v) w(u, v) du dv \right) \right. \right. \\ &\quad \left. \left. dP_{(x_i, y_i)}(x, y) \right] dt_1 dt_2 \right. \\ &\quad \left. + \int \left[ \int \sum_{m=1}^q a_m \mathbf{1}_{\mathcal{A}}(\lambda_{1m}, \lambda_{2m}) K \left( \frac{t_1 - x}{h_n}, \frac{t_2 - y}{h_n} \right) S_n(t_1 - \lambda_{1m}) R_n(t_2 - \lambda_{2m}) w(\lambda_{1m}, \lambda_{2m}) \right. \right. \\ &\quad \left. \left. dP_{(x_i, y_i)}(x, y) \right] dt_1 dt_2 \right\} \end{aligned}$$

Thus  $F_2 = F_{2,1} + F_{2,2} + F_{2,3} + F_{2,4}$ , where

$$F_{2,1} = \int \left[ \int S_n(t_1 - u) R_n(t_2 - v) \left( \int \frac{1}{h_n^2} K \left( \frac{t_1 - x}{h_n}, \frac{t_2 - y}{h_n} \right) f(x, y) dx dy \right) dt_1 dt_2 \right] \mathbf{1}_{\mathcal{A}}(u, v) w(u, v) f(u, v) du dv$$

$$F_{2,2} = \frac{1}{h_n^2} \int \int \left[ \sum_{m=1}^q a_m K \left( \frac{t_1 - \lambda_{1m}}{h_n}, \frac{t_2 - \lambda_{2m}}{h_n} \right) \times S_n(t_1 - u) R_n(t_2 - v) \mathbf{1}_{\mathcal{A}}(u, v) w(u, v) f(u, v) du dv \right] dt_1 dt_2$$

$$F_{2,3} = \frac{1}{(h_n^2)} \int \left[ \sum_{m=1}^q a_m \int S_n(t_1 - \lambda_{1m}) R_n(t_2 - \lambda_{2m}) w(\lambda_{1m}, \lambda_{2m}) K \left( \frac{t_1 - x}{h_n}, \frac{t_2 - y}{h_n} \right) dt_1 dt_2 \right] f(x, y) dx dy$$

$$F_{2,4} = \frac{1}{h_n^2} \sum_{m=1}^q a_m \sum_{m'=1}^q a_{m'} R(m, m'), \text{ where}$$

$$R(m, m') = \int S_n(t_1 - \lambda_{1m}) R_n(t_2 - \lambda_{2m}) w(\lambda_{1m}, \lambda_{2m}) K \left( \frac{t_1 - \lambda_{1m'}}{h_n}, \frac{t_2 - \lambda_{2m'}}{h_n} \right) dt_1 dt_2.$$

Let us now show for  $(x, y) \in \mathcal{A}$  and  $(z, t) \in \mathbf{R}^2$ , that  $\frac{1}{h_n^2} \int S_n(u - x) R_n(v - y) K \left( \frac{u - z}{h_n}, \frac{v - t}{h_n} \right) du dv$  converges to zero. Indeed,

$$\begin{aligned} &\frac{1}{h_n^2} \int S_n(u - x) R_n(v - y) K \left( \frac{u - z}{h_n}, \frac{v - t}{h_n} \right) du dv \\ &= \frac{M_n^{(2)}}{1 - \frac{M_n^{(2)}}{M_n^{(1)}}} \frac{L_n^{(2)}}{1 - \frac{L_n^{(2)}}{L_n^{(1)}}} \int \left\{ W^{(2)} \left[ M_n^{(2)}(u - x) \right] - W^{(1)} \left[ M_n^{(1)}(u - x) \right] \right\} \\ &\quad \times \left\{ W^{(4)} \left[ L_n^{(2)}(v - y) \right] - W^{(3)} \left[ L_n^{(1)}(v - y) \right] \right\} \frac{1}{h_n^2} K \left( \frac{u - z}{h_n}, \frac{v - t}{h_n} \right) du dv. \end{aligned}$$

Putting  $s = M_n^{(2)}(u - x)$  and  $r = L_n^{(2)}(v - y)$ . Using the fact  $W^{(1)}$  is vanishing outside  $[-1, 1]$ , we get

$$\int S_n(u - x) R_n(v - y) K \left( \frac{u - z}{h_n}, \frac{v - t}{h_n} \right) du dv$$

$$= \frac{1}{1 - \frac{M_n^{(2)}}{M_n^{(1)}}} \frac{1}{1 - \frac{L_n^{(2)}}{L_n^{(1)}}} \int_{\frac{M_n^{(2)}}{M^{(1)}}}^1 \int_{\frac{L_n^{(2)}}{L^{(1)}}}^1 W^{(2)}(s) W^{(4)}(r) K \left( \frac{x + \frac{s}{M_n^{(2)}} - z}{h_n}, \frac{y + \frac{r}{L_n^{(2)}} - t}{h_n} \right) ds dr$$

If  $x \neq z$  and  $y \neq t$ , since  $\int W^{(2)}(x) dx = 1$  and  $\frac{M_n^{(2)}}{M_n^{(1)}} ; \frac{L_n^{(2)}}{L_n^{(1)}}$  tend to zero, we have

$$\int S_n(u-x) R_n(v-y) K \left( \frac{u-z}{h_n}, \frac{v-t}{h_n} \right) dudv = O \left( K \left( \frac{1}{h_n}, \frac{1}{h_n} \right) \right).$$

If  $x = z$  and  $y \neq t$  since  $\frac{1}{M^{(1)}} < \frac{s}{M_n^{(2)}} < \frac{1}{M_n^{(2)}}$ ,  $K_2$  is bounded and the function  $K_1$  is decreasing, we obtain

$$\int S_n(u-x_2) R_n(v-y_2) K \left( \frac{u-x_1}{h_n}, \frac{v-y_1}{h_n} \right) dudv = O \left( K_1 \left( \frac{1}{h_n M_n^{(1)}} \right) \right). \quad (6)$$

If  $x = z$  and  $y = t$ , it is easy to show that  $\int S_n(u-x_2) R_n(v-y_2) K \left( \frac{u-x_1}{h_n}, \frac{v-y_1}{h_n} \right) dudv = O \left( K_1 \left( \frac{1}{h_n M_n^{(1)}} \right) K_2 \left( \frac{1}{h_n M_n^{(1)}} \right) \right)$ . Thus from the sum of the first term of  $F_1$  and the first term of  $F_2$ , we have  $E \left[ \hat{Q}_2(\hat{h}_H, f) \right]$  is asymptotically equal to  $E \left[ Q_2(\hat{h}_H, f) \right]$ .

ii) Let us show that the variance of  $\hat{Q}_2(\hat{h}_H, f)$  tends to zero.  $Var \left( \hat{Q}_2(\hat{h}_H, f) \right) = V_1 + V_2$ , where

$$\begin{aligned} V_1 &= \frac{1}{n^2} \sum_{j=1}^n \frac{1}{(n-1)^2 h_n^4} \left\{ E \left( \mathbf{1}_{\mathcal{A}^c}(x_j, y_j) \sum_{j \neq i}^n K \left( \frac{x_j - x_i}{h_n}, \frac{y_j - y_i}{h_n} \right) w(x_j, y_j) \right)^2 \right. \\ &\quad \left. - \left( E \left[ \mathbf{1}_{\mathcal{A}^c}(x_j, y_j) \sum_{j \neq i}^n K \left( \frac{x_j - x_i}{h_n}, \frac{y_j - y_i}{h_n} \right) w(x_j, y_j) \right] \right)^2 \right\} \\ V_2 &= \frac{1}{n^2} Var \left( \sum_{j=1}^n \mathbf{1}_{\mathcal{A}}(x_j, y_j) \int S_n(u-x_j) R_n(v-y_j) \hat{f}_{-j}(u, v) dudv \right). \end{aligned}$$

We can write  $V_1 = \frac{1}{n^2} \sum_{j=1}^n \frac{1}{(n-1)^2 h_n^4} (T_1 - T_2)$ . Since  $(\lambda_{1m}, \lambda_{2m}) \notin \mathcal{A}^c$ , we get

$$\begin{aligned} T_1 &= (n-1)^2 \int \mathbf{1}_{\mathcal{A}^c}(x_2, y_2) K^2 \left( \frac{x_2 - x_1}{h_n}, \frac{y_2 - y_1}{h_n} \right) w(x_2, y_2) f(x_2, y_2) f(x_1, y_1) dx_1 dy_1 dx_2 dy_2 \\ &\quad + (n-1)^2 \sum_{m=1}^q a_m \int_{\mathcal{A}^c} K^2 \left( \frac{x - \lambda_{1m}}{h_n}, \frac{y - \lambda_{2m}}{h_n} \right) w(x, y) f(x, y) dx dy. \end{aligned}$$

Let  $D(a, b) = \int K^2 \left( \frac{x-a}{h_n}, \frac{y-b}{h_n} \right) f(x, y) dx dy$ . Putting  $\frac{x-a}{h_n} = u$  and  $\frac{y-b}{h_n} = v$ , we have  $D(a, b) = h_n^2 \int K^2(u, v) f(a + u h_n, b + v h_n) dudv$ . Using the Taylor's formula we obtain

$$\begin{aligned} D(a, b) &\simeq h_n^2 f(a, b) \int K^2(u, v) dudv + h_n^3 \frac{df}{da}(a, b) \int u K^2(u, v) dudv \\ &\quad + h_n^3 \frac{df}{db}(a, b) \int v K^2(u, v) dudv, \end{aligned}$$

with  $\int uK^2(u, v) dudv < \infty$ . Thus, we obtain the following inequality

$$\begin{aligned} T_1 &\leq (n-1)^2 h_n^2 \int f^2(x_2, y_2) dx_2 dy_2 \int K^2(u, v) dudv \\ &+ (n-1)^2 h_n^3 \int \frac{df}{dx_2}(x_2, y_2) f(x_2, y_2) dx_2 dy_2 \int uK^2(u, v) dudv \\ &+ (n-1)^2 h_n^3 \int \frac{df}{dy_2}(x_2, y_2) f(x_2, y_2) dx_2 dy_2 \int vK^2(u, v) dudv \\ &+ (n-1)^2 \sum_{m=1}^q a_m h_n^2 f(\lambda_{1m}, \lambda_{2m}) \int K^2(u, v) dudv \\ &+ (n-1)^2 \sum_{m=1}^q a_m h_n^3 \frac{df}{dx}(\lambda_{1m}, \lambda_{2m}) \int uK^2(u, v) dudv \\ &+ (n-1)^2 \sum_{m=1}^q a_m h_n^3 \frac{df}{dy}(\lambda_{1m}, \lambda_{2m}) \int vK^2(u, v) dudv. \end{aligned}$$

Since the function  $w$  is bounded, we get  $\frac{1}{n^2} \sum_{i=1}^n \frac{1}{(n-1)^2 h_n^4} T_1 = O\left(\frac{1}{nh_n}\right)$ . Similarly, we show that  $\frac{1}{n^2} \sum_{i=1}^n \frac{1}{(n-1)^2 h_n^4} T_2 = O\left(\frac{1}{n}\right)$ . Thus,  $V_1 = O\left(\frac{1}{nh_n}\right)$ . On the other hand,

$$V_2 = \frac{1}{n^2} Var \left( \sum_{j=1}^n \mathbf{1}_{\mathcal{A}}(x_j, y_j) \int S_n(u - x_j) R_n(v - y_j) \hat{f}_{-j}(u, v) dudv \right)$$

Putting  $O_j = \mathbf{1}_{\mathcal{A}}(x_j, y_j) \int S_n(u - x_j) R_n(v - y_j) \sum_{i \neq j} K\left(\frac{u - x_i}{h_n}, \frac{v - y_i}{h_n}\right) dudv$  we have

$$\begin{aligned} V_2 &= \frac{1}{n^2} \frac{1}{(n-1)^2} \frac{1}{h_n^4} \sum_{j=1}^n E(O_j - E(O_j))^2 \\ &= \frac{1}{n^2} \sum_{j=1}^n E \left( \int \mathbf{1}_{\mathcal{A}}(x_j, y_j) S_n(u - x_j) R_n(v - y_j) \hat{f}_{-j}(u, v) dudv \right)^2 - (E(O_j))^2 \\ &= R_1 - R_2, \text{ where} \end{aligned}$$

$$\begin{aligned} R_1 &= \frac{1}{n} \frac{1}{h_n^4} \int \int \left( \int \mathbf{1}_{\mathcal{A}}(x_2, y_2) S_n(u - x_2) R_n(v - y_2) K\left(\frac{u - x_1}{h_n}, \frac{v - y_1}{h_n}\right) dudv \right)^2 \\ &\quad f(x_2, y_2) dx_2 dy_2 dP_{(x_i, y_i)}(x_1, y_1) \\ &\quad + \frac{1}{n} \frac{1}{h_n^2} \int \sum_{m=1}^q a_m \left( \int \mathbf{1}_{\mathcal{A}}(\lambda_{1m}, \lambda_{2m}) S_n(u - \lambda_{1m}) R_n(v - \lambda_{2m}) K\left(\frac{u - x_1}{h_n}, \frac{v - x_2}{h_n}\right) dudv \right)^2 \\ &\quad dP_{(x_i, y_i)}(x_1, y_1) \\ &= \frac{1}{n} \frac{1}{h_n^4} \int \left( \int \mathbf{1}_{\mathcal{A}}(x_2, y_2) S_n(u - x_2) R_n(v - y_2) K\left(\frac{u - x_1}{h_n}, \frac{v - y_1}{h_n}\right) dudv \right)^2 \\ &\quad f(x_2, y_2) dx_2 dy_2 f(x_1, y_1) dx_1 dy_1 \\ &\quad + \frac{1}{n} \frac{1}{h_n^4} \int \sum_{m=1}^q a_m \left( \int \mathbf{1}_{\mathcal{A}}(x_2, y_2) S_n(u - x_2) R_n(v - y_2) K\left(\frac{u - \lambda_{1m}}{h_n}, \frac{v - \lambda_{2m}}{h_n}\right) dudv \right)^2 \\ &\quad f(x_2, y_2) dx_2 dy_2 \end{aligned}$$

$$\begin{aligned}
 & + \frac{1}{n} \frac{1}{h_n^4} \int \sum_{m=1}^q a_m \left( \int S_n(u - \lambda_{1m}) R_n(v - \lambda_{2m}) K \left( \frac{u - x_1}{h_n}, \frac{v - y_1}{h_n} \right) dudv \right)^2 \\
 & f(x_1, y_1) dx_1 dy_1 \\
 & + \frac{1}{n} \frac{1}{h_n^4} \sum_{m'=1}^q a_{m'} \sum_{m=1}^q a_m \left( \int S_n(u - \lambda_{1m}) R_n(v - \lambda_{2m}) K \left( \frac{u - \lambda_{1m'}}{h_n}, \frac{u - \lambda_{2m'}}{h_n} \right) dudv \right)^2.
 \end{aligned}$$

Using (6), we get

$$R_1 = O \left\{ \frac{1}{nh_n^4} \left[ \sup \left( K_1 \left( \frac{1}{h_n M_n^{(1)}} \right) K_2 \left( \frac{1}{h_n L_n^{(1)}} \right); K_1 \left( \frac{1}{h_n M_n^{(1)}} \right); K_2 \left( \frac{1}{h_n L_n^{(1)}} \right) \right) \right]^2 \right\}.$$

Since  $R_2 = O \left( \frac{1}{n} \right)$ , we obtain

$$V_2 = O \left( \frac{1}{nh_n^4} \left[ \sup \left\{ K_1 \left( \frac{1}{h_n M_n^{(1)}} \right) K_2 \left( \frac{1}{h_n L_n^{(1)}} \right); K_1 \left( \frac{1}{h_n M_n^{(1)}} \right); K_2 \left( \frac{1}{h_n L_n^{(1)}} \right) \right\} \right]^2 \right).$$

## 5.2. Proof of the theorem 4.1

$LS(H)$  can be written as follows

$$LS(H) = \int \left( \hat{h}(x, y) \right)^2 w(x, y) dx dy - \frac{2}{n} \sum_{j=1}^n \hat{h}_{j-}(x_j, y_j) w(x_j, y_j).$$

Let  $\mathcal{P} = \mathcal{LS}(\mathcal{H}) - [\mathcal{I}(\hat{\mathcal{H}}, \{ \}) - \mathcal{T}]$ , from the definition of  $d_I(\hat{h}_H, f, P)$  becomes

$$\begin{aligned}
 \mathcal{P} &= -\frac{2}{n} \sum_{j=1}^n \hat{h}_{j-}(x_j, y_j) w(x_j, y_j) \\
 &+ 2 \int \hat{h}_H(x, y) f(x, y) w(x, y) dx dy \\
 &+ \frac{2}{n} \sum_{j=1}^n f(x_i, y_i) w(x_i, y_i) - 2E[f(x_i, y_i) w(x_i, y_i)]
 \end{aligned}$$

It is easy to show that  $|\mathcal{P}| = 2n^{-1}(n-1)^{-1} \left| \sum_{j=1}^n \sum_{i \neq j}^n U_{ij} \right|$ , where

$$\begin{aligned}
 U_{ij} &= -\frac{1}{h_n^2} K \left( \frac{x_j - x_i}{h_n}, \frac{y_j - y_i}{h_n} \right) \mathbf{1}_{\mathcal{A}^c}(x_j, y_j) w(x_j, y_j) \\
 &- \frac{1}{h_n^2} \int S_n(x_j - u) R_n(y_j - v) K \left( \frac{u - x_i}{h_n}, \frac{v - y_i}{h_n} \right) dudv \mathbf{1}_{\mathcal{A}}(x_j, y_j) w(x_j, y_j) \\
 &+ \frac{n-1}{nh_n^2} \int_{\mathcal{A}^c} K \left( \frac{x - x_i}{h_n}, \frac{y - y_i}{h_n} \right) w(x, y) f(x, y) \mathbf{1}_{\mathcal{A}^c}(x, y) dx dy \\
 &+ \frac{1}{nh_n^2} \int K \left( \frac{x - x_j}{h_n}, \frac{y - y_j}{h_n} \right) w(x, y) \mathbf{1}_{\mathcal{A}^c}(x, y) f(x, y) dx dy \\
 &+ \int_{\mathcal{A}} \frac{n-1}{nh_n^2} \left( \int S_n(u - x) R_n(v - y) K \left( \frac{u - x_i}{h_n}, \frac{v - y_i}{h_n} \right) dudv \right) w(x, y) f(x, y) \mathbf{1}_{\mathcal{A}^c}(x, y) dx dy \\
 &+ \int_{\mathcal{A}} \frac{1}{nh_n^2} \left( \int S_n(u - x) R_n(v - y) K \left( \frac{u - x_j}{h_n}, \frac{v - y_j}{h_n} \right) dudv \right) w(x, y) f(x, y) \mathbf{1}_{\mathcal{A}^c}(x, y) dx dy \\
 &+ f(x_j, y_j) w(x_j, y_j) - \int f^2(x, y) w(x, y) dx dy + \sum_{m=1}^q a_m f(\lambda_{1m}, \lambda_{2m}) w(\lambda_{1m}, \lambda_{2m}).
 \end{aligned}$$

For  $i \neq j$ , define  $V_{ij} = U_{ij} - E(U_{ij}/(x_j, y_j))$  Observe that

$$E[V_{ij}/(x_i, y_i)] = 0 \quad \text{and} \quad E[E(U_{ij}/(x_j, y_j))] = 0.$$

To finish the proof of the theorem 4.1 it enough to prove that

$$\sup_{H \in \Lambda_n} n^{-2} \left| \sum_{j=1}^n \sum_{i \neq j}^n V_{ij} \right| d_M^{-1}(\hat{h}_H, f) \rightarrow 0 \text{ a.s} \quad (19)$$

$$\text{and that} \quad \sup_{H \in \Lambda_n} n^{-1} \left| \sum_{j=1}^n E(U_{ij}/(x_j, y_j)) \right| d_M^{-1}(\hat{h}_H, f) \rightarrow 0 \text{ a.s.} \quad (20)$$

To verify (20), note that by the Borel-Cantelli lemma, it is enough to show that for  $\varepsilon > 0$

$$\sum_{n=1}^{\infty} \text{card}(\Lambda_n) \sup_{H \in \Lambda_n} P \left[ \left| n^{-1} \sum_{j=1}^n \alpha_j \right| > \varepsilon d_M(\hat{h}_H, f) \right] < \infty$$

$$E(U_{ij}/(x_j, y_j)) = B(x_j, y_j)w(x_j, y_j) - E[B(x_j, y_j)w(x_j, y_j)] + Q_m.$$

$$\begin{aligned} \text{where } Q_m &= \sum_{m'} a_{m'} \frac{1}{h_n^2} \int S_n(u - \lambda_{1m'}) R_n(v - \lambda_{2m'}) \left[ \int K\left(\frac{u-x}{h_n}, \frac{v-y}{h_n}\right) f(x, y) dx dy \right. \\ &\quad \left. - \sum_m a_m K\left(\frac{u - \lambda_{1m}}{h_n}, \frac{v - \lambda_{2m}}{h_n}\right) \right] dudv \end{aligned}$$

$$\text{Therefore } \text{Var}(E(U_{ij}/(x_j, y_j))) = \text{Var}(B(x_j, y_j)w(x_j, y_j) - E[B(x_j, y_j)w(x_j, y_j)]).$$

$$\begin{aligned} \text{Var}(E(U_{ij}/(x_j, y_j))) &= E[B^2(x_j, y_j)(w^2(x_j, y_j))] - (E[B(x_j, y_j)w(x_j, y_j)])^2 \\ &\leq \int B^2(x, y)w^2(x, y)f(x, y)dx dy + \sum_{m=1}^q a_m B^2(\lambda_{1m}, \lambda_{2m})w^2(\lambda_{1m}, \lambda_{2m}) \end{aligned}$$

From the assumptions (c), (d) and (e) we get

$$\begin{aligned} \text{Var}(E(U_{ij}/(x_j, y_j))) &\leq C^2 n^{-2\delta} \left( \int w^2(x, y)f(x, y)dx dy + \sum_m a_m w^2(\lambda_{1m}, \lambda_{2m}) \right) \\ &\leq cte n^{-2\delta}. \end{aligned}$$

From Bernstein's inequality(see [8]), we have

$$P \left[ \left| n^{-1} \sum_{j=1}^n \alpha_j \right| > \varepsilon d_M(\hat{h}_H, f) \right] \leq \exp \left\{ \frac{-n\varepsilon^2 d_M^2(\hat{h}_H, f)}{2 \left( \text{Var}(E(U_{ij}/(x_j, y_j))) + \frac{cn^{-\delta} \varepsilon d_M(\hat{h}_H, f)}{3} \right)} \right\}$$

Since  $d_M(\hat{h}_H, f) = \int Bias^2(x, y) dx dy + \int Var(\hat{h}_H(x, y)) dx dy$ , from b), f), e) and g), we have  $n^\delta \leq \int Var(\hat{h}_H(x, y)) dx dy \leq n^{1-\delta}$ . From (e), we have  $0 \leq B^2(x, y) \leq n^{-2\delta}$  and  $n^\delta \leq \int Var(\hat{h}_H(x, y)) dx dy \leq n^{1-\delta}$ , we obtain  $n^\delta \leq d_M(\hat{h}_H, f) \leq cte(n^{1-\delta} + n^{-2\delta})$ . Therefore

$$2Var(E(U_{ij}/(x_j, y_j))) + \left( \frac{cn^{-\delta} \varepsilon d_M(\hat{h}_H, f)}{3} \right) \leq cte(n^{-2\delta} + n^{-\delta}(n^{1-\delta} + n^{-2\delta}))$$

Then  $P\left[\left|n^{-1} \sum_{j=1}^n \alpha_j\right| > \varepsilon d_M(\hat{h}_H, f)\right] \leq \exp\{-Constant \cdot n^{4\delta}\}$ .

Thus  $\sum_{n=1}^{\infty} card(\lambda_n) \sup_{\lambda \in \Lambda_n} P\left[n^{-1} \sum_{j=1}^n |\alpha_j| > \varepsilon d_M(\hat{h}, f)\right] < \infty$ .

As in Marron [9], to verify (19), using the proof of (20) and the chebyshev inequality, it is enough to show that there is a constant  $\gamma > 0$ , such that for  $k = 1, 2, \dots$  there are constants  $C_k$  such that

$$\sup_{\lambda \in \Lambda_n} E \left[ n^{-2} \sum_{i \neq j} V_{ij} d_M(\hat{h}_H, f)^{-1} \right]^{2k} \leq C_k n^{-\gamma k}. \quad (21)$$

where  $cum_k$  is the  $k$ th order cumulant and  $\sum$  denotes the summation over  $i_1, j_1, \dots, i_k, j_k = 1, \dots, n$  subject to  $i_1 \neq j_1, \dots, i_k \neq j_k$ . From the inequality (7.5) in Marron [9] and the cumulant expansion of the  $2k$ th centred moment [see, for example, Kendall and Stuart [5], we obtain the inequality (21). This completes the proof of the theorem.

## 6. Example and application

A concrete application of this work to the one-dimensional case deals with the process of filling bottles of a 33cl volume each. To control the quality of this process, we check that this process is evenly distributed according to gaussian distribution. The experimenter has taken a sample of 115 bottles and measured the quantity of liquid contained in these bottles. The measures normally stand at around of 33cl. But on account of an abnormal disfunction due to a random slowing or acceleration of the motion of the rolling band. Thus the measure randomly increases to reach a value around of 37cl or decrease to a value around of 29cl. But apart from the problem of the disfunction of the rolling band, we would also like to verify whether the process of filling bottles respects the norm of 33cl with a standard deviation of 1cl.

We consider the random variable  $Y$  representing the quantity of liquid in the bottle. Let  $X = \frac{Y-\mu}{\sigma}$  where  $\mu = 33cl$  and  $\sigma = 0.8$ . testing if  $Y$  is gaussian variable is equivalent to tests if  $X$  is a standard gaussian variable. By using the histogram and density of the variable  $X$  we find that we have two points of discontinuity of the density: two peaks one in the interval  $[\alpha_1, \beta_2] = [-2.8, -2.5]$  and the second in the interval  $[\alpha_2, \beta_2] = [1.8, 2.2]$ . We propose the model:

$$d\mu = f(x)dx + a_1\delta_{\lambda_1} + a_2\delta_{\lambda_2},$$

where  $\lambda_1 \in [\alpha_1, \beta_1]$  and  $\lambda_2 \in [\alpha_2, \beta_2]$ .

The kernels  $K_1$  and  $K_2$  are chosen as follows:  $K_1(x) = K_2(x) = \frac{1}{\sqrt{2\pi}} \exp\{-\frac{x^2}{2}\}$ , and  $n = 1000$ .

First, choosing  $W^{(1)}$  as a nonnegative, even and integrable function. We propose:

$$W^{(1)}(t) = \begin{cases} \frac{64}{63}t + \frac{64}{63} & \text{if } t \in [-1, -1/8] \\ \frac{8}{9} & \text{if } t \in [-1/8, 1/8] \\ -\frac{64}{63}t + \frac{64}{63} & \text{if } t \in ]1/8, 1] \\ 0 & \text{otherwise} \end{cases}$$

It is easy to show that  $\int W^{(1)}(t)dt = 1$ . Choosing now a nonnegative, even and integrable function  $W^{(2)}$  such that  $W^{(1)}$  and  $W^{(2)}$  satisfying (4) and (5). We propose:

$$W^{(2)}(t) = \begin{cases} W^{(1)}\left(\frac{M_n^{(1)}}{M_n^{(2)}}t\right) & \text{if } t \in [-1/8, 1/8] \\ 4/7\left(1 - \frac{M_n^{(2)}}{M_n^{(1)}}\right) & \text{if } t \in [-1, -1/8] \cup [1/8, 1] \\ 0 & \text{if otherwise} \end{cases}$$

$$\int W^{(2)}(t)dt = \int_{-1}^{-1/8} W^{(2)}(t)dt + \int_{-1/8}^{1/8} W^{(2)}(t)dt + \int_{1/8}^1 W^{(2)}(t)dt.$$

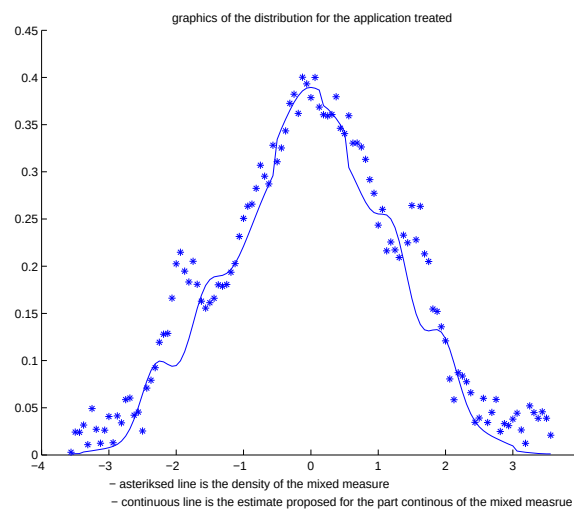
$$\int W^{(2)}(t)dt = \left(1 - \frac{M_n^{(2)}}{M_n^{(1)}}\right) + \int_{-1/8}^{1/8} W^{(2)}(t)dt$$

From the definition of  $W^{(2)}$ , we have  $\int_{-1/8}^{1/8} W^{(2)}(t)dt = 2 \int_0^{1/8} W^{(1)}\left(\frac{M_n^{(1)}}{M_n^{(2)}}t\right)dt$ . Putting  $u =$

$\frac{M_n^{(1)}}{M_n^{(2)}}t$ , we have  $\int_{-1/8}^{1/8} W^{(2)}(t)dt = 2 \left(\frac{M_n^{(2)}}{M_n^{(1)}}\right) \int_0^{1/8 \left(\frac{M_n^{(1)}}{M_n^{(2)}}\right)} W^{(1)}(u)du$ . Since  $\frac{M_n^{(1)}}{M_n^{(2)}} \rightarrow \infty$ , for  $n$  large enough we have  $\frac{M_n^{(1)}}{M_n^{(2)}} > 1$ . Thus, we obtain  $\int_{-1/8}^{1/8} W^{(2)}(t)dt = 2 \left(\frac{M_n^{(2)}}{M_n^{(1)}}\right) \int_0^1 W^{(1)}(u)du$ .  $W^{(1)}$  being even and  $\int W^{(1)}(t)dt = 1$ , we deduce  $\int_{-1/8}^{1/8} W^{(2)}(t)dt = \left(\frac{M_n^{(2)}}{M_n^{(1)}}\right)$ . Thus,  $\int W^{(2)}(t)dt = 1$ .

Let us show that (3) and (5) is satisfied. Indeed, let  $t$  a reel number belonging to  $\left]-\frac{1}{M_n^{(1)}}, \frac{1}{M_n^{(1)}}\right]$ , since  $\frac{M_n^{(2)}}{M_n^{(1)}}$  converges to zero, for  $n$  large enough, we have  $-1/8 < -\frac{M_n^{(2)}}{M_n^{(1)}}t \leq M_n^{(2)}t \leq \frac{M_n^{(2)}}{M_n^{(1)}} < 1/8$ .

Therefore,  $W^{(2)}(M_n^{(2)}t) = W^{(1)}\left(\frac{M_n^{(1)}}{M_n^{(2)}}M_n^{(2)}t\right) = W^{(1)}(M_n^{(1)}t)$ . We calculate the parameters which minimize  $LS(H)$  (cross-validation) by the function "fmns" implemented in Matlab with a respect the constraints and assumptions a)-f) given in the section 4. We calculate the estimate  $\hat{h}_H(x)$ , of the density function  $f$ . The graphics of the estimate  $\hat{h}_H(x)$  and the kernel density of the variable  $X$ :



## 7. Conclusions

We have presented in this paper some results about cross-validation method for choosing the smoothing parameters of density estimate when the measure has certain mixture. A criterion for find the smoothing parameter was given in order to have better convergence of the estimator of the density of the continuous part. The proposed methods can be extended to other applications in several sectors. Indeed, the control of the quality for a product manufactured in the auto industry use the measure of two variables: the consumption of diesel and the pollution. Their joint distribution can follow a continuous law except some observations which are taken when there is fog and reached the constant value (point of the jump). One example in economics, it is the observation of the variables: taxes on income and purchasing power can have a joint distribution contains some point of jumps due to exemption (disabled, former soldier, ...). In oceanography when we observe, by using a camera placed at a certain depth in water, two variables: the length of the fishes and their movement speed. The joint distribution may represent some jumps due to the acceleration of movement during the passage of a predator. In Astronomy the repeated passage of an object preventing the vision of stars (cloud, bird, ...) can create a jump of data.

## Acknowledgments

I would like to thank the anonymous referees for their interest in this paper and their valuable comments and suggestions.

## 8. References

- [1] M. I. Ashraf, Z. Zhao, C.A. Bourque, D. A. MacLean, F. Meng 2013 Integrating biophysical controls in forest growth and yield predictions with artificial intelligence technology, *Canadian Journal of Forest Research*, **43**, 12, 1162-1171,
- [2] A. W. Bowman, , P. Hall and D. M. Titterton 1984 Cross-validation in nonparametric estimation of probabilities and probability density , *Biometrika*, **71**, 341-352.
- [3] D. M. Hawkins , S.C. Basak and D. Mills 2003 Assessing Model Fit by Cross-Validation, *J. Chem. Inf. Comput. Sci.*, , **43**, 2, 579-586.
- [4] P. Hall 1983 Large sample optimality of least squares cross-validation in density estimation, *Ann. Statist.*, **11**, 1156-1174.
- [5] M.G. Kendall and A. Stuart 1977 *the Advanced Theory of statistics: Distribution theory*, ( 4 th ed. Macmillan, New York).
- [6] P. Hall and J.S. Marron 1988 Choice of Kernel Order in Density Estimation, *Ann. of Stat*, **16**, n 1 , 161-173.
- [7] W. Hardle and J.S. Marron 1985 Optimal Bandwidth Selection in Nonparametric Regression Function Estimation, *Ann.Statist*, **13**, n 4, 1465-1481.
- [8] W. Hoeffding 1963 Probability inequalities for sums of bounded random variables, *J. Amer. Statist. Assoc*, **58**, 13-30.
- [9] J.S. Marron 1987 A comparison of cross-validation techniques in density estimation, *Ann.Statist*, **15**, n 1, 152-162.
- [10] M. Rudemo 1982 Empirical choice of histogram and kernel density estimation, *Scand. J. Statist.* **9**, 65-78.
- [11] G. Wahba and J.Wendelberger 1994 Some New Mathematical Methods for Variational Objective Analysis Using Splines and Cross Validation , *C. R. Acad. Sci. Paris*, **319**,
- [12] R. Sabre 1995 Spectral density estimation for stationary stable random fields, *Journal Applications Mathematicae*, **23**, 2, 107-133.
- [13] R. Sabre 2003 Nonparametric estimation of the continuous part of mixed measure *Revisita Statistica* **LXIII**, n3, 441-467.
- [14] R. Sabre 2010 Tests and Asymptotic normality for mixed Bivariate measure *Revisita Statistica* **LXX**, n 2, 115-136.
- [15] A. Vehtari and J. Lampinen 2002 Bayesian Model Assessment and Comparison Using Cross-Validation Predictive Densities, *Neural Computation*, **14**, 10, 2439-2468.