NONPARAMETRIC DENSITY AND SURVIVAL FUNCTION ESTIMATION IN THE MULTIPLICATIVE CENSORING MODEL

E. BRUNEL, F. COMTE, AND V. GENON-CATALOT

Abstract. In this paper, we consider the multiplicative censoring model, given by $Y_i = X_i U_i$ where (X_i) are *i.i.d.* with unknown density f on $\mathbb{R}$, (U_i) are *i.i.d.* with uniform distribution $\mathcal{U}([0,1])$ and (U_i) and (X_i) are independent sequences. Only the sample $(Y_i)_{1\le i\le n}$ is observed. Nonparametric estimators of both the density f and the corresponding survival function $\bar{F}$ are proposed and studied. First, classical kernels are used and the estimators are studied from several points of view: pointwise risk bounds for the quadratic risk are given, upper and lower bounds for the rates in this setting are provided. Then, an adaptive non asymptotic bandwidth selection procedure in a global setting is proved to realize the bias-variance compromise in an automatic way. When the X_i's are nonnegative, using kernels fitted for $\mathbb{R}^+$-supported functions, we propose new estimators of the survival function which are proved to be adaptive. Simulation experiments allow us to check the good performances of the estimators and compare the two strategies.

Keywords. Adaptive procedure. Bandwidth selection. Kernel estimators. Multiplicative censoring model.

AMS 2000 subject classifications. 62G07 - 62N01

1. Introduction

In this paper, we consider the model

$$Y_i = X_i U_i, i = 1, \dots, n \tag{1}$$

under the assumptions: the U_i, $i = 1, \dots, n$ are independent and identically distributed (*i.i.d.*) with uniform distribution on $[0,1]$; the X_i, $i = 1, \dots, n$ are real valued, *i.i.d.*, with unknown density f and cumulative distribution function (c.d.f.) F; the sequences $(U_i)_{1\le i\le n}$ and $(X_i)_{1\le i\le n}$ are independent. We intend to propose estimation methods for f and F (or $\bar{F} = 1 - F$) when observing a sample $(Y_i)_{1\le i\le n}$ only.

Model (1) has been widely investigated mostly when the random variables X_i are nonnegative. In this case, Model (1) is usually called the *multiplicative censoring model* and was introduced in Vardi (1989), studied in more details in Vardi and Zhang (1992) and by Asgharian *et al.* (2012). As explained in Vardi (1989), the multicative censoring model unifies several important statistical problems (deconvolution of an exponential variable, estimation under decreasing density constraint or some estimation problems in renewal processes). However in the above quoted papers, authors assume that observations are composed of two independent samples, one of direct observations X with size m, in addition to the above Y n-sample. The statistical procedures for estimating the c.d.f. F rely on the fact that m tends to infinity and cannot be applied for

Date: March 5, 2015.

Université Montpellier, I3M UMR CNRS 5149. email: ebrunel@math.univ-montp2.fr.
Université Paris Descartes, MAP5, UMR CNRS 8145. email: fabienne.comte@parisdescartes.fr .
Université Paris Descartes, MAP5, UMR CNRS 8145. email: Valentine.genon-catalot@parisdescartes.fr.

$m = 0$. Let us mention that van Es *et al.* (2000) studied a survival analysis model involving both multiplicative censoring and length bias, in a parametric context.

The problem may be related to the moment problem: in Model (1), all moments of X can easily be estimated from the observations Y, so the question of distribution reconstruction from its moments as pointed out by Mnatsakanov (2006) can be addressed.

Another strategy is to take the logarithm of the squared model, as proposed in stochastic volatility models (see van Es *et al.* (2005), Comte and Genon-Catalot (2006)), and to apply deconvolution methods. In these papers, the U_i's are supposed to be Gaussian. But then, the estimated function is distorted and, in case of real random variables X_i, information about their sign is lost, when proceeding so.

Series expansion methods have been proposed in Andersen and Hansen (2001): they consider the problem as an inverse problem and apply Singular Value Decomposition in different bases to provide estimators. They obtain rates comparable to ours though on different regularity spaces, however their procedure is not adaptive and depends on the choice of a cutoff which is only empirically studied. Later on, wavelet methods have been applied by Abbaszadeh *et al.* (2012,2013) to estimate the density and its derivatives, considering a general L^p-risk, and in presence of additional bias. Their estimators are adaptive and reach the same rates as ours up to logarithmic terms (when $p = 2$). They do not provide lower bound, and consider neither global estimation of the density (wavelets are compactly supported) nor survival function estimation. Note that Chesneau (2013) studies the multiplicative censoring model when the sequence $(X_i)_{i\in\mathbb{N}}$ is α-mixing and the U_i's can be a product of independent uniform random variables. The dependence implies an important loss in the rate.

In this paper, we consider first the case where the X_i's are real-valued, and we investigate the pointwise nonparametric estimation on $\mathbb{R}$ of both f and the survival function $\bar{F}(x) = 1 - F(x)$. All nonparametric methods (likelihood, projection, kernel, ...) rely on relationships between the density f_Y and survival function $\bar{F}_Y = 1 - F_Y$ of Y_i and those of X_i, given by

$$\forall y \in \mathbb{R}, \quad f_Y(y) = \int_y^{+\infty} \frac{f(x)}{x} dx \; 1_{y\geq 0} + \int_{-\infty}^{y} \frac{f(x)}{|x|} dx \; 1_{y<0}, \tag{2}$$

$$\forall y \in \mathbb{R}, \quad \bar{F}_Y(y) + y f_Y(y) = \bar{F}(y), \tag{3}$$

which imply the following key property. Let $t : \mathbb{R} \to \mathbb{R}$ be bounded, derivable, with t' belonging to $L^2(\mathbb{R})$, and assume that $\mathbb{E}|X| < +\infty$, then

$$\mathbb{E}(t(Y) + Y t'(Y)) = \mathbb{E}t(X). \tag{4}$$

This relation allows us to propose adequate correction of the observation Y in order to get information about X, and yields simple kernel estimators of f and $\bar{F}$ (see Formulae (7) and (5)). We first study their pointwise L^2-risks properties. Under regularity assumptions, we can obtain rates of convergence for which lower bounds are provided. The study includes the classical case of nonnegative X_i's, for which pointwise kernel estimation of the density and the survival function is new.

Then we study the global risk for f on $\mathbb{R}$ or for $\bar{F}$ on $\mathbb{R}^+$ when the variables are nonnegative. An adaptive choice of the bandwidths is proposed following the Goldenshluger and Lepski (2011) theory and proved to lead to automatic bias-variance tradeoff for the resulting adaptive density or survival function estimators.

Next, still considering nonnegative X_i's, we use convolution power kernel estimators fitted to nonparametric estimation of functions on $\mathbb{R}^+$, proposed in Comte and Genon-Catalot (2012) for standard density estimation. We introduce estimators of $f, \bar{F}$ (now on $\mathbb{R}^+$), different from those based on classical kernels. The interest is to avoid boundary effects at 0. These kernels require the choice of an integer parameter m, for which a data driven procedure is proposed. The resulting estimator is proved to be adaptive.

The paper is organized as follows. Standard kernel estimators are described and studied in Section 2, and convolution power kernel method is explained in Section 3. Section 4 presents a simulation study that allows us to compare the two strategies. Lastly, proofs are gathered in Section 5.

2. Kernel estimation on the real line

2.1. **Definition of kernel estimators.** Let $K:\mathbb{R}\to\mathbb{R}$ be a kernel *i.e.* an integrable function with $\int K(u)du=1$, which is also assumed to be square integrable. We set for $h>0$, $K_h(u)=(1/h)K(u/h)$. Along the results hereafter, we possibly need additional conditions on K:

(A1) K is bounded.
(A2) K is an even function with $\lim_{u\to+\infty}K(u)=0$, K is derivable and K' is integrable.
(A3) $\int[K'(u)]^2du<+\infty$.
(A4) $\int|u|[K'(u)]^2du<+\infty$.
(A5) $\int[uK'(u)]^2du<+\infty$.

The estimator of $\bar{F}(x)$ is defined by:

$$\begin{aligned}\hat{\bar{F}}_h(x) &= \frac{1}{nh}\sum_{i=1}^n\left(\int K\left(\frac{u-x}{h}\right)\mathbb{1}_{Y_i\geq u}du+Y_iK\left(\frac{Y_i-x}{h}\right)\right)\\ &= K_h\star\hat{\bar{F}}_Y(x)+\frac{1}{n}\sum_{i=1}^n Y_iK_h\left(Y_i-x\right)\end{aligned}\tag{5}$$

where $s\star t(x)=\int s(x-u)t(u)du$ denotes the convolution product and

$$\hat{\bar{F}}_Y(x)=\frac{1}{n}\sum_{i=1}^n\mathbb{1}_{Y_i\geq x}.\tag{6}$$

For K satisfying **(A2)**, which implies **(A1)**, we define the estimator of $f(x)$ by:

$$\begin{aligned}\hat{f}_h(x) &= \frac{1}{nh}\sum_{i=1}^n\left\{\frac{Y_i}{h}K'\left(\frac{Y_i-x}{h}\right)+K\left(\frac{Y_i-x}{h}\right)\right\}\\ &= \frac{1}{n}\sum_{i=1}^n\left\{Y_iK_h'\left(Y_i-x\right)+K_h\left(Y_i-x\right)\right\}.\end{aligned}\tag{7}$$

With simple computations, we can prove:

Proposition 2.1. *Under* **(A2)***,*

$$(i)\ \int\hat{f}_h(x)dx=1,\ \ (ii)\ \lim_{x\to-\infty}\hat{\bar{F}}_h(x)=1,\ \ (iii)\ \lim_{x\to+\infty}\hat{\bar{F}}_h(x)=0,\ \ (iv)\ (\hat{\bar{F}}_h)'(x)=-\hat{f}_h(x).$$

2.2. **Pointwise risk.** Consider the Hölder ball:

$$\Sigma_I(\beta,R)=\{f:I\to\mathbb{R},f^{(\ell)}\text{exists for }\ell=\lfloor\beta\rfloor,|f^{(\ell)}(x)-f^{(\ell)}(x')|\leq R|x-x'|^{\beta-\ell},\forall x,x'\in I\}$$

where $\lfloor\beta\rfloor$ is the largest integer strictly smaller than β. Recall that K is a kernel of order ℓ if:

$$\int|u|^\ell|K(u)|du<\infty\ \text{and}\ \int u^jK(u)du=0\ \text{for}\ j=1,\dots,\ell.$$

The following proposition shows that the risk at x_0 has a different rate for $x_0\neq0$ and for $x_0=0$.

Proposition 2.2. *Assume that* $\mathbb{E}(|X_1|) < +\infty$.
Let $x_0 \in \mathbb{R}$. *Assume that* f *belongs to* $\Sigma_I(\beta, R)$ *for* I *a neighborhood of* x_0. *If the kernel* K *is of order* $\ell + 1$ *with* $\ell = \lfloor \beta \rfloor$ *and* $\int |u|^{\beta+1}|K(u)|du < +\infty$, *then under* **(A1)**,

$$\mathbb{E}[(\hat{\bar{F}}_h(x_0) - \bar{F}(x_0))^2] \leq C_1^2 h^{2(\beta+1)} + \frac{C_2}{nh} + \frac{C_3}{n}, \tag{8}$$

$$\mathbb{E}[(\hat{\bar{F}}_h(0) - \bar{F}(0))^2] \leq C_1^2 h^{2(\beta+1)} + \frac{C_4}{n}, \tag{9}$$

with $C_1 = R\int |u|^{\beta+1}|K(u)|du/(\ell+1)!$, $C_2 = 2\mathbb{E}(|X_1|)\|K\|^2$, $C_3 = 2\|K\|^2$ *and* $C_4 = 2\|K\|^2 + \int |u|K^2(u)du$.
If K *is of order* ℓ *with* $\ell = \lfloor \beta \rfloor$ *and* $\int |u|^{\beta}|K(u)|du < +\infty$, *then under* **(A2)-(A3)**, *for all* $h \in (0, 1)$, *we have*

$$\mathbb{E}[(\hat{f}_h(x_0) - f(x_0))^2] \leq C_5^2 h^{2\beta} + \frac{C_6}{nh^3} \tag{10}$$

with $C_5 = R\int |u|^{\beta}|K(u)|du/\ell!$ *and* $C_6 = 2\left(\mathbb{E}(|X_1|)\|K'\|^2 + \|K\|_\infty^2\right)$.
Under **(A2)**-**(A4)**, *for* $x_0 = 0$, *we have*

$$\mathbb{E}[(\hat{f}_h(0) - f(0))^2] \leq C_5^2 h^{2\beta} + \frac{C_6'}{nh^2}, \tag{11}$$

where $C_6' = \|K\|_\infty + \int |u|[K'(u)]^2du$.
Under **(A2)**-**(A3)** *and* **(A2)**, *if* $\mathbb{E}(1/|X|) = \|f_Y\|_\infty < +\infty$, $\|f\|_\infty < +\infty$, *and* $x_0 = 0$, *we have*

$$\mathbb{E}[(\hat{f}_h(0) - f(0))^2] \leq C_2^5 h^{2\beta} + \frac{C_6''}{nh}, \tag{12}$$

where $C_6'' = \|f\|_\infty\|K\|^2 + \|f_Y\|_\infty \int u^2[K'(u)]^2du$.

For h of order $n^{-1/(2\beta+3)}$, the estimator of $\bar{F}(x_0)$ has rate $O(n^{-2(\beta+1)/(2\beta+3)})$, except in 0, where choosing $h = n^{1/[2(\beta+1)]}$, gives the parametric rate. This is due to the fact that $\mathbb{P}(X > 0) = \mathbb{P}(Y > 0)$ and thus $\bar{F}(0) = \bar{F}_Y(0)$. For instance, $n^{-1}\sum_{i=0}^n \mathbb{1}_{Y_i \geq 0}$ is an estimator of $\bar{F}(0)$ with parametric rate.
For h of order $n^{-1/(2\beta+3)}$, the estimator of $f(x_0)$ has rate $O(n^{-2\beta/(2\beta+3)})$ when $x_0 \neq 0$. The rate is better at $x_0 = 0$ and of order $O(n^{-2\beta/(2\beta+2)})$ or $O(n^{-2\beta/(2\beta+1)})$, provided that $\|f_Y\|_\infty < +\infty$.

The next theorem states that the rates obtained for points $x_0 \neq 0$ are optimal-minimax.

Theorem 2.1. *Assume that* $x_0 \neq 0$, $x_0 \in I$ *and let* $\beta > 0$.
There exists a constant $c > 0$ *such that*

$$\liminf_{n\to+\infty} n^{2\beta/(2\beta+3)} \inf_{\hat{f}_n} \sup_{f\in\Sigma_I(\beta,R)} \mathbb{E}_f\left[(\hat{f}_n(x_0) - f(x_0))^2\right] \geq c \tag{13}$$

where $\inf_{\hat{f}_n}$ *denotes the infimum over all estimators of* f *based on* $(Y_j)_{1\leq j\leq n}$.
Moreover, for $\beta \geq 1$, *there exists a constant* $c > 0$ *such that*

$$\liminf_{n\to+\infty} n^{2\beta/(2\beta+1)} \inf_{\hat{\bar{F}}_n} \sup_{\hat{\bar{F}}\in\Sigma_I(\beta,R)} \mathbb{E}_f\left[(\hat{\bar{F}}_n(x_0) - \bar{F}(x_0))^2\right] \geq c \tag{14}$$

where $\inf_{\hat{\bar{F}}_n}$ *denotes the infimum over all estimators of* $\bar{F}$ *based on* $(Y_j)_{1\leq j\leq n}$.

2.3. **Global risk and bandwidth selection.** We denote by $\|\psi\| = (\int \psi^2(x)dx)^{1/2}$ the $\mathbb{L}^2$-norm and by $\|\psi\|_1 = \int |\psi(x)|dx$ the $\mathbb{L}^1$-norm of a function $\psi : \mathbb{R} \to \mathbb{R}$.

Let $f_h(x) = \int f(u)K((x-u)/h)/h du = f \star K_h(x)$ and $\bar{F}_h(x) = \bar{F} \star K_h(x)$. We can prove:

Proposition 2.3. *Assume that $\mathbb{E}(X_1^2) < +\infty$.*
If f belongs to $\mathbb{L}^2(\mathbb{R})$ and **(A2)-(A3)** *hold, then*

$$\mathbb{E}(\|\hat{f}_h - f\|^2) \le \|f - f_h\|^2 + \frac{\|K\|^2}{nh} + \frac{\mathbb{E}(Y_1^2)\|K'\|^2}{nh^3}. \tag{15}$$

If X is nonnegative, $\bar{F}$ belongs to $\mathbb{L}^2(\mathbb{R}_+)$ and K has compact support $[-1,1]$, then, for all $h \le 1$,

$$\mathbb{E}(\int_{\mathbb{R}^+} (\hat{\bar{F}}_h(x) - \bar{F}(x))^2 dx) \le \int_{\mathbb{R}^+} (\bar{F}_h(x) - \bar{F}(x))^2 dx + \frac{2\mathbb{E}(Y_1^2)\|K\|^2}{nh} + \frac{2\mathbb{E}(Y_1+1)\|K\|_1^2}{n}. \tag{16}$$

By considering Nikols'ki classes of functions (see Tsybakov (2009)) instead of Hölder classes, we may evaluate the bias order and deduce rates of convergence. As the regularity is not known, we rather propose a bandwidth selection strategy inspired from Goldenshluger and Lepski (2011), which yields a nonasymptotic risk bound result. To that aim, let

$$\hat{f}_{h,h'}(x) = K_{h'} \star \hat{f}_h(x), \text{ and } \hat{\bar{F}}_{h,h'} = K_{h'} \star \hat{\bar{F}}_h(x).$$

Note that, as the kernel is even, $\hat{f}_{h,h'} = \hat{f}_{h',h}$ and $\hat{\bar{F}}_{h,h'} = \hat{\bar{F}}_{h',h}$. Let $\mathcal{H}_n$ be a finite set of bandwidths. Then set

$$A(h) = \sup_{h' \in \mathcal{H}_n} \left(\|\hat{f}_{h'} - \hat{f}_{h,h'}\|^2 - V(h') \right)_+, \; B(h) = \sup_{h' \in \mathcal{H}_n} \left(\|\hat{\bar{F}}_{h'} - \hat{\bar{F}}_{h,h'}\|^2_{\mathbb{R}_+} - W(h') \right)_+,$$

with

$$V(h) = \kappa_1 \|K\|_1^2 \left(\frac{\|K\|^2}{nh} + \frac{\mathbb{E}(Y_1^2)\|K'\|^2}{nh^3} \right), \; W(h) = \kappa_2 \|K\|_1^2 \frac{\mathbb{E}(Y_1^2)\|K\|^2}{nh}, \tag{17}$$

where κ_1 and κ_2 are numerical constants.

For each estimator, the term $A(h)$ (resp. $B(h)$) approximates the square bias term while $V(h)$ (resp. $W(h)$) is proportional to the variance term. Therefore, the data-driven bandwidths are defined by:

$$\hat{h} = \arg\min_{h \in \mathcal{H}_n} \left(A(h) + V(h) \right), \quad h^\star = \arg\min_{h \in \mathcal{H}_n} (B(h) + W(h)). \tag{18}$$

The above definitions depend on the unknown moment $\mathbb{E}(Y_1^2)$, which should be replaced by $n^{-1}\sum_{i=1}^n Y_i^2$. This substitution is possible both in theory and in practice. Note that $\|K\|_1 \ge 1 = \int K$. The following holds:

Theorem 2.2. *Assume that f belongs to $\mathbb{L}^2(\mathbb{R})$, $\mathbb{E}(X_1^8) < +\infty$ and $\mathcal{H}_n$ is such that*
(i) $\mathrm{Card}(\mathcal{H}_n) \le n$,
(ii) $\forall a > 0$, $\exists \Sigma(a) > 0$ *such that* $\sum_{h \in \mathcal{H}_n} h^{-2}\exp(-a/h) < \Sigma(a) < \infty$,
(iii) $\forall h \in \mathcal{H}_n$, $1/(nh^3) \le 1$.
Then, under **(A2)-(A3)**, *there exists a numerical constant κ_1 in $V(h)$ defined by (17)such that*

$$\mathbb{E}(\|\hat{f}_{\hat{h}} - f\|^2) \le c \inf_{h \in \mathcal{H}_n} \left(\|K\|_1^2 \|f - f_h\|^2 + V(h) \right) + \frac{c'}{n}, \tag{19}$$

where c is a numerical constant and c' depends on K and f_Y.
If X is nonnegative, $\bar{F}$ belongs to $\mathbb{L}^2(\mathbb{R}^+)$, $\mathcal{H}_n$ satisfies (i) and (ii) and K is chosen with compact

support $[-1,1]$, *then there exists a numerical constant* κ_2 *in* $W(h)$ *defined by (17) such that,*

$$\mathbb{E}(\int_{\mathbb{R}^+}(\hat{\bar{F}}_{h^\star}(x)-\bar{F}(x))^2dx) \leq c_1 \inf_{h\in\mathcal{H}_n}\left(\|K\|_1^2\int_{\mathbb{R}^+}(\bar{F}_h(x)-\bar{F}(x))^2dx + W(h)\right)+\frac{c_1'}{n}, \quad (20)$$

where c_1 *is a numerical constant and* c_1' *depends on* K *and* f_Y*.*

The proof delivers numerical values for the constants κ_1, κ_2 which are too large. Finding the minimal values is theoretically difficult. This is why it is standard to calibrate their value by preliminary simulations (see Section 4).

For instance $\mathcal{H}_n = \{1/k, k=1,\dots,n\}$ or $\mathcal{H}_n = \{2^{-k},\ k=1,\dots,\log(n)/\log 2\}$ fulfill (i)-(ii). For (iii) the admissible values of k must be restricted to $n^{1/3}$ or $\log(n)/(3\log 2)$. Actually, (iii) can be replaced by $1/(nh^3)\leq C$ for a constant C.

3. Convolution power kernel estimation on $\mathbb{R}^+$

Now, we assume that the X_i's are nonnegative. The properties of the kernel estimators of f and the survival function $\bar{F}$ of the previous section are still valid for $x\geq 0$ by setting $f(x)=0$ for $x<0$. However, for estimating functions on $\mathbb{R}^+$, it is often better to use appropriate kernels so as to avoid the boundaries effects near 0. The convolution power kernels (Comte and Genon-Catalot (2012)) are well fitted to deal with this problem.

3.1. Definition of convolution power kernel estimators.

For k a density on $\mathbb{R}^+$ with expectation 1, we denote by k_m the density of $(E_1+\cdots+E_m)/m$ with E_i i.i.d. with density k, *i.e.*

$$k_m(u) = mk^{\star m}(mu), \quad u\geq 0 \quad (21)$$

where $k^{\star m} = k\star\cdots\star k$, m times and $\star$ denotes the convolution product. For h integrable, we denote by $h^*(t)=\int e^{itu}h(u)du, t\in\mathbb{R}$ its Fourier transform. The Fourier transform of k_m is given by

$$k_m^*(t) = (k^*(\frac{t}{m}))^m, \quad t\in\mathbb{R}.$$

For $\alpha_1,\dots,\alpha_L$ real numbers such that $\sum_{j=1}^L \alpha_j = 1$, $k^{(1)},\dots,k^{(L)}$ densities on $\mathbb{R}^+$ with expectation 1, we define the convolution power kernel (CPK) by

$$K_m = \sum_{j=1}^L \alpha_j k_m^{(j)}. \quad (22)$$

The following assumptions are required on the densities $k^{(j)}$, for $j=1,\dots,L$.

(B1) For $u\geq 0$, $k^{(j)}(u)\geq 0$, for $u<0$, $k^{(j)}(u)=0$, $\int_0^{+\infty}k^{(j)}(u)du=1$, $\int_0^{+\infty}(k^{(j)})^2(u)du<+\infty$, $\lim_{u\to+\infty}uk^{(j)}(u)=0$, and

$$\int_0^{+\infty}uk^{(j)}(u)du=1, \quad \exists\gamma\geq 4, \text{ such that } \int_0^{+\infty}|u-1|^\gamma k^{(j)}(u)du<+\infty$$

(B2) For m large enough, $\int_0^{+\infty}k_m^{(j)}(u)\frac{du}{u} = 1+O(\frac{1}{m})$.
(B3) There exists $m_0\geq 1$ such that the function $t[(k^{(j)})^*(t)]^{m_0}$ belongs to $\mathbb{L}^1(\mathbb{R})\cap\mathbb{L}^2(\mathbb{R})$.
Note that assumptions **(B2)** is not stringent as $k_m^{(j)}(u)du$ tends to δ_1 as m tends to infinity.

Now, we define for $x > 0$, the estimator of $\bar{F}(x)$ by:

$$\tilde{\bar{F}}_m(x) = \frac{1}{x}\int_0^{+\infty} K_m\left(\frac{u}{x}\right)\hat{\bar{F}}_Y(u)du + \frac{1}{nx}\sum_{i=1}^n Y_i K_m\left(\frac{Y_i}{x}\right). \tag{23}$$

Under **(B3)**, K_m is derivable, so we can define, for estimating $f(x)$ at $x > 0$,

$$\tilde{f}_m(x) = \frac{1}{nx}\sum_{i=1}^n \left[K_m\left(\frac{Y_i}{x}\right) + \frac{Y_i}{x}K'_m\left(\frac{Y_i}{x}\right)\right]. \tag{24}$$

Proposition 3.1. *Under* **(B1)***, we have* $\lim_{x\to 0^+}\tilde{\bar{F}}_m(x) = 1, \lim_{x\to+\infty}\tilde{\bar{F}}_m(x) = 0$.
Under **(B1)-(B3)***, we have* $\int_0^{+\infty}\tilde{f}_m(x)dx = 1 + O(\frac{1}{m})$.

3.2. **Examples of kernels yielding explicit formulae.** Examples of densities k leading to explicit formulae for k_m are the following.
Example 1. Uniform kernels and splines. Let $k(x) = (1/2)\mathbb{1}_{[0,2]}(x)$ the uniform density on $[0, 2]$. Then Formula (9) in Killmann and von Collani (2001) (see also Rényi (1970)) yields

$$k_m(x) = \frac{m}{(m-1)!2^m}\sum_{i=0}^{\lfloor mx/2\rfloor}(-1)^i\binom{m}{i}(mx-2i)^{m-1}\mathbb{1}_{[0,2]}(x)$$

Here, k is not continuous on $(0, +\infty)$ and successive convolutions increase the regularity. Thus, the exponent m plays clearly the role of regularity parameter. Assumptions **(B1)** and **(B3)** hold.

Example 2. Gamma kernels. For k the Gamma density $G(a, a)$, $a > 0$, **(B1)-(B3)** hold and:

$$k_m(u) = \frac{(am)^{am}}{\Gamma(am)}u^{am-1}e^{-mau}1_{u>0}.$$

For $m > 1/a$, $\int_0^{+\infty} u^{-1}k_m(u)du = (am)/(am-1) = 1 + O(1/m)$.

Example 3. Inverse Gaussian kernels. The inverse Gaussian distribution $IG(a, \theta)$ $a > 0, \theta > 0$, is defined as the distribution of the hitting time $T_a = \inf\{t \geq 0,\ \theta t + B_t = a\}$ where (B_t) is a standard Brownian motion. The density of an $IG(a, \theta)$ is $(a/\sqrt{2\pi t^3})e^{\theta a}e^{-(1/2)\left(\theta^2 t + a^2/t\right)}$. For $a = \theta$, the expectation is 1 and the variance is $v = 1/a^2$. For k the inverse Gaussian density $IG(a, a)$, **(B1)-(B3)** hold and k_m is the density of the law $IG(a\sqrt{m}, a\sqrt{m})$:

$$k_m(u) = \frac{a\sqrt{m}}{\sqrt{2\pi u^3}}e^{ma^2(1-\frac{1}{2}(\frac{1}{u}+u))}1_{u>0}, \quad \int_0^{+\infty} u^{-1}k_m(u)du = 1 + 1/(a^2m).$$

3.3. **Pointwise risk.** To evaluate the order of the bias term, we need to define the notion of convolution power kernel of order ℓ.

Definition 3.1. *We say that* $K_m = \sum_{j=1}^L \alpha_j k_m^{(j)}$ *defines a convolution power kernel of order* ℓ *if, for* $j = 1, \ldots, L$*, the density* $k^{(j)}$ *satisfies Assumptions* **(B1)**-**(B2)***, admits moments up to order* ℓ *and the coefficients* $\alpha_j, j = 1, \ldots, L$ *are such that* $\sum_{j=1}^L \alpha_j = 1$ *and for* $1 \leq k \leq \ell$ *and all* m *(at least large enough)*

$$\int_0^{+\infty}(u-1)^k K_m(u)du = \sum_{j=1}^L \alpha_j\int_0^{+\infty}(u-1)^k k_m^{(j)}(u)du = 0.$$

These relations allow to compute the α_j's as functions of the moments of the $k^{(j)}$'s (see Comte and Genon-Catalot (2012)). Note that a single convolution power kernel with $L = 1$ is of order one.
Now, we can prove the following result.

Proposition 3.2. *Let $x_0 > 0$. Consider the estimator (23) built with a kernel (22) satisfying* **(B1)**. *Set*

$$|\alpha|_1 := \sum_{i=1}^{L} |\alpha_i|, \quad v_j = \int_0^{+\infty} (u-1)^2 k^{(j)}(u)du, \quad j = 1, \dots L. \tag{25}$$

If $\bar{F}$ belongs to $\Sigma_I(\beta, R)$ for I a neighborhood of x_0, the kernel K_m is order $\ell = \lfloor \beta \rfloor$ in the sense of Definition 3.1 and for $j = 1, \dots, L$, $\int_0^{+\infty} |u-1|^\beta k^{(j)}(u)du < +\infty$, then for m, n large enough,

$$\mathbb{E}[(\tilde{\bar{F}}_m(x_0) - \bar{F}(x_0))^2] \le C_1(\beta)\frac{x_0^{2\beta}}{m^\beta} + 4|\alpha|_1 \sum_{j=1}^{L} \frac{|\alpha_j|}{\sqrt{2\pi v_j}} \frac{\sqrt{m}}{n} + 2\frac{|\alpha|_1^2}{n}$$

where $C_1(\beta)$ is a constant depending on R, β, the α_j's and the moments of the $k^{(j)}$'s.
Assume moreover that **(B2)** *and* **(B3)** *hold and f is bounded. Then,*

$$\mathbb{E}[(\tilde{f}_m(x_0) - f(x_0))^2] \le C_1(\beta)\frac{x_0^{2\beta}}{m^\beta} + \left(C_2' \|f\|_\infty \frac{\sqrt{m}}{n x_0} + C_2'' \frac{(\sqrt{m})^3}{n x_0^3} \right)$$

where $C_2' = 2\sum_{1\le i,j\le L} |\alpha_i\alpha_j|/\sqrt{2\pi(v_i+v_j)}$, $C_2'' = 2\sum_{1\le i,j\le L} |\alpha_i\alpha_j|/\sqrt{2\pi(v_i+v_j)^3)}$ and $C_1(\beta)$ is the same as above.

3.4. **Global risk and adaptation.** We prove a global result for $\tilde{\bar{F}}_m$.

Proposition 3.3. *Assume that* **(B1)** *holds and* $\mathbb{E}(X_1) < +\infty$. *Then the integrated risk satisfies:*

$$\mathbb{E}\left[\int_0^{+\infty} \left(\tilde{\bar{F}}_m(x) - \bar{F}(x)\right)^2 dx\right] \le \int_0^{+\infty} \left(\mathbb{E}\tilde{\bar{F}}_m(x) - \bar{F}(x)\right)^2 dx + C_2' \frac{10\sqrt{m}\,\mathbb{E}(Y_1)}{n}$$

where C_2' is defined in Proposition 3.2.

Below, we do not search to link the bias term with the regularity property of the function $\bar{F}$. We rather focus on finding a data-driven value of m without knowing the regularity of $\bar{F}$ on $\mathbb{R}^+$. From Propositions 3.2 and 3.3, $\sqrt{m}$ plays the role of the bandwidth.
For two functions s and t on $(0, +\infty)$, let us define, each time it exists on $(0, +\infty)$, the function

$$u \to s \odot t(u) = \int_0^{+\infty} s(u/v)t(v)dv/v.$$

If U_1, U_2 are nonnegative independent random variables with densities k_1, k_2 respectively, then the product U_1U_2 has density $k_1 \odot k_2(u)$. Now, we define

$$\mathcal{M}_n = \{m = k^2, \ \log(n) \le k \le n/\log(n)\} \tag{26}$$

as the set of possible indexes m and consider $K_m = \sum_{j=1}^{L} \alpha_j k_m^{(j)}$, where the densities $k^{(j)}$ satisfy **(B1)**. Set

$$\tilde{\bar{F}}_{m,m'}(x) = \frac{1}{x}\int_0^{+\infty} K_{m'} \odot K_m\left(\frac{u}{x}\right) \hat{\bar{F}}_Y(u)du + \frac{1}{nx}\sum_{i=1}^{n} Y_i K_{m'} \odot K_m\left(\frac{Y_i}{x}\right). \tag{27}$$

As $K_m \odot K_{m'} = K_{m'} \odot K_m$, we have $\tilde{\bar{F}}_{m,m'}(x) = \tilde{\bar{F}}_{m',m}(x)$. For κ a numerical constant and

$$C(K) = 2|\alpha|_1^3 (\sum_{i=1}^{L} |\alpha_i|/\sqrt{2\pi v_i}), \tag{28}$$

we set

$$Z(m) = \kappa C(K)\mathbb{E}(Y_1)\frac{\sqrt{m}}{n}, \qquad H(m) = \sup_{m' \in \mathcal{M}_n} \left(\|\tilde{\bar{F}}_{m'} - \tilde{\bar{F}}_{m,m'}\|^2 - Z(m') \right)_+ . \tag{29}$$

Note that, as $|\alpha|_1 \geq 1$, the constant C_2' of Proposition 3.3 satisfies

$$C_2' \leq 2|\alpha|_1 \sum_{j=1}^{L} \frac{|\alpha_j|}{\sqrt{2\pi v_j}} \leq C(K). \tag{30}$$

The adaptive estimator is then $\tilde{\bar{F}}_{\tilde{m}}$ with

$$\tilde{m} = \arg \min_{m \in \mathcal{M}_n} \left(H(m) + Z(m) \right). \tag{31}$$

As noted above, we should replace the unknown moment $\mathbb{E}(Y_1)$ by its empirical counterpart. This is no difficulty in the proofs. We can prove the following result:

Theorem 3.1. *Assume that $\bar{F}$ belongs to $\mathbb{L}^2((0, +\infty))$. Assume that* **(B1)** *holds and $\mathbb{E}(X_1^4) < +\infty$, then there exists a numerical constant κ such that*

$$\mathbb{E}\left[\int_0^{+\infty} \left(\tilde{\bar{F}}_{\tilde{m}}(x) - \bar{F}(x)\right)^2 dx\right] \leq C \inf_{m \in \mathcal{M}_n} \left\{ \int_0^{+\infty} \left[\mathbb{E}\bar{F}_m(x)) - \bar{F}(x)\right]^2 dx + Z(m) \right\} + \frac{C'}{n},$$

where C is a numerical constant and C' a constant depending on $\mathbb{E}(X_1^4)$ and on K.

Inverse Gaussian kernels (Example 3) are well fitted for practical implementation. Indeed if k is IG(1,1), then $k_m \odot k_{m'}$ has the following explicit density

$$k_m \odot k_{m'}(u) = \frac{\sqrt{mm'}}{\pi u^{3/2}} \exp(m + m')\tilde{K}_0 \left(\left(m^2 + (m')^2 + m\, m'(u + \frac{1}{u}) \right)^{1/2} \right) \tag{32}$$

where $\tilde{K}_0$ is the modified Bessel function of second kind with index 0 (available in R, library Bessel), see Proposition 3.6 in Comte and Genon-Catalot (2012).

4. Simulations

In this section, we implement our estimators on simulated data. We have selected the following distributions:

Model 1: a Gaussian density, $X \sim \mathcal{N}(2.5, 0.75)$,
Model 2: a mixture of Gaussian densities, $X \sim 0.5\mathcal{N}(-2, 1) + 0.5\mathcal{N}(2, 1)$,
Model 3: a Gamma distribution, $X \sim \Gamma(8, 4)$,
Model 4: a rescaled Beta distribution, $X = 5X'$, $X' \sim \beta(3, 3)$,
Model 5: an Exponential distribution $X \sim \exp(2)$,
Model 6: a mixture of Gamma distributions $X \sim 0.4\Gamma(1, 10) + 0.6\Gamma(40, 30)$.

	$n=$	Model 1 200	Model 1 500	Model 2 200	Model 2 500	Model 3 200	Model 3 500	Model 4 200	Model 4 500
Oracle	Mean	0.022	0.014	0.007	0.005	0.022	0.015	0.013	0.006
	(std)	(0.014)	(0.011)	(0.003)	(0.002)	(0.017)	(0.011)	(0.008)	(0.003)
GL	Mean	0.090	0.033	0.018	0.014	0.073	0.026	0.138	0.037
	(std)	(0.131)	(0.05)	(0.031)	(0.015)	(0.184)	(0.018)	(0.723)	(0.065)
	Med.	0.045	0.009	0.014	0.009	0.036	0.021	0.020	0.011
CV	Mean	0.403	0.269	0.215	0.109	0.297	0.126	0.509	0.314
	(std)	(0.705)	(0.378)	(0.448)	(0.222)	(0.518)	(0.182)	(1.031)	(0.386)
	Med.	0.067	0.047	0.011	0.009	0.044	0.034	0.031	0.094
CV on X	Mean	0.015	0.006	0.007	0.004	0.018	0.009	0.006	0.004
	(std)	(0.012)	(0.004)	(0.004)	(0.002)	(0.012)	(0.007)	(0.010)	(0.002)
	Med.	0.013	0.005	0.007	0.004	0.015	0.008	0.009	0.004
Oracle on X	Mean	0.004	0.002	0.003	0.002	0.004	0.002	0.003	0.002
	(std)	(0.003)	(0.001)	(0.002)	(0.001)	(0.003)	(0.001)	(0.002)	(0.001)

TABLE 1. Table of risks for density estimators and oracles.

4.1. **Density estimation.** We consider the estimator given by (7) for Models 1 to 4, where $K(x)$ is the standard Gaussian kernel. Bandwidths are selected between 0.1 and 1.5 in the set

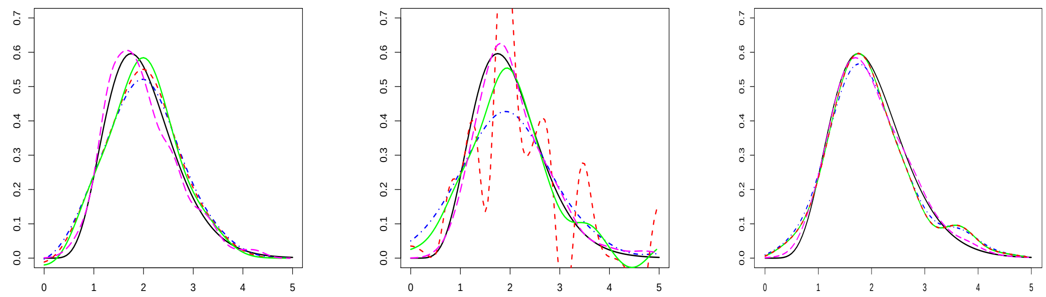
$$\mathcal{H}_n = \{0.1 + 0.05\,k, k = 0, 1, \dots, 28\}.$$

For each sample, we compute:
- first, the oracle $\hat{f}_{or} = \hat{f}_{h_{or}}$ where $h_{or} = \arg\min_{h\in\mathcal{H}_n} \|\hat{f}_h - f\|^2$,
- second, the Goldenshluger and Lepski estimator $\hat{f}_{\hat{h}}$ with $\hat{h}$ given by (18) and $\kappa_1 = 1.2$,
- third, the estimator $\hat{f}_{h_{CV}}$ where h_{CV} is selected by a cross validation criterion *i.e.* minimizing

$$CV(h) = \int \hat{f}_h^2(x)dx - \frac{2}{n}\sum_{i=1}^{n}[Y_i\hat{f}'_{h,[i]}(Y_i) + \hat{f}_{h,[i]}(Y_i)],$$

where $\hat{f}_{h,[i]}$ is the kernel estimator $\hat{f}_h$ computed on the sample without Y_i (leave-one-out),
- fourth, we compute the estimator of f based on the direct observations $X_1, \dots, X_n$, $\hat{f}^{(X)}_{h_{CV,X}}$ where $\hat{f}^{(X)}_h$ is the standard kernel estimator of f and $h_{CV,X}$ is the bandwidth selected by usual density cross-validation criterion (see e.g. Tsybakov (2009), Section 1.5),
- lastly, the oracle based on the direct observations $X_1, \cdots, X_n$, also using the standard kernel estimator.

We investigate two sample sizes $n = 200, 500$, and 50 repetitions. Table 1 gives the estimated $\mathbb{L}^2$-risks of all estimators, together with medians and standard deviations except for the oracles. As expected, the oracle on direct observations performs better than the one with censored data. The comparison between the GL method and the oracle on censored data shows that the GL method is stable and that the loss with respect to the oracle is stable. We stress that the GL method gives smaller risks than the CV method, much smaller for means, and still smaller for medians. Indeed, looking at both medians and standard deviations for the CV method, we can see that it is very unstable. This is the reason why we also experimented the CV method on the direct sample, but in this case, it behaves smartlier. We conclude that the estimator $\hat{f}_{\hat{h}}$ proposed in our paper works well. As an illustration, in Figure 1, we plot the oracle, the GL



(a) (b) (c)

FIGURE 1. True density (solid black), oracle $\hat{f}_{or}$ (green) and estimators $\hat{f}_{\hat{h}}$ (blue dash-dotted), $\hat{f}_{h_{CV}}$ (red dashed), $\hat{f}^{(X)}_{h_{CV,X}}$ (magenta long-dashed). $n = 200$ in (a) and (b), $n = 1000$ in (c).

		Model 3	Model 4	Model 5	Model 6
GL	Mean	0.025	0.043	0.006	0.013
	(std)	(0.028)	(0.038)	(0.004)	(0.012)
	Med.	0.015	0.027	0.005	0.012
Oracle GL	Mean	0.010	0.014	0.004	0.009
	(std)	(0.009)	(0.012)	(0.009)	(0.007)
	Med.	0.007	0.012	0.004	0.007
CPK	Mean	0.021	0.034	0.005	0.017
	(std)	(0.015)	(0.023)	(0.005)	(0.009)
	Med.	0.018	0.028	0.004	0.014
Oracle CPK	Mean	0.013	0.021	0.004	0.011
	(std)	(0.011)	(0.016)	(0.003)	(0.008)
	Med.	0.009	0.016	0.003	0.009

TABLE 2. Table of risks, $n = 100$ for survival function estimators and oracles.

estimator, the CV estimator on censored data and the CV estimator on direct data, for Model 3 with $n = 200$ (Figure 1 (a)-(b)), $n = 1000$ (Figure 1 (c)). When CV method works, it can be very competitive compared with GL method (see Figure 1 (a)), unfortunately, it sometimes completely fails as shown in Figure 1 (b). We can see on Figure 1 (c) that increasing n improves the estimators.

4.2. **Survival function estimation.** For survival function estimation, we investigate for Models 3 to 6 two couples of estimators:

- the Goldenshluger and Lepski-type kernel estimator $\hat{\bar{F}}_{h^\star}$ given by (5) with $h^\star$ given by (18) with $\kappa_2 = 1$, and the associated oracle $\hat{\bar{F}}^{(GL)}_{h_{or}}$, the bandwidths are selected among 30 equispaced values between 0.01 and 0.9.
- the convolution power estimator $\tilde{\bar{F}}_{\tilde{m}}$ given by (23) with $\kappa = 0.5$ and (31) with the inverse Gaussian kernel IG(1,1) of Example 3 (see also Formula (32), function 'besselK$(x, 0)$' of the R-package

		Model 3	Model 4	Model 5	Model 6
Oracle GL	Mean	0.004	0.005	0.001	0.003
	(std)	(0.003)	(0.004)	(0.001)	(0.002)
	Med.	0.003	0.004	0.001	0.002
Oracle CPK	Mean	0.005	0.008	0.001	0.004
	(std)	(0.003)	(0.005)	(0.001)	(0.002)
	Med.	0.004	0.007	0.001	0.004

TABLE 3. Table of risks ($n = 500$)

Bessel), and its associated oracle $\tilde{\bar{F}}_{m_{or}}$. The values of m are chosen among $\{5+3k, k = 0, \dots, 10\}$. This is not exactly consistent with the theoretical constraint but computationally more tractable, with good results.

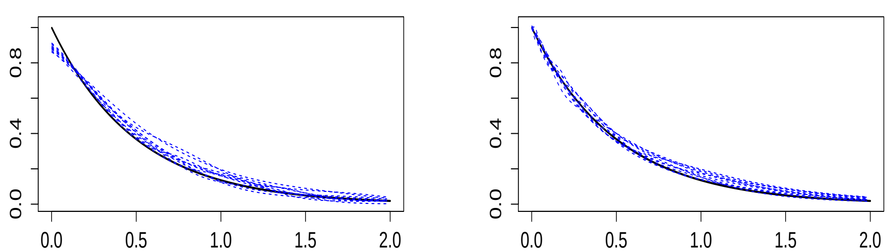


FIGURE 2. True survival function (solid black) and 10 estimators in dotted blue, $\hat{\bar{F}}_{h^\star}$ left (GL method), and $\tilde{\bar{F}}_{\tilde{m}}$ right (CPK method), for Model 5 and $n = 500$.

Table 2 gives the $\mathbb{L}^2$-risks for sample size $n = 100$ and 50 repetitions. Comparing $\mathbb{L}^2$-risks of the GL and the CPK estimators, we find that the methods behave similarly and are stable over the four models. The difference between estimators and oracles is less important for survival function estimators (Table 2) than for density estimators (Table 1). For both methods, the loss between estimators and oracles is very small for Models 5, 6. The oracle of the GL method is much better than the estimator itself for Models 3, 4. This is less true for the CPK method.

For $n = 500$ and 100 repetitions, the $\mathbb{L}^2$-risks of oracles are comparable (Table 3). In Figure 2, ten estimators of both methods for $n = 500$ are plotted together with the true function (bold), corresponding to Model 5. The GL method is on the left and the CPK on the right. Both methods yield convincing results and monotonic estimators. Although the CPK method is computationally slower, it provides better estimators, especially near zero (Figure 2).

Figure 3 concerns Model 6, with still 10 estimators and $n = 500$. On top left and right, the GL and CPK estimators. As they are not always monotonic, we have used (bottom left and right) the monotonic transformation of estimators defined by (see Chernozhukov *et al.* (2009), R-package 'quantreg'):

$$\bar{G} \mapsto \check{\bar{G}}(y) = \inf\{z; \int 1_{\bar{G}(u) \geq z} du \leq y\}.$$

One can prove that the risk of the monotonic version of any estimator on a bounded interval is smaller than the risk of the unmodified estimator. Finally, as the monotonic transformation leads to a step function, curves were smoothed using a method preserving monotonicity (R-function 'spline.fun' of the method “monoH.FC” from Fritsch and Carlson (1980)). Clearly, the monotonic transformation improves the curves. The CPK method is better near zero, and the GL method seems better for large abscissa.

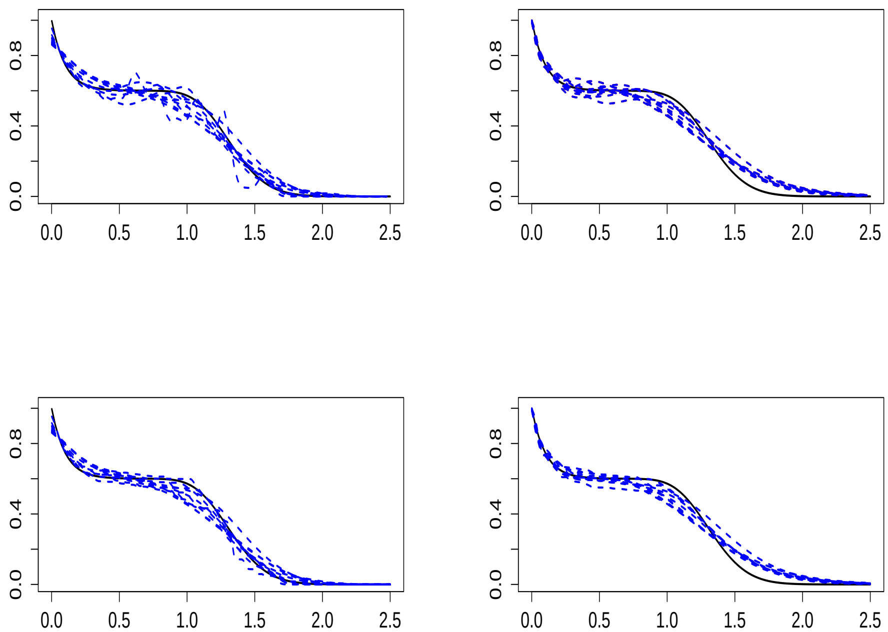


FIGURE 3. True survival function (solid black) and 10 estimators in dotted blue for Model 6 and $n = 500$. Top left: $\hat{\bar{F}}_{h^\star}$ (GL method). Top right $\tilde{\bar{F}}_{\tilde{m}}$ (CPK method). Bottom left: GL with monotonic transformation and smoothing. Bottom right: CPK with monotonic transformation and smoothing.

5. Proofs

We state a property used in proofs:

Lemma 5.1. *Let φ belong to $L^2(\mathbb{R})$, $\mathbb{E}(Y^2\varphi^2(Y)) \leq \mathbb{E}|X|\|\varphi\|^2$.*

5.1. **Proof of Equations (2)-(4) and of Lemma 5.1.** Equality (2) is elementary. For $y \geq 0$,

$$\begin{aligned}\bar{F}_Y(y) &= \int_y^{+\infty} f_Y(z)dz = \int_y^{+\infty}\int_z^{+\infty} \frac{f(x)}{x}dxdz = \int (\int_y^x dz)\frac{f(x)}{x}\mathbb{I}_{y\leq x}dx \\ &= \int_y^{+\infty}(x-y)\frac{f(x)}{x}dx = \int_y^{+\infty} f(x)dx - y\int_y^{+\infty}\frac{f(x)}{x}dx = \bar{F}(y) - yf_Y(y).\end{aligned}$$

For $y \leq 0$,

$$\begin{aligned} F_Y(y) &= \int_{-\infty}^{y} f_Y(z)dz = \int_{-\infty}^{y} dz \int_{-\infty}^{z} \frac{f(x)}{|x|}dx = \int (\int_{x}^{y} dz) \frac{f(x)}{|x|} \mathbb{1}_{x\leq y} dx \\ &= \int_{-\infty}^{y} (y-x)\frac{f(x)}{|x|}dx = y f_Y(y) + F(y) \end{aligned}$$

Thus, $\bar{F}_Y(y) = \bar{F}(y) - y f_Y(y)$, which is (3).

For (4), by (2), $y f_Y(y)$ tends to 0 as both y tends to $+\infty$ and $-\infty$. By (3), $\mathbb{E}Y^2(t'(Y))^2$ is finite. Integrating by parts yields

$$\begin{aligned} \int_{\mathbb{R}} f_Y(y)(t(y) + yt'(y))dy &= -\int_{\mathbb{R}} yt(y)(f_Y(y))'dy = -[\int_{0}^{+\infty} yt(y)(-\frac{f(y)}{y})dy + \int_{-\infty}^{0} yt(y)\frac{f(y)}{|y|}dy] \\ &= \int_{-\infty}^{+\infty} t(y)f(y)dy. \end{aligned}$$

Lemma 5.1 is immediate noting that $\mathbb{E}Y^2\varphi^2(Y) \leq \mathbb{E}X^2\varphi^2(UX)$. □

5.2. **Proof of proposition 2.1.** For (i), we use $\int K'(u)du = 0$ as K' is integrable and K is even, and $\int K(u)du = 1$. For (ii) and (iii), we write $\int K_h(u-x)\mathbb{1}_{Y_i\geq u}du = \int K(z)\mathbb{1}_{z\leq (Y_i-x)/h}dz$ for the first term and use that $\lim_{u\to+\infty} K(u) = 0$ for the second term. Lastly (iv) is straightforward. □

5.3. **Proof of Proposition 2.2.** First we study $\hat{f}_h$. Noting that

$$(yK(\frac{y-x}{h}))'_y = K(\frac{y-x}{h}) + \frac{y}{h}K(\frac{y-x}{h}),$$

Equation (4) yields $\mathbb{E}(\hat{f}_h(x)) = f_h(x) = f \star K_h(x)$. Thus, for all x,

$$\mathbb{E}[(\hat{f}_h(x) - f(x))^2] = (f(x) - f_h(x))^2 + \text{Var}\left[\hat{f}_h(x)\right].$$

As K is of order $\ell = \lfloor \beta \rfloor$, the assumption on f gives, at point x_0,$(f(x_0) - f_h(x_0))^2 \leq C_2^2 h^{2\beta}$ with $C_2 = R\int |u|^\beta |K(u)|du/\ell!$ (see Tsybakov (2009) Proposition 2.1).
Next, we have

$$\begin{aligned} \text{Var}(\hat{f}_h(x_0)) &\leq \frac{1}{nh^2}\mathbb{E}\left\{\left[\frac{Y_1}{h}K'\left(\frac{Y_1-x_0}{h}\right) + K\left(\frac{Y_1-x_0}{h}\right)\right]^2\right\} \\ &= \frac{1}{nh^2}\left\{\mathbb{E}\left[\frac{Y_1^2}{h^2}\left(K'\left(\frac{Y_1-x_0}{h}\right)\right)^2\right] + \mathbb{E}\left[K^2\left(\frac{X_1-x_0}{h}\right)\right]\right\} \end{aligned} \tag{33}$$

where the last equality follows from Equation (4). Then

$$\mathbb{E}\left[K^2\left(\frac{X_1-x_0}{h}\right)\right] \leq \min(h\|f\|_\infty\|K\|^2, \|K\|_\infty^2). \tag{34}$$

By Lemma 5.1 ,

$$\mathbb{E}\left[Y_1^2\left(K'\left(\frac{Y_1-x_0}{h}\right)\right)^2\right] \leq \mathbb{E}(|X_1|)\int (K'(\frac{u-x}{h}))^2du = h\mathbb{E}(|X_1|)\int (K')^2.$$

Therefore, $\text{Var}(\hat{f}_h(x)) \leq (nh^3)^{-1}\mathbb{E}(|X_1|)\int (K')^2 + (nh^2)^{-1}\|K\|_\infty^2$. This yields (10).

The special value $x_0 = 0$ leads to other bounds. As,$\forall z \in \mathbb{R}$, $|zf_Y(z)| \leq 1$,

$$\mathbb{E}\left[Y_1^2\left(K'\left(\frac{Y_1-x_0}{h}\right)\right)^2\right] = \int (x_0+uh)^2[K'(u)]^2 f_Y(x_0+uh)hdu \leq \int |x_0+uh|[K'(u)]^2 hdu.$$

Therefore, for $x_0 = 0$, we have

$$\mathbb{E}\left[Y_1^2\left(K'\left(\frac{Y_1}{h}\right)\right)^2\right] \leq h^2 \int |u|[K'(u)]^2 du. \tag{35}$$

Consequently,

$$\mathrm{Var}(\hat{f}_h(0)) \leq \frac{\|K\|_\infty + \int |u|[K'(u)]^2 du}{nh^2} := \frac{C_6'}{nh^2}.$$

If now f_Y is bounded and $\int u^2[K'(u)]^2 du < +\infty$, we get for $x_0 = 0$,

$$\mathbb{E}\left[Y_1^2\left(K'\left(\frac{Y_1}{h}\right)\right)^2\right] \leq h^3\|f_Y\|_\infty \int u^2[K'(u)]^2 du. \tag{36}$$

Thus if moreover $\|f\|_\infty < +\infty$, using (34),

$$\mathrm{Var}(\hat{f}_h(0)) \leq \frac{\|f\|_\infty\|K\|^2 + \|f_Y\|_\infty \int u^2[K'(u)]^2 du}{nh} := \frac{C_6''}{nh}.$$

This gives inequalities (11) and (12).

Now we study $\hat{\bar{F}}_h$ to prove (8). First we have $\mathbb{E}(\hat{\bar{F}}_h(x)) = \bar{F} \star K_h(x)$ so that the bias term can be studied using Proposition 2.1 in Tsybakov (2009). Hence the bias order at x_0. Next

$$\mathrm{Var}(\hat{\bar{F}}_h(x)) \quad \leq \quad \frac{2}{nh^2}\left\{\mathbb{E}\left[\int K\left(\frac{u-x}{h}\right)\mathbb{I}_{Y_1\geq u}du\right]^2 + \mathbb{E}\left[Y_1^2K^2\left(\frac{Y_1-x}{h}\right)\right]\right\}. \tag{37}$$

We have

$$\mathbb{E}\left[\int K\left(\frac{u-x}{h}\right)\mathbb{I}_{Y_1\geq u}du\right]^2 \leq \left[\int \left|K\left(\frac{u-x}{h}\right)\right|du\right]^2 = h^2\left(\int |K(v)|dv\right)^2$$

and

$$\mathbb{E}\left[Y_1^2K^2\left(\frac{Y_1-x}{h}\right)\right] \leq h\mathbb{E}|X_1|\|K\|^2.$$

Gathering terms gives (8).
Lastly, if $x_0 = 0$, inequality (35) applies with K' replaced by K and gives the result (9) and thus ends the proof of Proposition 2.2. □

5.4. **Proof of Theorem 2.1.** Proof of (13). To obtain lower bounds, we follow the reduction scheme described in Tsybakov (2009), chapter 2. We have to find two densities $f_{0,n}$, $f_{1,n}$ such that

(i) $f_{j,n} \in \Sigma_I(\beta, R)$, $j = 0, 1$,
(ii) $(f_{1,n}(x_0) - f_{0,n}(x_0))^2 \geq c\gamma_n^2$ where γ_n^2 is the desired rate,
(iii) $\chi^2 = \chi^2(P_{f_{1,n},Y}, P_{f_{0,n},Y}) \leq c/n$, where $P_{f,Y}$ is the law of Y when X has density f.

We only prove the result for $x_0 \in (0,1) = I$. Let h_n be small enough to have $[x_0 - h_n, x_0 + h_n] \subsetneq (0,1)$. We take $f_{0,n}(x) = \mathbb{I}_{[0,1]}(x)$ and

$$f_{1,n}(x) = f_{0,n}(x) + c\gamma_n\mathbb{L}(\frac{x-x_0}{h_n})$$

where $\mathbb{L}(v) = \mathbb{L}(v)\mathbb{1}_{[-1,1]}(v)$, $\mathbb{L} \in \Sigma_{\mathbb{R}}(\beta, R)$, $\mathbb{L}(0) \neq 0$ and $\int_{-1}^{1} \mathbb{L}(v)dv = 0$. We set $\gamma_n = n^{-\beta/(2\beta+3)}$ and $h_n = n^{-1/(2\beta+3)}$. We have $\int f_{1,n} = \int f_{0,n} = 1$ and we can choose c such that $f_{1,n}(x) \geq 0$, $\forall x \in [0,1]$, so that $f_{1,n}$ and $f_{0,n}$ are $[0,1]$-supported densities.

(i) The functions $f_{j,n}$, $j = 0,1$ are in $\Sigma_I(\beta, R)$ with $I = (0,1)$ as $\gamma_n/h_n^{\beta} = 1$.

(ii) $(f_{1,n} - f_{0,n})^2(x_0) = c^2\gamma_n^2\mathbb{L}^2(0)$ is of order $n^{-2\beta/(2\beta+3)} = \gamma_n^2$, the expected rate.

(iii) Then we must prove that $\chi^2 = \displaystyle\int_0^1 \frac{(g_1 - g_0)^2(x)}{g_0(x)}dx \leq c/n$ where $g_i(x) = \int_x^1 (f_{i,n}(u)/u)du$, for $i = 0,1$. We have

$$\chi^2 = c^2\gamma_n^2 \int_0^1 \frac{\left(\int_1^x \frac{\mathbb{L}(\frac{u-x_0}{h_n})}{u}\mathbb{1}_{[x_0-h_n, x_0+h_n]}du\right)^2}{|\log(x)|}dx := c^2\gamma_n^2(I_1 + I_2),$$

$$\text{with} \quad I_1 = \int_0^{x_0-h_n} \frac{\left(\int_{x_0-h_n}^{x_0+h_n} \frac{\mathbb{L}(\frac{u-x_0}{h_n})}{u}du\right)^2}{|\log(x)|}dx, \quad I_2 = \int_{x_0-h_n}^{x_0+h_n} \frac{\left(\int_x^{x_0+h_n} \frac{\mathbb{L}(\frac{u-x_0}{h_n})}{u}du\right)^2}{|\log(x)|}dx.$$

Using that $\int_{-1}^{1} \mathbb{L}(v)dv = 0$, we write

$$I_1 = \int_0^{x_0-h_n} \frac{\left(\int_{-1}^{1} \frac{\mathbb{L}(v)}{x_0+vh_n}h_n dv\right)^2}{|\log(x)|}dx = h_n^2 \int_0^{x_0-h_n} \frac{\left(\int_{-1}^{1} \frac{\mathbb{L}(v)}{x_0}\left(\frac{1}{1+vh_n/x_0} - 1\right)dv\right)^2}{|\log(x)|}dx$$

and thus we get

$$\begin{aligned} I_1 &\leq \frac{h^4}{x_0^2|\log(x_0)|}\int_0^{x_0-h}\left(\int_{-1}^{1}\mathbb{L}(v)\frac{v/x_0}{1+vh/x_0}dv\right)^2 dx \\ &\leq \frac{h^4}{x_0^2(x_0-h)^2|\log(x_0)|}\int_0^{x_0-h}\left(\int_{-1}^{1}|\mathbb{L}(v)|dv\right)^2 dx = \frac{\left(\int_{-1}^{1}|\mathbb{L}(v)|dv\right)^2}{x_0^2(x_0-h)|\log(x_0)|}h^4. \end{aligned}$$

Next

$$I_2 = h^2\int_{x_0-h}^{x_0+h}\frac{1}{|\log(x)|}\left(\int_{(x-x_0)/h}^{1}\frac{\mathbb{L}(v)}{x_0+vh}dv\right)^2 dx \leq \frac{2\left(\int_{-1}^{1}|\mathbb{L}(v)|dv\right)^2}{(x_0-h)^2|\log(x_0+h)|}h^3.$$

Therefore $\chi^2 \leq c(x_0)\gamma_n^2 h^3 = c(x_0)/n$ which is the desired result. □

Proof of (14). We seek a rate $\tau_n^2 = n^{-2\beta/(2\beta+1)}$. We build $S_{0,n}(x) = (1-x)$ for $x \in [0,1]$, the survival function associated to $f_{0,n}$ and

$$S_{1,n}(x) = S_{0,n}(x) + c\tau_n\mathcal{L}\left(\frac{x-x_0}{h_n}\right) \text{ for } x \in [0,1],$$

with $\mathcal{L}' = -\mathbb{L}$, $\mathcal{L}(x) = \int_x^1 \mathbb{L}(v)dv$ and $\mathbb{L}$ as above and $\mathcal{L}(0) \neq 0$. We take here $\tau_n = n^{-\beta/(2\beta+1)}$ and $h_n = n^{-1/(2\beta+1)}$. Note that $S_{1,n}$ is the survival function associated to $\tilde{f}_{1,n}(x) = f_{0,n}(x) + c(\tau_n/h_n)\mathbb{L}((x-x_0)/h_n)$. Indeed $\tau_n/h_n = n^{-(\beta-1)/(2\beta+1)}$ is $O(1)$ for $\beta \geq 1$ so that c can be chosen small enough to have $\tilde{f}_{1,n} \geq 0$.

(i) The functions $S_{0,n}$ and $S_{1,n}$ are survival functions belonging to $\Sigma_I(\beta, R)$ with $I = (0,1)$ as $\tau_n/h_n^\beta = 1$.

(ii) $(S_{0,n}(x_0) - S_{1,n}(x_0))^2 = c^2\tau_n^2\mathcal{L}^2(0)$.

(iii) For the χ^2 distance between the observations laws, it follows the same lines as previously and yields an order $(\tau_n^2/h_n^2) \times h_n^3$, i.e. $\tau_n^2 h_n = n^{-2\beta/(2\beta+1)} \times n^{-1/(2\beta+1)} = n^{-1}$.

5.5. **Proof of Proposition 2.3.** The integrated mean-square risk is decomposed as the integrate of the squared bias plus the integrate of the variance. We integrate equation (33) and easily obtain bound (15).

Now we turn to (16). We start from (37) and get

$$\int_{\mathbb{R}_+} \mathrm{Var}(\hat{\hat{F}}_h(x)dx \leq \frac{2}{n}\int_{\mathbb{R}_+} \mathbb{E}\left[\int (K_h\left(u-x\right)\mathbf{1}_{Y_1\geq u}du\right]^2 + \frac{2}{nh}\|K\|^2\mathbb{E}\left[Y_1^2\right].$$

Now we write

$$\int_{\mathbb{R}_+} \mathbb{E}\left[\int (K_h\left(u-x\right)\mathbf{1}_{Y_1\geq u}du\right]^2 dx = \mathbb{E}\left\{\int_{\mathbb{R}_+}\left[\int (K_h\left(u-x\right)\mathbf{1}_{Y_1\geq u}\mathbf{1}_{u\geq -1}du\right]^2 dx\right\}$$

by interchanging expectation and integral and using that as K has support $[-1,1]$, $u \in [x-h, x+h] \subset [-1,+\infty)$ for $x \geq 0$ and $h \leq 1$. Therefore

$$\int_{\mathbb{R}_+} \mathbb{E}\left[\int (K_h\left(u-x\right)\mathbf{1}_{Y_1\geq u}du\right]^2 dx = \mathbb{E}\left\{\|K_h \star g_{Y_1}\|^2\right\}$$

where $g_{Y_1}(u) = \mathbf{1}_{Y_1\geq u}\mathbf{1}_{u\geq -1}$. Applying the Young Inequality (55) for $p=1$, $r=q=2$, yields $\|K_h \star g_{Y_1}\|^2 \leq \|K_h\|_1^2\|g_{Y_1}\|^2 = \|K\|_1^2(Y_1+1)$. This implies

$$\int_{\mathbb{R}_+} \mathrm{Var}(\hat{\hat{F}}_h(x)dx \leq \frac{2}{n}\|K\|_1^2\mathbb{E}(Y_1+1) + \frac{2}{nh}\|K\|^2\mathbb{E}\left[Y_1^2\right],$$

and thus Inequality (16). □

5.6. **Proof of Theorem 2.2.** We start by proving (19). By using the definitions of $A(h)$, $V(h)$ and $\hat{h}$, we note that

$$\forall h, h' \in \mathcal{H}_n, \quad \|\hat{f}_{h,h'} - f_{h'}\|^2 \leq A(h) + V(h'),$$

and

$$\forall h \in \mathcal{H}_n, \quad A(\hat{h}) + V(\hat{h}) \leq A(h) + V(h).$$

Therefore, for all $h \in \mathcal{H}_n$,

$$\begin{aligned}\|\hat{f}_{\hat{h}} - f\|^2 &\leq 3\|\hat{f}_{\hat{h}} - \hat{f}_{h,\hat{h}}\|^2 + 3\|\hat{f}_{h,\hat{h}} - \hat{f}_h\|^2 + 3\|\hat{f}_h - f\|^2\\ &\leq 3(A(h) + V(\hat{h})) + 3(A(\hat{h}) + V(h)) + 3\|\hat{f}_h - f\|^2\\ &\leq 6A(h) + 6V(h) + 3\|\hat{f}_h - f\|^2.\end{aligned}$$

The term $\mathbb{E}(\|\hat{f}_h - f\|^2)$ is ruled by Inequality (15) and we only need to study $\mathbb{E}(A(h))$. Recall that $\hat{f}_{h,h'} = K_{h'} * \hat{f}_h$, and denote $f_h(x) = \mathbb{E}[\hat{f}_h(x)]$, $f_{h,h'}(x) = \mathbb{E}[\hat{f}_{h,h'}(x)]$. We split $\hat{f}_h := \hat{f}_h^{(1)} + \hat{f}_h^{(2)}$, $f_h := f_h^{(1)} + f_h^{(2)}$ with

$$\hat{f}_h^{(1)}(x) = \frac{1}{nh}\sum_{i=1}^n [Y_iK_h'(Y_i - x) + K_h(Y_i - x)]\mathbf{1}_{|Y_i|\leq c_n}, \quad f_h^{(1)}(x) = \mathbb{E}[\hat{f}_h^{(1)}(x)],$$

and analogously for $\hat f_{h,h'}$ and $f_{h,h'}$. Then using the definition of $A(h)$ we get

$$\begin{aligned} A(h) \le\, & 5 \sup_{h'\in\mathcal{H}_n} \left\{ \|\hat f^{(1)}_{h'} - f^{(1)}_{h'}\|^2 - V(h')/10 \right\}_+ + 5 \sup_{h'\in\mathcal{H}_n} \left\{ \|\hat f^{(1)}_{h,h'} - f^{(1)}_{h,h'}\|^2 - V(h')/10 \right\}_+ \\ & + 5 \sup_{h'\in\mathcal{H}_n} \|\hat f^{(2)}_{h'} - f^{(2)}_{h'}\|^2 + 5 \sup_{h'\in\mathcal{H}_n} \|\hat f^{(2)}_{h,h'} - f^{(2)}_{h,h'}\|^2 + 5 \sup_{h'\in\mathcal{H}_n} \|f_{h'} - f_{h,h'}\|^2 \\ & := 5(T_1+T_2+T_3+T_4+T_5). \end{aligned}$$

Using (55) with $p=1, q=r=2$, and $\|K_{h'}\|_1 = \|K\|_1$, we obtain

$$T_5 = \|f_{h'} - f_{h,h'}\|^2 = \|K_{h'} \star (f - K_h \star f)\|^2 \le (\|K\|_1)^2 \|f - K_h\star f\|^2.$$

For T_1, we write

$$T_1 = \sup_{h'\in\mathcal{H}_n} \left\{ \|\hat f^{(1)}_{h'} - f^{(1)}_{h'}\|^2 - V(h')/10 \right\}_+ \le \sum_{h\in\mathcal{H}_n} \left\{ \|\hat f^{(1)}_{h} - f^{(1)}_{h}\|^2 - V(h)/10 \right\}_+,$$

and note that

$$\|\hat f^{(1)}_h - f^{(1)}_h\|^2 = \sup_{t\in\mathbb{L}_2(\mathbb{R}), \|t\|=1} \langle \hat f^{(1)}_h - f^{(1)}_h, t\rangle^2 = \sup_{t\in\mathcal{B}(1)} \langle \hat f^{(1)}_h - f^{(1)}_h, t\rangle^2, \tag{38}$$

where $\mathcal{B}(1)$ denotes a countable dense subset of $\{t\in\mathbb{L}_2(\mathbb{R}), \|t\|=1\}$.

Now we introduce the centered empirical process

$$\nu_{n,h}(\psi_t) = \langle \hat f^{(1)}_h - f^{(1)}_h, t\rangle = \frac{1}{n}\sum_{i=1}^n [\psi_t(Y_i) - \mathbb{E}(\psi_t(Y_i))],$$

$$\begin{aligned} \psi_t(y) &:= \int \left\{ \frac{y}{h^2} K'\left(\frac{y-x}{h}\right) + \frac{1}{h} K\left(\frac{y-x}{h}\right) \right\} \mathbb{1}_{|y|\le c_n} t(x)dx \\ &= \left[yK'_h \star t(y) + K_h\star t(y)\right] \mathbb{1}_{|y|\le c_n}. \end{aligned}$$

Therefore,

$$\mathbb{E}[T_1] \le \sum_{h\in\mathcal{H}_n} \mathbb{E}\left[\left\{ \sup_{t\in\mathcal{B}(1)} \nu^2_{n,h}(\psi_t) - V(h)/10 \right\}_+ \right].$$

We bound the above expectation using the Talagrand inequality (see Appendix). To apply it, we compute H, M and v. Clearly, $H^2 = V(h)/\kappa_1$ suits. Next, we get

$$\begin{aligned} \sup_{t\in\mathcal{B}(1)} \sup_{u\in\mathbb{R}} |\psi_t(u)| &\le \frac{\sqrt{2}}{h} \sup_{u\in\mathbb{R}} \left[\int \frac{c_n^2}{h^2} (K')^2 \left(\frac{u-x}{h}\right) + K^2\left(\frac{u-x}{h}\right) dx \right]^{1/2} \\ &\le \frac{\sqrt{2}}{h} \left[\frac{c_n^2}{h} \|K'\|^2 + h\|K\|^2 \right]^{1/2} \le C(K) \frac{c_n}{h^{3/2}} := M. \end{aligned}$$

Lastly, we search for v.

$$\sup_{t\in\mathcal{B}(1)} \mathrm{Var}\left(\psi_t(Y_1)\right) \le \sup_{t\in\mathcal{B}(1)} \mathbb{E}\left(\psi_t^2(Y_1)\right).$$

Remark that

$$\psi_t^2(y) = \left\{ y^2 (K'_h\star t)^2(y) + \left[y(K_h\star t)^2(y)\right]' \right\} \mathbb{1}_{|y|\le c_n}.$$

Thus by Equation (4),

$$\mathbb{E}(\psi_t^2(Y_1)) \le \mathbb{E}\left[Y_1^2 (K'_h\star t)^2(Y_1)\right] + \mathbb{E}\left[(K_h\star t)^2(X_1)\right] := S_1 + S_2.$$

Next, by Lemma 5.1, Young's Inequality (55) and as $\|t\| = 1$, we get

$$S_1 \le \mathbb{E}(|X_1|)\|K'_h \star t\|^2 \le \mathbb{E}(|X_1|)\|K'_h\|_1^2\|t\|^2 = \mathbb{E}(|X_1|)\frac{\|K'\|_1^2}{h^2}.$$

For S_2, we write, applying twice the Young Inequality for $r = +\infty$, $p = q = 2$ and $p = 1$, $q = r = 2$

$$\begin{aligned} S_2 &= \mathbb{E}[(K_h \star t)^2(X_1)] = \int (K_h \star t)^2(x) f(x)dx \le \|K_h \star t\|_\infty \|K_h \star t\|\|f\| \\ &\le \|K_h\|\|t\| \, \|K\|_1\|t\| \, \|f\| = \frac{\|K\|\|K\|_1}{\sqrt{h}}\|f\|. \end{aligned}$$

Thus we get $v = c(K, f)/h^2$ where $c(K, f) = \|K'\|_1^2\mathbb{E}(|X_1|) + \|K\|_1\|K\|\|f\|$. Then, for $\kappa_1/10 = 3$ $(\epsilon = 1/2)$, we get

$$\mathbb{E}\left[\left\{\sup_{t\in\mathcal{B}(1)} \nu_{n,h}^2(\psi_t) - V(h)/10\right\}_+\right] \le \frac{C_1}{n}\left(\frac{1}{h^2}\exp(-C_2/h) + \frac{c_n^2}{nh^3}\exp\left(-C_3\frac{\sqrt{n}}{c_n}\right)\right).$$

By the definition of $\mathcal{H}_n$, we have $1/(nh^3) \le 1$, $\sum_{h\in\mathcal{H}_n} h^{-2}\exp(-C_2/h) < \Sigma(C_2) < \infty$, and $\text{Card}(\mathcal{H}_n) \le n$. So, choosing

$$c_n = C_3\sqrt{n}/(4\log(n)),$$

we obtain $\mathbb{E}[T_1] \le c/n$. The term T_2 is studied analogously, with additional factors $\|K\|_1^2$ due to an additional application of Young's Inequality.

For the terms T_3, T_4, rough bounds are used together with the definition of $\mathcal{H}_n$, in particular $1/(nh^3) \le 1$ to get $T_3 \le C(K)n\mathbb{E}(|Y_1|^{2+p}/c_n^p)$ for all $p > 0$, where $C(K)$ is a constant depending on the kernel. Thus, with the definition of c_n we obtain an order $1/n$ by choosing $p = 6$ with constraint $\mathbb{E}(|Y_1|^8) < +\infty$. Hence we get (19).

Now we turn to the proof of (20). The study follows the same line as previously, so we mainly give a sketch of proof. Here we can split in three parts $\hat{\bar{F}}_h = \hat{\bar{F}}_h^{(1)} + \hat{\bar{F}}_h^{(2)} + \hat{\bar{F}}_h^{(3)}$ with

$$\hat{\bar{F}}_h^{(1)}(x) = \frac{1}{nh}\sum_{i=1}^n Y_i\mathbf{1}_{Y_i<c_n}K\left(\frac{Y_i - x}{h}\right), \quad \hat{\bar{F}}_h^{(2)}(x) = \frac{1}{nh}\sum_{i=1}^n Y_i\mathbf{1}_{Y_i\ge c_n}K\left(\frac{Y_i - x}{h}\right),$$

$$\hat{\bar{F}}_h^{(3)}(x) = \int K_h(u - x)\hat{\bar{F}}_Y(u)du,$$

with $\bar{F}_h^{(i)} = \mathbb{E}[\hat{\bar{F}}_h^{(i)}]$ for $i = 1, 2, 3$, and analogously for $\hat{\bar{F}}_{h,h'}^{(i)}$, $i = 1, 2, 3$.

The first two terms are studied as previously T_1, T_2, T_3, T_4. There is also a term analogous to T_5. Let $G_Y(u) = (\hat{\bar{F}}_Y(u) - \bar{F}_Y(u))\mathbf{1}_{-1\le u}$. The additional new terms are

$$T_6 := \mathbb{E}\left(\sup_{h'\in\mathcal{H}_n}\int_0^{+\infty}[\hat{\bar{F}}_{h'}^{(3)}(x) - \bar{F}_{h'}^{(3)}(x)]^2dx\right) = \mathbb{E}\left(\sup_{h'\in\mathcal{H}_n}\int_{\mathbb{R}_+}[K_{h'} \star G_Y(x)]^2dx\right)$$

and its twin in h, h'. Thus using Inequality (55) as previously, we get

$$\begin{aligned} T_6 &\le \mathbb{E}\left(\sup_{h'\in\mathcal{H}_n}\|K_{h'}\|_1^2\|G_Y\|^2\right) \le \|K\|_1^2\mathbb{E}\left(\int(\hat{\bar{F}}_Y(u) - \bar{F}_Y(u))^2\mathbf{1}_{-1\le u}du\right) \\ &= \frac{\|K\|_1^2}{n}\int \text{Var}(\mathbf{1}_{Y_1\ge u})\mathbf{1}_{u\ge -1})du \le \frac{\|K\|_1^2\mathbb{E}(Y_1 + 1)}{n}. \end{aligned}$$

This ends the proof. □

5.7. **Proof of Proposition 3.1.** For sake of simplicity, we assume that $L = 1$ and $k^{(1)} = k$. By **(B1)**,

$$\frac{1}{x}\int_0^{+\infty} k_m(\frac{u}{x})\mathbb{1}_{Y_i \geq u} du = \int_0^{Y_i/x} k_m(v)dv \leq \int_0^{+\infty} k_m(v)dv = 1,$$

so that $\tilde{\tilde{F}}_m$ is well defined. Moreover, it is obvious from the formula above that

$$\lim_{x\to+\infty} \frac{1}{x}\int_0^{+\infty} k_m(\frac{u}{x})\hat{\hat{F}}_Y(u)du = 0, \quad \lim_{x\to 0^+} \frac{1}{x}\int_0^{+\infty} k_m(\frac{u}{x})\hat{\hat{F}}_Y(u)du = 1.$$

From Young's Inequality (see (55) with $r = +\infty$, $p = q = 2$) and **(B1)**, $\|k \star k\|_\infty \leq \|k\|^2$ so that for all $m \geq 2$, $\|k_m\|_\infty < \infty$. Consequently,

$$\lim_{x\to+\infty} \frac{1}{x} k_m(\frac{Y_i}{x}) = 0.$$

As $uk(u) \to 0$ when $u \to +\infty$, by induction we easily prove that $uk_m(u) \to 0$ when $u \to +\infty$. Therefore,

$$\lim_{x\to 0} \frac{1}{x} k_m(\frac{Y_i}{x}) = 0.$$

In summary, we proved that, under **(B1)**, $\lim_{x\to+\infty} \tilde{\tilde{F}}_m(x) = 0$ and $\lim_{x\to 0^+} \tilde{\tilde{F}}_m(x) = 1$. Without loss of generality, we assume that **(B3)** holds for $m_0 = 1$ and write that

$$k(x) = \frac{1}{2\pi}\int_{\mathbb{R}} e^{-itx}k^*(t)dt, \quad k'(x) = \frac{-i}{2\pi}\int_{\mathbb{R}} e^{-itx}tk^*(t)dt.$$

This implies that k and k' are continuous, tend to zero at $+\infty$, and $k(0) = k'(0) = 0$. As $k_m(y) = \int_0^y k(y-z)k_{m-1}(z)dz$ for $m > 1$, we have $k_m(0) = 0$, $\lim_{y\to+\infty} k_m(y) = 0$ and k_m is continuously derivable, with

$$(k_m')^*(t) = -itk_m^*(t). \tag{39}$$

Using **(B2)** and $\int_0^{+\infty} k_m'(v)dv = [k_m(v)]_0^{+\infty} = 0$ yields

$$\int_0^{+\infty} \tilde{f}_m(x)dx = \int_0^{+\infty} k_m(v)\frac{dv}{v} + \int_0^{+\infty} k_m'(v)dv = 1 + O(\frac{1}{m}).$$

5.8. **Proof of Proposition 3.2.**

Lemma 5.2. *Let $k, k^{(1)}, k^{(2)}$ satisfying* **(B1)**, *then for v, v_1, v_2 their variances(i.e. $v = \int_0^{+\infty}(u-1)^2k(u)du$), we have*

$$\|k_m\|^2 = \int_0^{+\infty} k_m^2(u)du = \sqrt{m}(1/2\sqrt{\pi v}(1+o(1)), \quad \|k_m\|_\infty \leq \sqrt{m}(1/\sqrt{2\pi v}(1+o(1)), \tag{40}$$

$$\langle k_m^{(1)}, k_m^{(2)}\rangle = \int_0^{+\infty} k_m^{(1)}(u)k_m^{(2)}(u)du = \sqrt{m}(1/\sqrt{2\pi(v_1+v_2)}(1+o(1)). \tag{41}$$

Let $k, k^{(1)}, k^{(2)}$ satisfying **(B1)** *and* **(B3)**, *then*

$$\|k_m'\|^2 = \frac{1}{4\sqrt{\pi}}\left(\frac{m}{v}\right)^{3/2}(1+o(1)), \quad \langle (k_m^{(1)})', (k_m^{(2)})'\rangle = \frac{1}{\sqrt{2\pi}}\left(\frac{m}{v_1+v_2}\right)^{3/2}(1+o(1)). \tag{42}$$

Proof of Lemma 5.2. Equalities (40) and (41) are proved in Lemma A.1. of Comte and Genon-Catalot (2012).

Thus we turn to (42). Under **(B1)** and **(B3)** with $m_0 = 1$, as $(k'_m)^*(t) = -itk^*_m(t)$,

$$\|k'_m\|^2 = \int_0^{+\infty} (k'_m(y))^2 dy = \frac{1}{2\pi}\int_{\mathbb{R}} t^2 |k^*_m(t)|^2 dt = \frac{(\sqrt{m})^3}{2\pi}\int_{\mathbb{R}} s^2 |k^*(\frac{s}{\sqrt{m}})|^{2m} ds.$$

Under the assumption $\int t^2|k^*(t)|^2 dt < +\infty$ (see **(B3)**), we can mimick the proof of Lemma A.1. (Comte and Genon-Catalot (2012)) to obtain:

$$\frac{1}{(\sqrt{m})^3}\int_{\mathbb{R}} t^2 |k^*_m(t)^2 dt \to \int s^2 e^{-vs^2} ds = \frac{\sqrt{\pi}}{2v^{3/2}}.$$

And for the case of two densities,

$$\langle (k_m^{(1)})', (k_m^{(2)})'\rangle = \int_0^{+\infty} (k_m^{(1)})'(y)(k_m^{(2)})'(y) dy \sim \frac{(\sqrt{m})^3}{2\pi}\int_{\mathbb{R}} s^2 e^{-(v_1+v_2)s^2/2} ds.$$

Hence (42). □

Now we turn to the proof of Proposition 3.2. For the bias order, we have

$$\begin{aligned}
\mathbb{E}(\tilde{\bar{F}}_m(x)) &= \frac{1}{x}\int_0^{+\infty} K_m(\frac{u}{x})\bar{F}_Y(u)du + \frac{1}{x}\mathbb{E}(Y_1 K_m(\frac{Y_1}{x}))\\
&= \frac{1}{x}\int_0^{+\infty} K_m(\frac{u}{x})\left(\bar{F}_Y(u) + uf_Y(u)\right)du = \frac{1}{x}\int_0^{+\infty} K_m(\frac{u}{x})\bar{F}(u)du\\
&= \int_0^{+\infty} K_m(v)\bar{F}(vx)dv.
\end{aligned}$$

The order of this term on $\Sigma_I(\beta, C)$, is given in Proposition 3.2 in Comte and Genon-Catalot (2012).

Now, we bound the variance term.

$$\begin{aligned}
\mathrm{Var}(\tilde{\bar{F}}_m(x)) &= \frac{1}{nx^2}\mathrm{Var}\left[\int_0^{+\infty} K_m(\frac{u}{x})\mathbf{1}_{Y_1\geq u}du + Y_1 K_m(\frac{Y_1}{x})\right]\\
&\leq \frac{2}{nx^2}\left(\mathbb{E}[(\int_0^{+\infty} K_m(\frac{u}{x})\mathbf{1}_{Y_1\geq u}du)^2] + \mathbb{E}[Y_1^2 K_m^2(\frac{Y_1}{x})]\right) := T_1(x) + T_2(x) \qquad (43)
\end{aligned}$$

$$T_1(x) \leq \frac{2}{nx^2}\left(\int_0^{+\infty} |K_m(\frac{u}{x})|du\right)^2 \leq \frac{2}{nx^2}\left(\sum_{j=1}^L |\alpha_j|\int_0^{+\infty} k_m^{(j)}(v)xdv\right)^2 = \frac{2}{n}|\alpha|_1^2,$$

$$\begin{aligned}
T_2(x) &= \frac{2}{nx^2}\iint ((uv)^2 K_m^2(\frac{uv}{x})\mathbf{1}_{[0,1]}(u)f(v)\mathbf{1}_{\mathbb{R}^+}(v)dudv = \frac{2}{nx^2}\int_0^{+\infty} v^2 f(v)\left(\int_0^1 u^2 K_m^2(\frac{uv}{x})du\right)dv\\
&= \frac{2}{nx^2}\int_0^{+\infty}\frac{f(v)}{v}x^3\left(\int_0^{v/x} z^2 K_m^2(z)dz\right)dv.
\end{aligned}$$

Thus using Lemma 5.2 namely $\int_0^{+\infty} z k_m^{(j)}(z)dz = 1$ and $\|k_m^{(j)}\|_\infty \leq 2\sqrt{m}/\sqrt{2\pi v_j})$, we get

$$\begin{aligned} T_2(x) &\leq \frac{2}{n}\int_0^{+\infty} f(v)\left(\int_0^{v/x} zK_m^2(z)dz\right)dv \leq \frac{2}{n}\int_0^{+\infty} zK_m^2(z)dz \\ &\leq \frac{2}{n}\|K_m\|_\infty \int_0^{+\infty} z|K_m(z)|dz \leq \frac{2}{n}\sum_{j=1}^L \frac{2|\alpha_j|\sqrt{m}}{\sqrt{2\pi v_j}}\sum_{j=1}^L |\alpha_j| \int zk_m^{(j)}(z)dz \\ &\leq 4|\alpha|_1 \frac{\sqrt{m}}{n}\sum_{j=1}^L \frac{|\alpha_j|}{\sqrt{2\pi v_j}}. \end{aligned}$$

Consequently we have

$$\mathrm{Var}(\tilde{\tilde{F}}_m(x)) \leq |\alpha|_1 \frac{2}{n}\left(|\alpha|_1 + 2\sqrt{m}\sum_{j=1}^L \frac{|\alpha_j|}{\sqrt{2\pi v_j}}\right).$$

Next, we study the estimator of f. As $(\partial/\partial y)(yK_m(\frac{y}{x})) = K_m(\frac{y}{x}) + (y/x)K_m'(\frac{y}{x})$, we have

$$\mathbb{E}\tilde{f}_m(x) = \frac{1}{x}\mathbb{E}K_m(\frac{X_1}{x}) = \int_0^{+\infty} K_m(v)f(xv)dv.$$

As for the bias term of $\tilde{\tilde{F}}_m(x)$, the study of the bias term of $\tilde{f}_m(x)$ is a direct application of Proposition 3.2 of Comte and Genon-Catalot (2012). For the variance term, we use that

$$\begin{aligned} \mathrm{Var}\tilde{f}_m(x) &\leq \frac{1}{nx^2}\mathbb{E}\left(K_m(\frac{Y_1}{x}) + \frac{Y_1}{x}K_m'(\frac{Y_1}{x})\right)^2 \\ &= \frac{1}{nx^2}\left[\mathbb{E}K_m^2(\frac{X_1}{x}) + \mathbb{E}\left(\frac{Y_1}{x}K_m'(\frac{Y_1}{x})\right)^2\right] \end{aligned}$$

We have

$$\mathbb{E}K_m^2(\frac{X_1}{x}) = x\int_0^{+\infty} K_m^2(v)f(xv)dv \leq x\|f\|_\infty \|K_m\|^2$$

where, by Lemma 5.2, the L^2-norm of K_m satisfies, using (30),

$$\|K_m\|^2 \leq \sqrt{m}\sum_{1\leq i,j\leq L} \frac{2\ |\alpha_i\alpha_j|}{\sqrt{2\pi(v_i+v_j)}} = C_2'\sqrt{m} \leq C(K)\sqrt{m}. \tag{44}$$

For the other term, we have

$$\begin{aligned} \mathbb{E}\left(\frac{Y_1}{x}K_m'(\frac{Y_i}{x})\right)^2 &= \int_{v\geq 0, 0\leq u\leq 1} \frac{(uv)^2}{x^2}(K_m'(\frac{uv}{x}))^2\ f(v)dudv \\ &= x\int_0^{+\infty}\frac{f(v)}{v}dv\left(\int_0^{\frac{v}{x}} t^2(K_m'(t))^2dt\right) \leq \frac{1}{x}\mathbb{E}(X_1)\int_0^{+\infty}(K_m'(t))^2dt. \end{aligned}$$

Now, we use (42) of Lemma 5.2

$$\|K_m'\|^2 \leq \frac{m^{3/2}}{\sqrt{2\pi}}\sum_{1\leq i,j\leq L}\frac{2|\alpha_i\alpha_j|}{(v_i+v_j)^{3/2}}.$$

The result follows. This ends the proof of Proposition 3.2. □

5.9. **Proof of Proposition 3.3.** Inequality (43) for $\mathrm{Var}(\tilde{\tilde{F}}_m(x))$ must be integrated over $\mathbb{R}^+$. The second term is the easiest:

$$\int_0^{+\infty} T_2(x)\,dx \le \frac{2}{n}\mathbb{E}(Y_1^2\int_0^{+\infty} K_m^2(\frac{Y_1}{x})\frac{dx}{x^2}) = \frac{2}{n}\,\mathbb{E}(Y_1)\,\|K_m\|^2. \tag{45}$$

For the term $T_1(x)$, we apply the generalized Minkowski inequality (see (54) in appendix):

$$\begin{aligned}\int_0^{+\infty} T_1(x)\,dx &\le \frac{2}{n}\mathbb{E}\int \mathbf{I}_{x\ge 0}\frac{dx}{x^2}\left(\int K_m(\frac{u}{x})\mathbf{I}_{Y_1\ge u}\mathbf{I}_{u\ge 0}du\right)^2\\ &\le \frac{2}{n}\mathbb{E}\left[\int \mathbf{I}_{Y_1\ge u}\mathbf{I}_{u\ge 0}du\left(\int K_m^2(\frac{u}{x})\mathbf{I}_{x\ge 0}\frac{dx}{x^2}\right)^{1/2}\right]^2\\ &\le \frac{2}{n}\mathbb{E}\left[\int \mathbf{I}_{Y_1\ge u\ge 0}\frac{1}{\sqrt{u}}du\left(\int K_m^2(v)\mathbf{I}_{v\ge 0}dv\right)^{1/2}\right]^2\\ &= \frac{2}{n}\mathbb{E}(2\sqrt{Y_1})^2)\int_0^{+\infty} K_m^2(v)dv = \frac{8}{n}\mathbb{E}Y_1\|K_m\|^2. \end{aligned}\tag{46}$$

Finally,

$$\int_0^{+\infty}\mathrm{Var}(\tilde{\tilde{F}}_m(x))dx \;\le\; \frac{10}{n}\mathbb{E}Y_1\|K_m\|^2$$

which is the announced result using (44). □

5.10. **Proof of Theorem 3.1.**

5.10.1. *Some preliminary Lemmas.* Let us set, for $m, m' > 0$,

$$B_m\bar{F}(x) = \mathbb{E}\tilde{\tilde{\bar{F}}}_m(x) - \bar{F}(x), \quad B_{m,m'}\bar{F}(x) = \mathbb{E}\tilde{\tilde{\bar{F}}}_{m,m'}(x) - \bar{F}(x). \tag{47}$$

Similarly to Lemma A.3 of Comte and Genon-Catalot (2012), the following relation between bias terms holds.

Lemma 5.3. *We have* $B_m\bar{F}(x) = \int_0^{+\infty} K_m(u)\bar{F}(xu)du - \bar{F}(x)$ *and*

$$B_{m,m'}\bar{F}(x) = B_{m'}\bar{F}(x) + \int_0^{+\infty} K_{m'}(u)B_m\bar{F}(xu)du.$$

We also state a result with useful bounds concerning the convolution power kernels.

Lemma 5.4. *Recall notations (25). Under assumptions* **(B1)**-**(B2)***, we have*
(o) $||K_m\odot K_{m'}||_\infty \le |\alpha|_1\sum_{j=1}^L \frac{|\alpha_j|}{\sqrt{2\pi v_j}}\sqrt{m\wedge m'}(1+o(1))$.
(i) $\|K_m\odot K_{m'}\|^2 \le C(K)\sqrt{m\wedge m'}$ *where $C(K)$ is defined by (28).*
(ii) $\int_0^{+\infty}(|K_m(z)|/\sqrt{z})dz \le 3|\alpha|_1$.
(iii) $\int_0^{+\infty}(|K_m\odot K_{m'}(z)|/\sqrt{z})dz \le 3|\alpha|_1^2$.

Proof of Lemma 5.4. For (o), see Lemma A4 of Comte and Genon-Catalot (2012).
For (i), we write

$$\int (K_m\odot K_{m'})^2(u)du \le \|K_m\odot K_{m'}\|_\infty\int_0^{+\infty}|K_m\odot K_{m'}|(u)du. \tag{48}$$

Now we know from (o) that $\|K_m \odot K_{m'}\|_\infty \le 2|\alpha|_1 \sum_{j=1}^L \frac{|\alpha_j|}{\sqrt{2\pi v_j}}\sqrt{m \wedge m'}$. Moreover

$$\begin{aligned}\int_0^{+\infty} |K_m \odot K_{m'}|(u)du &\le \int_0^{+\infty}\left(\int_0^{+\infty} |K_m(\frac{u}{v})||K_{m'}(v)|\frac{dv}{v}\right)du \\ &= \int_0^{+\infty} |K_{m'}(v)|dv \int_0^{+\infty} |K_m(z)|dz \le |\alpha|_1^2.\end{aligned}$$

Plugging these two bounds in (48) gives the first result.
For (ii), we simply split the integral

$$\begin{aligned}\int_0^{+\infty} \frac{|K_m(z)|}{\sqrt{z}}dz &= \int_0^1 \frac{|K_m(z)|}{\sqrt{z}}dz + \int_1^{+\infty} \frac{|K_m(z)|}{\sqrt{z}}dz \\ &\le \int_0^1 \frac{|K_m(z)|}{z}dz + \int_1^{+\infty} |K_m(z)|dz.\end{aligned}$$

By **(B2)**, $\int_0^{+\infty} z^{-1}|K_m(z)|dz \le 2|\alpha|_1$ and $\int_0^{+\infty} |K_m(z)|dz \le |\alpha|_1$. This ends the proof of (ii).
For (iii), we write

$$\int_0^{+\infty} \frac{K_m \odot K_{m'}(s)}{\sqrt{s}}ds = \int_0^{+\infty}\left(\int_0^{+\infty} K_{m'}(\frac{u}{s})\frac{ds}{s^{1/2}}\right)K_m(u)\frac{du}{u}$$

and this term is simply equal to $\int_0^{+\infty} K_m(u)/\sqrt{u}du \int K_{m'}(v)/vdv$, so that the result (iii) follows by applying (ii). Hence Lemma 5.4. □

Recall that $C_2' \le C(K)$. Thus, a straightforward consequence of Lemma 5.4 (i) and the bounds in Proposition 3.3 is the following Lemma.

Lemma 5.5. *Under assumptions* **(B1)**-**(B2)**, *we have*

$$\int_0^{+\infty} Var\tilde{\tilde{F}}_{m,m'}(x)\, dx \le 10\, \mathbb{E}(Y_1)\frac{\sqrt{m \wedge m'}}{n}C(K),$$

where the constant $C(K)$ (see (28)) does not depend on the density f.

5.10.2. *Proof of Theorem 3.1.* First note that the definition of $\hat{m}$ implies that $H(\tilde{m}) + Z(\tilde{m}) \le H(m) + Z(m)$ for all $m \in \mathcal{M}_n$. From now on, we extend all functions by setting them equal to 0 on $(-\infty, 0)$ so that $\|.\|$ is the L^2-norm on $\mathbb{R}^+$. Hence, for m any element of $\mathcal{M}_n$, we can write the decomposition

$$\begin{aligned}\|\tilde{\tilde{F}}_{\tilde{m}} - \bar{F}\|^2 &\le 3(\|\tilde{\tilde{F}}_{\tilde{m}} - \tilde{\tilde{F}}_{m,\tilde{m}}\|^2 + \|\tilde{\tilde{F}}_{m,\tilde{m}} - \tilde{\tilde{F}}_m\|^2 + \|\tilde{\tilde{F}}_m - \bar{F}\|^2) \\ &\le 3(H(m) + Z(\tilde{m})) + 3(H(\tilde{m}) + Z(m)) + 3\|\tilde{\tilde{F}}_m - \bar{F}\|^2 \\ &\le 6(H(m) + Z(m)) + 3\|\tilde{\tilde{F}}_m - \bar{F}\|^2.\end{aligned}$$

Therefore, $\mathbb{E}(\|\tilde{\tilde{F}}_{\tilde{m}} - \bar{F}\|^2) \le 3\mathbb{E}(\|\tilde{\tilde{F}}_m - \bar{F}\|^2) + 6Z(m) + 6\mathbb{E}(H(m))$. Let us study $H(m)$ (see (29)). Let $\mathbb{E}(\tilde{\tilde{F}}_m(x)) = \bar{F}_m(x)$ and $\mathbb{E}(\tilde{\tilde{F}}_{m,m'}(x)) = \bar{F}_{m,m'}(x)$. Then

$$\|\tilde{\tilde{F}}_{m'} - \tilde{\tilde{F}}_{m,m'}\|^2 \le 3\|\tilde{\tilde{F}}_{m'} - \bar{F}_{m'}\|^2 + 3\|\tilde{\tilde{F}}_{m,m'} - \bar{F}_{m,m'}\|^2 + 3\|\bar{F}_{m'} - \bar{F}_{m,m'}\|^2.$$

By Lemma 5.3, for all $m, m' \in \mathcal{M}_n$,

$$\|\bar{F}_{m'} - \bar{F}_{m,m'}\|^2 = \int_0^{+\infty}\left(\int_0^{+\infty} B_m\bar{F}(xu)K_{m'}(u)du\right)^2 dx.$$

Therefore, using that each $k_{m'}^{(i)}$ is a density, we obtain:

$$\begin{aligned}\|\bar{F}_{m'}-\bar{F}_{m,m'}\|^2 &\leq |\alpha|_1\int_0^{+\infty}\sum_{i=1}^L|\alpha_i|\left(\int_0^{+\infty}(B_m\bar{F})(xu)k_{m'}^{(i)}(u)du\right)^2dx\\ &\leq |\alpha|_1\sum_{i=1}^L|\alpha_i|\int\int(B_m\bar{F})^2(xu)k_{m'}^{(i)}(u)dudx\\ &\leq |\alpha|_1\int_0^{+\infty}(B_m\bar{F})^2(v)dv\sum_{i=1}^L|\alpha_i|\int_0^{+\infty}\frac{k_{m'}^{(i)}(u)}{u}du,\end{aligned}$$

having used Fubini and the change of variable $v=xu$. Now, $\int k_{m'}^{(i)}(u)/u\,du=1+O(1/m')\leq 2$. Therefore

$$\begin{aligned}H(m) &\leq 3\sup_{m'}\left(\|\tilde{\bar{F}}_{m'}-\bar{F}_{m'}\|^2-\frac{Z(m')}{6}\right)_+ +3\sup_{m'}\left(\|\tilde{\bar{F}}_{m,m'}-\bar{F}_{m,m'}\|^2-\frac{Z(m')}{6}\right)_+\\ &\quad +3\sup_{m'}\|\bar{F}_{m'}-\bar{F}_{m,m'}\|^2\\ &\leq 3\sup_{m'}\left(\|\tilde{\bar{F}}_{m'}-\bar{F}_{m'}\|^2-\frac{Z(m')}{6}\right)_+ +3\sup_{m'}\left(\|\tilde{\bar{F}}_{m,m'}-\bar{F}_{m,m'}\|^2-\frac{Z(m')}{6}\right)_+\\ &\quad +6|\alpha|_1^2\int_0^{+\infty}(B_m\bar{F})^2(v)dv.\end{aligned}$$

Now, we can prove the following Lemmas:

Lemma 5.6. *Under the assumptions of Theorem 3.1, we have*

$$\mathbb{E}\left(\sup_{m'}\left(\|\tilde{\bar{F}}_{m'}-\bar{F}_{m'}\|^2-\frac{Z(m')}{6}\right)_+\right)\leq\frac{C}{n}.$$

Lemma 5.7. *Under the assumptions of Theorem 3.1, we have*

$$\mathbb{E}\left(\sup_{m'}\left(\|\tilde{\bar{F}}_{m,m'}-\bar{F}_{m,m'}\|^2-\frac{Z(m')}{6}\right)_+\right)\leq\frac{C}{n}.$$

This yields that, $\forall m\in\mathcal{M}_n$,

$$\mathbb{E}(\|\tilde{\bar{F}}_{\tilde{m}}-\bar{F}\|^2)\leq 3\mathbb{E}(\|\tilde{\bar{F}}_m-\bar{F}\|^2)+6Z(m)+6|\alpha|_1^2\int(B_m\bar{F})^2(v)dv+\frac{6C}{n}.$$

As $\mathbb{E}(\|\tilde{\bar{F}}_m-\bar{F}\|^2)\leq C(Z(m)+\int(B_m\bar{F})^2(v)dv)$, the proof of Theorem 3.1 is complete. □

5.10.3. *Proof of Lemma 5.6.* First we write,

$$\mathbb{E}\left(\sup_{m'}\left(\|\tilde{\bar{F}}_{m'}-\bar{F}_{m'}\|^2-\frac{Z(m')}{6}\right)_+\right)\leq\sum_{m\in\mathcal{M}_n}\mathbb{E}\left(\left(\|\tilde{\bar{F}}_m-\bar{F}_m\|^2-\frac{Z(m)}{6}\right)_+\right). \tag{49}$$

Next, we split the estimator and its expectation in two parts,

$$\tilde{\bar{F}}_m-\bar{F}_m=(\tilde{\bar{F}}_m^{(1)}-\bar{F}_m^{(1)})+(\tilde{\bar{F}}_m^{(2)}-\bar{F}_m^{(2)})$$

where

$$\tilde{\bar{F}}_m^{(1)}(x)=\frac{1}{n}\sum_{i=1}^n\frac{1}{x}\int_0^{+\infty}K_m\left(\frac{u}{x}\right)\mathbb{I}_{Y_i\geq u}\mathbb{I}_{Y_i\leq c_n}(u)du+\frac{1}{nx}Y_i\mathbb{I}_{Y_i\leq c_n}K_m\left(\frac{Y_i}{x}\right),$$

$\tilde{\bar{F}}_m^{(2)} = \tilde{\bar{F}}_m - \tilde{\bar{F}}_m^{(1)}$, $\bar{F}_m^{(k)} = \mathbb{E}(\tilde{\bar{F}}_m^{(k)})$ for $k = 1, 2$ and for $\mathtt{a}$ a numerical constant

$$c_n = \frac{n\mathbb{E}(Y_1)}{\mathtt{a}\log^2(n)}. \tag{50}$$

We get

$$\mathbb{E}\left(\left(\|\tilde{\bar{F}}_m - \bar{F}_m\|^2 - \frac{Z(m)}{6}\right)_+\right) \leq 2\mathbb{E}\left(\left(\|\tilde{\bar{F}}_m^{(1)} - \bar{F}_m^{(1)}\|^2 - \frac{Z(m)}{12}\right)_+\right) + 2\mathbb{E}\left(\|\tilde{\bar{F}}_m^{(2)} - \bar{F}_m^{(2)}\|^2\right)$$

and from the variance bound, for $\sqrt{m} \leq n$,

$$\mathbb{E}\left(\|\tilde{\bar{F}}_m^{(2)} - \bar{F}_m^{(2)}\|^2\right) \leq \frac{C\mathbb{E}(Y_1 \mathbf{1}_{Y_1 > c_n})\sqrt{m}}{n} \leq C\frac{\mathbb{E}(Y_1^{p+1})}{c_n^p}.$$

With c_n given by (50), $p = 3$ and $\mathrm{card}\mathcal{M}_n \leq n/\log(n)$, we get

$$\sum_{m\in\mathcal{M}_n} \mathbb{E}\left(\|\tilde{\bar{F}}_m^{(2)} - \bar{F}_m^{(2)}\|^2\right) \leq C\mathtt{a}^3 \frac{\mathbb{E}(Y_1^4)}{(\mathbb{E}Y_1)^3}\frac{\log^5(n)}{n^2} \leq \frac{C'}{n},$$

provided that $\mathbb{E}(Y_1^4) < +\infty$, which makes this term negligible.

Next, we note that $\|\tilde{\bar{F}}_m^{(1)} - \bar{F}_m^{(1)}\|^2 = \sup_{t,\|t\|=1}\langle \tilde{\bar{F}}_m^{(1)} - \bar{F}_m^{(1)}, t\rangle^2$, and the supremum can be taken over a dense countable family of functions t such that $\|t\| = 1$; we denote by $\mathcal{B}(1)$ this set.

Thus, setting

$$\theta_t(y) = \int_0^{+\infty} \frac{1}{x}\left[\int_0^{+\infty} K_m(\frac{u}{x})(\mathbf{1}_{y\geq u}\mathbf{1}_{y\leq c_n}du + yK_m(\frac{y}{x})\mathbf{1}_{y\leq c_n}\right] t(x)dx := \theta_t^{(1)}(y) + \theta_t^{(2)}(y)$$

with obvious splitting into two terms, we introduce the centered empirical process

$$\nu_n(\theta_t) = \langle \tilde{\bar{F}}_m^{(1)} - \bar{F}_m^{(1)}, t\rangle = \frac{1}{n}\sum_{i=1}^n \left[\theta_t(Y_i) - \mathbb{E}\theta_t(Y_i)\right]. \tag{51}$$

We can apply the Talagrand inequality (see Appendix). For this, we search for H, v, M such that:

$$\mathbb{E}(\sup_{t\in\mathcal{B}(1)} \nu_n^2(\theta_t)) \leq H^2, \ \sup_{t\in\mathcal{B}(1)} \mathrm{Var}\left(\theta_t(Y_1)\right) \leq v \text{ and } \sup_{t\in\mathcal{B}(1)}\sup_y |\theta_t(y)| \leq M.$$

It follows from the definition of ν_n and Proposition 3.3 that

$$\mathbb{E}(\sup_{t\in\mathcal{B}(1)} \nu_n^2(\theta_t)) \leq \mathbb{E}(\|\tilde{\bar{F}}_m - \bar{F}_m\|^2) \leq 10C(K)\mathbb{E}(Y_1)\sqrt{m}/n := H^2$$

where $C(K)$ is defined in (28). Next, for $\|t\| = 1$, we have by (44)

$$\begin{aligned}
\theta_t^{(1)}(y) &= \int_0^y \left(\int_0^{+\infty} \frac{1}{x}K_m(\frac{u}{x})t(x)dx\right) du\mathbf{1}_{y\leq c_n} \leq \int_0^{c_n}\left(\int_0^{+\infty}\frac{1}{x^2}K_m^2(\frac{u}{x})dx\right)^{1/2} du \\
&= \int_0^{c_n}\frac{1}{\sqrt{u}}\left(\int_0^{+\infty} K_m^2(v)dv\right)^{1/2} du = 2\|K_m\|\sqrt{c_n} \leq 2\sqrt{C(K)}m^{1/4}\sqrt{c_n},
\end{aligned}$$

and

$$\begin{aligned}
|\theta_t^{(1)}(y)| &\leq \left(\int_0^{\infty}(\frac{y}{x}K_m(\frac{y}{x}))^2\mathbf{1}_{y\leq c_n}dx\right)^{1/2} \\
&= \left(\mathbf{1}_{y\leq c_n}\int_0^{\infty}(K_m(u))^2 ydu\right)^{1/2} \leq \sqrt{c_n}\|K_m\| \leq \sqrt{C(K)}m^{1/4}\sqrt{c_n}
\end{aligned}$$

Hence, we can take $M = 3\sqrt{C(K)}\sqrt{c_n}m^{1/4}$. Now, we study

$$\begin{aligned}\mathrm{Var}(\theta_t^{(1)}(Y_1)) &\leq \mathbb{E}\left[\left(\theta_t^{(1)}(Y_1)\right)^2\right] \\ &\leq \int_{[0,+\infty[^5} \frac{1}{x}K_m(\frac{u}{x})\mathbf{1}_{y\geq u}t(x)\frac{1}{z}K_m(\frac{v}{z})\mathbf{1}_{y\geq v}t(z)f_Y(y)dudvdxdydz \\ &\leq \int_{[0,+\infty[^5} K_m(s)\mathbf{1}_{y\geq sx}t(x)K_m(w)\mathbf{1}_{y\geq wz}t(z)f_Y(y)dsdwdxdydz\end{aligned}$$

Writing that $\int_0^{+\infty} t(x)\mathbf{1}_{y\geq sx}dx \leq \|t\|(\int_0^{+\infty}\mathbf{1}_{y\geq sx}dx)^{1/2} = \sqrt{y/s}$, we get

$$\begin{aligned}\mathrm{Var}(\theta_t^{(1)}(Y_1)) &\leq \int_{[0,+\infty[^3} |K_m(s)|\sqrt{\frac{y}{s}}|K_m(w)|\sqrt{\frac{y}{w}}f_Y(y)dsdwdy \\ &\leq \mathbb{E}(Y_1)\left(\int_0^{+\infty}\frac{|K_m(s)|}{\sqrt{s}}ds\right)^2 = 9|\alpha|_1^2\mathbb{E}(Y_1),\end{aligned}$$

by (ii) of Lemma 5.4. Therefore, $\mathrm{Var}(\theta_t^{(1)}(Y_1))$ is bounded by a constant independent of m, n. Next we consider $\mathrm{Var}(\theta_t^{(2)}(Y_1))$.

$$\mathrm{Var}(\theta_t^{(2)}(Y_1)) \leq \int_{(0,+\infty)^3} f_Y(y)(y/x)K_m(y/x)t(x)(y/z)K_m(y/z)t(z)dxdydz.$$

First, $\int(y^2/xz)f_Y(y)K_m(y/x)K_m(y/z)dy = (x^2/z)\int u^2K_m(u)K_m(xu/z)f_Y(xu)du$. Hence,

$$\mathrm{Var}(\theta_t^{(2)}(Y_1)) \leq \int_0^{+\infty} u^2|K_m(u)|\left(\int_0^{+\infty} x^2|t(x)|f_Y(xu)(\int_0^{+\infty}|t(z)K_m(xu/z)|\frac{dz}{z})dx\right)du. \tag{52}$$

Next, with $v = xu/z$, we get

$$\int_0^{+\infty}|t(z)K_m(xu/z)|\frac{dz}{z} \leq \left[\int_0^{+\infty}(K_m(xu/z)(1/z))^2dz\right]^{1/2} \leq \frac{\|K_m\|^2}{xu} \leq \frac{C(K)\sqrt{m}}{xu}.$$

This yields:

$$\mathrm{Var}(\theta_t^{(2)}(Y_1)) \leq \int_0^{+\infty} u^2|K_m(u)|\frac{1}{\sqrt{u}}du\left(\int_0^{+\infty}x^2|t(x)|f_Y(xu)\frac{1}{\sqrt{x}}dx\right)C(K)^{1/2}m^{1/4}.$$

Then, as $yf_Y(y) \leq 1$, we get $\int_0^{+\infty} y^3f_Y^2(y)dy \leq \mathbb{E}(Y_1^2)$ and

$$\int_0^{+\infty} x^2|t(x)|f_Y(xu)\frac{1}{\sqrt{x}}dx \leq \left[\int_0^{+\infty}x^3f_Y^2(xu)dx\right]^{1/2} \leq \frac{1}{u^2}\left(\int_0^{+\infty}y^3f_Y^2(y)dy\right)^{1/2} \leq \frac{\sqrt{\mathbb{E}(Y_1^2)}}{u^2}.$$

Finally,

$$\mathrm{Var}(\theta_t^{(2)}(Y_1)) \leq C(K)^{1/2}\sqrt{\mathbb{E}(Y_1^2)}m^{1/4}\int_0^{+\infty}|K_m(u)|\frac{1}{\sqrt{u}}du \leq 3|\alpha|_1C(K)^{1/2}\sqrt{\mathbb{E}(Y_1^2)}m^{1/4}.$$

Thus, we can take $v = 3|\alpha|_1C(K)^{1/2}\sqrt{\mathbb{E}(Y_1^2)}m^{1/4}$. Lastly,

$$\frac{nH}{M} = \frac{\sqrt{10\,\mathtt{a}}}{3}\log(n), \quad \frac{nH^2}{v} = B\,m^{1/4}, \quad B = \frac{10\sqrt{C(K)}\mathbb{E}(Y_1}{3|\alpha|_1\sqrt{\mathbb{E}(Y_1^2)}}$$

This yields, choosing $\epsilon^2 = 1/2$, using (50) and taking $\mathtt{a}$ such that $2K_1C(\epsilon^2)\epsilon\sqrt{10\,\mathtt{a}}/(21\sqrt{2}) = 2$, and using $m \le n^2$ for any m in $\mathcal{M}_n$ we get

$$\mathbb{E}\left[\sup_{t\in\mathcal{B}(1)}(\nu_n^2(t) - 4H^2)_+\right] \le C_1\left[\frac{m^{1/4}}{n}e^{-Bm^{1/4}} + \frac{1}{\log^2 n}e^{-2\log(n)}\right].$$

Now, reminding of (49) and $\mathrm{Card}(\mathcal{M}_n) \le n$, we get

$$\mathbb{E}\left(\sup_{m'}\left(\|\tilde{\bar{F}}_{m'} - \bar{F}_{m'}\|^2 - \frac{Z(m')}{6}\right)_+\right) \le \frac{C_1}{n}\sum_{m\in\mathcal{M}_n} m^{1/4}e^{-Bm^{1/4}} + \frac{C_2}{n} \le \frac{C}{n}$$

This ends the proof of Lemma 5.6. □.

5.10.4. *Proof of Lemma 5.7.* The proof of Lemma (5.7) follows the same line as previously with K_m replaced by $K_m \odot K_{m'}$, where m is fixed and the sum is now taken over m' in $\mathcal{M}_n$.

(53)
$$\mathbb{E}\left(\sup_{m'}\left(\|\tilde{\bar{F}}_{m,m'} - \bar{F}_{m,m'}\|^2 - \frac{Z(m')}{6}\right)_+\right) \le \sum_{m'\in\mathcal{M}_n}\mathbb{E}\left(\left(\|\tilde{\bar{F}}_{m,m'} - \bar{F}_{m,m'}\|^2 - \frac{Z(m')}{6}\right)_+\right).$$

The truncation of the Y_i's by c_n is done as previously, and the bound given in Lemma 5.5 leads to the same result. Therefore, we can work as if the Y_i's were bounded by c_n.

Thus, we apply the Talagrand inequality to the empirical process

$$\nu_n^*(\theta_t^{(m,m')}) = \langle\tilde{\bar{F}}_{m,m'} - \bar{F}_{m,m'}, t\rangle^2$$

where $\theta_t^{(m,m')}$ is the analogous of θ_t with $K_m \odot K_{m'}$ instead of K_m. We have to find the three quantities H, v, M. This reduces to using Lemma 5.5 for H and inequalities given in (i) and (iii) of Lemma 5.4. The bounds being the same as for Lemma 5.6, the conclusion is also analogous. □

6. Appendix

6.1. **Auxiliary result.** We recall the generalized Minkowski inequality. The proof of the following inequality can be found in *e.g.* Tsybakov (2004, p. 161). For all Borel function g on $\mathbb{R}\times\mathbb{R}$, we have

$$\int_{\mathbb{R}}\left(\int_{\mathbb{R}} g(u,x)du\right)^2 dx \le \left(\int_{\mathbb{R}}\left(\int_{\mathbb{R}} g^2(u,x)dx\right)^{1/2} du\right)^2. \tag{54}$$

The Young inequality. (see [13]). Let f be a function belonging to $\mathbb{L}^p(\mathbb{R})$ and g belonging to $\mathbb{L}^q(\mathbb{R})$, let p, q, r be real numbers in $[1, +\infty]$ and such that

$$\frac{1}{p} + \frac{1}{q} = \frac{1}{r} + 1.$$

Then

$$\|f \star g\|_r \le \|f\|_p\,\|g\|_q. \tag{55}$$

where $f \star g$ is the convolution product and $\|f\|_p^p = \int |f(x)|^p dx$. In particular, for $p = 1$, $r = q = 2$, we have $\|f \star g\|_2 \le \|f\|_1\,\|g\|_2$.

The Talagrand inequality. The result below follows from the Talagrand concentration inequality given in Klein and Rio (2005) and arguments in Birgé and Massart (1998) (see the proof of their Corollary 2 page 354).

Lemma 6.1. *(Talagrand Inequality) Let $Y_1, \ldots, Y_n$ be independent random variables, let $\nu_{n,Y}(f) = (1/n)\sum_{i=1}^n[f(Y_i) - \mathbb{E}(f(Y_i))]$ and let $\mathcal{F}$ be a countable class of uniformly bounded measurable functions. Then for $\epsilon^2 > 0$*

$$\mathbb{E}\Big[\sup_{f\in\mathcal{F}}|\nu_{n,Y}(f)|^2 - 2(1+2\epsilon^2)H^2\Big]_+ \le \frac{4}{K_1}\left(\frac{v}{n}e^{-K_1\epsilon^2\frac{nH^2}{v}} + \frac{98M^2}{K_1n^2C^2(\epsilon^2)}e^{-\frac{2K_1C(\epsilon^2)\epsilon}{7\sqrt{2}}\frac{nH}{M}}\right),$$

with $C(\epsilon^2) = \sqrt{1+\epsilon^2} - 1$, $K_1 = 1/6$, *and*

$$\sup_{f\in\mathcal{F}} \|f\|_\infty \le M, \quad \mathbb{E}\Big[\sup_{f\in\mathcal{F}} |\nu_{n,Y}(f)|\Big] \le H, \; \sup_{f\in\mathcal{F}} \frac{1}{n}\sum_{k=1}^n \mathrm{Var}(f(Y_k)) \le v.$$

By standard density arguments, this result can be extended to the case where $\mathcal{F}$ is a unit ball of a linear normed space, after checking that $f \mapsto \nu_n(f)$ is continuous and $\mathcal{F}$ contains a countable dense family.

References

[1] Abbaszadeh, M, Chesneau, C. and Doosti, H. (2013) Multiplicative censoring: estimation of a density and its derivatives under the Lp-risk. *REVSTAT* **11**, 255-276.

[2] Abbaszadeh, M., Chesneau, C. and Doosti, H. (2012) Nonparametric estimation of density under bias and multiplicative censoring via wavelet methods. *Statist. Probab. Lett.* **82**, 932-941.

[3] Asgharian, M., Carone, M., Fakoor, V. (2012) Large-sample study of the kernel density estimators under multiplicative censoring. *Ann. Statist.* **40**, 159-187.

[4] Andersen, K. E. and Hansen, M. B. (2001) Multiplicative censoring: density estimation by a series expansion approach. *J. Statist. Plann. Inference* **98**, 137-155.

[5] Chesneau, C. (2013) Wavelet estimation of a density in a GARCH-type model. *Comm. Statist. Theory Methods* **42**, 98-117.

[6] Chernozhukov, V., Fernandez-Val, I. and Galichon, A. (2009) Improving point and interval estimators of monotone functions by rearrangement. *Biometrika* **96**, 559-575.

[7] Comte, F. and Genon-Catalot, V. (2006) Penalized projection estimator for volatility density. *Scand. J. Statist.* **33**, 875-893.

[8] Comte, F. and Genon-Catalot, V. (2012) Convolution power kernels for density estimation. *J. Statist. Plann. Inference* **142**, 1698-1715.

[9] van Es, B., Klaassen, C. A. J. and Oudshoorn, K. (2000). Survival analysis under cross-sectional sampling: length bias and multiplicative censoring. Prague Workshop on Perspectives in Modern Statistical Inference: Parametrics, Semi-parametrics, Non-parametrics (1998). *J. Statist. Plann. Inference* **91**, 295-312.

[10] Van Es, B., Spreij, P. and van Zanten, H. (2005) Nonparametric volatility density estimation for discrete time models. *J. Nonparametr. Stat.* **17**, 237-251.

[11] Fritsch, F. N. and Carlson, R. E. (1980). Monotone piecewise cubic interpolation. *SIAM Journal on Numerical Analysis,* **17**, 238-246

[12] Goldenshluger A. and Lepski, O. (2011). Bandwidth selection in kernel density estimation: oracle inequalities and adaptive minimax optimality. *Ann. Statist.* **39**, no. 3, 1608-1632.

[13] Hirsch, F. and Lacombe, G. (1999). *Elements of functional analysis.* Graduate Texts in Mathematics, 192. Springer-Verlag, New York.

[14] Killmann, F. and von Collani, E. (2001). A note on the convolution of the uniform and related distributions and their use in quality control. *Economic Quality Control* **16**, no. 1, 17-41.

[15] Mnatsakanov, R. M. (2008) Hausdorff moment problem: reconstruction of distributions. *Statist. Probab. Lett.* **78**, 1612-1618.

[16] Rényi, A. (1970). *Probability Theory.* North-Holland, Amsterdam.

[17] Tsybakov, A. B. (2009). *Introduction to nonparametric estimation.* Revised and extended from the 2004 French original. Springer Series in Statistics. Springer, New York.

[18] Vardi, Y. (1989) Multiplicative censoring, renewal processes, deconvolution and decreasing density: nonparametric estimation. *Biometrika* **76**, 751-761.

[19] Vardi, Y. and Zhang, C.-H. (1992) Large sample study of empirical distributions in a random-multiplicative censoring model. *Ann. Statist.* **20**, 1022-1039.