



HAL
open science

New developments on (non-)normalized continuous associated-kernel density estimators

Youssef Esstafa, Célestin C Kokonendji, Thi-Bao-Trâm Ngô

► **To cite this version:**

Youssef Esstafa, Célestin C Kokonendji, Thi-Bao-Trâm Ngô. New developments on (non-)normalized continuous associated-kernel density estimators. 2023. hal-04112846

HAL Id: hal-04112846

<https://hal.science/hal-04112846>

Preprint submitted on 1 Jun 2023

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

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

NEW DEVELOPMENTS ON (NON-)NORMALIZED CONTINUOUS ASSOCIATED-KERNEL DENSITY ESTIMATORS

BY YOUSSEF ESSTAF^a AND THI-BAO-TRÂM NGÔ^b

*Le Mans Université, Laboratoire Manceau de Mathématiques,
Avenue Olivier Messiaen, 72085 Le Mans Cedex 09, France.*

^aYoussef.Esstafa@univ-lemans.fr; ^bThi-Bao-Tran.Ngo@univ-lemans.fr

BY CÉLESTIN C. KOKONENDJI ^{c,d}

*Université de Franche-Comté, Laboratoire de Mathématiques de Besançon
UMR 6623 CNRS-UFC, 16 route de Gray, 25000 Besançon, France.*

*& Université de Bangui, Laboratoire de mathématiques et connexes de Bangui,
B.P. 908 Avenue des Martyrs, Bangui, Centrafrique.*

^ccelestin.kokonendji@univ-fcomte.fr

^d<http://lmb.univ-fcomte.fr/Kokonendji-Celestin-C>

We consider the general modern notion of the so-called associated kernels for smoothing density function on a given support $\mathbb{T} \subseteq \mathbb{R}$. According to the recent and global properties of normalized discrete associated-kernel estimators, we here investigate the continuous associated-kernel contexts in a completely different way. Diverse and numerous in the literature, the standard (non-)normalized density estimators by non-classical kernels have great interests, including modified versions for reducing the possible boundary bias. We first show, under specific assumptions such the asymptotic unimodality on the continuous associated kernel, that the normalizing random variable also converges in mean square to 1. We then deduce the consistency of the considered estimator. The comparison in favour of the standard normalized estimator is obtained by the mean squared error. We conclude by providing, for the first time, the general asymptotic normalities through some regularity assumptions for both (un)normalized associated-kernel density estimators. The Gumbel, Weibull, gamma, lognormal, and other associated kernels are required for illustrating theoretically and numerically some of our results with an application to original data of automobile claim amounts from Covéa Affinity.

1. Introduction. The associated kernels are the contemporary version of the so-called classical kernels. Since the origin, classical kernels generally indicate continuous and symmetric smoothers for functions on the whole real line, not depending on the bandwidth and the target point, where the estimated probability density function (pdf) is evaluated; see, for instance, Parzen [20], Epanechnikov [6], Scott [25], Wand and Jones [29] and Zougab et al. [34]. The associated kernels also contain the family of non-classical (or asymmetric) kernels which can be discrete or continuous smoothers according to the support of the unknown function to be estimated, without edge effect. Hence, it reveals a choice of the associated kernels with respect to the support (un)bounded $\mathbb{T} \subseteq \mathbb{R}$ of the pdf f to be estimated. More generally, one can select an associated kernel for each $x \in \mathbb{T}$, a target point inside or at an edge, to obtain a combination ones. For discrete kernels, we need to refer to Esstafa et al. [7]. They

MSC2020 subject classifications: Primary 62G07, 62G20; secondary 62G99.

Keywords and phrases: Asymmetric kernel, convergence, limit distribution, normalizing constant, probability density function.

have recently established that their standard normalized estimators of a probability mass function admit very good behaviour under some general conditions than the unnormalized ones. However, the classical kernels do not exist for discrete setup for doing extensions.

With respect to continuous ones, multiple authors have worked on the topic in these 2000s such as Chen [2, 3] for beta and gamma kernels, Jin and Kawczak [13] for the Birnbaum-Saunders and lognormal kernels, Scaillet [24] for inverse and reciprocal inverse Gaussian kernels, Marchant et al. [18] for generalized Birnbaum-Saunders kernel, Salha et al. [22, 23] for the Weibull and Erlang kernels, Hirukawa and Sakudo [11] for a family of generalized gamma kernels, Libengué Dobélé-Kpoka and Kokonendji [17] for extended beta and lognormal kernels built by the mode-dispersion method with some general comments, Ziane et al. [33] for the Birnbaum-Saunders power-exponential kernel, Geenens [8] for the Mellin-Meijer kernel, Khan and Akbar [14] for the Gumbel kernel, and Ouimet and Tolosana-Delgado [19] for the Dirichlet kernel. See also Hirukawa [10], Igarashi [12] and Doho et al. [5] for their uses and properties among many others. They are not too many general and unified (asymptotic) results, including all the family of the classical kernels which are always provided normalized estimators of pdf. It is in this sense that we propose the current new developments.

In this paper, and almost similar but completely different to the discrete situation studied by Esstafa et al. [7], we mainly focus on the normalized phenomenon of the continuous non-classical kernel estimator for any univariate pdf f , which has support $\mathbb{T} \subseteq \mathbb{R}$, bounded or not. It is very common but rarely considered in practice, except in Wansouwé et al. [30], Kokonendji and Somé [16], Somé and Kokonendji [26] and Somé et al. [27]. However, theoretical and comparative results are completely missing so far. Let us first redefine the so-called (second-order of) continuous associated kernel from, for instance, Kokonendji and Libengué Dobélé-Kpoka [15] and, then, recall two basical asymptotic properties in Theorem 1.2 to be completed later.

DEFINITION 1.1. *Let $\mathbb{T} \subseteq \mathbb{R}$ be the support of the pdf f to be estimated, $x \in \mathbb{T}$ a target point and $h > 0$ a bandwidth. A parameterized pdf $K_{x,h}(\cdot)$ on the support \mathbb{S}_x or $\mathbb{S}_{x,h} \subseteq \mathbb{R}$ is called "continuous associated kernel" if the following conditions are satisfied:*

$$(1.1) \quad x \in \mathbb{S}_{x,h}, \quad \mathbb{E}[Z_{x,h}] - x =: A(x,h) \xrightarrow{h \rightarrow 0} 0 \quad \text{and} \quad \text{Var}(Z_{x,h}) =: B(x,h) \xrightarrow{h \rightarrow 0} 0,$$

where $Z_{x,h}$ denotes the continuous random variable with pdf $K_{x,h}(\cdot)$.

Hence, to our knowledge, all continuous associated kernels satisfy Definition 1.1 even if their authors have made them through different artisanal approaches. See Libengué Dobélé-Kpoka and Kokonendji [17] for the sophisticated mode-dispersion method, which is the basical idea that several authors ignore or do not want to reveal. In particular, a classical or symmetric continuous kernel K is derived as an associated kernel in the following sense for holding (1.1):

$$(1.2) \quad \mathbb{S}_{x,h} = x + h\mathbb{S}_K \quad \text{and} \quad K_{x,h}(\cdot) = \frac{1}{h}K\left(\frac{\cdot - x}{h}\right),$$

where $\mathbb{S}_K \subseteq \mathbb{R}$ designates the symmetric support in 0 of K ; e.g. Scott [25] and Wand and Jones [29] for a list of classical kernels K as the standard Gaussian density. From (1.1), one easily checks that the corresponding $A(x,h) = 0$ and $B(x,h) = h^2 \int_{\mathbb{S}_K} u^2 K(u) du$ for (1.2).

Now, we consider the arbitrary sequence of positive smoothing parameters $(h_n)_{n \geq 1}$ which satisfies $\lim_{n \rightarrow \infty} h_n = 0$, while $K_{x,h_n}(\cdot)$ stands for a suitably chosen continuous kernel function. Also, the n -sample (X_1, X_2, \dots, X_n) of independent and identically distributed (i.i.d.) continuous random variables has a pdf f on $\mathbb{T} \subseteq \mathbb{R}$. With reference to Esstafa et al. [7] but not Glad et al. [9], we express the standard normalized estimator

$$(1.3) \quad \widehat{f}_n(x) = \frac{\widetilde{f}_n(x)}{C_n},$$

with

$$(1.4) \quad \widetilde{f}_n(x) = \frac{1}{n} \sum_{i=1}^n K_{x,h_n}(X_i) \quad \text{and} \quad C_n = \int_{x \in \mathbb{T}} \widetilde{f}_n(x) dx > 0.$$

In general, and except for classical associated kernels (1.2) with $\widetilde{f}_n = \widehat{f}_n$ or $C_n = 1$, any basical estimator \widetilde{f}_n of f in (1.4) seems improper density estimate by using a non-classical kernel or $0 < \int \widetilde{f}_n(x) dx \neq 1$. Indeed, from (1.4), it should be written

$$(1.5) \quad C_n = \frac{1}{n} \sum_{i=1}^n \int_{x \in \mathbb{T}} K_{x,h_n}(X_i) dx$$

and C_n generally fails to be equal to 1 since the integral $\int_{x \in \mathbb{T}} K_{x,h_n}(X_i) dx$ works on a domain of the intrinsic parameter x of the pdf $K_{x,h_n}(\cdot)$. See, e.g., Kokonendji and Libengué Dobélé-Kpoka [15] and Wansouwé et al. [30] for numerical results. But, how to explain that C_n is always around 1 from (1.5)? An attempt of understand comes from inequalities of concentration. Alas, it is not possible in our knowledge. The random variables X_i could be there of some things to follow. This phenomenon leads to consider (1.3) the normalized estimator $\widehat{f}_n = \widetilde{f}_n/C_n$ for getting a true pdf estimate of f , with the so-called random associated kernel $\widetilde{K}_{x,h_n}(\cdot) = K_{x,h_n}(\cdot)/C_n$.

Under general conditions, we shall prove the mean square convergence of C_n to 1 and, therefore, discuss some asymptotic properties of \widehat{f}_n and \widetilde{f}_n with the L^2 -comparison between them. It is to be noted that, in the literature many results are very specific to a given family of non-classical associated kernels; except the following one that we shall complete and improve below.

THEOREM 1.2. [15, Theorem 2.2] *Let $f \in C^2(\mathbb{T})$. For any $x \in \mathbb{T}$ and under conditions (1.1), one has $\widetilde{f}_n(x) \xrightarrow{a.s.} f(x)$ as $n \rightarrow \infty$; furthermore, if there is a real largest number $r_2 = r_2(K_{x,h_n}) > 0$ such that*

$$h_n^{r_2} \int_{\mathbb{S}_{x,h_n} \cap \mathbb{T}} K_{x,h_n}^2(u) du \leq c_2(x) < \infty \quad \text{and} \quad \lim_{n \rightarrow \infty} n h_n^{r_2} = \infty$$

then $\widetilde{f}_n(x) \xrightarrow{L^2} f(x)$ as $n \rightarrow \infty$, where “ $\xrightarrow{L^2}$ ” (resp. “ $\xrightarrow{a.s.}$ ”) stands for “mean square” (resp. “almost surely”) convergence.

Another important fact (and already solved) in the continuous non-classical kernels for estimating a pdf (with bounded support) corresponds to the problem of

boundary bias. In fact, from (1.1) and for $f \in C^2(\mathbb{T})$, the pointwise bias of \widetilde{f}_n is represented by

$$\text{Bias}\{\widetilde{f}_n(x)\} = A(x, h)f'(x) + \frac{1}{2}\{A^2(x, h) + B(x, h)\}f''(x) + o(h^2);$$

see, e.g., [15, Proposition 2.1]. Hence, for any non-classical kernel with $A(x, h) \neq 0$, we need to eliminate the term $A(x, h)$ in the largest region of the support \mathbb{T} of f for obtaining a second version of \widetilde{f}_n or the modified one, which is denoted by \widetilde{f}_n^* . We refer for that, for instance, to Chen [2, 3], Zhang [31], Zhang and Karunamuni [32] and, Kokonendji and Libengué Dobélé-Kpoka [15, Section 3.2] for a few and general treatments. Also, we can consider again its modified normalized version as Eq. (1.3) by $\widehat{f}_n^* = \widetilde{f}_n^*/C_n^*$ with $C_n^* = \int \widetilde{f}_n^*(x)dx > 0$; see Somé et al. [27, Section 4] for a recent practical use of both standard and modified normalized gamma-kernel density estimates. Consequently and through our new assumptions, we shall show that all results of this paper hold too for both standard and modified (non-)normalized estimators. Without loss of generality, we only retain the standard (non-)normalized estimators from (1.3) and (1.4) for all the theoretical results below; especially, when h tends to 0, then this region of the boundary becomes insignificant [15, Section 3.2].

The rest of the paper is laid out as follows. Section 2 presents basic assumptions on all continuous associated kernels that we illustrate on some examples of (non-)classical kernels. In Section 3, we show our main asymptotic results for both (non-)normalized associated-kernel estimators; in particular, pointwise consistence and central limit theorem are proven with L^2 -comparison of both estimators. Section 4 of numerical results points out some practical behaviour of our general results with an illustrative application to the original dataset from Covéa Affinity, which is the expert entity in affinity insurance of the Covéa group (a French mutual insurance company). Finally, Section 5 is devoted to the remaining detailed proofs of the main results.

2. Assumptions and settings. In what follows, we assume that the pdf f on $\mathbb{T} \subseteq \mathbb{R}$ to be estimated satisfies:

(A0) $f \in C^1(\mathbb{T})$ with absolute integrable derivative f' on \mathbb{T} and for any $x \in \mathbb{T}$, there exists positive constants $r_{1,x}$ and $r_{2,x}$ such that $x - r_{1,x}, x + r_{2,x} \in \mathbb{T}$ and f is increasing on $(-\infty, x - r_{1,x}] \cap \mathbb{T}$ and decreasing on $[x + r_{2,x}, \infty) \cap \mathbb{T}$.

REMARK 2.1. From Assumption (A0) and for a fixed target $x \in \mathbb{T}$, it obviously appears that the two quantities $f(x + r_{2,x})$ and $f(x - r_{1,x})$ are finite. Moreover, since f and f' are continuous on \mathbb{T} , the interval $[x - r_{1,x}, x + r_{2,x}]$ is included in \mathbb{T} and, therefore, $\max_{z \in [x - r_{1,x}, x + r_{2,x}]} |f'(z)|$ is bounded.

In addition, for any target point $x \in \mathbb{T}$, we suppose on the asymptotic properties of the continuous associated random variable Z_{x, h_n} having the density K_{x, h_n} , that:

(A1) Assume that, for some $\alpha > 0$, for any $x \in \mathbb{T}$, we have

$$(2.1) \quad x \in \mathbb{S}_x, \quad |\mathbb{E}[Z_{x, h_n}] - x| = O(h_n^\alpha) \quad \text{and} \quad \text{Var}(Z_{x, h_n}) = O(h_n^{2\alpha});$$

$$(2.2) \quad \max_{y \in \mathbb{S}_x \cap \mathbb{T}} K_{x, h_n}(y) = O(h_n^{-\alpha});$$

$$(2.3) \quad \int_{y \in \mathbb{S}_x \cap \mathbb{T}} K_{x, h_n}^2(y) dy = O(h_n^{-\alpha}) \quad \text{and} \quad \int_{y \in \mathbb{S}_x \cap \mathbb{T}} K_{x, h_n}^3(y) dy = O(h_n^{-2\alpha}).$$

Notice that (2.1) is a refined version of Condition (1.1) in Definition 1.1.

Also, we shall need the following ones on K_{x,h_n} when the normalized associated-kernel estimator shall be explored.

(A2) Assume that, for a fixed $y \in \mathbb{T}$, there exists $\alpha > 0$ such that

$$\int_{z \in \Lambda_y \cap \mathbb{T}} K_{z,h_n}(y) dz < \infty \quad \text{and} \quad \int_{z \in \Lambda_y \cap \mathbb{T}} K_{z,h_n}^2(y) dz = O(h_n^{-\alpha}),$$

with $\Lambda_y = \{z \in \mathbb{T} \mid y \in \mathbb{S}_z\}$.

Eight examples of continuous associated kernels are proposed including classical as well as non-classical with new ones.

EXAMPLE 2.2. The following continuous associated kernels satisfy our Assumptions (A1) and (A2) with $\alpha = 1$.

1. *Lognormal kernel*: for $x \in \mathbb{T} = \mathbb{R}_+$ and $h_n > 0$, the mode-dispersion lognormal kernel on $\mathbb{S}_x = \mathbb{R}_+$ yields

$$K_{x,h_n}(y) = \frac{1}{yh_n \sqrt{2\pi}} \exp \left\{ -\frac{1}{2} \left(\frac{1}{h_n} \log(y/x) - h_n \right)^2 \right\}, \quad y \in \mathbb{S}_x = \mathbb{R}_+.$$

Then, it is easy to see that $\int_{y \in \mathbb{S}_x} K_{x,h_n}(y) dy = 1$ and that Equation (2.2) holds since

$\max_{y \in \mathbb{S}_x \cap \mathbb{T}} K_{x,h_n}(y) = K_{x,h_n}(x) = (xh_n \sqrt{2\pi})^{-1} \exp(-h_n^2/2)$. In addition, one has

$$\mathbb{E}[Z_{x,h_n}] = x + (e^{3h_n^2/2} - 1)x \quad \text{and} \quad \text{Var}(Z_{x,h_n}) = x^2 e^{3h_n^2} (e^{h_n^2} - 1);$$

therefore, the conditions in (2.1) are also satisfied. Finally, we calculate

$$\int_{y \in \mathbb{S}_x \cap \mathbb{T}} K_{x,h_n}^2(y) dy = \frac{e^{-3h_n^2/4}}{2h_n x \sqrt{\pi}} \quad \text{and} \quad \int_{y \in \mathbb{S}_x \cap \mathbb{T}} K_{x,h_n}^3(y) dy = \frac{e^{-4h_n^2/3}}{2\pi \sqrt{3} h_n^2 x^2}$$

which fulfill (2.3). With regard to (A2), one obtains

$$\int_{z \in \mathbb{R}_+} K_{z,h_n}(y) dz = e^{-h_n^2/2} \quad \text{and} \quad \int_{z \in \mathbb{R}_+} K_{z,h_n}^2(y) dz = \frac{e^{-3h_n^2/4}}{2h_n y \sqrt{\pi}}, \quad \forall y \in \mathbb{R}_+;$$

and, therefore, one here gets the normalizing constant $C_n = e^{-h_n^2/2} \neq 1$ by (1.5). Similar results of $C_n \neq 1$ hold for many other non-classical kernel estimators such for (generalized) gamma, (extended) beta, (reciprocal) inverse Gaussian, (generalized) Birnbaum-Saunders, Weibull, Erlang and Mellin-Meijer.

2. *Weibull kernel*: for $x > 0$ and $h_n > 0$, the mode-dispersion Weibull kernel is directly defined on $\mathbb{S}_x = \mathbb{R}_+ = \mathbb{T}$ by

$$K_{x,h_n}(y) = \frac{1}{xh_n} \left(\frac{y}{x} \right)^{1/h_n - 1} \exp \left\{ -\left(\frac{y}{x} \right)^{1/h_n} \right\}, \quad y \in \mathbb{S}_x = \mathbb{R}_+.$$

Then, it is clear that $\int_{y \in \mathbb{S}_x} K_{x,h_n}(y) dy = 1$ and that (2.2) is satisfied since $\max_{y \in \mathbb{S}_x \cap \mathbb{T}} K_{x,h_n}(y) =$

$K_{x,h_n}((h_n - 1)^{h_n} x) = (xh_n)^{-1} (h_n - 1)^{1-h_n} e^{-(h_n-1)}$. Further, one checks that

$$\mathbb{E}[Z_{x,h_n}] = x\Gamma(1 + h_n) \quad \text{and} \quad \text{Var}(Z_{x,h_n}) = x^2[\Gamma(1 + 2h_n) - \Gamma^2(1 + h_n)];$$

consequently, the conditions in (2.1) are also satisfied and, where the gamma function is known to be the improper integral:

$$\Gamma(v) = \int_0^{\infty} s^{v-1} \exp(-s) ds, \quad v > 0.$$

Hence, we have

$$\int_{y \in \mathbb{S}_x \cap \mathbb{T}} K_{x,h_n}^2(y) dy = \frac{\Gamma(2-h_n)2^{h_n-2}}{h_n x} \quad \text{and} \quad \int_{y \in \mathbb{S}_x \cap \mathbb{T}} K_{x,h_n}^3(y) dy = \frac{\Gamma(3-2h_n)3^{2h_n-3}}{h_n^2 x^{h_n}},$$

which fulfills (2.3). With respect to (A2), one obtains

$$\int_{z \in \mathbb{R}_+} K_{z,h_n}(y) dz = \Gamma(1-h_n) \quad \text{and} \quad \int_{z \in \mathbb{R}_+} K_{z,h_n}^2(y) dz = \frac{2^{h_n} \Gamma(2-h_n)}{4h_n}, \quad \forall y \in \mathbb{R}_+;$$

and, thus, one here has the normalizing constant $C_n = \Gamma(1-h_n) \neq 1$ through (1.5).

3. *Gamma kernel*: for $x > 0$ and $h_n > 0$, the well-known gamma kernel is given on $\mathbb{S}_x = \mathbb{R}_+$ as

$$K_{x,h_n}(y) = \frac{1}{\Gamma(1+x/h_n)h_n^{1+x/h_n}} y^{x/h_n} \exp(-y/h_n), \quad y \in \mathbb{S}_x = \mathbb{R}_+.$$

Hence, it is obvious to have $\int_{y \in \mathbb{S}_x} K_{x,h_n}(y) dy = 1$ and Equation (2.2) holds since

$$\max_{y \in \mathbb{S}_x \cap \mathbb{T}} K_{x,h_n}(y) = K_{x,h_n}(x) = \frac{1}{\Gamma(1+x/h_n)h_n^{1+x/h_n}} x^{x/h_n} \exp(-x/h_n). \quad \text{Moreover, one has}$$

$$\mathbb{E}[Z_{x,h_n}] = x + h_n \quad \text{and} \quad \text{Var}(Z_{x,h_n}) = (x + h_n)h_n;$$

thus, the conditions in (2.1) are also verified. Finally, we calculate

$$\int_{y \in \mathbb{S}_x \cap \mathbb{T}} K_{x,h_n}^2(y) dy = \frac{2^{-1-2x/h_n} \Gamma(1+2x/h_n)}{h_n \Gamma^2(1+x/h_n)}$$

and

$$\int_{y \in \mathbb{S}_x \cap \mathbb{T}} K_{x,h_n}^3(y) dy = \frac{3^{-1-3x/h_n} \Gamma(1+3x/h_n)}{h_n^2 \Gamma^3(1+x/h_n)},$$

which fulfill (2.3) from the Stirling formula. Concerning (A2), one gets

$$\int_{z \in \mathbb{R}_+} K_{z,h_n}(y) dz = e^{-y/h_n} \int_0^{\infty} \frac{(y/h_n)^t}{\Gamma(1+t)} dt < \infty$$

and

$$\int_{z \in \mathbb{R}_+} K_{z,h_n}^2(y) dz = \frac{e^{-2y/h_n}}{2h_n} \int_0^{\infty} \frac{(y/h_n)^t}{\Gamma^2(1+t/2)} dt < O(h_n^{-1}), \quad \forall y \in \mathbb{R}_+,$$

since these integrals are finite by using the Stirling asymptotic formula

$$\frac{(y/h_n)^t}{\Gamma(1+t)} \sim \frac{1}{\sqrt{2\pi t} e^{t \log\{ye/(th_n)\}}} \quad \text{and} \quad \frac{(y/h_n)^t}{\Gamma^2(1+t/2)} \sim \frac{1}{\pi t} e^{t \log\{2ye/(th_n)\}}$$

as $t \rightarrow \infty$. Hence, one cannot explicitly get the normalizing constant C_n , which is not equal to 1 via (1.5).

4. *Gumbel kernel*: given $x \in \mathbb{T} = \mathbb{R}$ and $h_n > 0$, the Gumbel kernel has been proposed on $\mathbb{S}_x = \mathbb{R}$ by

$$K_{x,h_n}(y) = \frac{1}{h_n} \exp \left[-\exp \left(-\frac{y-x}{h_n} \right) - \frac{y-x}{h_n} \right], \quad y \in \mathbb{S}_x = \mathbb{R}.$$

Then, we easily check $\int_{y \in \mathbb{S}_x} K_{x,h_n}(y) dy = 1$ and also Equation (2.2) holds, because $\max_{y \in \mathbb{S}_x \cap \mathbb{T}} K_{x,h_n}(y) = K_{x,h_n}(x) = 1/(eh_n)$. Next, we have

$$\mathbb{E}[Z_{x,h_n}] = x + h_n \gamma \quad \text{and} \quad \text{Var}(Z_{x,h_n}) = \frac{\pi^2}{6} h_n^2,$$

where $\gamma = 0.577215\dots$ is the Euler-Mascheroni constant. One also verifies that

$$\int_{y \in \mathbb{S}_x \cap \mathbb{T}} K_{x,h_n}^2(y) dy = \frac{1}{4h_n} (1 + 2e^{(x-y)/h_n}) e^{-2 \exp[(x-y)/h_n]} \Big|_{-\infty}^{\infty} = \frac{1}{4h_n}$$

and

$$\int_{y \in \mathbb{S}_x \cap \mathbb{T}} K_{x,h_n}^3(y) dy = \frac{1}{27h_n^2} (2 + 6e^{(x-y)/h_n} + 9e^{2(x-y)/h_n}) e^{-3 \exp[(x-y)/h_n]} \Big|_{-\infty}^{\infty} = \frac{2}{27h_n^2}.$$

Assumption (A2) holds since we substitute x to $-y$ in both

$$\int_{z \in \mathbb{R}} K_{z,h_n}(y) dz = 1 \quad \text{and} \quad \int_{z \in \mathbb{R}} K_{z,h_n}^2(y) dz = \frac{1}{4h_n}, \quad \forall y \in \mathbb{R};$$

and, then, one easily obtains $C_n = 1$ by (1.5). The Gumbel kernel corresponds to an asymmetric one and has its mode at x . It is useful to estimate extreme value distribution on $\mathbb{T} = \mathbb{R}$.

5. *Generalized extreme value (GVE) kernel*: for the position $x \in \mathbb{T} = \mathbb{R}$, the scale $h_n > 0$ and the shape $\xi = -1$, the GVE kernel is defined on $\mathbb{S}_{x,h_n} = (-\infty, x + h_n]$ as

$$K_{x,h_n}(y) = \frac{1}{h_n} \exp \left[-\left(1 + \frac{x-y}{h_n} \right) \right], \quad y \in \mathbb{S}_{x,h_n} = (-\infty, x + h_n].$$

Then, one has $\int_{y \in \mathbb{S}_{x,h_n}} K_{x,h_n}(y) dy = 1$ and the condition (2.2) is satisfied, since $\max_{y \in \mathbb{S}_{x,h_n} \cap \mathbb{T}} K_{x,h_n}(y) = K_{x,h_n}(x + h_n) = 1/h_n$. Also, we have

$$\mathbb{E}[Z_{x,h_n}] = x \quad \text{and} \quad \text{Var}(Z_{x,h_n}) = h_n^2,$$

and, one can check that

$$\int_{y \in \mathbb{S}_{x,h_n} \cap \mathbb{T}} K_{x,h_n}^2(y) dy = \frac{1}{2h_n} \quad \text{and} \quad \int_{y \in \mathbb{S}_{x,h_n} \cap \mathbb{T}} K_{x,h_n}^3(y) dy = \frac{1}{3h_n^2}.$$

Notice that $y \in \mathbb{S}_{x,h_n} \cap \mathbb{T} \Leftrightarrow x \geq y - h_n$. Thus, Assumption (A2) here holds because of

$$\int_{[y-h_n, \infty)} K_{z,h_n}(y) dz = 1 \quad \text{and} \quad \int_{[y-h_n, \infty)} K_{z,h_n}^2(y) dz = \frac{1}{2h_n}, \quad \forall y \in \mathbb{R};$$

and, one here deduces $C_n = 1$ by (1.5). The GVE kernel appears to be an asymmetric one and has its mode at x . It can be used to estimate extreme value distribution from the extreme left $-\infty$ to the target point x on $\mathbb{T} = \mathbb{R}$.

6. *Logistic kernel*: let $x \in \mathbb{T} = \mathbb{R}$ and $h_n > 0$, we introduce the logistic kernel on $\mathbb{S}_x = \mathbb{R}$ as

$$K_{x,h_n}(y) = \frac{\left(1 + \exp\left(-\frac{y-x}{h_n}\right)\right)^{-2}}{h_n} \exp\left(-\frac{y-x}{h_n}\right), \quad y \in \mathbb{S}_x = \mathbb{R}.$$

Thus, one easily has $\int_{y \in \mathbb{S}_x} K_{x,h_n}(y) dy = 1$ and Equation (2.2) is satisfied, since $\max_{y \in \mathbb{S}_x \cap \mathbb{T}} K_{x,h_n}(y) = K_{x,h_n}(x) = 1/(4h_n)$. Also, we have

$$\mathbb{E}[Z_{x,h_n}] = x \quad \text{and} \quad \text{Var}(Z_{x,h_n}) = \frac{\pi^2}{3} h_n^2.$$

Furthermore, direct calculations give

$$\int_{y \in \mathbb{S}_x \cap \mathbb{T}} K_{x,h_n}^2(y) dy = \frac{1}{6h_n} \quad \text{and} \quad \int_{y \in \mathbb{S}_x \cap \mathbb{T}} K_{x,h_n}^3(y) dy = \frac{1}{30h_n^2}.$$

Therefore, Assumption (A2) holds too since we have here

$$\int_{z \in \mathbb{R}} K_{z,h_n}(y) dz = 1 \quad \text{and} \quad \int_{z \in \mathbb{R}} K_{z,h_n}^2(y) dz = \frac{1}{6h_n}, \quad \forall y \in \mathbb{R},$$

and also $C_n = 1$ via (1.5). However, the logistic kernel is a symmetric one and has its mode at x . It is in the family of classical kernels (1.2) such as the below Gaussian one.

7. *Gaussian kernel*: as the most common example of (1.2), let $x \in \mathbb{T} = \mathbb{R}$ and $h_n > 0$. The classical associated Gaussian kernel is defined on $\mathbb{S}_x = \mathbb{R}$ by

$$K_{x,h_n}(y) = \frac{1}{h_n \sqrt{2\pi}} \exp\left(-\frac{(y-x)^2}{2h_n^2}\right), \quad y \in \mathbb{S}_x = \mathbb{R}.$$

Then, it is trivial that $\int_{y \in \mathbb{S}_x} K_{x,h_n}(y) dy = 1$ and that (2.2) is satisfied, with $\max_{y \in \mathbb{S}_x \cap \mathbb{T}} K_{x,h_n}(y) = K_{x,h_n}(x) = 1/(h_n \sqrt{2\pi})$. Next, one has

$$\mathbb{E}[Z_{x,h_n}] = x \quad \text{and} \quad \text{Var}(Z_{x,h_n}) = h_n^2;$$

and, one obtains

$$\int_{y \in \mathbb{S}_x \cap \mathbb{T}} K_{x,h_n}^2(y) dy = \frac{1}{2h_n \sqrt{\pi}} \quad \text{and} \quad \int_{y \in \mathbb{S}_x \cap \mathbb{T}} K_{x,h_n}^3(y) dy = \frac{1}{2\sqrt{3}\pi h_n^2}.$$

Concerning (A2), we here have

$$\int_{z \in \mathbb{R}} K_{z,h_n}(y) dz = 1 \quad \text{and} \quad \int_{z \in \mathbb{R}} K_{z,h_n}^2(y) dz = \frac{1}{2\sqrt{\pi}h_n}, \quad \forall y \in \mathbb{R},$$

and finally $C_n = n^{-1} \sum_{i=1}^n \int_{z \in \mathbb{R}} K_{z,h_n}(X_i) dz = 1$ by (1.5).

8. *Associated kernel generated from a classical kernel of (1.2)*: because of boundary biases we consider $\mathbb{T} = \mathbb{R}$ and, for $x \in \mathbb{T} = \mathbb{R}$ and $h_n > 0$, $\mathbb{S}_{x,h_n} = x + h_n \mathbb{S}_K$, where $\mathbb{S}_K \subseteq \mathbb{R}$ is the symmetric support around 0 of the classical kernel K with

$$K_{x,h_n}(y) = \frac{1}{h_n} K\left(\frac{y-x}{h_n}\right), \quad \forall y \in \mathbb{S}_{x,h_n},$$

and it satisfies

$$\int_{\mathbb{R}} K^2(u)du < \infty \quad \text{and} \quad \int_{\mathbb{R}} K^3(u)du < \infty.$$

These conditions are fulfilled by most of the well-known kernels such as Epanechnikov or Parabolic, Cosine, Triangular, Quartic (Biweight), Triweight, Tricube, Gaussian, Logistic, Double Exponentielle and Sylverman. Then, it is easy to see that $\int_{y \in \mathbb{S}_{x,h_n}} K_{x,h_n}(y)dy = 1$ and that (2.2) is satisfied from $\max_{y \in \mathbb{S}_{x,h_n} \cap \mathbb{T}} K_{x,h_n}(y) = K_{x,h_n}(x) = K(0)/h_n$. Next, we have

$$\mathbb{E}[Z_{x,h_n}] = x \quad \text{and} \quad \text{Var}(Z_{x,h_n}) = h_n^2 \int_{\mathbb{R}} u^2 K(u)du = O(h_n^2).$$

One gets

$$\int_{y \in \mathbb{S}_{x,h_n} \cap \mathbb{T}} K_{x,h_n}^2(y)dy = \frac{1}{h_n} \int_{\mathbb{R}} K^2(u)du = O(h_n^{-1})$$

and

$$\int_{y \in \mathbb{S}_{x,h_n} \cap \mathbb{T}} K_{x,h_n}^3(y)dy = \frac{1}{h_n^2} \int_{\mathbb{R}} K^3(u)du = O(h_n^{-2}).$$

Since the classical kernel K is symmetric, Assumption (A2) is automatically satisfied with $C_n = n^{-1} \sum_{i=1}^n \int_{z \in \mathbb{R}} K_{z,h_n}(X_i)dz = 1$ by (1.5).

3. Results on (non-)normalized estimators. We here provide our general results of the continuous associated-kernel estimators, which are first unnormalized and then normalized with a comparison.

3.1. *Unnormalized associated-kernel estimator.* This section consists of two main parts. In the first one, we prove the convergence in mean square of the non-normalized estimator \tilde{f}_n to the unknown pdf f . From this result, the second main part is dedicated to demonstrating its corresponding central limit theorem. Here, it should be noticed that, from now on, we denote c as some generic constant that can change from line to line.

First of all, and using Assumption (A1), we need the following lemma. It will be very important for showing many results below, since the behaviour of the continuous associated kernels (implicitly) can be built by the mode-dispersion method of Libengué Dobélé-Kpoka and Kokonendji [17].

LEMMA 3.1. *Under Parts (2.1) and (2.2) of Assumption (A1) and considering the target point $x \in \mathbb{T}$, we have that for any compact subset $C_x \subset \mathbb{S}_x \setminus \{x\}$*

$$(3.1) \quad \sup_{y \in C_x} K_{x,h_n}(y) \xrightarrow{n \rightarrow \infty} 0.$$

PROOF. It is straightforward from the fact that, for any $\beta > 0$, we have

$$\begin{aligned} \mathbb{P}(|Z_{x,h_n} - x| > \beta) &\leq \frac{1}{\beta^2} \mathbb{E}[|Z_{x,h_n} - x|^2] \\ &= \frac{1}{\beta^2} \left\{ \text{Var}(Z_{x,h_n}) + (\mathbb{E}[Z_{x,h_n} - x])^2 \right\} \xrightarrow{n \rightarrow \infty} 0. \end{aligned}$$

This entails that Z_{x,h_n} converges in distribution to x . It means that we get the result (3.1) from the equicontinuity of K_{x,h_n} ; see, e.g., [1, Lemma 1] or [28]. \square

Now, proceeding is the same ideas as in [15], we provide the next technical proposition.

PROPOSITION 3.2. *Under Assumptions (A0) and (A1), we have*

$$\mathbb{E} \left[\widetilde{f}_n(x) - f(x) \right] = \mathcal{O}(\max\{f(x+r_{2,x}), f(x-r_{1,x}), \max_{z \in [x-r_{1,x}, x+r_{2,x}]} |f'(z)|\}).$$

PROOF. Using the i.i.d. property of $(X_i)_{1 \leq i \leq n}$, we get

$$(3.2) \quad \mathbb{E} \left[\widetilde{f}_n(x) \right] = \frac{1}{n} \sum_{i=1}^n \mathbb{E} [K_{x,h_n}(X_i)] = \mathbb{E} [K_{x,h_n}(X_1)] =: A_x^n + B_x^n$$

with

$$A_x^n := \int_{y \in \mathbb{S}_{x,h_n} \cap \mathbb{T}} K_{x,h_n}(y) f(y) \mathbb{1}_{\{y \in [x-r_{1,x}, x+r_{2,x}]\}} dy$$

and

$$B_x^n := \int_{y \in \mathbb{S}_{x,h_n} \cap \mathbb{T}} K_{x,h_n}(y) f(y) \mathbb{1}_{\{y \in \mathbb{R} \setminus [x-r_{1,x}, x+r_{2,x}]\}} dy.$$

For A_x^n : from the mean value theorem, we successively obtain

$$(3.3) \quad \begin{aligned} A_x^n &\leq \int_{y \in \mathbb{S}_{x,h_n} \cap \mathbb{T} \setminus \{x\}} K_{x,h_n}(y) |f(y) - f(x)| \mathbb{1}_{\{y \in [x-r_{1,x}, x+r_{2,x}]\}} dy + f(x) \\ &\leq (r_{2,x} + r_{1,x}) \int_{y \in \mathbb{S}_x \cap \mathbb{T} \setminus \{x\}} K_{x,h_n}(y) \max_{z \in [x-r_{1,x}, x+r_{2,x}] \cap \mathbb{T}} |f'(z)| \mathbb{1}_{\{y \in [x-r_{1,x}, x+r_{2,x}]\}} dy + f(x) \\ &\leq (r_{2,x} + r_{1,x}) \max_{z \in [x-r_{1,x}, x+r_{2,x}]} |f'(z)| \int_{y \in \mathbb{S}_x} K_{x,h_n}(y) dy + f(x) \\ &= (r_{2,x} + r_{1,x}) \max_{z \in [x-r_{1,x}, x+r_{2,x}]} |f'(z)| + f(x). \end{aligned}$$

For B_x^n : since f is increasing on $(-\infty, x-r_{1,x}]$ under (A0), and from (A1), we have

$$(3.4) \quad \begin{aligned} \int_{y \in (-\infty, x-r_{1,x}] \cap \mathbb{S}_x \cap \mathbb{T}} K_{x,h_n}(y) f(y) dy &= \mathbb{E} \left[f(Z_{x,h_n}) \mathbb{1}_{\{Z_{x,h_n} \leq x-r_{1,x}\}} \right] \\ &\leq \mathbb{E} \left[f(x-r_{1,x}) \mathbb{1}_{\{Z_{x,h_n} < x-r_{1,x}\}} \right] = f(x-r_{1,x}) \mathbb{P}(x - Z_{x,h_n} \geq r_{1,x}) \leq f(x-r_{1,x}). \end{aligned}$$

In the same way, f decreasing on $[x+r_{2,x}, \infty)$ provides

$$(3.5) \quad \begin{aligned} \int_{y \in [x+r_{2,x}, \infty) \cap \mathbb{S}_x \cap \mathbb{T}} K_{x,h_n}(y) f(y) dy &= \mathbb{E} \left[f(Z_{x,h_n}) \mathbb{1}_{\{Z_{x,h_n} \geq x+r_{2,x}\}} \right] \\ &\leq \mathbb{E} \left[f(x+r_{2,x}) \mathbb{1}_{\{Z_{x,h_n} \geq x+r_{2,x}\}} \right] = f(x+r_{2,x}) \mathbb{P}(Z_{x,h_n} - x \geq r_{2,x}) \leq f(x+r_{2,x}). \end{aligned}$$

The proof is now completed with (3.2)-(3.5). \square

In what follows, we state that $\widetilde{f}_n(x)$ is consistent and its proof is reported in Section 5. It improves a part of Theorem 1.2.

THEOREM 3.3. *Under Assumptions (A0)-(A1) with some $\alpha \geq 1$, if the sequence $(h_n)_{n \geq 1}$ is chosen such that $nh_n^\alpha \xrightarrow{n \rightarrow \infty} \infty$, then, for any $x \in \mathbb{T}$, we have*

$$\widetilde{f}_n(x) \xrightarrow[n \rightarrow \infty]{\mathbb{L}^2} f(x).$$

Eventually, we state the following central limit theorem and its proof is established in Section 5. That is a novelty and a first general result in the direction of continuous associated-kernel estimators.

THEOREM 3.4. *Under Assumptions (A0)-(A1) with some $\alpha \geq 1$, if the sequence $(h_n)_{n \geq 1}$ is chosen such that $\sqrt{nh_n^{(3/2)\alpha}} \xrightarrow{n \rightarrow \infty} 0$ and $nh_n^\alpha \xrightarrow{n \rightarrow \infty} \infty$, then, for any $x \in \mathbb{T}$, we have*

$$\sqrt{nh_n^\alpha} (\widetilde{f}_n(x) - f(x)) \xrightarrow[n \rightarrow \infty]{\mathcal{L}} \mathcal{N}\left(0, f(x) \lim_{n \rightarrow \infty} h_n^\alpha \int_{y \in \mathbb{T} \cap \mathbb{S}_x} K_{x, h_n}^2(y) dy\right),$$

where “ $\xrightarrow{\mathcal{L}}$ ” stands for “convergence in law”.

3.2. Standard normalized associated-kernel estimator. Relying on the discrete associated-kernel case treated by Esstafa et al. [7], we demonstrate in Section 5 the following proposition related on the normalizing random variable for continuous estimator (1.3) from a non-classical associated kernel.

PROPOSITION 3.5. *Under Assumptions (A0)-(A2) for some $\alpha \geq 1$, assume that $\mathbb{T} \subseteq \mathbb{S}_x$ and $nh_n^\alpha \rightarrow \infty$, then the normalizing random variable C_n converges in mean square to 1.*

In what follows, we show that $\widehat{f}_n(x)$ is consistent.

THEOREM 3.6. *Under Assumptions (A0)-(A2), for any $x \in \mathbb{T}$, we have*

$$\widehat{f}_n(x) \xrightarrow[n \rightarrow \infty]{\mathbb{P}} f(x),$$

where “ $\xrightarrow{\mathbb{P}}$ ” indicate “convergence in probability”.

PROOF. To begin with, we note that for $x \in \mathbb{T}$, the error is performed by

$$\widehat{f}_n(x) - f(x) = \frac{1}{C_n} \left\{ (\widetilde{f}_n(x) - f(x)) + (1 - C_n)f(x) \right\}.$$

From Theorem 3.3, we recall that $\widetilde{f}_n(x)$ converges in mean square to $f(x)$. Therefore, Proposition 3.5 and the Slutsky theorem complete the proof. \square

In addition, with our hypotheses, we are able to obtain a general central limit theorem for this normalized estimator $\widehat{f}_n(x)$ from the unnormalized one.

THEOREM 3.7. *Under Assumptions (A0)-(A2) with some $\alpha \geq 1$, considering $\mathbb{T} \subseteq \mathbb{S}_x$ and if the sequence $(h_n)_{n \geq 1}$ is chosen such that $\sqrt{nh_n}^{(3/2)\alpha} \xrightarrow{n \rightarrow \infty} 0$ and $nh_n^\alpha \xrightarrow{n \rightarrow \infty} \infty$, then, for any $x \in \mathbb{T}$, we have*

$$\sqrt{nh_n^\alpha} (\widehat{f}_n(x) - f(x)) \xrightarrow[n \rightarrow \infty]{\mathcal{L}} \mathcal{N} \left(0, f(x) \lim_{n \rightarrow \infty} h_n^\alpha \int_{y \in \mathbb{T} \cap \mathbb{S}_x} K_{x,h_n}^2(y) dy \right).$$

PROOF. Let us rewrite our normalized error as follows

$$(3.6) \quad \sqrt{nh_n^\alpha} (\widehat{f}_n(x) - f(x)) = \frac{1}{C_n} \left(\frac{\widetilde{f}_n(x) - \mathbb{E}[\widetilde{f}_n(x)]}{\sqrt{\text{Var}(\widetilde{f}_n(x))}} \right) \sqrt{nh_n^\alpha \text{Var}(\widetilde{f}_n(x))} \\ + \frac{1}{C_n} \sqrt{nh_n^\alpha} \{ \mathbb{E}[\widetilde{f}_n(x)] - f(x) \} + \frac{f(x)}{C_n} \sqrt{nh_n^\alpha} (1 - C_n).$$

Then, the convergence of the first term in the r.h.s. of (3.6) is direct by Proposition 3.2, Inequation (5.12) from the proof of the central limit theorem of the non-normalized error, and Slutsky's theorem. More precisely, we get

$$\frac{1}{C_n} \left(\frac{\widetilde{f}_n(x) - \mathbb{E}[\widetilde{f}_n(x)]}{\sqrt{\text{Var}(\widetilde{f}_n(x))}} \right) \sqrt{nh_n^\alpha \text{Var}(\widetilde{f}_n(x))} \xrightarrow[n \rightarrow \infty]{\mathcal{L}} \mathcal{N} \left(0, f(x) \lim_{n \rightarrow \infty} h_n^\alpha \int_{y \in \mathbb{T} \cap \mathbb{S}_x} K_{x,h_n}^2(y) dy \right).$$

The convergence to zero of the second term in the r.h.s. of (3.6) is also straightforward through Proposition 3.2 and (5.13). Therefore, it is left to show the convergence in probability to zero of the last term in the r.h.s. of (3.6). First, on the one hand, let us denote

$$S_n = \sqrt{nh_n^\alpha} (1 - C_n) = \sum_{i=1}^n Y_{i,n}, \quad Y_{i,n} = \frac{\sqrt{h_n^\alpha}}{\sqrt{n}} \int_{\mathbb{T}} \{f(x) - K_{x,h_n}(X_i)\} dx.$$

From Assumption (A2) with $\mathbb{T} \subseteq \mathbb{S}_x$, there exists a positive constant c such that $|Y_{i,n}| < c$. Then, by Hoeffding's inequality, we get for any $\varepsilon > 0$,

$$\mathbb{P}(|S_n - \mathbb{E}[S_n]| \geq \varepsilon) \leq c \exp\left(-\frac{\varepsilon^2}{n}\right) \xrightarrow{n \rightarrow \infty} 0.$$

On the other side, as $\sqrt{nh_n}^{(3/2)\alpha} \xrightarrow{n \rightarrow \infty} 0$, by (5.2) and (5.3), we obtain

$$(3.7) \quad \mathbb{E}[S_n] \leq \sqrt{nh_n^\alpha} |\mathbb{E}[1 - C_n]| \leq h_n^{(3/2)\alpha} \sqrt{n} \int_{x \in \mathbb{T}} g(x) dx \xrightarrow{n \rightarrow \infty} 0,$$

where the integrable function g is defined for any $x \in \mathbb{T}$ as

$$g(x) = \max \left\{ \max_{z \in [x-r_{1,x}, x+r_{2,x}]} |f'(z)|, \frac{\max\{|f(x-r_{1,x}) - f(x)|, |f(x+r_{2,x}) - f(x)|, f(x)\}}{(\min\{r_{1,x}, r_{2,x}\})^2} \right\}.$$

Thus, we deduce the convergence in probability to zero of $\sqrt{nh_n^\alpha} (C_n - 1)$ and also the one of the last term in the r.h.s. of (3.6). This completes the proof. \square

3.3. L^2 -comparison. For making a global comparison between \widehat{f}_n and \widetilde{f}_n , we here use the L^2 error instead of L^1 in Esstafa et al. [7] of discrete cases. The next proposition highlights the fact that our estimator outperforms, in the sense of the L^2 error, the unnormalized one. In the continuous framework of estimating a pdf with classical kernels, similar conclusions have been reached; see, for example, Devroye and Lugosi [4, section 5.6] and, also, Glad et al. [9] for L^2 criterion.

Notice that we need the convergence in L^4 of C_n to 1 for getting the following similar comparison between \widehat{f}_n and \widetilde{f}_n .

LEMMA 3.8. *Under Assumptions (A0)-(A2), assume that $\mathbb{T} \subseteq \mathbb{S}_x$, we have*

$$\mathbb{E}[(C_n - 1)^4] \xrightarrow{n \rightarrow \infty} 0.$$

PROOF. First, by (A2) and from the definition of C_n , one has $\mathbb{P}(|C_n| \leq M) = 1$. In view of the almost sure absolute boundedness of C_n when n is large enough, one gets $\mathbb{E}[|C_n|^k] \leq M^k$ for any $k \in \mathbb{N}^*$. Combined with the Cauchy-Schwarz inequality, we successively have

$$\begin{aligned} \text{Var}(C_n^2) &= \mathbb{E}[C_n^4 - 1] - \left\{ (\mathbb{E}[C_n^2])^2 - 1 \right\} \\ &\leq \left| \mathbb{E}[(C_n - 1)(C_n + 1)(C_n^2 + 1)] \right| + (\mathbb{E}[C_n^2] + 1) |\mathbb{E}[(C_n - 1)(C_n + 1)]| \\ &\leq \sqrt{\mathbb{E}[(C_n - 1)^2]} \left\{ \sqrt{\mathbb{E}[(C_n + 1)^2(C_n^2 + 1)^2]} + (\mathbb{E}[C_n^2] + 1) \sqrt{\mathbb{E}[(C_n + 1)^2]} \right\} \\ &\leq c \sqrt{\mathbb{E}[(C_n - 1)^2]} \end{aligned}$$

and

$$\text{Cov}(C_n^2, C_n) = \mathbb{E}[C_n^2(C_n - \mathbb{E}[C_n])] \leq \sqrt{\mathbb{E}[C_n^4] \text{Var}(C_n)} \leq c \sqrt{\text{Var}(C_n)},$$

where c denotes our generic constant that can change from line to line. Write again the considering fourth-moment as follows

$$\begin{aligned} \mathbb{E}[(C_n - 1)^4] &= \text{Var}\{(C_n - 1)^2\} + (\mathbb{E}[(C_n - 1)^2])^2 \\ &= \text{Var}(C_n^2) + 4\text{Var}(C_n) - 4\text{Cov}(C_n^2, C_n) + (\mathbb{E}[(C_n - 1)^2])^2. \end{aligned}$$

From the above arguments, when n is large enough, we obtain

$$\mathbb{E}[(C_n - 1)^4] \leq c \sqrt{\mathbb{E}[(C_n - 1)^2]} + 4\text{Var}(C_n) + 4c \sqrt{\text{Var}(C_n)} + (\mathbb{E}[(C_n - 1)^2])^2.$$

Hence, using the results from Proposition 3.5 that $\text{Var}(C_n)$ and $\mathbb{E}[(C_n - 1)^2]$ converge to zeros as n approaches infinity, we complete the proof. \square

Departing from Lemma 3.8, we obtain easily the comparison between the mean integrated squared error of the normalized continuous estimator \widehat{f}_n with the one of the unnormalized associated version \widetilde{f}_n in the below proposition.

PROPOSITION 3.9. *Under Assumptions (A0)-(A2), assume that \mathbb{T} is compact, for any $\varepsilon > 0$, there exists $N \in \mathbb{N}$ such that for all $n \geq N$,*

$$\mathbb{E} \left[\int_{\mathbb{T}} |\widehat{f}_n(x) - f(x)|^2 dx \right] < \mathbb{E} \left[\int_{\mathbb{T}} |\widetilde{f}_n(x) - f(x)|^2 dx \right] + \varepsilon.$$

PROOF. First, we successively have

$$\begin{aligned}
\mathbb{E} \left[\int_{\mathbb{T}} |\widetilde{f}_n(x) - f(x)|^2 dx \right] &= \mathbb{E} \left[\int_{\mathbb{T}} \left| \frac{\widetilde{f}_n(x)}{C_n} - f(x) \right|^2 dx \right] \\
&= \mathbb{E} \left[\int_{\mathbb{T}} \left| \frac{\widetilde{f}_n(x)}{C_n} - \frac{f(x)}{C_n} + \frac{f(x)}{C_n} - f(x) \right|^2 dx \right] \\
&= \mathbb{E} \left[\frac{1}{C_n^2} \int_{\mathbb{T}} |\widetilde{f}_n(x) - f(x)|^2 dx \right] + \int_{\mathbb{T}} f^2(x) dx \mathbb{E} \left[\left| \frac{C_n - 1}{C_n} \right|^2 \right] \\
&\quad + 2\mathbb{E} \left[C_n^{-2} \int_{\mathbb{T}} (\widetilde{f}_n(x) - f(x))(1 - C_n)f(x) dx \right] \\
&= \mathbb{E} \left[\int_{\mathbb{T}} |\widetilde{f}_n(x) - f(x)|^2 dx \right] + \mathbb{E} \left[\frac{1 - C_n^2}{C_n^2} \int_{\mathbb{T}} |\widetilde{f}_n(x) - f(x)|^2 dx \right] \\
&\quad + \mathbb{E} \left[\left| \frac{C_n - 1}{C_n} \right|^2 \right] \int_{\mathbb{T}} f^2(x) dx + 2\mathbb{E} \left[C_n^{-2} \int_{\mathbb{T}} (\widetilde{f}_n(x) - f(x))(1 - C_n)f(x) dx \right].
\end{aligned}$$

Since f is bounded on \mathbb{T} , one can use successively the Jensen and Cauchy-Schwarz inequalities combined with the fact that $\int_{\mathbb{T}} \widetilde{f}_n^2(x) dx \leq c C_n^2$ in order to obtain:

(3.8)

$$\begin{aligned}
&\mathbb{E} \left[\int_{\mathbb{T}} |\widetilde{f}_n(x) - f(x)|^2 dx \right] \\
&\leq \mathbb{E} \left[\int_{\mathbb{T}} |\widetilde{f}_n(x) - f(x)|^2 dx \right] + c \mathbb{E} \left[\frac{|1 - C_n^2|(C_n^2 + 1)}{C_n^2} \right] + c \mathbb{E} \left[\left| \frac{C_n - 1}{C_n} \right|^2 \right] \\
&\quad + c \mathbb{E} \left[C_n^{-2} |1 - C_n|(C_n + 1) \right] \\
&\leq \mathbb{E} \left[\int_{\mathbb{T}} |\widetilde{f}_n(x) - f(x)|^2 dx \right] + c \mathbb{E} \left[\left| \frac{C_n - 1}{C_n} \right|^2 \right] + c \mathbb{E} \left[\left| \frac{(1 - C_n^2)(C_n^2 + 1)}{C_n^2} \right| \right] \\
&\leq \mathbb{E} \left[\int_{\mathbb{T}} |\widetilde{f}_n(x) - f(x)|^2 dx \right] + c \mathbb{E} \left[\left| \frac{C_n - 1}{C_n} \right|^2 \right] + c \mathbb{E} \left[\left| \frac{(1 - C_n)(C_n + 1)(C_n^2 + 1)}{C_n^2} \right| \right] \\
&\leq \mathbb{E} \left[\int_{\mathbb{T}} |\widetilde{f}_n(x) - f(x)|^2 dx \right] + c \sqrt{\mathbb{E} [(C_n - 1)^4] \mathbb{E} [C_n^{-4}]} \\
&\quad + c \sqrt{\mathbb{E} [(1 - C_n)^2] \mathbb{E} \left[\frac{(C_n + 1)^2 (C_n^2 + 1)^2}{C_n^4} \right]} \\
&\leq \mathbb{E} \left[\int_{\mathbb{T}} |\widetilde{f}_n(x) - f(x)|^2 dx \right] + c \sqrt{\mathbb{E} [(C_n - 1)^4] \mathbb{E} [C_n^{-4}]} \\
&\quad + c \sqrt{\mathbb{E} [(1 - C_n)^2] \left\{ \mathbb{E} [C_n^2] + \mathbb{E} [C_n] + 1 + \mathbb{E} [C_n^{-1}] + \mathbb{E} [C_n^{-2}] + \mathbb{E} [C_n^{-3}] + \mathbb{E} [C_n^{-4}] \right\}}.
\end{aligned}$$

Then, on the one hand, from the almost sure absolute boundedness of C_n and C_n^{-1} combined with Proposition 3.2, there exists $n_0 \in \mathbb{N}$ such that for all $n \geq n_0$, $\mathbb{E}(C_n^j)$ with

$j \in \{-4, -3, \dots, 2\}$ are bounded by a finite constant $M > 0$. On the other hand, for some arbitrarily small $\varepsilon > 0$, from the proof of Theorem 3.5 and Lemma 3.8, there exists $n_1 \in \mathbb{N}$ such that for all $n \geq n_1$, we respectively have $\mathbb{E}[(C_n - 1)^2] < \varepsilon^2/(28Mc^2)$ and $\mathbb{E}[(C_n - 1)^4] < \varepsilon^2/(4c^2M)$. Therefore, taking $N = \max\{n_0, n_1\}$, for all $n \geq N$, the upper bound (3.8) provides the desired result of the proposition. \square

4. Simulation studies and an illustrative application. In this section, all numerical studies are performed in the context with standard and (non-)normalized associated-kernel estimators. Thus, we provide simulation results which are conducted for evaluating the performance of the proposed approaches. Computations have been done by using the R software of R Core Team [21]. The following numerical studies have two objectives with respect to the simulation schemes and to an application to original real data. We shall use the support of lognormal and Weibull kernel estimators.

4.1. *Numerical simulations.* Three following scenarios which are denoted by A, B and C are considered for the simulation studies.

- Scenario A is generated by using the gamma distribution

$$f_A(x) = xe^{-x}, \quad x \in \mathbb{R}_+;$$

- Scenario B comes from the Weibull distribution

$$f_B(x) = \frac{3}{2} \left(\frac{x}{2}\right)^2 \exp\left(-\left(\frac{x}{2}\right)^3\right), \quad x \in \mathbb{R}_+;$$

- Scenario C is from the lognormal distribution

$$f_C(x) = \frac{1}{0.25x\sqrt{2\pi}} \exp(-8(\log x)^2), \quad x \in \mathbb{R}_+.$$

We evaluate the performances of these continuous associated-kernel estimators with the cross-validation selections of the optimal bandwidth parameter. In fact, for all the considering associated-kernel estimators, $\alpha = 1$ and the optimal bandwidth of h by the cross-validation method is determined by

$$\tilde{h}_{cv} = \arg \min_{h>0} \left[\int_{x \in \mathbb{T}} \{\tilde{f}_n(x)\}^2 dx - \frac{2}{n} \sum_{i=1}^n \tilde{f}_{n,h,-i}(X_i) \right],$$

where $\tilde{f}_{n,h,-i}(X_i) = (n-1)^{-1} \sum_{\ell=1, \ell \neq i}^n K_{X_i, h}(X_\ell)$ is being computed as $\tilde{f}_n(X_i)$ without the observation X_i .

The efficiency of these estimators is examined via the empirical estimates of \widehat{C}_n , the integrated squared errors \widehat{ISE}_n (in normalized case) with

$$\widehat{C}_n := \frac{1}{N_{sim}} \sum_{t=1}^{N_{sim}} \int_{x \in \mathbb{T}} \tilde{f}_n(x) dx$$

and

$$\widehat{ISE}_n := \frac{1}{N_{sim}} \sum_{t=1}^{N_{sim}} \int_{x \in \mathbb{T}} \{\widehat{f}_n(x) - f(x)\}^2 dx,$$

where N_{sim} represents the number of replications and n corresponds to the sample size which shall be chosen to be small, medium and large.

Table 1 reports some empirical mean values of \widehat{C}_n and \widehat{ISE}_n with their standard deviations using $N_{sim} = 750$ replications from Scenarios A, B and C to the corresponding sample sizes $n = 100, 500, 1000$ having the support of lognormal and Weibull kernel estimators. For each given subsample and the continuous associated-kernel, we need to compute the related bandwidth \widetilde{h}_{cv} through the cross-validation method in advance and use it to produce $\widetilde{f}_n, \widehat{C}_n, \widehat{f}_n$ and, therefore, \widehat{ISE}_n . With regard to the three scenarios, the estimate of \widehat{C}_n converges to 1 and \widehat{ISE}_n is decreasing when the sample size increases. However, it is to be noticed that the lognormal kernel estimator is more accurate than the one based on the Weibull one; especially, when the sample size is relatively large.

TABLE 1
Empirical mean values of \widehat{C}_n and \widehat{ISE}_n with their standard deviations in parentheses for lognormal and the Weibull kernels under Scenarios A, B and C.

	n	$\widehat{C}_{n,Lognormal}$	$\widehat{ISE}_{n,Lognormal}$	$\widehat{C}_{n,Weibull}$	$\widehat{ISE}_{n,Weibull}$
A	100	0.9504 (0.0197)	0.0106 (0.0083)	1.1774 (0.0585)	0.0161 (0.0098)
	500	0.9754 (0.0091)	0.0031 (0.0035)	1.0893 (0.0331)	0.0060 (0.0033)
	1000	0.9818 (0.0081)	0.0020 (0.0034)	1.0669 (0.0285)	0.0038 (0.0019)
B	100	0.9866 (0.0065)	0.0159 (0.0145)	1.0721 (0.0202)	0.0233 (0.0145)
	500	0.9938 (0.0030)	0.0043 (0.0032)	1.0406 (0.0129)	0.0080 (0.0037)
	1000	0.9956 (0.0026)	0.0026 (0.0020)	1.0321 (0.0111)	0.0050 (0.0022)
C	100	0.9879 (0.0065)	0.0337 (0.0299)	1.0364 (0.0113)	0.0554 (0.0396)
	500	0.9930 (0.0028)	0.0093 (0.0073)	1.0207 (0.0069)	0.0186 (0.0106)
	1000	0.9940 (0.0025)	0.0055 (0.0042)	1.0160 (0.0054)	0.0116 (0.0059)

Now, we illustrate both empirical distributions of $\sqrt{nh_n}\{\widetilde{f}_n(3) - f_A(3)\}$ (Figure 1) and $\sqrt{nh_n}\{\widehat{f}_n(3) - f_A(3)\}$ (Figure 2) for the lognormal and Weibull kernels over 5,000 replications of Scenario A with the sample size $n = 1,000$ and the choice of bandwidth $h_n = n^{-2/3}$. It is to observe that we have similarity of behavior of the pointwise empirical distribution to both normalized and non-normalized estimators. This confirms the results of Theorems 3.4 and 3.7.

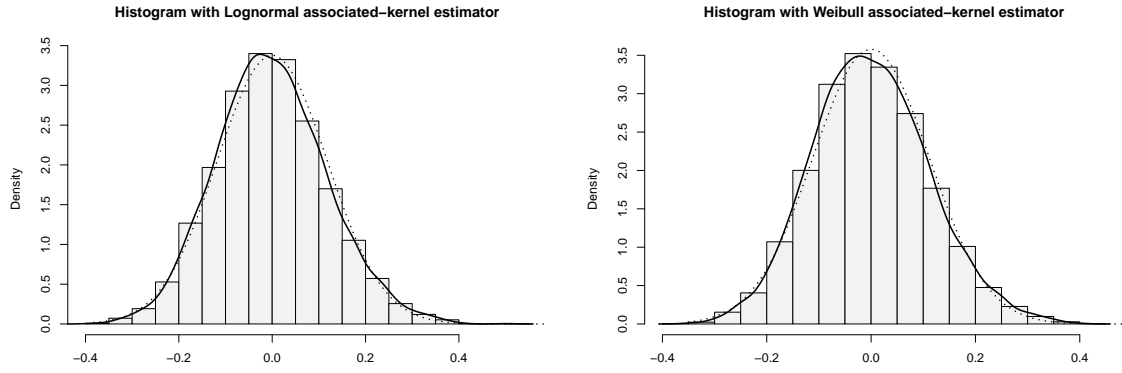


Fig 1: Empirical distributions of $\sqrt{nh_n}\{\tilde{f}_n(3) - f_A(3)\}$ using the non-normalized lognormal (left) and Weibull (right) kernel estimators. The smoothed kernel density is displayed in full line, and the centered Gaussian density with the limit variance of Theorem 3.4 is plotted in dotted line.

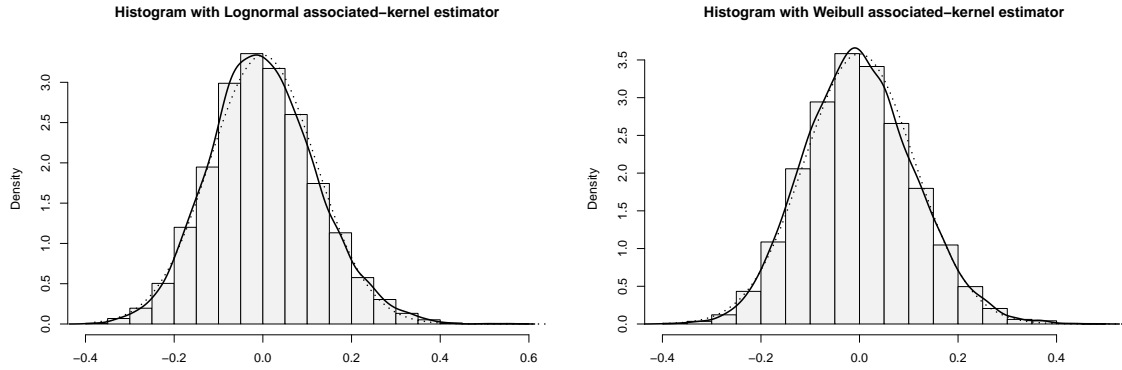


Fig 2: Empirical distributions of $\sqrt{nh_n}\{\widehat{f}_n(3) - f_A(3)\}$ using the normalized lognormal (left) and Weibull (right) kernel estimators. The smoothed kernel density is displayed in full line, and the centered Gaussian density with the limit variance of Theorem 3.7 is plotted in dotted line.

4.2. *An example of application on original data.* In this subsection, we compare the performance of our estimators on data of automobile claim amounts occurring between 2016 and 2020 provided by Covéa Affinity. This dataset is composed of $n = 76,414$ observations distributed between 3.6 and 6,985.51 EUR.

In Figure 3, we fit the empirical distribution of claim amounts using the normalized estimators, which are computed from the lognormal (solid line) and Weibull (dotted line) kernels. The optimal bandwidth h is selected by the cross-validation method.

It is pointed out that the lognormal kernel is more appropriate and provides a very satisfactory adjustment of the distribution of claim amounts (see also the below panel of Figure 3 for a zoom on the distribution of claim amounts between 3.6 and 800 EUR).

Table 2 confirms this finding for which other kernels (given in Section 2) have been implemented. Here, practical performances are examined via the cross-validation method and the empirical criterion of *ISE*: $\widehat{ISE}_0 := \int_0^\infty \{\widehat{f}_n(\mathbf{x}) - f_0(\mathbf{x})\}^2 dx$, where $f_0(\cdot)$ is the empirical or naive estimator. Recall that the estimators corresponding to the last three kernels of Table 2 are normalized by construction (their C_n is equal to 1). We notice by referring to their \widehat{ISE}_0 that they remain less efficient than the estimator with lognormal kernel but do slightly better than the one computed from the Weibull kernel.

Relying upon practical reasons, we evaluate the right tail of the distribution (from 1,000 EUR) of claim amounts using the different kernels of Table 2. These tail probabilities are estimated to be 0.1331, 0.1321, 0.1615, 0.1342, 0.1337 and 0.1344 for the empirical estimator f_0 , lognormal, Weibull, Gaussian, Gumbel and Laplace kernel estimations, respectively. Again, the lognormal kernel appears to be competitive and provides a very good fit to the tail distribution of this data set.

TABLE 2
Empirical values of \widehat{C}_n and \widehat{ISE}_0 of the normalized estimator of the distribution of claims amounts using different associated kernels.

Associated kernel	\widehat{C}_n	$\widehat{ISE}_0 (\times 10^4)$
Lognormal	0.9967	8.5482
Weibull	1.0700	9.1646
Gaussian	1.0000	9.0724
Gumbel	1.0000	9.0822
Laplace	1.0000	9.0383

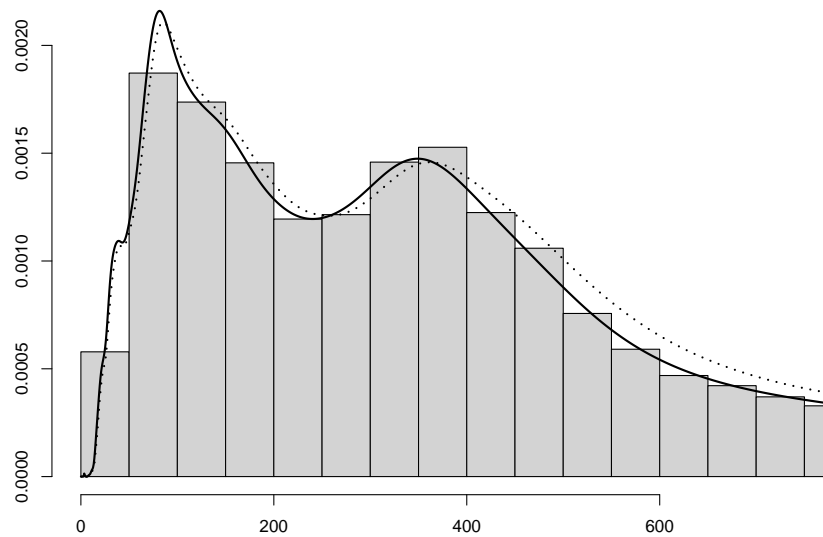
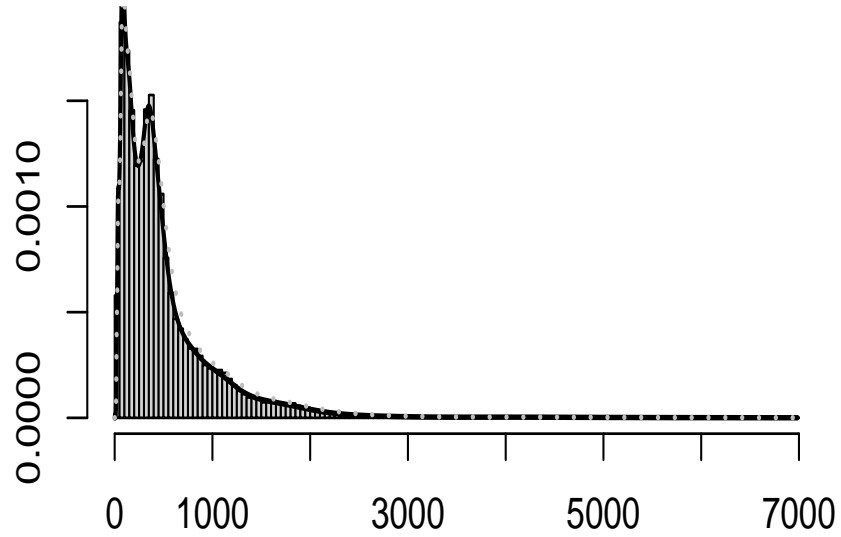


Fig 3: Empirical distribution of Covéa data and its estimates using both normalized lognormal (solid line) and Weibull (dotted line) kernel estimators. The graph on the below depicts a zoom of the region of claims amounts from 3.6 to 800 EUR of the graph on the top.

5. Proofs of Theorems 3.3 and 3.4 and of Proposition 3.5.

PROOF OF THEOREM 3.3. In order to obtain the theorem, for any $x \in \mathbb{T}$, we demonstrate the pointwise convergence to zero of the following mean-squared error

$$(5.1) \quad \mathbb{E} \left[\left(\widetilde{f}_n(x) - f(x) \right)^2 \right] = \left(\mathbb{E} \left[\widetilde{f}_n(x) - f(x) \right] \right)^2 + \text{Var} \left(\widetilde{f}_n(x) \right).$$

Considering the first term in the r.h.s. of (5.1): using $[x - r_{1,x}, x + r_{2,x}] \subset \mathbb{T}$ from Remark 2.1, we have

$$\mathbb{E} \left[\widetilde{f}_n(x) - f(x) \right] =: \mathbf{C}_{1,n} + \mathbf{C}_{2,n},$$

with

$$\mathbf{C}_{1,n} := \int_{y \in \mathbb{S}_x \cap [x - r_{1,x}, x + r_{2,x}] \setminus \{x\}} K_{x,h_n}(y) (f(y) - f(x)) dy$$

and

$$\mathbf{C}_{2,n} := \int_{y \in \mathbb{S}_x \cap \mathbb{T} \setminus \{x\} \cap [x - r_{1,x}, x + r_{2,x}]^c} K_{x,h_n}(y) (f(y) - f(x)) dy + f(x) \int_{y \in \mathbb{S}_x \cap \mathbb{T}^c} K_{x,h_n}(y) dy.$$

For the term $\mathbf{C}_{1,n}$: using the mean-value theorem and the Cauchy-Schwarz inequality, one has

$$\begin{aligned} |\mathbf{C}_{1,n}| &\leq \max_{z \in [x - r_{1,x}, x + r_{2,x}]} |f'(z)| \int_{y \in \mathbb{S}_x \cap [x - r_{1,x}, x + r_{2,x}] \setminus \{x\}} K_{x,h_n}(y) |y - x| dy \\ &\leq \max_{z \in [x - r_{1,x}, x + r_{2,x}]} |f'(z)| \mathbb{E} [|Z_{x,h_n} - x|] \\ &\leq \max_{z \in [x - r_{1,x}, x + r_{2,x}]} |f'(z)| \sqrt{\mathbb{E} [(Z_{x,h_n} - x)^2]} \\ (5.2) \quad &\leq \max_{z \in [x - r_{1,x}, x + r_{2,x}]} |f'(z)| \sqrt{\text{Var}(Z_{x,h_n}) + (\mathbb{E} [Z_{x,h_n}] - x)^2}. \end{aligned}$$

For the term $\mathbf{C}_{2,n}$: we successively use Remark 2.1, the fact that f is monotone on $[x - r_{1,x}, x + r_{2,x}]^c$ from Assumption (A0) and the Markov inequality to obtain

$$\begin{aligned} |\mathbf{C}_{2,n}| &\leq \int_{y \in \mathbb{S}_x \cap \mathbb{T} \setminus \{x\} \cap [x - r_{1,x}, x + r_{2,x}]^c} K_{x,h_n}(y) |f(y) - f(x)| dy + f(x) \int_{y \in \mathbb{S}_x \cap [x - r_{1,x}, x + r_{2,x}]^c} K_{x,h_n}(y) dy \\ &\leq \max\{|f(x - r_{1,x}) - f(x)|, |f(x + r_{2,x}) - f(x)|, f(x)\} \int_{y \in \mathbb{S}_x \cap [x - r_{1,x}, x + r_{2,x}]^c} K_{x,h_n}(y) dy \\ &\leq \max\{|f(x - r_{1,x}) - f(x)|, |f(x + r_{2,x}) - f(x)|, f(x)\} \mathbb{P}(|Z_{x,h_n} - x| > \min\{r_{1,x}, r_{2,x}\}) \\ (5.3) \quad &\leq \frac{\max\{|f(x - r_{1,x}) - f(x)|, |f(x + r_{2,x}) - f(x)|, f(x)\}}{(\min\{r_{1,x}, r_{2,x}\})^2} \left\{ \text{Var}(Z_{x,h_n}) + (\mathbb{E} [Z_{x,h_n}] - x)^2 \right\}. \end{aligned}$$

Therefore, through Remark 2.1 from Assumption (A0), we get

$$\begin{aligned} \left| \mathbb{E} \left[\widetilde{f}_n(x) - f(x) \right] \right| &\leq |\mathbf{C}_{1,n}| + |\mathbf{C}_{2,n}| \\ (5.4) \quad &\leq c \left\{ \sqrt{\text{Var}(Z_{x,h_n}) + (\mathbb{E} [Z_{x,h_n}] - x)^2} + \text{Var}(Z_{x,h_n}) + (\mathbb{E} [Z_{x,h_n}] - x)^2 \right\}. \end{aligned}$$

It follows

$$(5.5) \quad \mathbb{E}[\widetilde{f}_n(x)] = f(x) + \mathcal{O}(h_n^\alpha)$$

from Part (2.1) of Assumption (A1), which leads to the convergence to zero of the bias.

Considering the second term in the r.h.s. of (5.1): we have

$$\text{Var}(\widetilde{f}_n(x)) = \frac{1}{n} \text{Var}(K_{x,h_n}(X_1)) \leq \frac{1}{n} \mathbb{E}[K_{x,h_n}^2(X_1)].$$

In order to prove the convergence of this variance term, we demonstrate that

$$(5.6) \quad \mathbb{E}[K_{x,h_n}^2(X_1)] = f(x) \int_{y \in \mathbb{T} \cap \mathbb{S}_x} K_{x,h_n}^2(y) dy + \mathcal{O}(1).$$

Indeed, from Remark 2.1, we can write

$$\mathbb{E}[K_{x,h_n}^2(X_1)] - f(x) \int_{y \in \mathbb{T} \cap \mathbb{S}_x} K_{x,h_n}^2(y) dy =: \mathbf{D}_{1,n} + \mathbf{D}_{2,n}$$

with

$$\mathbf{D}_{1,n} := \int_{y \in \mathbb{S}_x \cap [x-r_{1,x}, x+r_{2,x}] \setminus \{x\}} K_{x,h_n}^2(y) (f(y) - f(x)) dy$$

and

$$\mathbf{D}_{2,n} := \int_{y \in \mathbb{S}_x \cap \mathbb{T} \cap [x-r_{1,x}, x+r_{2,x}]^c} K_{x,h_n}^2(y) (f(y) - f(x)) dy.$$

For the term $\mathbf{D}_{1,n}$: using the mean-value theorem, Part (2.2) of Assumption (A1) and the Cauchy-Schwarz inequality, we have

$$\begin{aligned} |\mathbf{D}_{1,n}| &\leq \max_{z \in [x-r_{1,x}, x+r_{2,x}]} |f'(z)| \max_{y \in \mathbb{S}_x \cap \mathbb{T}} K_{x,h_n}(y) \int_{y \in \mathbb{S}_x \cap [x-r_{1,x}, x+r_{2,x}] \setminus \{x\}} K_{x,h_n}(y) |y-x| dy \\ &\leq ch_n^{-\alpha} \mathbb{E}[|Z_{x,h_n} - x|] \\ &\leq ch_n^{-\alpha} \sqrt{\text{Var}(Z_{x,h_n}) + (\mathbb{E}[Z_{x,h_n}] - x)^2} \leq c. \end{aligned}$$

The last inequality is obtained through Part (2.1) of Assumption (A1).

For the term $\mathbf{D}_{2,n}$: Part (2.2) of Assumption (A1) and the Markov inequality entail that

$$\begin{aligned} |\mathbf{D}_{2,n}| &\leq \max_{y \in \mathbb{S}_x \cap \mathbb{T}} K_{x,h_n}(y) \int_{y \in \mathbb{S}_x \cap \mathbb{T} \setminus \{x\} \cap [x-r_{1,x}, x+r_{2,x}]^c} K_{x,h_n}(y) |f(y) - f(x)| dy \\ &\leq \max\{|f(x-r_{1,x}) - f(x)|, |f(x+r_{2,x}) - f(x)|\} h_n^{-\alpha} \int_{y \in \mathbb{S}_x \cap [x-r_{1,x}, x+r_{2,x}]^c} K_{x,h_n}(y) dy \\ &\leq ch_n^{-\alpha} \mathbb{P}(|Z_{x,h_n} - x| > \min\{r_{1,x}, r_{2,x}\}) \\ &\leq ch_n^{-\alpha} \frac{1}{(\min\{r_{1,x}, r_{2,x}\})^2} \{\text{Var}(Z_{x,h_n}) + (\mathbb{E}[Z_{x,h_n}] - x)^2\} \leq ch_n^\alpha. \end{aligned}$$

Once Again, the last inequality holds via Part (2.1) of Assumption (A1).

Therefore, Equation (5.6) is verified. And from Part (2.3) of Assumption (A1), one gets

$$\frac{1}{n} \mathbb{E} \left[K_{x,h_n}^2(X_1) \right] = \frac{1}{n} f(x) \int_{y \in \mathbb{T} \cap \mathbb{S}_x} K_{x,h_n}^2(y) dy + O(1/n) \leq c \frac{\max(h_n^{-\alpha}, 1)}{n},$$

which leads to the convergence to zero of the variance. The proof is completed. \square

PROOF OF THEOREM 3.4. To begin with, let us rewrite our error as follows

$$(5.7) \quad \sqrt{nh_n^\alpha} (\tilde{f}_n(x) - f(x)) = \left(\frac{\tilde{f}_n(x) - \mathbb{E}[\tilde{f}_n(x)]}{\sqrt{\text{Var}(\tilde{f}_n(x))}} \right) \sqrt{nh_n^\alpha \text{Var}(\tilde{f}_n(x))} + \sqrt{nh_n^\alpha} \{ \mathbb{E}[\tilde{f}_n(x)] - f(x) \}.$$

Concerning the first term in the r.h.s. of (5.7): first, let us identify the limit of the normalized variance using the results in (5.6) and (5.5). In particular, we have

$$(5.8) \quad \begin{aligned} nh_n^\alpha \text{Var}(\tilde{f}_n(x)) &= h_n^\alpha \text{Var}(K_{x,h_n}(X_1)) \\ &= h_n^\alpha \{ \mathbb{E} [K_{x,h_n}^2(X_1)] - (\mathbb{E} [K_{x,h_n}(X_1)])^2 \} \\ &= f(x) h_n^\alpha \int_{y \in \mathbb{T} \cap \mathbb{S}_x} K_{x,h_n}^2(y) dy + O(h_n^\alpha) - h_n^\alpha (\mathbb{E} [\tilde{f}_n(x)])^2 \\ &= f(x) h_n^\alpha \int_{y \in \mathbb{T} \cap \mathbb{S}_x} K_{x,h_n}^2(y) dy + O(h_n^\alpha). \end{aligned}$$

Therefore, Part (2.3) of Assumption (A1) implies

$$(5.9) \quad nh_n^\alpha \text{Var}(\tilde{f}_n(x)) \xrightarrow{n \rightarrow \infty} f(x) \lim_{n \rightarrow \infty} h_n^\alpha \int_{y \in \mathbb{T} \cap \mathbb{S}_x} K_{x,h_n}^2(y) dy < \infty.$$

Now, let $(Y_{n,i})$ be the rowwise i.i.d. triangular array defined by

$$Y_{n,i} = \frac{K_{x,h_n}(X_i) - \mathbb{E} [K_{x,h_n}(X_i)]}{\sqrt{n \text{Var}(K_{x,h_n}(X_i))}}, \quad i \in \{1, \dots, n\}.$$

It is clear that for all $n \geq 1$ and any $i \in \{1, \dots, n\}$,

$$\mathbb{E} [Y_{n,i}] = 0 \text{ and } \sum_{i=1}^n \mathbb{E} [Y_{n,i}^2] = 1.$$

Moreover, the Jensen inequality allows to write

$$(5.10) \quad \begin{aligned} \sum_{i=1}^n \mathbb{E} [|Y_{n,i}|^3] &= \frac{\mathbb{E} \left[|K_{x,h_n}(X_1) - \mathbb{E}(K_{x,h_n}(X_1))|^3 \right]}{n^{1/2} \{ \text{Var}(K_{x,h_n}(X_1)) \}^{3/2}} \\ &\leq \frac{c \mathbb{E} [|K_{x,h_n}(X_1)|^3]}{n^{1/2} \{ \text{Var}(K_{x,h_n}(X_1)) \}^{3/2}} = \frac{ch_n^{3\alpha/2} \mathbb{E} [K_{x,h_n}^3(X_1)]}{n^{1/2} \{ nh_n^\alpha \text{Var}(\tilde{f}_n(x)) \}^{3/2}}. \end{aligned}$$

Similar arguments as in (5.6) lead to

$$(5.11) \quad \mathbb{E} \left[K_{x,h_n}^3(X_1) \right] = f(x) \int_{y \in \mathbb{T} \cap \mathbb{S}_x} K_{x,h_n}^3(y) dy + \mathcal{O}(\max(h_n^{-2\alpha}, 1)).$$

Hence, we can write

$$\mathbb{E} \left[K_{x,h_n}^3(X_1) \right] - f(x) \int_{y \in \mathbb{T} \cap \mathbb{S}_x} K_{x,h_n}^3(y) dy =: \mathbf{E}_{1,n} + \mathbf{E}_{2,n}$$

with

$$\mathbf{E}_{1,n} := \int_{y \in \mathbb{S}_x \cap [x-r_{1,x}, x+r_{2,x}] \setminus \{x\}} K_{x,h_n}^3(y) (f(y) - f(x)) dy$$

and

$$\mathbf{E}_{2,n} := \int_{y \in \mathbb{S}_x \cap \mathbb{T} \cap [x-r_{1,x}, x+r_{2,x}]^c} K_{x,h_n}^3(y) (f(y) - f(x)) dy.$$

For the term $\mathbf{E}_{1,n}$: by the mean-value theorem, Part (2.2) of Assumption (A1) and the Cauchy-Schwarz inequality, we have

$$\begin{aligned} |\mathbf{E}_{1,n}| &\leq \max_{z \in [x-r_{1,x}, x+r_{2,x}]} |f'(z)| \left(\max_{y \in \mathbb{S}_x \cap \mathbb{T}} K_{x,h_n}(y) \right)^2 \int_{y \in \mathbb{S}_x \cap [x-r_{1,x}, x+r_{2,x}] \setminus \{x\}} K_{x,h_n}(y) |y-x| dy \\ &\leq c h_n^{-2\alpha} \mathbb{E} [|Z_{x,h_n} - x|] \\ &\leq c h_n^{-2\alpha} \sqrt{\text{Var}(Z_{x,h_n}) + (\mathbb{E} [Z_{x,h_n}] - x)^2} \leq c h_n^{-\alpha}. \end{aligned}$$

The last inequality is obtained via Part (2.1) of Assumption (A1).

For the term $\mathbf{E}_{2,n}$: Part (2.2) of Assumption (A1) and the Markov inequality allow to infer

$$\begin{aligned} |\mathbf{E}_{2,n}| &\leq \left\{ \max_{y \in \mathbb{S}_x \cap \mathbb{T}} K_{x,h_n}(y) \right\}^2 \int_{y \in \mathbb{S}_x \cap \mathbb{T} \setminus \{x\} \cap [x-r_{1,x}, x+r_{2,x}]^c} K_{x,h_n}(y) |f(y) - f(x)| dy \\ &\leq \max\{|f(x-r_{1,x}) - f(x)|, |f(x+r_{2,x}) - f(x)|\} h_n^{-2\alpha} \int_{y \in \mathbb{S}_x \cap [x-r_{1,x}, x+r_{2,x}]^c} K_{x,h_n}(y) dy \\ &\leq c h_n^{-2\alpha} \mathbb{P}(|Z_{x,h_n} - x| > \min\{r_{1,x}, r_{2,x}\}) \\ &\leq c h_n^{-2\alpha} \frac{1}{(\min\{r_{1,x}, r_{2,x}\})^2} \left\{ \text{Var}(Z_{x,h_n}) + (\mathbb{E} [Z_{x,h_n}] - x)^2 \right\} \leq c. \end{aligned}$$

Then, from Equation (5.11) and Part (2.3) of Assumption (A1), we get $\mathbb{E}[K_{x,h_n}^3(X_1)] = \mathcal{O}(\max(h_n^{-2\alpha}, 1))$ which is combined with (5.9) and $nh_n^\alpha \rightarrow \infty$ give

$$\frac{h_n^{3\alpha/2} \mathbb{E} \left[K_{x,h_n}^3(X_1) \right]}{n^{1/2} \left\{ nh_n^\alpha \text{Var}(\tilde{f}_n(x)) \right\}^{3/2}} \xrightarrow{n \rightarrow \infty} 0$$

and, the Lyapunov condition in (5.10) holds. According to the Lindeberg theorem, we obtain

$$\frac{\tilde{f}_n(x) - \mathbb{E}[\tilde{f}_n(x)]}{\sqrt{\text{Var}(\tilde{f}_n(x))}} \xrightarrow[n \rightarrow \infty]{\mathcal{L}} \mathcal{N}(0, 1).$$

Consequently, from (5.9) and using the Slutsky theorem, we get

$$(5.12) \quad \left(\frac{\widetilde{f}_n(x) - \mathbb{E}[\widetilde{f}_n(x)]}{\sqrt{\text{Var}(\widetilde{f}_n(x))}} \right) \sqrt{nh_n^\alpha \text{Var}(\widetilde{f}_n(x))} \xrightarrow[n \rightarrow \infty]{\mathcal{L}} \mathcal{N}\left(0, f(x) \lim_{n \rightarrow \infty} h_n^\alpha \int_{y \in \mathbb{T} \cap \mathcal{S}_x} K_{x,h_n}^2(y) dy\right).$$

As for the second term in the r.h.s. of (5.7): directly followed from (5.5) and the assumption that $\sqrt{nh_n^{(3\alpha)/2}} \rightarrow 0$ as $n \rightarrow \infty$, we have

$$(5.13) \quad \sqrt{nh_n^\alpha} \left| \mathbb{E}[\widetilde{f}_n(x) - f(x)] \right| \leq c \sqrt{nh_n^{(3\alpha)/2}} \xrightarrow[n \rightarrow \infty]{} 0,$$

which concludes the convergence to zero for the second term in the r.h.s. of (5.7) when n goes to infinity. This completes the proof. \square

PROOF OF PROPOSITION 3.5. To prove this proposition, we rewrite the considered mean square error as the sum of two terms

$$(5.14) \quad \mathbb{E}[(C_n - 1)^2] = \text{Var}(C_n) + (\mathbb{E}[C_n] - 1)^2.$$

Considering the first term $\text{Var}(C_n)$ in (5.14), we have the following development under the i.i.d. property of $(X_i)_{1 \leq i \leq n}$:

$$(5.15) \quad \begin{aligned} \text{Var}(C_n) &= \mathbb{E} \left[\int_{x,y \in \mathbb{T}} \widetilde{f}_n(x) \widetilde{f}_n(y) dx dy \right] - \mathbb{E} \left[\int_{x \in \mathbb{T}} \widetilde{f}_n(x) dx \right] \mathbb{E} \left[\int_{y \in \mathbb{T}} \widetilde{f}_n(y) dy \right] \\ &= \int_{x,y \in \mathbb{T}} \left\{ \mathbb{E}[\widetilde{f}_n(x) \widetilde{f}_n(y)] - \mathbb{E}[\widetilde{f}_n(x)] \mathbb{E}[\widetilde{f}_n(y)] \right\} dx dy \\ &= \frac{1}{n^2} \int_{x,y \in \mathbb{T}} \text{Cov} \left(\sum_{i=1}^n K_{x,h_n}(X_i), \sum_{i=1}^n K_{y,h_n}(X_i) \right) dx dy \\ &= \frac{1}{n^2} \int_{x,y \in \mathbb{T}} \sum_{i=1}^n \sum_{j=1}^n \text{Cov}(K_{x,h_n}(X_i), K_{y,h_n}(X_j)) dx dy \\ &= \frac{1}{n^2} \int_{x,y \in \mathbb{T}} \sum_{i=1}^n \text{Cov}(K_{x,h_n}(X_i), K_{y,h_n}(X_i)) dx dy \\ &= \frac{1}{n} \int_{x,y \in \mathbb{T}} \text{Cov}(K_{x,h_n}(X_1), K_{y,h_n}(X_1)) dx dy \\ &= \frac{1}{n} \int_{x \in \mathbb{T}} \text{Var}(K_{x,h_n}(X_1)) dx + \frac{1}{n} \int_{\substack{x,y \in \mathbb{T} \\ x \neq y}} \text{Cov}(K_{x,h_n}(X_1), K_{y,h_n}(X_1)) dx dy. \end{aligned}$$

Using Assumption (A2), the first term in (5.15) satisfies

$$(5.16) \quad \begin{aligned} \frac{1}{n} \int_{x \in \mathbb{T}} \text{Var}(K_{x,h_n}(X_1)) dx &\leq \frac{1}{n} \int_{y \in \mathbb{T}} \int_{x \in \mathbb{T}} K_{x,h_n}^2(y) dx f(y) dy \\ &\leq \frac{1}{nh_n^\alpha} \int_{y \in \mathbb{T}} f(y) dy \leq \frac{c}{nh_n^\alpha}, \end{aligned}$$

where c denotes our generic constant that can change from line to line. Similarly, using Assumption (A2), one easily see that the second term in (5.15) is bounded as follows

$$(5.17) \quad \begin{aligned} \frac{1}{n} \left| \int_{\substack{x,y \in \mathbb{T} \\ x \neq y}} \text{Cov}(K_{x,h_n}(X_1), K_{y,h_n}(X_1)) dx dy \right| &\leq \frac{1}{n} \int_{z \in \mathbb{T}} \int_{\substack{x,y \in \mathbb{T} \\ x \neq y}} K_{x,h_n}(z) K_{y,h_n}(z) dx dy f(z) dz \\ &\quad + \frac{1}{n} \int_{z \in \mathbb{T}} \int_{z' \in \mathbb{T}} \int_{\substack{x,y \in \mathbb{T} \\ x \neq y}} K_{x,h_n}(z) K_{y,h_n}(z') dx dy f(z') dz' f(z) dz \\ &\leq \frac{c}{n} \int_{z \in \mathbb{T}} f(z) dz + \frac{c}{n} \int_{z \in \mathbb{T}} \int_{z' \in \mathbb{T}} f(z') f(z) dz' dz \leq \frac{c}{n}. \end{aligned}$$

Therefore, we conclude from (5.16) and (5.17) that $\text{Var}(C_n) \xrightarrow[n \rightarrow \infty]{} 0$.

As for the second term $(\mathbb{E}[C_n] - 1)^2$ in the right hand side of (5.14): on the one hand, our analyses above gives that for any $x \in \mathbb{T}$ and from (5.5), $\mathbb{E}[\tilde{f}_n(x)] - f(x)$ converges to zero as n tends to infinity. On the other hand, from Proposition 3.2, $|\mathbb{E}[\tilde{f}_n(x) - f(x)]|$ is bounded by $c \max\{f(x + r_{2,x}), f(x - r_{1,x}), \max_{z \in [x-r_{1,x}, x+r_{2,x}]} |f'(z)|\}$ for any $x \in \mathbb{T}$ and $n \in \mathbb{N}$.

Furthermore, this upper bound is integrable on \mathbb{T} with respect to variable x from Assumption (A0). Therefore, we get

$$|\mathbb{E}[C_n - 1]| \leq \int_{x \in \mathbb{T}} |\mathbb{E}[\tilde{f}_n(x) - f(x)]| dx \xrightarrow[n \rightarrow \infty]{} 0,$$

where the last convergence to zero is obtained from the dominated convergence theorem. This completes the proof. \square

Acknowledgments. For the first two authors, this research benefited from the support of the ANR project "Efficient inference for large and high-frequency data" (ANR-21-CE40-0021), the "Chair Risques Émergents ou Atypiques en Assurance", under the aegis of Fondation du Risque, a joint initiative by Le Mans University and MMA company, member of Covéa group and the "Chair Impact de la Transition Climatique en Assurance", under the aegis of Fondation du Risque, a joint initiative by Le Mans University and Groupama Centre-Manche company, member of Groupama group.

For the third coauthor, this work is supported by the EIPHI Graduate School (contract ANR-17-EURE-0002). Part of this paper is done during the visit of the third author to the Department of Statistics, Universidade Federal do Rio Grande do Norte, Natal, RN, Brazil, which is supported by grant from "Brazilian-French Network in Mathematics".

REFERENCES

- [1] BOOS, D. D. (1985). A converse to Scheffe's theorem. *Annals of Statistics* **13** 423–427.
- [2] CHEN, S. X. (1999). Beta kernel estimators for density functions. *Computational Statistics and Data Analysis* **31** 131–145.
- [3] CHEN, S. X. (2000). Gamma kernel estimators for density functions. *Annals of the Institute of Statistical Mathematics* **52** 471–480.
- [4] DEVROYE, L. and LUGOSI, G. (2001). *Combinatorial Methods in Density Estimation*. Springer, New York.
- [5] DOHO, L. R. A., SOMÉ, S. M. and BANTO, J.-M. (2023). Inflation and west African sectoral stock price indices: an asymmetric kernel method analysis. *Emerging Markets Review* **54** 100987.
- [6] EPANECHNIKOV, V. (1969). Nonparametric estimates of a multivariate probability density. *Theory of Probability and Its Applications* **14** 153–158.
- [7] ESSTAFI, Y., KOKONENDJI, C. C. and SOMÉ, S. M. (2023). Asymptotic properties of the normalised discrete associated-kernel estimator for probability mass function. *Journal of Nonparametric Statistics* **35** 355–372.
- [8] GEENENS, G. (2021). Mellin-Meijer kernel density estimation on \mathbb{R}^+ . *Annals of the Institute of Statistical Mathematics* **73** 953–977.
- [9] GLAD, I. K., HJORT, N. L. and USHAKOV, N. G. (2003). Correction of density estimators that are not densities. *Scandinavian Journal of Statistics* **30** 415–427.
- [10] HIRUKAWA, M. (2018). *Asymmetric Kernel Smoothing - Theory and Applications in Economics and Finance*. Springer Briefs in Statistics, Singapore.
- [11] HIRUKAWA, M. and SAKUDO, M. (2015). Family of the generalised gamma kernels: a generator of asymmetric kernels for nonnegative data. *Journal of Nonparametric Statistics* **27** 41–63.
- [12] IGARASHI, G. (2023). A nonparametric discontinuity test of density using a beta kernel. *Journal of Nonparametric Statistics* in press (DOI: 10.1080/10485252.2022.2150766) pp.1–32.
- [13] JIN, X. and KAWCZAK, J. (2003). Birnbaum-Saunders and lognormal kernel estimators for modelling durations in high frequency financial data. *Annals of Economics and Finance* **4** 103–124.
- [14] KHAN, J. A. and AKBAR, A. (2021). Density estimation using Gumbel kernel estimator. *Open Journal of Statistics* **11** 319–328.
- [15] KOKONENDJI, C. C. and LIBENGUÉ DOBÉLÉ-KPOKA, F. G. B. (2018). Asymptotic results for continuous associated kernel estimators of density functions. *African Diaspora Journal of Mathematics* **21** 87–97.
- [16] KOKONENDJI, C. C. and SOMÉ, S. M. (2018). On multivariate associated kernels to estimate general density functions. *Journal of the Korean Statistical Society* **47** 112–126.
- [17] LIBENGUÉ DOBÉLÉ-KPOKA, F. G. B. and KOKONENDJI, C. C. (2017). The mode-dispersion approach for constructing continuous associated kernels. *Afrika Statistika* **12** 1417–1446.
- [18] MARCHANT, C., BERTIN, K., LEIVA, V. and SAULO, H. (2013). Generalized Birnbaum-Saunders kernel density estimators and an analysis of financial data. *Computational Statistics and Data Analysis* **63** 1–15.
- [19] OUIMET, F. and TOLOSANA-DELGADO, R. (2022). Asymptotic properties of Dirichlet kernel density estimators. *Journal of Multivariate Analysis* **187** 104832.
- [20] PARZEN, E. (1962). On estimation of a probability density function and Mode. *Annals of Mathematical Statistics* **33** 1065–1076.
- [21] R CORE TEAM (2022). R: A Language and Environment for Statistical Computing, Vienna, Austria: R Foundation for Statistical Computing. Available at <http://cran.r-project.org/>
- [22] SALHA, R. B., AHMED, H. I. E. S. and ALHOUBI, I. M. (2014a). Hazard rate function estimation using Weibull kernel. *Open Journal of Statistics* **4** 650–661.
- [23] SALHA, R. B., AHMED, H. I. E. S. and ALHOUBI, I. M. (2014b). Hazard rate function estimation using Erlang kernel. *Pure Mathematical Sciences* **3** 141–152.
- [24] SCAILLET, O. (2004). Density estimation using inverse and reciprocal inverse Gaussian kernels. *Journal of Nonparametric Statistics* **16** 217–226.
- [25] SCOTT, D. W. (1992). *Multivariate Density Estimation - Theory, Practice, and Visualization*. Wiley, New York.
- [26] SOMÉ, S. M. and KOKONENDJI, C. C. (2022). Bayesian selector of adaptive bandwidth for multivariate gamma kernel estimator on $[0, \infty)^d$. *Journal of Applied Statistics* **49** 1692–1713.
- [27] SOMÉ, M. S., KOKONENDJI, C. C., MAMODE KHAN, N. A., ADJABI, S. and BEDDEK, S. (2023). Multiple combined gamma kernel estimations for nonnegative data with Bayesian adaptive bandwidths. *Computational Statistics* in press (DOI :10.1007/s00180-023-01327-7) pp.1-33.
- [28] SWEETING, T. J. (1986). On a converse to Scheffe's theorem. *Annals of Statistics* **14** 1252–1256.

- [29] WAND, M. P. and JONES, M. C. (1994). *Kernel Smoothing*, Chapman & Hall, London.
- [30] WANSOUWÉ, W. E., SOMÉ, S. M. and KOKONENDJI, C. C. (2016). Ake: An R package for discrete and continuous associated kernel estimations. *The R Journal* **8** 258–276.
- [31] ZHANG, S. (2010). A note on the performance of the gamma kernel estimators at the boundary. *Statistics and Probability Letters* **80** 548–557.
- [32] ZHANG, S. and KARUNAMUNI, R. J. (2009). Boundary performance of the beta-kernel estimator. *Journal of Nonparametric Statistics* **22** 81–104.
- [33] ZIANE, Y., ZOUGAB, N. and ADJABI, S. (2018). Birnbaum-Saunders power-exponential kernel density estimation and Bayes local bandwidth selection for nonnegative heavy tailed data. *Computational Statistics* **33** 299–318.
- [34] ZOUGAB, N., ADJABI, S. and KOKONENDJI, C. C. (2014). Bayesian estimation of adaptive bandwidth matrices in multivariate kernel density estimation. *Computational Statistics and Data Analysis* **75** 28–38.

CONTENTS

1	Introduction	1
2	Assumptions and settings	4
3	Results on (non-)normalized estimators	9
	3.1 Unnormalized associated-kernel estimator	9
	3.2 Standard normalized associated-kernel estimator	11
	3.3 L^2 -comparison	13
4	Simulation studies and an illustrative application	15
	4.1 Numerical simulations	15
	4.2 An example of application on original data	17
5	Proofs of Theorems 3.3 and 3.4 and of Proposition 3.5	20
	Acknowledgments	25
	References	26