

Article

Not peer-reviewed version

Asymmetric Kernel Density Estimation for Biased Data

[Yoshihide Kakizawa](#) *

Posted Date: 28 June 2023

doi: 10.20944/preprints202306.1944.v1

Keywords: Biased data; nonparametric density estimation; boundary bias; asymmetric kernel



Preprints.org is a free multidiscipline platform providing preprint service that is dedicated to making early versions of research outputs permanently available and citable. Preprints posted at Preprints.org appear in Web of Science, Crossref, Google Scholar, Scilit, Europe PMC.

Copyright: This is an open access article distributed under the Creative Commons Attribution License which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Article

Asymmetric Kernel Density Estimation for Biased Data

Yoshihide Kakizawa ^{1,*},†¹ Graduate School of Economics and Business, Hokkaido University; kakizawa@econ.hokudai.ac.jp

* Correspondence: kakizawa@econ.hokudai.ac.jp

† Current address: Nishi 7, Kita 9, Kita-ku, Sapporo 060-0809, Japan.

Abstract: Nonparametric density estimation for nonnegative data is considered in a situation where a random sample is not directly available but the data are instead observed from the length-biased sampling. Due to the boundary bias problem of the location-scale kernel, the approach in this paper is an application of asymmetric kernel. Two nonparametric density estimators are proposed. The mean integrated squared error, strong consistency, and asymptotic normality of the estimators are investigated. Some simulations illustrate the finite sample performance of the estimators.

Keywords: Biased data; nonparametric density estimation; boundary bias; asymmetric kernel

1. Introduction

Nonparametric density estimation is a major issue of statistics and econometrics. Although the kernel density estimator (KDE) of location-scale type, originally proposed by Rosenblatt (1956) and Parzen (1962), is perhaps the most popular in the literature, its boundary bias problem is an important matter of concern when the support of the density to be estimated is not the whole real line \mathbb{R} . Various remedies for avoiding the boundary bias problem have been discussed, on the basis of renormalization, reflection, and generalized jackknifing (Jones (1993)), transformation (Marron and Ruppert (1994)), advanced reflection (Zhang et al. (1999)), and so on. Recently, there has been a vast literature on this subject, especially, with the renewed interest in revisiting an asymmetric kernel method after Chen (1999, 2000).

For the standard statistical inference problem of nonnegative data, it is assumed that a random sample $\{X_1, \dots, X_n\}$ of a size n is drawn from a population with density $f(x)$, $x \geq 0$. Practically, one encounters many situations where the available data are observed only under a certain biased sampling scheme (see, e.g., Patil and Rao (1978)). Then, it may be reasonable to regard a functional of the biased distribution as the inferential target. But, some analyses need the estimation referring to the original distribution. Cox (1969) considered estimating the population mean $\mu = \int_0^\infty tf(t) dt (> 0)$ and cumulative distribution function $F(x) = \int_0^x f(t) dt$, $x > 0$. An overview of nonparametric functional estimation under the biased sampling scheme is found in Cristóbal and Alcalá (2001).

Nonparametric density estimation from the biased data started in the late 1980s. Two important papers often quoted are Bhattacharyya et al. (1988) and Jones (1991) (see also Richardson et al. (1991) and Guillaumon et al. (1998)). Jones (1991) studied a kernel smoothing on the basis of Cox's (1969) distribution estimator, without cares about the boundary bias problem of the classical Rosenblatt–Parzen location-scale KDE. The main contribution of this paper is to revisit the nonparametric density estimation under the biased sampling scheme, using asymmetric kernel method which enables us to avoid the boundary bias problem and then have desirable asymptotic properties. Our approach is different from Mnatsakanov and Ruymgaart (2003, 2006) on moment-type density estimation motivated by the so-called moment problem, and Chaubey et al. (2010) using Hill's lemma (Feller (1971; (1.5) of page 220)).

The rest of this paper is organized as follows. Section 2 describes the length-biased (LB) distribution and illustrates, in detail, the boundary behavior of a convolution integral. After a brief introduction of the asymmetric kernel method, two density estimators from the LB data are proposed, in parallel with those of Bhattacharyya et al. (1988) and Jones (1991). Sections 3 and 4 state the required



assumptions and main results of this paper. All proofs are given in the appendix. In Section 5, some simulations illustrate the finite sample performance of the density estimators.

As usual, we use the notation $\|h\|_{\mathcal{S}} = \sup_{x \in \mathcal{S}} |h(x)|$ for any bounded function h on \mathcal{S} . We write, for $j \in \mathbb{N}$, $h^{(j)}(x) = (d/dx)^j h(x)$ (if it exists), and $h^{(0)}(x) = h(x)$. For an estimator $\hat{f}(x)$ of $f(x)$, where $x \geq 0$, the mean squared error (MSE) is defined by

$$MSE[\hat{f}(x)] = E[\{\hat{f}(x) - f(x)\}^2] = Bias^2[\hat{f}(x)] + V[\hat{f}(x)],$$

and the mean integrated squared error (MISE); $MISE[\hat{f}] = \int_0^\infty MSE[\hat{f}(x)] dx$ is a global measure of discrepancy of \hat{f} from f .

2. Preliminaries

2.1. LB Density

Nonparametrically, we wish to estimate the density f with nonnegative support, in a situation where a random sample $\{X_1, \dots, X_n\}$ is not directly available but a sample $\{Y_1, \dots, Y_n\} = \mathcal{Y}_n$ (say) is instead observed from the LB distribution having the density

$$f_{LB}(x) = \frac{xf(x)}{\mu}, \quad x \geq 0.$$

Throughout this paper, if there is no confusion, the X_i 's are iid copies of the random variable X having the density f , whereas the Y_i 's are iid copies of the random variable Y having the LB density f_{LB} . We repeatedly use the fact that, for a measurable function G and a real number r ,

$$E[Y^r G(Y)] = \int_0^\infty G(t) t^r f_{LB}(t) dt = \frac{1}{\mu} \int_0^\infty G(t) t^{r+1} f(t) dt = \frac{1}{\mu} E[X^{r+1} G(X)] \quad (\text{if it exists}).$$

2.2. Boundary Bias Problem and Asymmetric Kernel Method

When $\text{supp}(g) = [0, \infty)$, it is well known that a usual approximation (near the origin) of a certain smooth nonnegative function g as the convolution integral does not hold when $g(0) > 0$. We emphasize that the following fact is a starting point in the present paper: Even for the case $g(0) = 0$, the convergence rate near the origin $x = 0$ is slower when $g'(0) \neq 0$; a typical example is $g(x) = xe^{-x}$.

More precisely, if k is a symmetric density on $[-1, 1]$ (say) and $h = h_n > 0$ is a bandwidth which tends to zero as $n \rightarrow \infty$ (hereafter, we will omit the phrase "as the sample size n tends to the infinity" unless otherwise stated, and we denote by x the location where the density estimation is made), then, as shown in, e.g., Jones (1993),

$$\begin{aligned} & \int_0^\infty \frac{1}{h} k\left(\frac{x-s}{h}\right) g(s) ds - g(x) \\ &= \int_{-1}^{\min(x/h, 1)} k(t) g(x - ht) dt - g(x) \\ &= \begin{cases} \int_{-1}^1 k(t) g(x - ht) dt - g(x) \approx h^2 \frac{g''(x)}{2} \int_{-1}^1 t^2 k(t) dt, \\ \int_{-1}^p k(t) g(x - ht) dt - g(x) \approx -g(0) \int_p^1 k(t) dt + hg'(0) \int_p^1 (-p+t) k(t) dt, \end{cases} \end{aligned}$$

according to $x \geq h$ or $x = hp$ ($0 \leq p \leq 1$), because the kernel $k((x - \cdot)/h)/h$ creates a mass outside $[0, \infty)$ when the location x is at or near the origin. This is the motivation that, instead of the location-scale kernel $k((x - \cdot)/h)/h$, we focus on an application of an asymmetric kernel $k(\cdot; \beta, x)$, whose support matches the support of the target function, where $\beta = \beta_n > 0$ is a smoothing parameter, with $\beta \rightarrow 0$. It should be remarked that, after Chen (2000), the notation β (rather than h) is common in

the asymmetric kernel method; indeed, the parameter β for the asymmetric kernel under consideration has no meaning of the bandwidth h as in the classical Rosenblatt–Parzen location-scale kernel with a compact support, although β corresponds to h^2 and controls the bias-variance trade-off. This is why we refer to β as a smoothing parameter (rather than a bandwidth).

We formally require that

$$\int_0^\infty k(s; \beta, x) g(s) ds = g(x) + O(\beta) \quad \text{for any } x \geq 0,$$

if $\text{supp}(g) = [0, \infty)$. To the best of our knowledge, Silverman (1986; page 28) first mentioned a fairly simple idea of using gamma and log-normal (LN) kernels for the nonnegative data, where the kernel shape varies according to (β, x) . Perhaps, Chen's (2000) gamma kernel, defined by¹

$$k^{(G)}(s; \beta, x) = \frac{1}{\beta} \left(\frac{s}{\beta} \right)^{x/\beta} \frac{e^{-s/\beta}}{\Gamma(x/\beta)}, \quad s, x \geq 0,$$

plays a central role in this area; however, some simulations reveal that when the target density at the origin is zero, the gamma kernel is disadvantage compared to the LN kernel (e.g., Igarashi (2016)). Subsequent authors have discussed various kernels like LN, inverse Gaussian (IG), reciprocal IG (RIG), Birnbaum–Saunders (BS), and inverse gamma. The analysts can now choose what they like, among many options available for the kernel with support $[0, \infty)$. According to Igarashi and Kakizawa (2020) (see also Kakizawa (2021)), let us choose $k(\cdot; \beta, x)$ in the following form:

Definition 1 (Igarashi and Kakizawa (2020)). Given a baselined density $p(\cdot; \cdot)$, we set

$$k(s; \beta, x) = \frac{1}{\beta} p\left(\frac{s}{\beta}; \frac{x}{\beta}\right), \quad s, x \geq 0,$$

where the functional form of p , with nonnegative support, is independent of β and x .

To make this formulation clear, let

$$f_s^{(qBS)}(s; \theta_1, \theta_2, \theta_3) = C_g g\left(\left\{\frac{a^{(q)}(s/\theta_2)}{\theta_1} - \theta_3\right\}^2\right) \frac{A^{(q)}(s/\theta_2)}{\theta_1 \theta_2}, \quad s \geq 0,$$

be a symmetrical-based qBS density, associated with a density generator g , where $\theta_1, \theta_2 > 0, \theta_3 \in \mathbb{R}$,

$$a^{(q)}(t) = \begin{cases} \frac{1}{2q} (t^q - t^{-q}), & q \neq 0, \\ \log t, & q = 0, \end{cases} \quad \text{and } A^{(q)}(t) = \frac{1}{2} (t^{q-1} + t^{-(q+1)})$$

(due to the fact $(t^q - t^{-q})/q = (t^{|q|} - t^{-|q|})/|q|$, it is enough to take $q \geq 0$). Note that the density $C_g g(u^2)$ on \mathbb{R} is symmetric about the origin, where $1/C_g = \int_0^\infty y^{-1/2} g(y) dy$. Here, it is common to standardize g so that $\int_{-\infty}^\infty u^2 C_g g(u^2) du = 1$, without loss of generality. Indeed, as long as $\int_{-\infty}^\infty u^2 C_g g(u^2) du = J$ for some constant $J = J_g > 0$, this standardization can always be imposed with the replacement of $g(y)$ by $J^{1/2} g(Jy)$; the normalizing constant C_g is then invariant, i.e., $\int_{-\infty}^\infty J^{1/2} g(Ju^2) du = \int_{-\infty}^\infty g(t^2) dt = 1/C_g$.

¹ Clearly, the kernel is a rescaled version of the (standard) gamma density $s^{\theta-1} e^{-s} / \Gamma(\theta)$, with substitution of $x/\beta + 1$ for the shape parameter θ . Here, the shape θ is limited to be greater than or equal to 1, so as to ensure that the resulting kernel is bounded.

Given constants $q \geq 0$ and $c > 0$, a family of symmetrical-based non-central qBS kernels (Kakizawa (2021))² is defined by

$$k_g^{(qBS)}(s; \beta, x) = f_g^{(qBS)}\left(s; \sqrt{1/(x/\beta + c)}, \beta(x/\beta + c), \theta \sqrt{1/(x/\beta + c)}\right), \quad s, x \geq 0,$$

where $\theta \in \mathbb{R}$. Kakizawa (2018) considered the central case $\theta = 0$. Such a family $k_g^{(qBS)}$ is flexible via the (infinite-dimensional) density generator g , as well as the parameter $q \geq 0$. In some numerical studies of Section 5, we will put $q = 1/2$ (symmetrical-based BS kernel) or $q = 0$ (log-symmetrical (LS) kernel), with $(\theta, c) = (0, 1)$, for simplicity, and use the power exponential (PE) generator; $g_{PE[p]}(y) = \exp(-\lambda_p y^p)$, $p \geq 1/2$, where the particular choice $\lambda_p = \{\Gamma(3/(2p))/\Gamma(1/(2p))\}^p$ ensures that the PE density has the variance 1, like the standard normal density. Other generators (Kotz-type, generalized Pearson-type VII, and generalized logistic-type III) are found in Kakizawa (2018,2021). A symmetrical-based central qBS kernel belongs to a family of symmetrical-based qMIG kernels (MIG is an abbreviation of a mixture of IG and RIG), which is enlarged, linking to a class of skew-BS type kernels. See Kakizawa (2018,2021).

Remark 1. Chaubey and Li (2013) applied Scaillet's (2004) RIG kernel. As pointed out by Igarashi and Kakizawa (2014), an RIG KDE, however, also suffered from the boundary bias problem, so that the re-formulated RIG kernel should be applied.

2.3. Two density estimators under LB sampling scheme

Using $E[Y^{-1}] = \int_0^\infty t^{-1} f_{LB}(t) dt = 1/\mu$, as well as the relation

$$f(x) = \frac{\mu}{x} f_{LB}(x) = \frac{x^{-1} f_{LB}(x)}{E[Y^{-1}]}, \quad x > 0,$$

our first estimator is defined in parallel with that of Bhattacharyya et al. (1988), as follows:

$$\tilde{f}_{\beta, \epsilon}(x) = \frac{(nx)^{-1} \sum_{i=1}^n k(Y_i; \beta, x)}{n^{-1} \sum_{i=1}^n (Y_i + \epsilon)^{-1} + \epsilon}, \quad x > 0.$$

Except for a technical issue that needs to take a small $\epsilon \propto n^{-1/2}$, this estimator may be natural in the sense that $f_{LB}(x)$, $x > 0$, is consistently estimable by the asymmetric KDE $n^{-1} \sum_{i=1}^n k(Y_i; \beta, x)$ (Bhattacharyya et al. (1988) used the classical Rosenblatt-Parzen KDE $n^{-1} \sum_{i=1}^n (1/h)k((x - Y_i)/h)$). On the other hand, Jones's (1991) idea;

$$\frac{E[Y^{-1}(1/h)k((x - Y)/h)]}{E[Y^{-1}]} = \int_0^\infty \frac{1}{h} k\left(\frac{x-s}{h}\right) f(s) ds$$

(see Cox (1969)) is also reasonable. However, in order to solve its boundary bias problem unless $f(0) = f'(0) = 0$ (see Introduction), our second estimator

$$\hat{f}_{\beta, \epsilon}(x) = \frac{n^{-1} \sum_{i=1}^n Y_i^{-1} k(Y_i; \beta, x)}{n^{-1} \sum_{i=1}^n (Y_i + \epsilon)^{-1} + \epsilon}, \quad x \geq 0$$

² It is easy to see that

$$\begin{aligned} & \{k_g^{(qBS)}(s; \beta, x)\}^2 \\ &= \frac{(C_g^2/C_{g^2})}{\{\beta(x+c\beta)\}^{1/2}} f_{g^2}^{(qBS)}\left(s; \sqrt{1/(x/\beta + c)}, \beta(x/\beta + c), \theta \sqrt{1/(x/\beta + c)}\right) A^{(q)}\left(\frac{s}{\beta(x/\beta + c)}\right), \quad s, x \geq 0, \end{aligned}$$

provided that g^2 is also a density generator.

is proposed, for which some asymptotic properties will be studied in Subsection 4.1.

Before we proceed with description of our required assumptions (Section 3) and specific asymptotic results (Section 4), we highlight a novelty, compared with Jones (1991). Suppose that $f(0) = 0$ but $f'(0) \neq 0$ (we have an example $f(x) = xe^{-x}$, $x \geq 0$). Then,

- Jones's (1991) estimator $\hat{f}_{h,\text{Jones}}$, based on the location-scale kernel with the support $[-1, 1]$ (say), suffers from the boundary bias problem, i.e., $\text{Bias}[\hat{f}_{h,\text{Jones}}(x)] = O(h^2)$ for $x \geq h$ and $\text{Bias}[\hat{f}_{h,\text{Jones}}(x)] = O(h)$ for $0 \leq x \leq h$, and, as a result, it is shown that

$$\text{MISE}[\hat{f}_{h,\text{Jones}}] = O(h^3 + (nh)^{-1}) (= O(n^{-3/4}));$$

- our estimator $\hat{f}_{\beta,\epsilon}$ achieves the convergence rate $n^{-4/5}$ of the MISE (see Theorem 4).

We notice that $\hat{f}_{\beta,\epsilon}$ is more preferable, since $\tilde{f}_{\beta,\epsilon}$ has the factor x^{-1} (it is numerically unstable near the origin); besides, a rigorous error analysis for the (unweighted) MISE of $\tilde{f}_{\beta,\epsilon}$ seems to be hard technically (or it might be impossible), although the pointwise MSE (for $x > 0$) and weighted MISE of $\tilde{f}_{\beta,\epsilon}$ are tractable (see Subsection 4.2).

3. Assumptions

We use the notation $f_{-j}(\cdot) = f(\cdot) / (\cdot)^j$. In order to prove asymptotic properties of $\hat{f}_{\beta,\epsilon}$ (Subsection 4.1), the following set of assumptions, labeled as F, is imposed for the density f to be estimated:

- F. (i) 1. f is a twice continuously differentiable function on $[0, \infty)$, where f, f' , and f'' are bounded;
 2. f'' is a Hölder-continuous function (with exponent $0 < \eta \leq 1$) on $[0, \infty)$, i.e., there exists a constant $L > 0$, such that $|f''(u) - f''(v)| \leq L|u - v|^\eta$ for any $u, v \geq 0$.
 (ii) 1. f_{-1} is a bounded function on $[0, \infty)$;
 2. f_{-1} is a Hölder-continuous function (with exponent $0 < \eta' \leq 1$) on $[0, \infty)$.
 (iii) the inverse moment of X ; $E[X^{-1}] = \int_0^\infty t^{-1}f(t) dt = \int_0^\infty f_{-1}(t) dt = \mu'_{-1}$ (say) exists (note that $E[Y^{-2}] = \mu^{-1}E[X^{-1}]$).
 (iii') $E[X^{-(1+q)}] = \int_0^\infty f_{-(1+q)}(t) dt = \mu'_{-(1+q)}$ (say) exists for some constant $q > 0$.
 (iv) $\int_0^\infty \{f'(t)\}^2 dt$, $\int_0^\infty \{tf''(t)\}^2 dt$, and $\int_0^\infty f_{-3/2}(t) dt$ exist.

Remark 2. Under the boundedness of f_{-1} , given in F(ii.1), the density f to be estimated must have a constraint $f(0) = 0$ (note that F(iii') for some $q > 1$ implies $f_{-1}(0) = 0$). However, as illustrated earlier, even in the case $f(0) = 0$, Jones's (1991) estimator suffers from the boundary bias problem when $f'(0) \neq 0$ (we have an example $f(x) = xe^{-x}$, $x \geq 0$).

On the other hand, for $\tilde{f}_{\beta,\epsilon}$ (Subsection 4.2), we additionally make some assumptions on the corresponding LB density, labeled as F[†]:

- F[†]. 1. In addition to F(i.1), $f_{\text{LB}}, f'_{\text{LB}}$, and f''_{LB} are bounded functions³ on $[0, \infty)$;
 2. f''_{LB} is a Hölder-continuous function (with exponent $0 < \eta'' \leq 1$) on $[0, \infty)$.

Lastly, high-level conditions on the kernel $k(\cdot; \beta, x)$, labeled as A, are needed, whose details will be given in the top of Appendix. For notational simplicity, we write

$$B_g(x) = \zeta_{1,1}g'(x) + \frac{\zeta_{2,1}}{2}xg''(x), \quad V_g(x) = \zeta \frac{g(x)}{x^{1/2}},$$

³ It automatically means that f_{LB} is a Lipschitz-continuous function (i.e., Hölder-continuous function, with exponent $\eta = 1$) on $[0, \infty)$, with $f_{\text{LB}}(0) = 0$, and that f'_{LB} and f''_{LB} are continuous functions on $[0, \infty)$; besides, under F(ii.1), $f(0) = 0$ and $f'_{\text{LB}}(0) = 0$. Note that $f_{\text{LB}}^{(j)}(t) = \mu^{-1}\{jf^{(j-1)}(t) + tf^{(j)}(t)\}$, $j = 1, 2$.

where the constants $\zeta_{1,1}, \zeta_{2,1}$, and ζ ($\zeta_{2,1}, \zeta > 0$) appear in Assumption A1.2–3 of Appendix. It is convenient for us to define $J_{\beta}^{(k,g)}(x) = \int_0^{\infty} k(s; \beta, x) g(s) ds$, $B_{\beta}^{(k,g)}(x) = J_{\beta}^{(k,g)}(x) - g(x)$, and $J_{\beta}^{(k^2,g)}(x) = \int_0^{\infty} k^2(s; \beta, x) g(s) ds$. Obviously, the following inequalities hold:

$$J_{\beta}^{(k,g)}(x) \leq \|g\|_{[0,\infty)}, \quad |B_{\beta}^{(k,g)}(x)| \leq 2\|g\|_{[0,\infty)}, \quad J_{\beta}^{(k^2,g)}(x) \leq \left\{ \sup_{s \geq 0} k(s; \beta, x) \right\} J_{\beta}^{(k,g)}(x).$$

It should be remarked that most of the items in Assumption F (or F^{\dagger}) are needed to approximate $J_{\beta}^{(\cdot, \cdot)}(x)$, whose error analyse is found in Appendix. Roughly speaking, we obtain $J_{\beta}^{(k,f)}(x) \approx f(x) + \beta B_f(x)$ under F(i) and $\mu J_{\beta}^{(k^2, f^{-1})}(x) \approx \beta^{-1/2} \mu V_{f^{-1}}(x)$ under F(ii), for the estimator $\hat{f}_{\beta, \epsilon}$. Also, for the estimator $\tilde{f}_{\beta, \epsilon}$, it is shown that, under F^{\dagger} ,

$$\begin{aligned} \frac{\mu}{x} J_{\beta}^{(k, f_{LB})}(x) &\approx \frac{\mu}{x} [f_{LB}(x) + \beta B_{f_{LB}}(x)] = f(x) + \beta B(x), \\ \left(\frac{\mu}{x}\right)^2 J_{\beta}^{(k^2, f_{LB})}(x) &\approx \left(\frac{\mu}{x}\right)^2 \frac{V_{f_{LB}}(x)}{\beta^{1/2}} = \frac{\mu V_{f^{-1}}(x)}{\beta^{1/2}}, \end{aligned}$$

where $B(x) = (\mu/x) B_{f_{LB}}(x)$. Note that

$$B(x) = \zeta_{1,1} \{x^{-1} f(x) + f'(x)\} + \zeta_{2,1} \left\{ f'(x) + \frac{x}{2} f''(x) \right\} = B_f(x) + \zeta_{1,1} f_{-1}(x) + \zeta_{2,1} f'(x).$$

We can see that, under F(i.1 and iv), $\int_0^{\infty} B_f^2(x) dx = I^{B_f^2}$ (say) and $\int_0^{\infty} V_{f^{-1}}(x) dx = I^{V_{f^{-1}}}$ (say) are well-defined; besides⁴, $\int_0^{\infty} B^2(x) dx = I^{B^2}$ (say) is well-defined (we assume F(ii.1)).

4. Main Results

We assume that

B(i). $\beta = n^{-\ell} \ell(n)$, where ℓ is a (positive) slowly varying function.

Note that all powers of $\log y$ and a function $L(y)$ approaching a positive limit vary slowly. For the achievement of the optimal rate of the M(I)SE, $\beta = C n^{-2/5}$ must be feasible, at least, where $C > 0$ is a constant, independent of n .

In what follows, let $\epsilon = C' n^{-1/2}$, where $C' > 0$ is a constant, independent of n . We write

$$\begin{aligned} \omega_{\beta, \eta}(x) &= \beta^{3/2} x^{-1/2} + \beta^2 + (\beta x)^{1+\eta/2}, \\ \omega_{\beta, \eta}(x; g) &= \beta^{-1/2} V_g(x) (\beta x^{-1})^{1/2} + \chi_{\{0 < \eta < 1\}} (\beta x)^{(\eta-1)/2}, \end{aligned}$$

where χ_S is the indicator function of the set S .

⁴ Note that

$$\int_0^{\infty} f_{-1}(t) dt = \left(\int_0^1 + \int_1^{\infty} \right) f_{-1}(t) dt \leq \int_0^1 f_{-3/2}(t) dt + \int_1^{\infty} f(t) dt < \int_0^{\infty} f_{-3/2}(t) dt + 1.$$

This, together with F(ii.1), implies that $\int_0^{\infty} f_{-1}^2(t) dt$ exists.

4.1. Asymptotic Properties of $\hat{f}_{\beta,\epsilon}$

Theorem 1. Suppose that Assumptions A1, A2.1($\nu = 0, 1$); see Appendix, and F(i–iii) hold. Under $B(0 < \iota < 1)$, given constants $c_L > 0$ and $0 < \tau < 1$, we have

1. $\sup_{0 \leq x \leq c_L \beta^\tau} |\text{Bias}[\hat{f}_{\beta,\epsilon}(x)]| = O(\beta^\tau + n^{-1/2}),$
 $\sup_{0 \leq x \leq c_L \beta^\tau} V[\hat{f}_{\beta,\epsilon}(x)] = O(n^{-1}\beta^{-1} + n^{-1/2}\beta^{2\tau});$
2. for $x \geq c_L \beta^\tau,$
 $\text{Bias}[\hat{f}_{\beta,\epsilon}(x)] = \beta B_f(x) + R_\beta^{\text{Bias}}(x), \quad V[\hat{f}_{\beta,\epsilon}(x)] = n^{-1}\beta^{-1/2}\mu V_{f_{-1}}(x) + R_\beta^V(x),$

with

$$|R_\beta^{\text{Bias}}(x)| \leq M \left[\omega_{\beta,\eta}(x) + n^{-1/2} + n^{-1/2}\beta |B_f(x)| \right],$$

$$|R_\beta^V(x)| \leq M' \left[n^{-1} \{ \omega_{\beta,\eta'}(x; f_{-1}) + 1 + V_{f_{-1}}(x) \} + n^{-1/2} \{ \beta^2 B_f^2(x) + \omega_{\beta,\eta}^2(x) \} \right],$$

where $M, M' > 0$ are constants, independent of n, β , and x . Also, we have

$$\text{Bias}[\hat{f}_{\beta,\epsilon}(0)] = \beta f'(0) \int_0^\infty up(u;0) du + O(n^{-1/2} + \beta^2),$$

$$V[\hat{f}_{\beta,\epsilon}(0)] = n^{-1}\beta^{-1}\mu f_{-1}(0) \int_0^\infty p^2(u;0) du$$

$$+ O(n^{-1}\beta^{\eta'-1} + n^{-1} + n^{-3/2}\beta^{-1} + n^{-1/2}\beta^2).$$

Remark 3. (i) The asymptotic bias and variance of $\hat{f}_{\beta,\epsilon}(x)$ when x is near the origin; $x/\beta \rightarrow \kappa$, where $\kappa \geq 0$ is finite, can be obtained, as in Kakizawa (2018,2021). The details are omitted.

(ii) The pointwise MSE of $\hat{f}_{\beta,\epsilon}$ is a corollary of Theorem 1(2), as follows: For fixed $x > 0$,

$$\text{MSE}[\hat{f}_{\beta,\epsilon}(x)] = \text{AMSE}_x[\beta] + o(\beta^2 + n^{-1}\beta^{-1/2}),$$

where

$$\text{AMSE}_x[\beta] = \beta^2 B_f^2(x) + n^{-1}\beta^{-1/2}\mu V_{f_{-1}}(x)$$

$$\geq \frac{5}{4^{4/5}} \{B_f^2(x)\}^{1/5} \{\mu V_{f_{-1}}(x)\}^{4/5} n^{-4/5} \quad \text{if } B_f(x) \neq 0$$

(the equality holds iff $\beta = [\{\mu V_{f_{-1}}(x)\} / \{4B_f^2(x)\}]^{2/5} n^{-2/5}$).

(iii) We have $\text{MSE}[\hat{f}_{\beta,\epsilon}(0)] = O(\beta^2 + n^{-1}\beta^{-1}) (= O(n^{-2/3})$ if $f'(0)f_{-1}(0) \neq 0$).

Theorem 2. Suppose that Assumptions A1, A2.1($\nu = 0, 1$); see Appendix, and F(i–iii) hold. Under $B(0 < \iota < 1/2)$, we have $\hat{f}_{\beta,\epsilon}(x) \xrightarrow{a.s.} f(x)$ for fixed $x > 0$ (note that $\hat{f}_{\beta,\epsilon}(0) \xrightarrow{a.s.} f(0) = 0$).

Theorem 3. Suppose that Assumptions A1, A2.2($\nu = 0, 1$); see Appendix, and F(i–iii') hold.

(i) Under $B(0 < \iota < \min\{2q/(2+q), 1\})$, we have

$$(n\beta^{1/2})^{1/2} \{\hat{f}_{\beta,\epsilon}(x) - E[\hat{f}_{\beta,\epsilon}(x)]\} \xrightarrow{d} N(0, \mu V_{f_{-1}}(x)) \quad \text{for fixed } x > 0.$$

(ii) Under $B(0 < \iota < q/(2+q))$, we have

$$(n\beta)^{1/2} \{\hat{f}_{\beta,\epsilon}(0) - E[\hat{f}_{\beta,\epsilon}(0)]\} \xrightarrow{d} N\left(0, \mu f_{-1}(0) \int_0^\infty p^2(u;0) du\right).$$

For fixed $x > 0$, the replacement of $E[\hat{f}_{\beta,\epsilon}(x)]$ by $f(x)$ is a routine by combining Theorem 3(i) with a bias of Theorem 1(2): If Assumption F(iii') holds for some $q > 1/2$, and if $2/5 \leq \iota < \min\{2q/(2+q), 1\}$ (for the extreme case $\iota = 2/5$, assume $n\beta^{5/2} \rightarrow 0$), then,

$$(n\beta^{1/2})^{1/2}\{\hat{f}_{\beta,\epsilon}(x) - f(x)\} \xrightarrow{d} N(0, \mu V_{f-1}(x)) \quad \text{for fixed } x > 0.$$

Theorem 4. Suppose that Assumptions A1, A2.1($v = 0, 1$), A3($H = 6/\eta + 1 + \delta_0$); see Appendix, and F hold, where $\int_0^\infty t^{2(3/\eta+1)+\delta_0} f(t) dt$ exists for some constant $\delta_0 > 0$. Under B($0 < \iota < 1$), we have

$$MISE[\hat{f}_{\beta,\epsilon}] = AMISE[\beta] + o(\beta^2 + n^{-1}\beta^{-1/2}),$$

where

$$\begin{aligned} AMISE[\beta] &= \beta^2 I_f^{B_f^2} + n^{-1}\beta^{-1/2} \mu I^{V_{f-1}} \\ &\geq \frac{5}{4^{4/5}} (I_f^{B_f^2})^{1/5} (\mu I^{V_{f-1}})^{4/5} n^{-4/5} \quad \text{if } B_f(x) \neq 0 \end{aligned}$$

(the equality holds iff $\beta = \{(\mu I^{V_{f-1}})/(4I_f^{B_f^2})\}^{2/5} n^{-2/5} = \beta_{\text{opt}}$ (say)).

4.2. Asymptotic Properties of $\tilde{f}_{\beta,\epsilon}$

Due to the presence of the factor x^{-1} in the estimator $\tilde{f}_{\beta,\epsilon}$, the case $x = 0$ is excluded throughout this subsection; besides, in Theorem 8 (see also Remark 5), we will consider a truncated MISE of $\tilde{f}_{\beta,\epsilon}$, for the global performance.

Theorem 5. Suppose that Assumptions A1, A2.1($v = 0$); see Appendix, F(ii.1 and iii), and F^\dagger hold. Under B($0 < \iota < 1$), given constants $c_L > 0$ and $0 < \tau < 1$, we have, for $x \geq c_L \beta^\tau$,

$$\text{Bias}[\tilde{f}_{\beta,\epsilon}(x)] = \beta B(x) + R_\beta^{\dagger \text{Bias}}(x), \quad V[\tilde{f}_{\beta,\epsilon}(x)] = n^{-1}\beta^{-1/2} \mu V_{f-1}(x) + R_\beta^{\dagger V}(x),$$

with

$$\begin{aligned} |R_\beta^{\dagger \text{Bias}}(x)| &\leq M^\dagger \left[x^{-1} \omega_{\beta,\eta''}(x) + n^{-1/2}(1+x^{-1}) + n^{-1/2} \beta |B(x)| \right], \\ |R_\beta^{\dagger V}(x)| &\leq M^{\dagger'} \left[n^{-1} \{ \omega_{\beta,\eta'}(x; f_{-1}) + 1 + x^{-2} + V_{f-1}(x) \} \right. \\ &\quad \left. + n^{-1/2} \{ \beta^2 B^2(x) + x^{-2} \omega_{\beta,\eta''}^2(x) \} \right], \end{aligned}$$

where $M^\dagger, M^{\dagger'} > 0$ are constants, independent of n, β , and x .

Remark 4. As a corollary of Theorem 5, we have, for fixed $x > 0$,

$$MSE[\tilde{f}_{\beta,\epsilon}(x)] = AMSE_x^\dagger[\beta] + o(\beta^2 + n^{-1}\beta^{-1/2}),$$

where

$$\begin{aligned} AMSE_x^\dagger[\beta] &= \beta^2 B^2(x) + n^{-1}\beta^{-1/2} \mu V_{f-1}(x) \\ &\geq \frac{5}{4^{4/5}} \{B^2(x)\}^{1/5} \{\mu V_{f-1}(x)\}^{4/5} n^{-4/5} \quad \text{if } B(x) \neq 0 \end{aligned}$$

(the equality holds iff $\beta = [\{\mu V_{f-1}(x)\} / \{4B^2(x)\}]^{2/5} n^{-2/5}$).

Theorem 6. Suppose that Assumptions A1, A2.1($v = 0$); see Appendix, F(ii.1 and iii), and F^+ hold. Under $B(0 < \iota < 1)$, we have $\tilde{f}_{\beta,\epsilon}(x) \xrightarrow{a.s.} f(x)$ for fixed $x > 0$.

Theorem 7. Suppose that Assumptions A1, A2.2($v = 0$); see Appendix, F(ii.1 and iii), and F^+ hold. Under $B(0 < \iota < 1)$, we have

$$(n\beta^{1/2})^{1/2}\{\tilde{f}_{\beta,\epsilon}(x) - E[\tilde{f}_{\beta,\epsilon}(x)]\} \xrightarrow{d} N(0, \mu V_{f-1}(x)) \quad \text{for fixed } x > 0.$$

For fixed $x > 0$, the replacement of $E[\tilde{f}_{\beta,\epsilon}(x)]$ by $f(x)$ is a routine by combining Theorem 7 with a bias of Theorem 6: If $2/5 \leq \iota < 1$ (for the extreme case $\iota = 2/5$, assume $n\beta^{5/2} \rightarrow 0$), then,

$$(n\beta^{1/2})^{1/2}\{\tilde{f}_{\beta,\epsilon}(x) - f(x)\} \xrightarrow{d} N(0, \mu V_{f-1}(x)) \quad \text{for fixed } x > 0.$$

Theorem 8. Suppose that Assumptions A1, A2.1($v = 0$), A3($H = 2/\eta'' + 1 + \delta_0$); see Appendix, F(ii.1, iii, and iv), and F^+ hold, where $\int_0^\infty t^{2(1/\eta''+1)+\delta_0} f_{LB}(t) dt$ exists for some constant $\delta_0 > 0$. Under $B(0 < \iota < 1)$, we have, for every $0 < \tau < 1/2$,

$$\int_{\beta^\tau}^\infty \text{MSE}[\tilde{f}_{\beta,\epsilon}(x)] dx = \text{AMISE}^+[\beta] + o(\beta^2 + n^{-1}\beta^{-1/2}),$$

where

$$\begin{aligned} \text{AMISE}^+[\beta] &= \beta^2 I^{B^2} + n^{-1}\beta^{-1/2}\mu I^{V_{f-1}} \\ &\geq \frac{5}{4^{4/5}} (I^{B^2})^{1/5} (\mu I^{V_{f-1}})^{4/5} n^{-4/5} \quad \text{if } B(x) \not\equiv 0 \end{aligned}$$

(the equality holds iff $\beta = \{(\mu I^{V_{f-1}})/(4I^{B^2})\}^{2/5} n^{-2/5}$).

Remark 5. Whether or not there exists a $0 < \tau < 1/2$, such that $\int_0^{\beta^\tau} \text{MSE}[\tilde{f}_{\beta,\epsilon}(x)] dx$ is order $o(\beta^2 + n^{-1}\beta^{-1/2})$ (under some additional unnecessary stronger conditions) would be rather technical. We do not pursue the issue any more.

5. Simulation Studies

To demonstrate the finite sample performance of the proposed density estimator $\hat{f}_{\beta,\epsilon}$, we generated 1000 random samples of size $n = 200, 300, 500$ from the LB density $f_{LB}(x) = x^2 e^{-x}/2$, and computed the PE[p]-based BS/log KDEs ($p = 1, 3/2$) and gamma KDE for the original density $f(x) = x e^{-x}$. In the simulation, we used the least squared cross-validated (LSCV) smoothing parameter, for each sample. Then, the average integrated squared errors (ISEs), $(1/1000) \sum_{\ell=1}^{1000} \int_0^\infty \{\hat{f}_{\beta,\epsilon,[\ell]}(x) - f(x)\}^2 dx$, were reported in Table 1, where $\hat{f}_{\beta,\epsilon,[\ell]}$ is computed from the ℓ th sample.

Table 1. Average ISEs $\times 1000$ of the estimators using the LSCV selected smoothing parameter. The value in the parenthesis stands for the standard deviation.

n	PE[p]-based BS KDE		PE[p]-based log KDE		Gamma KDE
	$p = 1$	$p = 3/2$	$p = 1$	$p = 3/2$	
200	12.882	12.440	12.678	12.128	12.698
	(14.863)	(14.960)	(15.083)	(14.326)	(13.601)
300	9.767	9.699	9.809	9.833	9.939
	(11.067)	(11.587)	(11.414)	(12.160)	(10.339)
500	6.896	6.703	7.123	6.886	7.567
	(7.873)	(7.376)	(8.257)	(7.973)	(8.373)

As expected, all average ISEs decreased as the sample size n increased, which is in agreement with the MISE result. The BS/LN KDEs ($p = 1$) were, overall, improved by the estimators with $p = 3/2$; such a tendency can be illustrated via the AMISE relative efficiency index

$$\frac{\text{AMISE}_{\text{opt}}(p)}{\text{AMISE}_{\text{opt}}(1)} = \left\{ 2^{1-1/(2p)} p \sqrt{\pi} \frac{\Gamma^{1/2}(3/(2p))}{\Gamma^{3/2}(1/(2p))} \right\}^{4/5}$$

(see Kakizawa (2018,2021)), since

$$\text{AMISE}_{\text{opt}}(p) = \frac{5}{4} (I^{B_f^2})^{1/5} \left[\left(\frac{C_{\mathcal{S}_{\text{PE}[p]}^2}}{C_{\mathcal{S}_{\text{PE}[p]}^2}} \right) \mu \int_0^\infty f_{-3/2}(x) dx \right]^{4/5} n^{-4/5}.$$

Needless to say, the best implemented smoothing parameter β_{opt} , given in Theorem 4, depends on the unknown f , so that the data-driven procedure is crucial. We tried to conduct the LSCV smoothing parameter selection⁵. Unlike the direct sample (Kakizawa (2018,2021)), the present LB setting, for small sample size $n = 100$ (not being reported here), produced multiple local minima for the LSCV score (in many cases, it was rather unstable numerically), whereas such an undesirable behavior seemed to be fixed when $n = 300$. A further issue of considering a plug-in selection with a pilot estimator is left in future.

6. Discussion

Our asymptotic results under the LB sampling can be extended to more general biased sampling, i.e., a weighted distribution for a known (positive) weight function ω , given by $f_\omega(x) \propto \omega(x)f(x)$. The LB density is a special case of $\omega(x) = x$, and another example is $\omega(x) = x^2$ (the area-biased density). Also, the d -variate weighted density is defined by $f_\omega(\mathbf{x}) \propto \omega(\mathbf{x})f(\mathbf{x})$, $\mathbf{x} = (x_1, \dots, x_d)'$. Ahmad (1995) extended Jones's (1991) estimator to the d -variate case. Note that the product-kernel-method, using the product asymmetric kernel $\prod_{j=1}^d k(X_{ij}; \beta_j, x_j)$, $\mathbf{x} \in [0, \infty)^d$, instead of $\prod_{j=1}^d k((x_j - X_{ij})/h_j)/h_j$, or the non-product kernel (Igarashi (2018) and Kakizawa (2022)), can be straightforwardly applied to solve the boundary bias problem of Ahmad's (1995) estimator.

Funding: The author has been supported in part by the Japan Society for the Promotion of Science; Grant-in-Aid for Scientific Research (C), 20K11700 and 23K11002.

Data Availability Statement: Not available.

Acknowledgments: Some preliminary results were first announced, without face-to-face talking (due to the pandemic of covid-19), at Japanese Joint Statistical Meeting 2021 (Japanese federation of statistical science associations), autumn meeting 2021 (Mathematical Society of Japan), and the 5th International Conference on Econometrics and Statistics (EcoStat2022).

Conflicts of Interest: The author declares no conflict of interest.

Abbreviations

Abbreviations

The following abbreviations are used in this manuscript:

⁵ The rule of thumb (ROT) procedure, with the gamma or LN reference, was also considered. But, these results were not reported here, since, not surprisingly, the ROT was non-robust for the misspecification, although the computational speed of the ROT was very high.

KDE	kernel density estimator
LB	length-biased
MSE	mean squared error
MISE	mean integrated squared error
LN	log-normal
IG	inverse Gaussian
RIG	reciprocal inverse Gaussian
BS	Birnbaum–Saunders
LS	log-symmetrical
PE	power exponential
MIG	mixture of IG and RIG
LSCV	least squared cross-validated
ISE	integrated squared error
ROT	rule of thumb

Appendix A

Appendix A.1 Technical Conditions on $k(\cdot; \beta, x)$

There are three indispensable requirements on the kernel $k(s; \beta, x)$, $s, x \geq 0$:

- (I) approximations of $\mu_j(k(\cdot; \beta, x)) = \int_0^\infty (s-x)^j k(s; \beta, x) ds$ and $\int_0^\infty k^2(s; \beta, x) ds$;
- (II) uniform/nonuniform bounds of $\sup_{s \geq 0} k(s; \beta, x)$;
- (III) behavior of the tail integral of $k(s; \beta, x)$ with respect to x (the regularity as $x \rightarrow \infty$ will be required only for dealing with an asymptotic expansion of the MISE rigorously).

More precisely, we assume (e.g., Igarashi and Kakizawa (2020) and Kakizawa (2021)):

A1. In addition to Definition 1, there exists a density $p(\cdot; \cdot)$, such that

1. $\int_0^\infty u^2 p(u; 0) du$ exists and, for any $y \geq 0$, $\int_0^\infty u p(u; y) du (\leq \tilde{C}(1+y))$ exists, where $\tilde{C} > 0$ is a constant, independent of y ;
2. given constants $0 \leq \tilde{\eta} < 1$ and $c_L > 0$, for all sufficiently small $\beta > 0$, $x \geq c_L \beta^{\tilde{\eta}}$ implies that

$$\mu_j(k(\cdot; \beta, x)) = \begin{cases} \beta \zeta_{1,1} + r_{1,\beta}(x), & j = 1, \\ \beta \zeta_{2,1} x + r_{2,\beta}(x), & j = 2, \\ r_{4,\beta}(x), & j = 4, \end{cases}$$

with $|r_{1,\beta}(x)| \leq \tilde{M}_1 \beta^{3/2} / x^{1/2}$, $|r_{2,\beta}(x)| \leq \tilde{M}_2 \beta^2$, $0 < r_{4,\beta}(x) \leq \tilde{M}_4 \beta^2 (x + \beta)^2$, where $\zeta_{1,1}, \zeta_{2,1}$ ($\zeta_{2,1} > 0$) and $\tilde{M}_1, \tilde{M}_2, \tilde{M}_4 > 0$ are constants, independent of β and x ;

3. $\int_0^\infty p^2(u; 0) du$ exists, and, given constants $0 < \tilde{\eta} < 1$ and $c_L > 0$, for all sufficiently small $\beta > 0$, $x \geq c_L \beta^{\tilde{\eta}}$ implies that

$$\left| \int_0^\infty k^2(s; \beta, x) ds - \frac{\zeta}{(\beta x)^{1/2}} \right| \leq \frac{\tilde{M}}{(\beta x)^{1/2}} \left(\frac{\beta}{x + \beta} \right)^{1/2},$$

where $\zeta, \tilde{M} > 0$ are constants, independent of β and x .

A2(ν). $u_{\beta,\nu}(x) = \sup_{s \geq 0} \{(\beta/s)^\nu k(s; \beta, x)\}$ satisfies:

1. $\sup_{x \geq 0} u_{\beta,\nu}(x) \leq L_{K,\nu} \beta^{-1}$, where $L_{K,\nu} > 0$ is a constant, independent of β ;
2. for $x > 0$, $u_{\beta,\nu}(x) \leq L'_{K,\nu} (\beta x)^{-1/2}$, where $L'_{K,\nu} > 0$ is a constant, independent of β and x .

A3(H). Given a constant $\tau > 0$, and for all sufficiently small $\beta > 0$,

$$\int_0^\infty \left(\int_{\beta^{-\tau}}^\infty k(s; \beta, x) g(s) dx \right) ds = O(\beta^{\tau(H+1)})$$

(assume that $\int_0^\infty s^{H+1} g(s) ds$ exists).

Most of the existing asymmetric kernels satisfy Assumptions A1, A2(0), and A3(H); see Igarashi and Kakizawa (2020) and Kakizawa (2021). For instance, the constants $\zeta_{1,1}$, $\zeta_{2,1}$, and ζ (given in A1.2–3), associated with $k_g^{(qBS)}(\cdot; \beta, x)$, are given by $\zeta_{1,1} = c + \theta + J_g/2$, $\zeta_{2,2} = J_g$, and $\zeta = C_g^2/C_{g^2}$, independent of $q \geq 0$, where $J_g = \int_{-\infty}^{\infty} u^2 C_g g(u^2) du$. Of course, we need to impose a set of requirements on the density generator g , under which A1, A2(0) and A3(H) hold for the asymmetric kernel $k_g^{(qBS)}(\cdot; \beta, x)$; see Kakizawa (2018,2021). For simplicity, we assume that there exist constants $M_g, B > 0$ such that $g(y) \leq M_g e^{-By}$ for every $y \geq 0$. It remains to discuss A2($\nu > 0$). Note that A2($\nu = 0, 1$) is technically required to prove Theorems 1–4 (i.e., (A8) and (A11) under A2($\nu = 1$)); indeed, (A11) is crucial for Lemma A3. On the other hand, A2($\nu = 0$) is enough for the proofs of Theorems 5–8.

Property A1. The kernel $k_g^{(qBS)}(\cdot; \beta, x)$ satisfies Assumption A2(ν) for any $\nu \geq 0$, with

$$u_{\beta,\nu}^{(qBS)}(x) \leq \begin{cases} \frac{C_g \exp\left\{\frac{\nu+1}{c} \max(-\theta, 0)\right\}}{c^\nu \{\beta(x + \beta c)\}^{1/2}} \sup_{u \in \mathbb{R}} \left\{ \exp\left(\frac{\nu+1}{c^{1/2}} |u|\right) g(u^2) \right\}, & q = 0, \\ \frac{C_g \left\{ \frac{8q^2}{c} \left(1 + \frac{\theta^2}{c}\right) + 2 \right\}^{(q+\nu+1)/(2q)}}{c^\nu \{\beta(x + \beta c)\}^{1/2}} M_{g,\nu}(q), & q > 0, \end{cases}$$

where $M_{g,\nu}(q) = \sup_{y \geq 0} [(y+1)^{(q+\nu+1)/(2q)} g(y)]$.

Proof. Recall that, with $\alpha(y) = 1/(y+c)^{1/2}$ (note that $\sup_{y \geq 0} \alpha(y) = 1/c^{1/2}$),

$$\begin{aligned} k_g^{(qBS)}(s; \beta, x) &= \frac{C_g}{\{\beta(x + \beta c)\}^{1/2}} g\left(\left\{ \frac{a^{(q)}(\alpha^2(x/\beta)(s/\beta))}{\alpha(x/\beta)} - \theta \alpha(x/\beta) \right\}^2\right) A^{(q)}(\alpha^2(x/\beta)(s/\beta)). \end{aligned}$$

This, together with $\beta/s \leq (x + \beta c)/(cs) = \beta/\{c\alpha^2(x/\beta)\}$, yields

$$u_{\beta,\nu}^{(qBS)}(x) \leq \frac{C_g}{c^\nu \{\beta(x + \beta c)\}^{1/2}} \sup_{t \geq 0} \left[g\left(\left\{ \frac{a^{(q)}(t)}{\alpha(x/\beta)} - \theta \alpha(x/\beta) \right\}^2\right) \frac{A^{(q)}(t)}{t^\nu} \right], \quad \nu \geq 0.$$

It suffices to bound $t^{-\nu} A^{(q)}(t) = (1/2)t^{-\nu}(t^{q-1} + t^{-(q+1)})$, $q \geq 0$, in the same manner as Kakizawa (2018). \square

Appendix A.2 Auxiliary Lemmas

We mention (without proof) the following basic lemma; (ii) is a slight modification of Kakizawa (2021):

Lemma A1. (i) Let g be a twice continuously differentiable function on $[0, \infty)$, where g, g' , and g'' are bounded; besides, g'' is Hölder-continuous with exponent $0 < \eta \leq 1$. Under Assumption A1.1–2, given constants $c_L > 0$ and $0 < \tau < 1$, we have, for all sufficiently small $\beta > 0$,

1. $\sup_{0 \leq x \leq c_L \beta^\tau} |B_\beta^{(k,g)}(x)| = O(\beta^\tau);$
2. $B_\beta^{(k,g)}(0) = \beta g'(0) \int_0^\infty up(u; 0) du + O(\beta^2);$
3. for $x \geq c_L \beta^\tau$, $B_\beta^{(k,g)}(x) = \beta B_g(x) + \mathcal{E}_\beta(x),$

with $|\mathcal{E}_\beta(x)| \leq M_g \omega_{\beta,\eta}(x)$, where $M_g > 0$ is a constant, independent of β and x .

(ii) Let g be a bounded and Hölder-continuous function (with exponent $0 < \eta' \leq 1$) on $[0, \infty)$. Under Assumptions A1 and A2($\nu = 0$), given constants $c_L > 0$ and $0 < \tau < 1$, we have, for all sufficiently small $\beta > 0$,

1. $J_\beta^{(k^2, g)}(0) = \beta^{-1}g(0) \int_0^\infty p^2(u; 0) du + O(\beta^{\eta'-1})$;
2. for $x \geq c_L \beta^\tau$, $J_\beta^{(k^2, g)}(x) = \beta^{-1/2}V_g(x) + \mathcal{E}'_\beta(x)$,

with $|\mathcal{E}'_\beta(x)| \leq M'_g \{\omega_{\beta, \eta'}(x; g) + 1\}$, where $M'_g > 0$ is a constant, independent of β and x .

We can verify that, according to, e.g., Igarashi and Kakizawa (2020) and Kakizawa (2021),

$$\int_0^\infty \{B_\beta^{(k, g)}(x)\}^2 dx = \beta^2 I_g^{B_g^2} + o(\beta^2) \quad \text{and} \quad \int_0^\infty J_\beta^{(k^2, g)}(x) dx = \beta^{-1/2} I_g^{V_g} + o(\beta^{-1/2})$$

if $\int_0^\infty t^{2\{3/\min(\eta, \eta') + 1\} + \delta_0} g(t) dt$ exists for some constant $\delta_0 > 0$, as follows:

1. (i) For $2/3 < \tau < 1$, $\int_0^{\beta^\tau} \{B_\beta^{(k, g)}(x)\}^2 dx = O(\beta^{3\tau}) = o(\beta^2)$ using Lemma A1(i.1).
 (ii) Take a constant $0 < \tau' < \eta/(3 + \eta) (\leq 1/4)$. Using Lemma A1(i.3), we have, for any $0 < \tau < 1$,

$$\begin{aligned} & \left| \int_{\beta^\tau}^{\beta^{-\tau'}} \{B_\beta^{(k, g)}(x)\}^2 dx - \beta^2 I_g^{B_g^2} \right| \\ & \leq \beta^2 \left(\int_0^{\beta^\tau} + \int_{\beta^{-\tau'}}^\infty \right) B_g^2(x) dx + 2\beta \left[I_g^{B_g^2} \int_{\beta^\tau}^{\beta^{-\tau'}} \mathcal{E}_\beta^2(x) dx \right]^{1/2} + \int_{\beta^\tau}^{\beta^{-\tau'}} \mathcal{E}_\beta^2(x) dx \\ & = o(\beta^2). \end{aligned}$$

- (iii) Under A3($H = 6/\eta + 1 + \delta_0$), we have, for $2/(H + 1) < \tau' < \eta/(3 + \eta)$,

$$\begin{aligned} \int_{\beta^{-\tau'}}^\infty \{B_\beta^{(k, g)}(x)\}^2 dx & \leq 2\|g\|_{[0, \infty)} \left[\int_0^\infty \left(\int_{\beta^{-\tau'}}^\infty k(s; x, \beta) g(s) dx \right) ds + \int_{\beta^{-\tau'}}^\infty g(x) dx \right] \\ & = O(\beta^{\tau'(H+1)}) = o(\beta^2). \end{aligned}$$

2. (i) For $\tau > 1/2$, $\int_0^{\beta^\tau} J_\beta^{(k^2, g)}(x) dx \leq L_{K,0} \beta^{\tau-1} \|g\|_{[0, \infty)} = o(\beta^{-1/2})$ by A2.1($\nu = 0$).
 (ii) Take a constant $0 < \tau' < \eta'/(1 + \eta') (\leq 1/2)$. Using Lemma A1(ii.2), we have, for any $0 < \tau < 1$,

$$\begin{aligned} & \left| \int_{\beta^\tau}^{\beta^{-\tau'}} J_\beta^{(k^2, g)}(x) dx - \beta^{-1/2} I_g^{V_g} \right| \leq \beta^{-1/2} \left(\int_0^{\beta^\tau} + \int_{\beta^{-\tau'}}^\infty \right) V_g(x) dx + \int_{\beta^\tau}^{\beta^{-\tau'}} |\mathcal{E}'_\beta(x)| dx \\ & = o(\beta^{-1/2}). \end{aligned}$$

- (iii) Under A3($H = 6/\eta' + 1 + \delta_0$), we have, for $1/\{2(H + 1)\} < \tau' < \eta'/(1 + \eta')$,

$$\begin{aligned} \int_{\beta^{-\tau'}}^\infty J_\beta^{(k^2, g)}(x) dx & \leq L_{K,0} \beta^{-1} \int_0^\infty \left(\int_{\beta^{-\tau'}}^\infty k(s; x, \beta) g(s) dx \right) ds \quad (\text{by A2.1}(\nu = 0)) \\ & = O(\beta^{\{\tau'(H+1)-1/2\}-1/2}) = o(\beta^{-1/2}). \end{aligned}$$

Appendix B

Define

$$\zeta_\epsilon = \frac{1}{n} \sum_{i=1}^n \mu(Y_i + \epsilon)^{-1} + \mu\epsilon - 1.$$

Then,

$$\hat{f}_{\beta,\epsilon}(x) = \frac{f(x) + \zeta_\beta(x)}{1 + \zeta_\epsilon}, \text{ with } \zeta_\beta(x) = \frac{1}{n} \sum_{i=1}^n \mu Y_i^{-1} k(Y_i; \beta, x) - f(x). \quad (\text{A1})$$

Also,

$$\tilde{f}_{\beta,\epsilon}(x) = \frac{f(x) + (\mu/x)\zeta_\beta^+(x)}{1 + \zeta_\epsilon}, \text{ with } \zeta_\beta^+(x) = \frac{1}{n} \sum_{i=1}^n k(Y_i; \beta, x) - f_{\text{LB}}(x). \quad (\text{A2})$$

Appendix B.1 Some basic results for ζ_ϵ

For ease of reference, we first mention the tail probability/moment bounds of ζ_ϵ . Rewrite

$$\zeta_\epsilon = \bar{\Delta}_\epsilon + \delta_\epsilon, \quad (\text{A3})$$

where $\delta_\epsilon = \mu E[(Y + \epsilon)^{-1}] + \mu\epsilon - 1$, and $\bar{\Delta}_\epsilon$ is the average of zero-mean independent random variables

$$\Delta_{i,\epsilon} = \mu(Y_i + \epsilon)^{-1} - \mu E[(Y_i + \epsilon)^{-1}], \quad i = 1, \dots, n,$$

with $|\Delta_{i,\epsilon}| \leq \mu\epsilon^{-1}$, $V[\Delta_{i,\epsilon}] \leq \mu^2 E[(Y + \epsilon)^{-2}] \leq \mu\epsilon^{-1}$ (we also have $V[\Delta_{i,\epsilon}] \leq \mu^2 E[Y^{-2}]$). Then, Bernstein's inequality yields the exponential bound of the tail probability

$$P[|\bar{\Delta}_\epsilon| \geq t] \leq 2 \exp \left\{ -\frac{n\epsilon t^2}{2(1+t/3)\mu} \right\} \quad \text{for all } t > 0 \quad (\text{A4})$$

(hence, $\bar{\Delta}_\epsilon \xrightarrow{a.s.} 0$ if $(n\epsilon)/\log n \rightarrow \infty$).

Suppose that $E[Y^{-2}]$ exists. Using $\mu(Y + \epsilon)^{-1} - \mu Y^{-1} = -\mu\epsilon Y^{-1}(Y + \epsilon)^{-1}$, we have

$$|\delta_\epsilon| \leq \mu\epsilon(1 + E[Y^{-2}]). \quad (\text{A5})$$

Furthermore, it is easy to see that

$$\begin{aligned} V[\bar{\Delta}_\epsilon] &= n^{-1} \mu^2 V[(Y + \epsilon)^{-1}] \leq n^{-1} \mu^2 E[Y^{-2}], \\ E[\bar{\Delta}_\epsilon^4] &= \frac{1}{n^3} E[\Delta_{1,\epsilon}^4] + \frac{3(n-1)}{n^3} (V[\Delta_{1,\epsilon}])^2 \leq \mu^4 \left\{ (n\epsilon^2)^{-1} + 3E[Y^{-2}] \right\} n^{-2} E[Y^{-2}]. \end{aligned}$$

It follows that

$$E[\zeta_\epsilon^2] = \delta_\epsilon^2 + V[\bar{\Delta}_\epsilon] = O(n^{-1}), \quad E[\zeta_\epsilon^4] \leq 8(\delta_\epsilon^4 + E[\bar{\Delta}_\epsilon^4]) = O(n^{-2}). \quad (\text{A6})$$

Appendix B.2 Some preliminary results for $\zeta_\beta(x)$

We next list some facts about $\zeta_\beta(x)$, including the tail probability/moment bounds and asymptotic normality. Rewrite

$$\zeta_\beta(x) = \bar{\Delta}_\beta(x) + B_\beta^{(k,f)}(x), \quad (\text{A7})$$

where $\bar{\Delta}_\beta(x)$ is the average of zero-mean independent random variables

$$\Delta_{i,\beta}(x) = \mu Y_i^{-1} k(Y_i; \beta, x) - \mu E[Y_i^{-1} k(Y_i; \beta, x)], \quad i = 1, \dots, n,$$

with $|\Delta_{i,\beta}(x)| \leq \beta^{-1}u_{\beta,1}(x)\mu$, $V[\Delta_{i,\beta}(x)] \leq \mu^2 E[\{Y^{-1}k(Y; \beta, x)\}^2] \leq \beta^{-1}u_{\beta,1}(x)\mu J_{\beta}^{(k,f)}(x)$ (we also have $V[\Delta_{i,\beta}(x)] \leq u_{\beta,0}(x)\mu J_{\beta}^{(k,f-1)}(x)$). Assumption A2.1($\nu = 1$) and Bernstein's inequality yield the exponential bound of the tail probability

$$P[|\bar{\Delta}_{\beta}(x)| \geq t] \leq 2 \exp \left\{ -\frac{n\beta^2 t^2}{2(\|f\|_{[0,\infty)} + t/3)L_{K,1}\mu} \right\} \quad \text{for all } t > 0 \quad (\text{A8})$$

(hence, $\bar{\Delta}_{\beta}(x) \xrightarrow{a.s.} 0$ if $(n\beta^2)/\log n \rightarrow \infty$, which is implied by, e.g., $\beta \propto n^{-\iota}$ for some constant $0 < \iota < 1/2$). On the other hand, F(iii'), A1.3, and A2.2($\nu = 0$) imply that, for fixed $x > 0$ (assume $f(x) > 0$), we have, under $B(0 < \iota < \min\{2q/(2+q), 1\})$ (note that $\iota = 2/5$ is feasible when F(iii') holds for some $q > 1/2$),

$$\begin{aligned} \frac{1}{n^{2+q}} \sum_{i=1}^n \frac{E[|\Delta_{i,\beta}(x)|^{2+q}]}{(V[\bar{\Delta}_{\beta}(x)])^{1+q/2}} &\leq \mu^{1+q} \mu'_{-(1+q)} \left[\frac{n\{2L'_{K,0}(\beta x)^{-1/2}\}^{2+q}}{(n^2 V[\bar{\Delta}_{\beta}(x)])^{1+q/2}} \right] \\ &= O(n^{-q/2} \beta^{-(1+q/2)/2}) = o(1), \end{aligned}$$

hence,

$$\frac{\bar{\Delta}_{\beta}(x)}{(V[\bar{\Delta}_{\beta}(x)])^{1/2}} \xrightarrow{d} N(0, 1), \text{ i.e., } (n\beta^{1/2})^{1/2} \bar{\Delta}_{\beta}(x) \xrightarrow{d} N(0, \mu V_{f-1}(x)) \quad (\text{A9})$$

using Lyapunov's theorem (for triangular arrays), together with $n\beta^{1/2}V[\bar{\Delta}_{\beta}(x)] \rightarrow \mu V_{f-1}(x)$ for fixed $x > 0$, which will be shown in (A18) below. Similarly, F(iii'), A1.3, and A2.1($\nu = 0$) imply that, for the case $f_{-1}(0) > 0$, we have, under $B(0 < \iota < q/(2+q))$ (note that $\iota = 1/3$ is feasible when F(iii') holds for some $q > 1$ (in this case, $f_{-1}(0) = 0$), i.e., $\iota = 1/3$ is unfortunately infeasible for $0 < q \leq 1$),

$$\begin{aligned} \frac{1}{n^{2+q}} \sum_{i=1}^n \frac{E[|\Delta_{i,\beta}(0)|^{2+q}]}{(V[\bar{\Delta}_{\beta}(0)])^{1+q/2}} &\leq \mu^{1+q} \mu'_{-(1+q)} \left[\frac{n(2L_{K,0}\beta^{-1})^{2+q}}{(n^2 V[\bar{\Delta}_{\beta}(0)])^{1+q/2}} \right] \\ &= O(n^{-q/2} \beta^{-(1+q/2)}) = o(1), \end{aligned}$$

hence,

$$\frac{\bar{\Delta}_{\beta}(0)}{(V[\bar{\Delta}_{\beta}(0)])^{1/2}} \xrightarrow{d} N(0, 1), \text{ i.e., } (n\beta)^{1/2} \bar{\Delta}_{\beta}(0) \xrightarrow{d} N\left(0, \mu f_{-1}(0) \int_0^{\infty} p^2(u; 0) du\right) \quad (\text{A10})$$

using Lyapunov's theorem, together with $n\beta V[\bar{\Delta}_{\beta}(0)] \rightarrow \mu f_{-1}(0) \int_0^{\infty} p^2(u; 0) du$, which will be shown in (A14) below.

It is easy to see that

$$\begin{aligned} V[\bar{\Delta}_{\beta}(x)] &= n^{-1} \left[\mu^2 E[\{Y^{-1}k(Y; \beta, x)\}^2] - \{\mu E[Y^{-1}k(Y; \beta, x)]\}^2 \right] \\ &= n^{-1} \left[\mu J_{\beta}^{(k^2, f-1)}(x) - \{J_{\beta}^{(k, f)}(x)\}^2 \right] (\leq n^{-1} \mu J_{\beta}^{(k^2, f-1)}(x)), \\ E[\zeta_{\beta}^2(x)] &= \{B_{\beta}^{(k, f)}(x)\}^2 + V[\bar{\Delta}_{\beta}(x)] \\ &\leq \{B_{\beta}^{(k, f)}(x)\}^2 + n^{-1} \mu J_{\beta}^{(k^2, f-1)}(x) = D_{\beta}(x) \quad (\text{say}), \\ E[\bar{\Delta}_{\beta}^4(x)] &= \frac{1}{n^3} E[\Delta_{1,\beta}^4(x)] + \frac{3(n-1)}{n^3} (V[\Delta_{1,\beta}(x)])^2 \\ &\leq \{(n\beta)^{-1} L_{K,1} \beta^{-1} \mu\}^2 D_{\beta}(x) + 3D_{\beta}^2(x) \quad (\text{by A2.1}(\nu = 1)). \end{aligned} \quad (\text{A11})$$

In addition to $\sup_{x \geq 0} |B_\beta^{(k,f)}(x)| \leq 2\|f\|_{[0,\infty)}$,

$$\sup_{x \geq 0} V[\bar{\Delta}_\beta(x)] \leq n^{-1} L_{K,0} \beta^{-1} \mu \|f_{-1}\|_{[0,\infty)} \quad (\text{by A2.1}(v=0)), \quad (\text{A12})$$

$$B_\beta^{(k,f)}(0) = \beta f'(0) \int_0^\infty u p(u;0) du + O(\beta^2) \quad (\text{by Lemma A1(i.2)}), \quad (\text{A13})$$

$$\begin{aligned} V[\bar{\Delta}_\beta(0)] &= n^{-1} \beta^{-1} \mu f_{-1}(0) \int_0^\infty p^2(u;0) du \\ &\quad + O(n^{-1} \beta^{\eta'-1} + n^{-1}) \quad (\text{by Lemma A1(ii.1)}), \end{aligned} \quad (\text{A14})$$

Lemma A1 implies that, given a constant $0 < \tau < 1$,

$$\sup_{0 \leq x \leq \beta^\tau} |B_\beta^{(k,f)}(x)| = O(\beta^\tau) \quad (\text{A15})$$

(obviously, $\int_0^{\beta^\tau} \{B_\beta^{(k,f)}(x)\}^2 dx = O(\beta^{3\tau})$ and $\int_0^{\beta^\tau} V[\bar{\Delta}_\beta(x)] dx = O(n^{-1} \beta^{\tau-1})$), and that, for $x \geq \beta^\tau$,

$$|B_\beta^{(k,f)}(x) - \beta B_f(x)| \leq \mathcal{M}_f \omega_{\beta,\eta}(x), \quad (\text{A16})$$

$$|J_\beta^{(k^2,f_{-1})}(x) - \beta^{-1/2} V_{f_{-1}}(x)| \leq \mathcal{M}'_{f_{-1}} \{\omega_{\beta,\eta'}(x; f_{-1}) + 1\}, \quad (\text{A17})$$

$$|V[\bar{\Delta}_\beta(x)] - n^{-1} \beta^{-1/2} \mu V_{f_{-1}}(x)| \leq n^{-1} [\mathcal{M}'_{f_{-1}} \mu \{\omega_{\beta,\eta'}(x; f_{-1}) + 1\} + \|f\|_{[0,\infty)}^2]. \quad (\text{A18})$$

Remark A1. If $\int_0^\infty t^{2\{3/\min(\eta,\eta')+1\}+\delta_0} f(t) dt$ exists for some constant $\delta_0 > 0$ (this ensures $\int_0^\infty t^{2\{3/\min(\eta,\eta')+1\}+\delta_0} f_{-1}(t) dt \leq \{\int_0^\infty t^{2\{3/\min(\eta,\eta')+1\}+\delta_0} f(t) dt\}^{\frac{6/\min(\eta,\eta')+1+\delta_0}{2\{3/\min(\eta,\eta')+1\}+\delta_0}}$ exists), in line with, e.g., Kakizawa (2021), Assumption A3($H = 6/\min(\eta,\eta') + 1 + \delta_0$) about the behavior of the tail integral of $k(s; \beta, x)$ with respect to x is crucial for proving the negligibility of the integral $\int_{\beta^{-\tau'}}^\infty E[\zeta_\beta^2(x)] dx (\leq \int_{\beta^{-\tau'}}^\infty D_\beta(x) dx)$, as follows: We have, for any constant $\tau' > 2/(H+1)$,

$$\begin{aligned} \int_{\beta^{-\tau'}}^\infty D_\beta(x) dx &\leq 2\|f\|_{[0,\infty)} \left\{ \int_0^\infty \left(\int_{\beta^{-\tau'}}^\infty k(s; x, \beta) f(s) dx \right) ds + \int_{\beta^{-\tau'}}^\infty f(x) dx \right\} \\ &\quad + n^{-1} L_{K,0} \beta^{-1} \mu \int_0^\infty \left(\int_{\beta^{-\tau'}}^\infty k(s; x, \beta) f_{-1}(s) dx \right) ds \quad (\text{by A2.1}(v=0)) \\ &= O(\beta^{\tau'(H+1)} (1 + n^{-1} \beta^{-1})) = o(\beta^2 + n^{-1} \beta^{-1/2}). \end{aligned}$$

Then, the approximation

$$\int_0^\infty E[\zeta_\beta^2(x)] dx = \beta^2 I^{B_f^2} + n^{-1} \beta^{-1/2} \mu I^{V_{f_{-1}}} + o(\beta^2 + n^{-1} \beta^{-1/2})$$

can be verified, since, by taking $2/(H+1) < \tau' < \min\{\eta/(3+\eta), \eta'/(1+\eta')\}$,

$$\begin{aligned} \int_0^{\beta^{-\tau'}} E[\zeta_\beta^2(x)] dx &= O(\beta^{3\tau} + n^{-1} \beta^{\tau-1}) + \int_{\beta^\tau}^{\beta^{-\tau'}} [D_\beta(x) - n^{-1} \{J_\beta^{(k,f)}(x)\}^2] dx \\ &= \beta^2 I^{B_f^2} + n^{-1} \beta^{-1/2} \mu I^{V_{f_{-1}}} + o(\beta^2 + n^{-1} \beta^{-1/2}) \end{aligned}$$

(we also take $2/3 < \tau < 1$; see the argument after Lemma A1).

Appendix B.3 Some Preliminary Results for $\zeta_\beta^\dagger(x)$

We here list some facts about $\zeta_\beta^\dagger(x)$, including the tail probability/moment bounds and asymptotic normality. Rewrite

$$\zeta_\beta^\dagger(x) = \overline{\Delta}_\beta^\dagger(x) + B_\beta^{\langle k, f_{LB} \rangle}(x), \quad (\text{A19})$$

where $\overline{\Delta}_\beta^\dagger(x)$ is the average of zero-mean independent random variables

$$\Delta_{i,\beta}^\dagger(x) = k(Y_i; \beta, x) - E[k(Y_i; \beta, x)], \quad i = 1, \dots, n,$$

with $|\Delta_{i,\beta}^\dagger(x)| \leq u_{\beta,0}(x)$, $V[\Delta_{i,\beta}^\dagger(x)] \leq E[k^2(Y; \beta, x)] \leq u_{\beta,0}(x)J_\beta^{\langle k, f_{LB} \rangle}(x)$. As in, e.g., Igarashi and Kakizawa (2020) and Kakizawa (2021), by Assumption A2.1($\nu = 0$), an application of Bernstein's inequality yields the exponential bound of the tail probability

$$P[|\overline{\Delta}_\beta^\dagger(x)| \geq t] \leq 2 \exp \left\{ -\frac{n\beta t^2}{2(\|f_{LB}\|_{[0,\infty)} + t/3)L_{K,0}} \right\} \quad \text{for all } t > 0 \quad (\text{A20})$$

(hence, $\overline{\Delta}_\beta^\dagger(x) \xrightarrow{a.s.} 0$ if $(n\beta)/\log n \rightarrow \infty$), whereas, if $\beta \rightarrow 0$ and $n\beta^{1/2} \rightarrow \infty$, then, A2.2($\nu = 0$) implies that, for fixed $x > 0$ (assume $f_{LB}(x) > 0$),

$$\frac{1}{n^{2+p}} \sum_{i=1}^n \frac{E[|\Delta_{i,\beta}^\dagger(x)|^{2+p}]}{(V[\overline{\Delta}_\beta^\dagger(x)])^{1+p/2}} = \left\{ \frac{(L'_{K,0})^2(\beta x)^{-1}}{n(nV[\overline{\Delta}_\beta^\dagger(x)])} \right\}^{p/2} = O((n\beta^{1/2})^{-p/2}) \quad \text{for any } p > 0,$$

hence,

$$\frac{\overline{\Delta}_\beta^\dagger(x)}{(V[\overline{\Delta}_\beta^\dagger(x)])^{1/2}} \xrightarrow{d} N(0, 1), \quad \text{i.e., } (n\beta^{1/2})^{1/2} \frac{\overline{\Delta}_\beta^\dagger(x)}{x} \xrightarrow{d} N(0, \mu V_{f_{-1}}(x)) \quad (\text{A21})$$

using Lyapunov's theorem, together with $n\beta^{1/2}(\mu/x)^2 V[\overline{\Delta}_\beta^\dagger(x)] \rightarrow \mu V_{f_{-1}}(x)$ for fixed $x > 0$, which will be shown in (A26) below.

It is easy to see that

$$\begin{aligned} V[\overline{\Delta}_\beta^\dagger(x)] &= n^{-1} \left[J_\beta^{\langle k^2, f_{LB} \rangle}(x) - \{J_\beta^{\langle k, f_{LB} \rangle}(x)\}^2 \right] (\leq n^{-1} J_\beta^{\langle k^2, f_{LB} \rangle}(x)), \\ E[\{\zeta_\beta^\dagger(x)\}^2] &= \{B_\beta^{\langle k, f_{LB} \rangle}(x)\}^2 + V[\overline{\Delta}_\beta^\dagger(x)] \\ &\leq \{B_\beta^{\langle k, f_{LB} \rangle}(x)\}^2 + n^{-1} J_\beta^{\langle k^2, f_{LB} \rangle}(x) = D_\beta^\dagger(x) \quad (\text{say}), \\ E[\{\overline{\Delta}_\beta^\dagger(x)\}^4] &\leq \frac{1}{n^3} E[\{\Delta_{1,\beta}^\dagger(x)\}^4] + \frac{3(n-1)}{n^3} (V[\Delta_{1,\beta}^\dagger(x)])^2 \\ &\leq (n^{-1} L_{K,0} \beta^{-1})^2 D_\beta^\dagger(x) + 3\{D_\beta^\dagger(x)\}^2 \quad (\text{by A2.1}(\nu = 0)). \end{aligned} \quad (\text{A22})$$

In addition to $\sup_{x \geq 0} |B_\beta^{\langle k, f_{LB} \rangle}(x)| \leq 2\|f_{LB}\|_{[0,\infty)}$ and

$$V[\overline{\Delta}_\beta^\dagger(x)] \leq n^{-1} L_{K,0} \beta^{-1} \|f_{LB}\|_{[0,\infty)} \quad (\text{by A2.1}(\nu = 0)), \quad (\text{A23})$$

Lemma A1 implies that, given a constant $0 < \tau < 1$, we have, for $x \geq \beta^\tau$,

$$\left| \frac{\mu}{x} B_{\beta}^{(k, f_{LB})}(x) - \beta B(x) \right| \leq \mathcal{M}_{f_{LB}} \frac{\mu}{x} \omega_{\beta, \eta''}(x), \quad (\text{A24})$$

$$\left| \left(\frac{\mu}{x} \right)^2 J_{\beta}^{(k^2, f_{LB})}(x) - \beta^{-1/2} \mu V_{f-1}(x) \right| \leq \mathcal{M}'_{f_{LB}} \left\{ \mu \omega_{\beta, 1}(x; f_{-1}) + \left(\frac{\mu}{x} \right)^2 \right\}, \quad (\text{A25})$$

$$\begin{aligned} \left| \left(\frac{\mu}{x} \right)^2 V[\overline{\Delta^+}_{\beta}(x)] - n^{-1} \beta^{-1/2} \mu V_{f-1}(x) \right| &\leq n^{-1} \left[\mathcal{M}'_{f_{LB}} \mu \omega_{\beta, 1}(x; f_{-1}) \right. \\ &\quad \left. + \left(\frac{\mu}{x} \right)^2 \{ \mathcal{M}'_{f_{LB}} + \|f_{LB}\|_{[0, \infty)}^2 \} \right] \\ &\leq n^{-1} \mathcal{M}''_{f_{LB}} \{ \omega_{\beta, 1}(x; f_{-1}) + x^{-2} \}, \end{aligned} \quad (\text{A26})$$

where $\mathcal{M}''_{f_{LB}} = \mathcal{M}'_{f_{LB}} \mu + (\mathcal{M}'_{f_{LB}} + \|f_{LB}\|_{[0, \infty)}^2) \mu^2$.

Remark A2. As mentioned in Theorem 8, we take an arbitrary constant $0 < \tau < 1/2$ and consider a weighted MISE, with a weight function $w(t) = \chi_{[\beta^\tau, \infty)}(t)$ (say). In line with, e.g., Kakizawa (2021), if $\int_0^\infty t^{2(1+\eta'')/\eta'' + \delta_0} f_{LB}(t) dt$ exists for some constant $\delta_0 > 0$, Assumption A3($H = 2/\eta'' + 1 + \delta_0$), together with $\sup_{x \geq \beta^{-\tau'}} (\mu/x) \leq 1$ (say) for any constant $\tau' > 0$, is crucial for proving the negligibility of $\int_{\beta^{-\tau'}}^\infty (\mu/x)^2 E[\{\zeta_{\beta}^+(x)\}^2] dx (\leq \int_{\beta^{-\tau'}}^\infty D_{\beta}^+(x) dx)$, as follows: We have, for any constant $2/(H+1) < \tau' < \eta''/(1+\eta'')$,

$$\begin{aligned} \int_{\beta^{-\tau'}}^\infty D_{\beta}^+(x) dx &\leq 2\|f_{LB}\|_{[0, \infty)} \left\{ \int_0^\infty \left(\int_{\beta^{-\tau'}}^\infty k(s; x, \beta) f_{LB}(s) dx \right) ds + \int_{\beta^{-\tau'}}^\infty f_{LB}(x) dx \right\} \\ &\quad + n^{-1} L_{K,0} \beta^{-1} \int_0^\infty \left(\int_{\beta^{-\tau'}}^\infty k(s; x, \beta) f_{LB}(s) dx \right) ds \quad (\text{by A2.1}(\nu = 0)) \\ &= O(\beta^{\tau'(H+1)} (1 + n^{-1} \beta^{-1})) = o(\beta^2 + n^{-1} \beta^{-1/2}). \end{aligned}$$

We can verify that

$$\int_{\beta^\tau}^\infty \left(\frac{\mu}{x} \right)^2 E[\{\zeta_{\beta}^+(x)\}^2] dx = \beta^2 I^{B^2} + n^{-1} \beta^{-1/2} \mu I^{V_{f-1}} + o(\beta^2 + n^{-1} \beta^{-1/2})$$

for every $0 < \tau < 1/2$, since

$$\begin{aligned} &\left| \int_{\beta^\tau}^{\beta^{-\tau'}} \left\{ \frac{\mu}{x} B_{\beta}^{(k, f_{LB})}(x) \right\}^2 dx - \beta^2 I^{B^2} \right| \\ &\leq \beta^2 \left(\int_0^{\beta^\tau} + \int_{\beta^{-\tau'}}^\infty \right) B^2(x) dx \\ &\quad + 2\beta \mathcal{M}_{f_{LB}} \left[I^{B^2} \int_{\beta^\tau}^{\beta^{-\tau'}} \left(\frac{\mu}{x} \right)^2 \omega_{\beta, \eta''}^2(x) dx \right]^{1/2} + \mathcal{M}_{f_{LB}}^2 \int_{\beta^\tau}^{\beta^{-\tau'}} \left(\frac{\mu}{x} \right)^2 \omega_{\beta, \eta''}^2(x) dx \\ &= o(\beta^2), \\ &\left| \int_{\beta^\tau}^{\beta^{-\tau'}} \left(\frac{\mu}{x} \right)^2 V[\overline{\Delta^+}_{\beta}(x)] dx - n^{-1} \beta^{-1/2} \mu I^{V_{f-1}} \right| \\ &\leq n^{-1} \beta^{-1/2} \mu \left(\int_0^{\beta^\tau} + \int_{\beta^{-\tau'}}^\infty \right) V_{f-1}(x) dx + \mathcal{M}''_{f_{LB}} \int_{\beta^\tau}^{\beta^{-\tau'}} \{ \omega_{\beta, 1}(x; f_{-1}) + x^{-2} \} dx \\ &= o(n^{-1} \beta^{-1/2}). \end{aligned}$$

Appendix B.4 Proofs of Main Results

Before proving Theorems 1–4, we prepare two lemmas (Lemmas A2 and A3):

Lemma A2. Suppose that $E[Y^{-2}]$ exists. Then,

$$E[\zeta_\beta^2(x)\zeta_\epsilon^2] \leq M_1 n^{-1} D_\beta(x),$$

where $M_1 > 0$ is a constant, independent of n, β , and x .

Proof. It is easy to see that

$$\begin{aligned} E[\bar{\Delta}_\beta^2(x)\bar{\Delta}_\epsilon^2] &= \frac{1}{n^3} E[\Delta_{1,\beta}^2(x)\Delta_{1,\epsilon}^2] + \frac{n-1}{n^3} \left[V[\Delta_{1,\beta}(x)]V[\Delta_{1,\epsilon}] + 2\{Cov[\Delta_{1,\beta}(x), \Delta_{1,\epsilon}]\}^2 \right] \\ &\leq \frac{1}{n^3} E[\Delta_{1,\beta}^2(x)\Delta_{1,\epsilon}^2] + \frac{3}{n^2} \mu_\beta^{(k^2, f^{-1})}(x) \mu^2 E[Y^{-2}], \end{aligned}$$

where

$$\begin{aligned} E[\Delta_{1,\beta}^2(x)\Delta_{1,\epsilon}^2] &\leq 4 \left[\mu^4 E[(Y+\epsilon)^{-2} Y^{-2} k^2(Y; \beta, x)] + 3\mu^2 E[Y^{-2} k^2(Y; \beta, x)] \mu^2 E[(Y+\epsilon)^{-2}] \right] \\ &\leq 4\mu^3 (\epsilon^{-2} + 3E[Y^{-2}]) J_\beta^{(k^2, f^{-1})}(x). \end{aligned}$$

The result follows from

$$E[\zeta_\beta^2(x)\zeta_\epsilon^2] \leq 4\{E[\bar{\Delta}_\beta^2(x)\bar{\Delta}_\epsilon^2] + V[\bar{\Delta}_\beta(x)]\delta_\epsilon^2\} + 2\{B_\beta^{(k, f)}(x)\}^2 E[\zeta_\epsilon^2].$$

□

Now, we rewrite (A1) as

$$\hat{f}_{\beta, \epsilon}(x) = f(x) \left(1 - \zeta_\epsilon + \frac{\zeta_\epsilon^2}{1 + \zeta_\epsilon}\right) + \zeta_\beta(x) \left(1 - \frac{\zeta_\epsilon}{1 + \zeta_\epsilon}\right) = f(x) + L_{\beta, \epsilon}(x) + R_{\beta, \epsilon}(x), \quad (\text{A27})$$

where

$$L_{\beta, \epsilon}(x) = \zeta_\beta(x) - f(x)\zeta_\epsilon, \quad R_{\beta, \epsilon}(x) = f(x) \frac{\zeta_\epsilon^2}{1 + \zeta_\epsilon} - \zeta_\beta(x) \frac{\zeta_\epsilon}{1 + \zeta_\epsilon}.$$

We need to evaluate $E[R_{\beta, \epsilon}^2(x)]$.

Lemma A3. Suppose that Assumptions A2.1($v = 1$) and B hold, and that $E[Y^{-2}]$ exists. Under the boundedness of f , we have

$$E[R_{\beta, \epsilon}^2(x)] \leq \mathcal{M}[4n^{-2} + n^{-1} D_\beta(x)],$$

where $\mathcal{M} > 0$ is a constant, independent of n, β , and x

Proof. Considering an event $\mathcal{S}_n = \{\mathcal{Y}_n : |\zeta_\epsilon| < 1/2\}$ (say), we have

$$\begin{aligned} |R_{\beta, \epsilon}(x)| \chi_{\mathcal{S}_n} &\leq 2\{f(x)\zeta_\epsilon^2 + |\zeta_\beta(x)\zeta_\epsilon|\} \chi_{\mathcal{S}_n}, \\ |R_{\beta, \epsilon}(x)|(1 - \chi_{\mathcal{S}_n}) &\leq \left[(\mu\epsilon)^{-1} \{f(x) + |\zeta_\beta(x)|\} + f(x) + \{f(x)|\zeta_\epsilon| + |\zeta_\beta(x)|\} \right] (1 - \chi_{\mathcal{S}_n}) \\ &\leq \left[(\mu\epsilon)^{-1} \{f(x) + |\zeta_\beta(x)|\} + 2\{3f(x)\zeta_\epsilon^2 + |\zeta_\beta(x)\zeta_\epsilon|\} \right] (1 - \chi_{\mathcal{S}_n}). \end{aligned}$$

Using $\{\mathcal{Y}_n : |\zeta_\epsilon| \geq 1/2\} \subset \{\mathcal{Y}_n : |\bar{\Delta}_\epsilon| \geq 1/4\}$ (say) for all sufficiently large n , it can be shown that

$$\begin{aligned} E[R_{\beta, \epsilon}^2(x)] &\leq 16 \left[9f^2(x)E[\zeta_\epsilon^4] + E[\zeta_\beta^2(x)\zeta_\epsilon^2] \right] \\ &\quad + 4(\mu\epsilon)^{-2} \left[\{f^2(x) + D_\beta(x)\} P[|\bar{\Delta}_\epsilon| \geq 1/4] + \{E[\bar{\Delta}_\beta^4(x)] P[|\bar{\Delta}_\epsilon| \geq 1/4]\}^{1/2} \right]. \end{aligned}$$

Then, the result follows from (A4), (A6), (A11), and Lemma A2, under Assumption B (recall that $\epsilon = C'n^{-1/2}$; in this case, $P[|\bar{\Delta}_\epsilon| \geq 1/4] \leq 2e^{-qn^{1/2}}$, where $q > 0$ is a constant, independent of n). \square

We are ready to prove Theorems 1–4.

Proof of Theorem 1. We start with

$$E[\hat{f}_{\beta,\epsilon}(x)] = f(x) + B_\beta^{(k,f)}(x) - f(x)\delta_\epsilon + E[R_{\beta,\epsilon}(x)], \quad (\text{A28})$$

$$\hat{f}_{\beta,\epsilon}(x) - E[\hat{f}_{\beta,\epsilon}(x)] = \bar{\Delta}_\beta(x) - f(x)\bar{\Delta}_\epsilon + R_{\beta,\epsilon}(x) - E[R_{\beta,\epsilon}(x)], \quad (\text{A29})$$

where

$$V[\hat{f}_{\beta,\epsilon}(x)] = V[\bar{\Delta}_\beta(x)] + 2\text{Cov}[\bar{\Delta}_\beta(x), -f(x)\bar{\Delta}_\epsilon + R_{\beta,\epsilon}(x)] + V[-f(x)\bar{\Delta}_\epsilon + R_{\beta,\epsilon}(x)]. \quad (\text{A30})$$

Also,

$$\text{Cov}[\bar{\Delta}_\beta(x), \bar{\Delta}_\epsilon] = n^{-1} \left[\mu^2 E[(Y + \epsilon)^{-1} Y^{-1} k(Y; \beta, x)] - \mu E[Y^{-1} k(Y; \beta, x)] \mu E[(Y + \epsilon)^{-1}] \right],$$

hence,

$$|\text{Cov}[\bar{\Delta}_\beta(x), \bar{\Delta}_\epsilon]| \leq n^{-1} [\mu J_\beta^{(k,f-1)}(x) + J_\beta^{(k,f)}(x)] \leq n^{-1} [\mu \|f_{-1}\|_{[0,\infty)} + \|f\|_{[0,\infty)}].$$

It is shown from (A28) and (A30) that

$$\begin{aligned} |\text{Bias}[\hat{f}_{\beta,\epsilon}(x)] - B_\beta^{(k,f)}(x)| &\leq \|f\|_{[0,\infty)} \delta_\epsilon + \{E[R_{\beta,\epsilon}^2(x)]\}^{1/2}, \\ |V[\hat{f}_{\beta,\epsilon}(x)] - V[\bar{\Delta}_\beta(x)]| &\leq M_2 \left[n^{-1} + \{n^{-1} \mu J_\beta^{(k^2,f-1)}(x) E[R_{\beta,\epsilon}^2(x)]\}^{1/2} + E[R_{\beta,\epsilon}^2(x)] \right], \end{aligned}$$

where $M_2 > 0$ is a constant, independent of n, β , and x . Using Lemma A3 and

$$\begin{aligned} \{n^{-1} \mu J_\beta^{(k^2,f-1)}(x) E[R_{\beta,\epsilon}^2(x)]\}^{1/2} &\leq \mathcal{M}^{1/2} [2n^{-1} D_\beta^{1/2}(x) + n^{-1/2} D_\beta(x)] \\ &\leq \mathcal{M}^{1/2} [n^{-1} + (n^{-1} + n^{-1/2}) D_\beta(x)], \end{aligned}$$

we have

$$\begin{aligned} |\text{Bias}[\hat{f}_{\beta,\epsilon}(x)] - B_\beta^{(k,f)}(x)| &\leq M_3 \left[n^{-1/2} + n^{-1/2} |B_\beta^{(k,f)}(x)| \right] \quad (\text{assume } n^{-1} \beta^{-1} = o(1)), \\ |V[\hat{f}_{\beta,\epsilon}(x)] - V[\bar{\Delta}_\beta(x)]| &\leq M_4 \left[n^{-1} + n^{-1/2} D_\beta(x) \right], \end{aligned}$$

where $M_3, M_4 > 0$ are constants, independent of n, β , and x . Then, using (A12)–(A18), the proof is completed. \square

Proof of Theorem 2. Recall (A1) (also (A3) and (A7)). The strong consistency follows from (A4), (A5), and (A8), together with $B_\beta^{(k,f)}(x) = O(\beta)$ for fixed $x > 0$ (we also have $B_\beta^{(k,f)}(0) = O(\beta)$). See Lemma A1(i.2–3). \square

Proof of Theorem 3. By Lemma A3, we notice that $R_{\beta,\epsilon}(x) - E[R_{\beta,\epsilon}(x)] = o_p((n\beta^{1/2})^{-1/2})$ for fixed $x > 0$, and that $R_{\beta,\epsilon}(0) - E[R_{\beta,\epsilon}(0)] = o_p((n\beta)^{-1/2})$. Recalling (A29), where $\bar{\Delta}_\epsilon = O_p(n^{-1/2})$, we have

$$\begin{aligned} (n\beta^{1/2})^{1/2} (\hat{f}_{\beta,\epsilon}(x) - E[\hat{f}_{\beta,\epsilon}(x)]) &= (n\beta^{1/2})^{1/2} \bar{\Delta}_\beta(x) + o_p(1) \quad \text{for fixed } x > 0, \\ (n\beta)^{1/2} (\hat{f}_{\beta,\epsilon}(0) - E[\hat{f}_{\beta,\epsilon}(0)]) &= (n\beta)^{1/2} \bar{\Delta}_\beta(0) + o_p(1). \end{aligned}$$

The results (i) and (ii) follow from (A9) and (A10), respectively. \square

Proof of Theorem 4. In the same way as the argument of Remark A1, a rigorous derivation of the MISE is made by splitting the integral into three parts (we set $H = 6/\eta + 1 + \delta_0$), as follows:

- (i) Theorem 1 yields $\int_0^{\beta^\tau} \text{MSE}[\hat{f}_{\beta,\epsilon}(x)] dx = O(\beta^{3\tau} + n^{-1}\beta^{\tau-1}) = o(\beta^2 + n^{-1}\beta^{-1/2})$ for $2/3 < \tau < 1$.
- (ii) Take a constant $\tau' > 2/(H+1)$. The inequality

$$\begin{aligned} \epsilon^{-2} \int_{\beta^{-\tau'}}^{\infty} E[\zeta_\beta^2(x) \zeta_\epsilon^2] dx &\leq O(n^{-1}) \int_0^{\infty} \left[\int_{\beta^{-\tau'}}^{\infty} k(s; \beta, x) \{f_{-1}(s) + f(s)\} dx \right] ds \\ &\quad + O(1) \int_{\beta^{-\tau'}}^{\infty} D_\beta(x) dx \quad (\text{Lemma A2}), \end{aligned}$$

together with (A6), enables us to see that, by (A27),

$$\begin{aligned} &\int_{\beta^{-\tau'}}^{\infty} E[\{\hat{f}_{\beta,\epsilon}(x) - f(x)\}^2] dx \\ &\leq 4 \left[\int_{\beta^{-\tau'}}^{\infty} D_\beta(x) dx + \|f\|_{[0,\infty)} E[\zeta_\epsilon^2] \right. \\ &\quad \left. + (\mu\epsilon)^{-2} \left\{ \|f\|_{[0,\infty)} E[\zeta_\epsilon^4] + \int_{\beta^{-\tau'}}^{\infty} E[\zeta_\beta^2(x) \zeta_\epsilon^2] dx \right\} \right] \\ &= o(\beta^2 + n^{-1}\beta^{-1/2}) \quad (\text{see the argument in Remark A1}). \end{aligned}$$

- (iii) Taking $2/3 < \tau < 1$ and $2/(H+1) < \tau' < \min\{\eta/(3+\eta), \eta'/(1+\eta')\}$, Theorem 1 yields

$$\begin{aligned} &\left| \int_{\beta^\tau}^{\beta^{-\tau'}} \{ \text{Bias}^2[\hat{f}_{\beta,\epsilon}(x)] + V[\hat{f}_{\beta,\epsilon}(x)] \} dx - (\beta^2 I_f^{B_f^2} + n^{-1}\beta^{-1/2} \mu I^{V_{f-1}}) \right| \\ &\leq \beta^2 \left(\int_0^{\beta^\tau} + \int_{\beta^{-\tau'}}^{\infty} \right) B_f^2(x) dx \\ &\quad + 2\beta \left[I_f^{B_f^2} \int_{\beta^\tau}^{\beta^{-\tau'}} \{R_\beta^{\text{Bias}}(x)\}^2 dx \right]^{1/2} + \int_{\beta^\tau}^{\beta^{-\tau'}} \{R_\beta^{\text{Bias}}(x)\}^2 dx \\ &\quad + n^{-1}\beta^{-1/2} \mu \left(\int_0^{\beta^\tau} + \int_{\beta^{-\tau'}}^{\infty} \right) V_{f-1}(x) dx + \int_{\beta^\tau}^{\beta^{-\tau'}} |R_\beta^V(x)| dx \\ &= o(\beta^2 + n^{-1}\beta^{-1/2}). \end{aligned}$$

□

Before proving Theorems 5–8, we prepare two lemmas (Lemmas A4 and A5):

Lemma A4. Suppose that $E[Y^{-2}]$ exists. Then,

$$E[\{\zeta_\beta^+(x)\}^2 \zeta_\epsilon^2] \leq M_1^\dagger n^{-1} D_\beta^\dagger(x),$$

where $M_1^\dagger > 0$ is a constant, independent of n, β , and x .

Proof. It is easy to see that

$$\begin{aligned} &E[\{\bar{\Delta}_\beta^\dagger(x)\}^2 \bar{\Delta}_\epsilon^2] \\ &= \frac{1}{n^3} E[\{\Delta_{1,\beta}^\dagger(x)\}^2 \Delta_{1,\epsilon}^2] + \frac{n-1}{n^3} [V[\Delta_{1,\beta}^\dagger(x)] V[\Delta_{1,\epsilon}] + 2\{\text{Cov}[\Delta_{1,\beta}^\dagger(x), \Delta_{1,\epsilon}]\}^2] \\ &\leq \frac{1}{n^3} E[\{\Delta_{1,\beta}^\dagger(x)\}^2 \Delta_{1,\epsilon}^2] + \frac{3}{n^2} J_\beta^{(k^2, f_{\text{LB}})}(x) \mu^2 E[Y^{-2}], \end{aligned}$$

where

$$\begin{aligned} E[\{\Delta_{1,\beta}^+(x)\}^2 \Delta_{1,\epsilon}^2] &\leq 4 \left[\mu^2 E[(Y + \epsilon)^{-2} k^2(Y; \beta, x)] + 3E[k^2(Y; \beta, x)] \mu^2 E[(Y + \epsilon)^{-2}] \right] \\ &\leq 4\mu^2(\epsilon^{-2} + 3E[Y^{-2}]) J_{\beta}^{(k^2, f_{LB})}(x). \end{aligned}$$

The result follows from

$$E[\{\zeta_{\beta}^+(x)\}^2 \zeta_{\epsilon}^2] \leq 4\{E[\{\bar{\Delta}_{\beta}^+(x)\}^2 \bar{\Delta}_{\epsilon}^2] + V[\bar{\Delta}_{\beta}^+(x)] \delta_{\epsilon}^2\} + 2\{B_{\beta}^{(k, f_{LB})}(x)\}^2 E[\zeta_{\epsilon}^2].$$

□

Now, (A2) can be rewritten as

$$\tilde{f}_{\beta,\epsilon}(x) = f(x) \left(1 - \zeta_{\epsilon} + \frac{\zeta_{\epsilon}^2}{1 + \zeta_{\epsilon}}\right) + \frac{\mu}{x} \zeta_{\beta}^+(x) \left(1 - \frac{\zeta_{\epsilon}}{1 + \zeta_{\epsilon}}\right) = f(x) + \frac{\mu}{x} \{L_{\beta,\epsilon}^+(x) + R_{\beta,\epsilon}^+(x)\}, \quad (\text{A31})$$

where

$$L_{\beta,\epsilon}^+(x) = \zeta_{\beta}^+(x) - f_{LB}(x) \zeta_{\epsilon}, \quad R_{\beta,\epsilon}^+(x) = f_{LB}(x) \frac{\zeta_{\epsilon}^2}{1 + \zeta_{\epsilon}} - \zeta_{\beta}^+(x) \frac{\zeta_{\epsilon}}{1 + \zeta_{\epsilon}}.$$

We need to evaluate $E[\{R_{\beta,\epsilon}^+(x)\}^2]$.

Lemma A5. Suppose that Assumptions A2.1($v = 0$) and B hold, and that $E[Y^{-2}]$ exists. Under the boundedness of f_{LB} , we have

$$E[\{R_{\beta,\epsilon}^+(x)\}^2] \leq \mathcal{M}^+[n^{-2} + n^{-1} D_{\beta}^+(x)],$$

where $\mathcal{M}^+ > 0$ is a constant, independent of n, β , and x .

Proof. We have

$$\begin{aligned} |R_{\beta,\epsilon}^+(x)| \chi_{S_n} &\leq 2\{f_{LB}(x) \zeta_{\epsilon}^2 + |\zeta_{\beta}^+(x) \zeta_{\epsilon}|\} \chi_{S_n}, \\ |R_{\beta,\epsilon}^+(x)| (1 - \chi_{S_n}) &\leq \left[(\mu\epsilon)^{-1} \{f_{LB}(x) + |\zeta_{\beta}^+(x)|\} \right. \\ &\quad \left. + f_{LB}(x) + \{f_{LB}(x) |\zeta_{\epsilon}| + |\zeta_{\beta}^+(x)|\} \right] (1 - \chi_{S_n}) \\ &\leq \left[(\mu\epsilon)^{-1} \{f_{LB}(x) + |\zeta_{\beta}^+(x)|\} + 2\{3f_{LB}(x) \zeta_{\epsilon}^2 + |\zeta_{\beta}^+(x) \zeta_{\epsilon}|\} \right] (1 - \chi_{S_n}). \end{aligned}$$

Using $\{\mathcal{Y}_n : |\zeta_{\epsilon}| \geq 1/2\} \subset \{\mathcal{Y}_n : |\bar{\Delta}_{\epsilon}| \geq 1/4\}$ (say) for all sufficiently large n , it can be shown that

$$\begin{aligned} E[\{R_{\beta,\epsilon}^+(x)\}^2] &\leq 16 \left[9f_{LB}^2(x) E[\zeta_{\epsilon}^4] + E[\{\zeta_{\beta}^+(x)\}^2 \zeta_{\epsilon}^2] \right] \\ &\quad + 4(\mu\epsilon)^{-2} \left[\{f_{LB}^2(x) + D_{\beta}^+(x)\} P[|\bar{\Delta}_{\epsilon}| \geq 1/4] + \{E[\{\bar{\Delta}_{\beta}^+(x)\}^4] P[|\bar{\Delta}_{\epsilon}| \geq 1/4]\}^{1/2} \right]. \end{aligned}$$

Then, the result follows from (A4), (A6), (A22), and Lemma A4, under Assumption B (recall that $\epsilon = C'n^{-1/2}$; in this case, $P[|\bar{\Delta}_{\epsilon}| \geq 1/4] \leq 2e^{-qn^{1/2}}$, where $q > 0$ is a constant, independent of n). □

We are ready to prove Theorems 5–8.

Proof of Theorem 5. We start with

$$E[\tilde{f}_{\beta,\epsilon}(x)] = f(x) + \frac{\mu}{x} \{B_{\beta}^{(k, f_{LB})}(x) - f_{LB}(x) \delta_{\epsilon} + E[R_{\beta,\epsilon}^+(x)]\}, \quad (\text{A32})$$

$$\tilde{f}_{\beta,\epsilon}(x) - E[\tilde{f}_{\beta,\epsilon}(x)] = \frac{\mu}{x} \{\bar{\Delta}_{\beta}^+(x) - f_{LB}(x) \bar{\Delta}_{\epsilon} + R_{\beta,\epsilon}^+(x) - E[R_{\beta,\epsilon}^+(x)]\}, \quad (\text{A33})$$

where

$$V[\tilde{f}_{\beta,\epsilon}(x)] = \left(\frac{\mu}{x}\right)^2 \left\{ V[\bar{\Delta}_{\beta}^+(x)] + 2Cov[\bar{\Delta}_{\beta}^+(x), -f_{LB}(x)\bar{\Delta}_{\epsilon} + R_{\beta,\epsilon}^+(x)] + V[-f_{LB}(x)\bar{\Delta}_{\epsilon} + R_{\beta,\epsilon}^+(x)] \right\}. \quad (A34)$$

Also,

$$Cov[\bar{\Delta}_{\beta}^+(x), \bar{\Delta}_{\epsilon}] = n^{-1} \left[\mu E[(Y + \epsilon)^{-1} k(Y; \beta, x)] - E[k(Y; \beta, x)] \mu E[(Y + \epsilon)^{-1}] \right],$$

hence,

$$|Cov[\bar{\Delta}_{\beta}^+(x), \bar{\Delta}_{\epsilon}]| \leq n^{-1} [J_{\beta}^{(k,f)}(x) + J_{\beta}^{(k,f_{LB})}(x)] \leq n^{-1} [\|f\|_{[0,\infty)} + \|f_{LB}\|_{[0,\infty)}].$$

It is shown from (A32) and (A34) that

$$\begin{aligned} |Bias[\tilde{f}_{\beta,\epsilon}(x)] - \frac{\mu}{x} B_{\beta}^{(k,f_{LB})}(x)| &\leq \frac{\mu}{x} [\|f_{LB}\|_{[0,\infty)} \delta_{\epsilon} + \{E[\{R_{\beta,\epsilon}^+(x)\}^2]\}^{1/2}], \\ |V[\tilde{f}_{\beta,\epsilon}(x)] - \left(\frac{\mu}{x}\right)^2 V[\bar{\Delta}_{\beta}^+(x)]| &\leq M_2' \left(\frac{\mu}{x}\right)^2 \left[n^{-1} + \{n^{-1} J_{\beta}^{(k^2,f_{LB})}(x) E[\{R_{\beta,\epsilon}^+(x)\}^2]\}^{1/2} + E[\{R_{\beta,\epsilon}^+(x)\}^2] \right], \end{aligned}$$

where $M_2' > 0$ is a constant, independent of n, β , and x . Using Lemma A5 and

$$\begin{aligned} \{n^{-1} J_{\beta}^{(k^2,f_{LB})}(x) E[\{R_{\beta,\epsilon}^+(x)\}^2]\}^{1/2} &\leq (\mathcal{M}^+)^{1/2} [2n^{-1} \{D_{\beta}^+(x)\}^{1/2} + n^{-1/2} D_{\beta}^+(x)] \\ &\leq (\mathcal{M}^+)^{1/2} [n^{-3/2} + 2n^{-1/2} D_{\beta}^+(x)], \end{aligned}$$

we have

$$\begin{aligned} |Bias[\tilde{f}_{\beta,\epsilon}(x)] - \frac{\mu}{x} B_{\beta}^{(k,f_{LB})}(x)| &\leq M_3^+ [n^{-1/2} (1 + x^{-1}) + n^{-1/2} \frac{\mu}{x} |B_{\beta}^{(k,f_{LB})}(x)|] \\ &\quad (\text{assume } n^{-1} \beta^{-1} = o(1)), \\ |V[\tilde{f}_{\beta,\epsilon}(x)] - \left(\frac{\mu}{x}\right)^2 V[\bar{\Delta}_{\beta}^+(x)]| &\leq M_4^+ [n^{-1} (1 + x^{-2}) + n^{-1/2} \left(\frac{\mu}{x}\right)^2 D_{\beta}^+(x)], \end{aligned}$$

where $M_3^+, M_4^+ > 0$ are constants, independent of n, β , and x . Then, using (A24)–(A26), the proof is completed. \square

Proof of Theorem 6. Recall (A2) (also (A3) and (A19)). The strong consistency follows from (A4), (A5), and (A20), together with $(\mu/x) B_{\beta}^{(k,f_{LB})}(x) = O(\beta)$ for fixed $x > 0$ (see (A24)). \square

Proof of Theorem 7. Recall (A33), where $\bar{\Delta}_{\epsilon} = O_p(n^{-1/2})$. For fixed $x > 0$, Lemma A5 implies that $R_{\beta,\epsilon}^+(x) - E[R_{\beta,\epsilon}^+(x)] = o_p((n\beta^{1/2})^{-1/2})$, i.e.,

$$(n\beta^{1/2})^{1/2} (\tilde{f}_{\beta,\epsilon}(x) - E[\tilde{f}_{\beta,\epsilon}(x)]) = (n\beta^{1/2})^{1/2} \frac{\mu}{x} \bar{\Delta}_{\beta}^+(x) + o_p(1).$$

The asymptotic normality follows from (A21). \square

Proof of Theorem 8. We assume that $\int_0^{\infty} t^{2(1/\eta''+1)+\delta_0} f_{LB}(t) dt = I$ (say) exists for some constant $\delta_0 > 0$ (then, $\int_0^{\infty} t^{2(1/\eta''+1)+\delta_0} f(t) dt = \mu \int_0^{\infty} t^{2/\eta''+1+\delta_0} f_{LB}(t) dt \leq \mu I^{\frac{2/\eta''+1+\delta_0}{2(1/\eta''+1)+\delta_0}}$ exists). In the same way as the argument of Remark A2, a rigorous derivation is made by splitting the integral into two parts (we set $H = 2/\eta'' + 1 + \delta_0$), as follows:

(i) Take a constant $\tau' > 2/(H+1)$. The inequality

$$\begin{aligned} \epsilon^{-2} \int_{\beta^{-\tau'}}^{\infty} E[\{\zeta_{\beta}^{+}(x)\}^2 \zeta_{\epsilon}^2] dx &\leq O(n^{-1}) \int_0^{\infty} \left[\int_{\beta^{-\tau'}}^{\infty} k(s; \beta, x) \{f(s) + f_{LB}(s)\} dx \right] ds \\ &\quad + O(1) \int_{\beta^{-\tau'}}^{\infty} D_{\beta}^{+}(x) dx \quad (\text{Lemma A4}), \end{aligned}$$

together with (A6) and $\sup_{x \geq \beta^{-\tau'}} (\mu/x) \leq 1$ (say), enables us to see that, by (A31),

$$\begin{aligned} &\int_{\beta^{-\tau'}}^{\infty} E[\{\tilde{f}_{\beta, \epsilon}(x) - f(x)\}^2] dx \\ &\leq 4 \left[\int_{\beta^{-\tau'}}^{\infty} D_{\beta}^{+}(x) dx + \|f\|_{[0, \infty)} E[\zeta_{\epsilon}^2] \right. \\ &\quad \left. + (\mu\epsilon)^{-2} \left\{ \|f\|_{[0, \infty)} E[\zeta_{\epsilon}^4] + \int_{\beta^{-\tau'}}^{\infty} E[\{\zeta_{\beta}^{+}(x)\}^2 \zeta_{\epsilon}^2] dx \right\} \right] \\ &= o(\beta^2 + n^{-1} \beta^{-1/2}) \quad (\text{repeat the same argument in Remark A2}). \end{aligned}$$

(ii) Taking $2/(H+1) < \tau' < \eta''/(1+\eta'')$, Theorem 5 enables us to see that, for every $0 < \tau < 1/2$,

$$\begin{aligned} &\left| \int_{\beta^{\tau}}^{\beta^{-\tau'}} \{ \text{Bias}^2[\tilde{f}_{\beta, \epsilon}(x)] + V[\tilde{f}_{\beta, \epsilon}(x)] \} dx - (\beta^2 I^{B^2} + n^{-1} \beta^{-1/2} \mu I^{V_{f-1}}) \right| \\ &\leq \beta^2 \left(\int_0^{\beta^{\tau}} + \int_{\beta^{-\tau'}}^{\infty} \right) B^2(x) dx \\ &\quad + 2\beta \left[I^{B^2} \int_{\beta^{\tau}}^{\beta^{-\tau'}} \{ R_{\beta}^{+ \text{Bias}}(x) \}^2 dx \right]^{1/2} + \int_{\beta^{\tau}}^{\beta^{-\tau'}} \{ R_{\beta}^{+ \text{Bias}}(x) \}^2 dx \\ &\quad + n^{-1} \beta^{-1/2} \mu \left(\int_0^{\beta^{\tau}} + \int_{\beta^{-\tau'}}^{\infty} \right) V_{f-1}(x) dx + \int_{\beta^{\tau}}^{\beta^{-\tau'}} |R_{\beta}^{+ V}(x)| dx \\ &= o(\beta^2 + n^{-1} \beta^{-1/2}). \end{aligned}$$

□

References

- Ahmad, I.A. 1995. On multivariate kernel estimation for samples from weighted distributions. *Statist.Probab.Lett.* **22** 121–129.
- Bhattacharyya, B.B., Franklin, L.A., and Richardson, G.D. 1988. A comparison of nonparametric unweighted and length-biased density estimation of fibres. *Comm.Statist.Theory Methods* **17** 3629–3644.
- Chaubey, Y.P. and Li, J. 2013. Asymmetric kernel density estimator for length-biased data. In *Contemporary Topics in Mathematics and Statistics with Applications*; Adhikari, A., Adhikari, M.R., and Chaubey, Y.P., Eds.; New Delhi: Asian Books Private Ltd, pp.1–28.
- Chaubey, Y.P., Sen, P.K., and Li, J. 2010. Smooth density estimation for length-biased data. *J.Indian Soc.Agricultural Statist.* **64** 145–155.
- Chen, S.X. 1999. Beta kernel estimators for density functions. *Comput.Statist.Data.Anal.* **31** 131–145.
- Chen, S.X. 2000. Probability density function estimation using gamma kernels. *Ann.Inst.Stat.Math.* **52** 471–480.
- Cox, D.R. 1969. Some sampling problems in technology. In *New Developments in Survey Sampling*; Johnson, N.L. and Smith, H.Jr., Eds.; New York: Wiley, pp.506–527.
- Cristóbal, J.A. and Alcalá, J.T. 2001. An overview of nonparametric contributions to the problem of functional estimation from biased data. *Test* **10** 309–332.
- Feller, W. 1971. *An Introduction to Probability Theory and its Applications*, Vol. II, 2nd edn.; New York: Wiley.
- Guillamón, A., Navarro, J., and Ruiz, J.M. 1998. Kernel density estimation using weighted data. *Comm.Statist.Theory Methods* **27** 2123–2135.
- Igarashi, G. 2016. Weighted log-normal kernel density estimation. *Comm.Statist.Theory Methods* **45** 6670–6687.

- Igarashi, G. 2018. Multivariate density estimation using a multivariate weighted log-normal kernel. *Sankyā* **80** 247–266.
- Igarashi, G. and Kakizawa, Y. 2014. Re-formulation of inverse Gaussian, reciprocal inverse Gaussian, and Birnbaum–Saunders kernel estimators. *Statist.Probab.Lett.* **84** 235–246.
- Igarashi, G. and Kakizawa, Y. 2020. Multiplicative bias correction for asymmetric kernel density estimators revisited. *Comput.Statist.Data.Anal.* **141** 40–61.
- Jones, M.C. 1991. Kernel density estimation for length biased data. *Biometrika* **78** 511–519.
- Jones, M.C. 1993. Simple boundary correction for kernel density estimation. *Stat.Comput.* **3** 135–146.
- Kakizawa, Y. 2018. Nonparametric density estimation for nonnegative data, using symmetrical-based inverse and reciprocal inverse Gaussian kernels through dual transformation. *J.Statist.Plann.Inference* **193** 117–135.
- Kakizawa, Y. 2021. A class of Birnbaum–Saunders type kernel density estimators for nonnegative data. *Comput.Statist.Data.Anal.* **161** Article 107249, 1–18.
- Kakizawa, Y. 2022. Multivariate elliptical-based Birnbaum–Saunders kernel density estimation for nonnegative data. *J.Multivariate.Anal.* **187** Article 104834, 1–20.
- Marron, J.S. and Ruppert, D. 1994. Transformations to reduce boundary bias in kernel density estimation. *J.R.Stat.Soc. B* **56** 653–671.
- Mnatsakanov, R.M. and Ruymgaart, F.H. 2003. Some properties of moment-empirical cdf's with application to some inverse estimation problems. *Math.Methods Statist.* **12** 478–495.
- Mnatsakanov, R.M. and Ruymgaart, F.H. 2006. On moment-density estimation in some biased models. In *Optimality: The Second Erich L. Lehmann Symposium*; Institute of Mathematical Statistics; pp.322–333.
- Parzen, E. 1962. On estimation of a probability density function and mode. *Ann.Math.Statist.* **33** 1065–1076.
- Patil, G.P. and Rao, C.R. 1978. Weighted distributions and size-biased sampling with applications to wildlife populations and human families. *Biometrics* **34** 179–189.
- Richardson, G.D., Kazempour, M.K., and Bhattacharyya, B.B. 1991. Length biased density estimation of fibres. *J.Nonparametr.Stat.* **1** 127–141.
- Rosenblatt, M. 1956. Remarks on some nonparametric estimates of density functions. *Ann.Math.Statist.* **27** 832–837.
- Scaillet, O. 2004. Density estimation using inverse and reciprocal inverse Gaussian kernels. *J.Nonparametr.Stat.* **16** 217–226.
- Silverman, B.W. 1986. *Density Estimation for Statistics and Data Analysis*; London: Chapman and Hall.
- Zhang, S., Karunamuni, R.J., and Jones, M.C. 1999. An improved estimator of the density function at the boundary. *J.Amer.Statist.Assoc.* **94** 1231–1241.

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.