Department of Policy and Planning Sciences

Discussion Paper Series

No.1358

Higher-order bias corrections for kernel type density estimators on the unit or semi-infinite interval

by

Gaku IGARASHI and Yoshihide KAKIZAWA

Apr 2019

UNIVERSITY OF TSUKUBA
Tsukuba, Ibaraki 305-8573
JAPAN

Higher-order bias corrections for kernel type density estimators on the unit or semi-infinite interval

Gaku Igarashi[1] and Yoshihide Kakizawa[2]

[1] Department of Policy and Planning Sciences, Faculty of Engineering, Information and Systems, University of Tsukuba, 1-1-1 Tennodai, Tsukuba, Ibaraki 305-8573, Japan
[2]Faculty of Economics, Hokkaido University, Nishi 7, Kita 9, Kita-ku, Sapporo 060-0809, Japan

Abstract

For the data of size n from the unit or semi-infinite interval, several asymmetric kernel density estimators (KDEs), having the mean integrated squared errors (MISEs) of order $O(n^{-4/5})$ or $O(n^{-8/9})$, have been studied over the last two decades. In this paper, we develop more higher-order bias-corrected asymmetric KDEs, achieving the order $O(n^{-4p/(4p+1)})$, where $p \geq 2$ is a given integer; these higher-order bias correction methods can be also applied to the classical Rosenblatt–Parzen KDEs. We illustrate the finite sample performance of the higher-order bias-corrected asymmetric KDEs through the simulations.

Keywords: nonparametric density estimation; boundary bias problem; asymmetric kernel; higher-order bias correction;
MSC: 62G07; 62G20

1. Introduction

The kernel density estimator (KDE), $\widehat{f}_h^{(K)}(x) = (nh)^{-1}\sum_{i=1}^n K((x - X_i)/h)$, developed by Rosenblatt (1956) and Parzen (1962), is a popular nonparametric estimator, where $\{X_1, \dots, X_n\}$ is a random sample drawn from an unknown density f with support $\mathbb{R}$, $h > 0$ is a bandwidth, and K is a symmetric kernel. If f is $2p$ times continuously differentiable for some $p \in \mathbb{N}$, using a $2p$th-order kernel $K_{[2p]}$, i.e., $\int_{-\infty}^{\infty} K_{[2p]}(s)ds = 1$, $\int_{-\infty}^{\infty} s^\ell K_{[2p]}(s)ds = 0$, $\ell = 1, \dots, 2p-1$, and $\int_{-\infty}^{\infty} s^{2p}K_{[2p]}(s)ds \neq 0$, the bias and variance of the $2p$th-order KDE $\widehat{f}_h^{(K_{[2p]})}$ are $O(h^{2p})$ and $O(n^{-1}h^{-1})$, respectively, hence, with $h \propto n^{-1/(4p+1)}$, the mean squared error (MSE) and mean integrated squared error (MISE) are $O(n^{-4p/(4p+1)})$. The use of higher-order kernels enables us to get the faster convergence rate of the M(I)SE. Schucany and Sommers (1977) and Jones and Foster (1993) addressed how to generate a reasonable $K_{[4]}$ from a given $K_{[2]}$, in a variety of ways. One attractive and simple answer is to produce a class of the fourth-order kernels

$$K_{[4],(1,a)}(s) = \begin{cases} \dfrac{1}{1-a^2}\{K_{[2]}(s) - a^3K_{[2]}(as)\}, & a \neq 1, \\ \dfrac{1}{2}\{3K_{[2]}(s) + sK'_{[2]}(s)\}, & a = 1. \end{cases}$$

Email: g-igarashi@sk.tsukuba.ac.jp (G. Igarashi), kakizawa@econ.hokudai.ac.jp (Y. Kakizawa).

The authors reported some preliminary results at the Japanese Joint Statistical Meeting (2018, September) and the Mathematical Society of Japan Autumn Meeting (2018, September).

The first author has been supported by the Japan Society for the Promotion of Science (JSPS); Grant-in-Aid for Young Scientists (B) [Grant Number 17K13714]. The second author has been supported by the JSPS; Grant-in-Aid for Scientific Research (C) [Grant Number 17K00041].

However, by definition, the $2p$th-order KDE $\widehat{f}_h^{(K_{[2p]})}$ necessarily loses the nonnegativity unless $p = 1$, so that nonnegative bias correction methods were discussed by Terrell and Scott (1980), Jones and Foster (1993), and Jones et al. (1995).

Unfortunately, if $supp(f) \neq \mathbb{R}$, the classical Rosenblatt–Parzen KDE has, in general, the boundary bias which is $O(1)$ near the boundary of $supp(f)$. Various remedies were studied, e.g., renormalization, reflection, generalized jackknifing (Jones (1993)), transformation (Marron and Ruppert (1994)), and advanced reflection (Zhang et al. (1999)). On the other hand, instead of the location-scale type $K((x - \cdot)/h)/h$, applying several asymmetric kernels whose support match $supp(f)$ has attracted considerable attention over the last two decades. Among many papers are Chen (1999, 2000), Jin and Kawczak (2003), Scaillet (2004), Marchant et al. (2013), Saulo et al. (2013), Igarashi and Kakizawa (2014b), Igarashi (2016b), Kakizawa and Igarashi (2017), and Kakizawa (2018). Note that "good" asymmetric KDEs have the MISEs of order $O(n^{-4/5})$. To achieve the order $O(n^{-8/9})$, some bias correction methods have been further discussed in recent years, even when $supp(f) = [0, 1]$ or $[0, \infty)$. See Hirukawa (2010; correction 2016), Leblanc (2010), Hirukawa and Sakudo (2014, 2015), Igarashi and Kakizawa (2014a, 2015, 2018a,b,c), Igarashi (2016a), and Zougab and Adjabi (2016).

The objective of this paper is to develop more higher-order bias-corrected density estimation, by generalizing novel ideas of the bias correction methods due to Schucany and Sommers (1977), Terrell and Scott (1980), and Jones and Foster (1993). In Section 2, we describe basic asymptotic properties of the asymmetric KDE (without bias corrections). In Section 3, we establish that the proposed higher-order bias-corrected asymmetric KDEs attain the convergence rate $n^{-4p/(4p+1)}$ of the MISE. It is revealed, however, that, for some asymmetric KDEs, the convergence rates after the bias corrections are at most $n^{-(4p^*+2)/(4p^*+3)}$, with an integer p^*. In Section 4, we provide examples of kernels with support $[0, \infty)$ or $[0, 1]$. Section 5 presents simulation studies to illustrate the finite sample performance of the bias-corrected estimators. Some general comments are given in Section 6. The proofs are postponed to the Appendix.

Notation The dependency on the sample size n is suppressed (e.g., the smoothing parameter is denoted by β, instead of β_n), but, unless otherwise stated, the limits will be taken as $n \to \infty$. We write $\mathcal{S} = supp(f)$ for simplicity, and, as usual, use the notation $||h||_{\mathcal{S}} = \sup_{x \in \mathcal{S}} |h(x)|$ for any bounded function h on $\mathcal{S}$. We denote by $h^{(j)}(x) = (d/dx)^j h(x)$ the jth derivative of h (if it exists), and write $h^{(0)}(x) = h(x)$. Further, χ_A denotes the indicator function of a set A, and $\lceil y \rceil$ denotes the smallest integer greater than or equal to y. Conventionally, the empty sum (e.g., $\sum_{k=1}^{0}$) is defined to equal zero. The bias, MSE, and MISE for an estimator $\widehat{f}(x)$ of $f(x)$, $x \in \mathcal{S}$, are denoted by $Bias[\widehat{f}(x)] = E[\widehat{f}(x)] - f(x)$, $MSE[\widehat{f}(x)] = E[\{\widehat{f}(x) - f(x)\}^2]$, and $MISE[\widehat{f}] = \int_{\mathcal{S}} MSE[\widehat{f}(x)]dx$.

2. Preliminaries

We assume that $\mathcal{X}^{(n)} = \{X_1, \ldots, X_n\}$ is a random sample drawn from an unknown density f with support $\mathcal{S} = [0, \infty)$ or $[0, 1]$. Let $\beta > 0$ be a smoothing parameter, such that $\beta \to 0$ and $n\beta \to \infty$, unless otherwise stated. We construct an estimator in the form of

$$\widehat{f}_\beta(x) = \frac{1}{n}\sum_{i=1}^{n} K(X_i; x, \beta), \quad x \in \mathcal{S} \tag{1}$$

(referred to as an asymmetric KDE throughout this paper), where $K(\cdot; x, \beta)(\geq 0)$ is a density with support $\mathcal{S}$, such that the kernel $K(s; x, \beta)$ concentrates around $s = x$ as $\beta \to 0$.

Before proceeding to complete description of assumptions, we briefly mention what kinds of assumptions are required here. According to our previous works (e.g., Igarashi and Kakizawa (2015, 2018a,c) and Igarashi (2016a)), additional properties on $K(\cdot; x, \beta)$, i.e.,

- the uniform/nonuniform bounds of $\sup_{s \in \mathcal{S}} K(s; x, \beta)$,
- the tractability of the product kernel $K(s; x, \beta/a_0)K(s; x, \beta/a_0')$ for any $a_0, a_0' > 0$, and
- when $\mathcal{S} = [0, \infty)$, the asymptotic behaviour of $\int_{\beta^{-\tau}}^{\infty} K(s; x, \beta)dx$ for any $\tau \in (0, 1)$

(Assumptions A1–A3) are indispensable. The jth moment around $x \in \mathcal{S}$ is denoted by

$$\mu_j(K(\cdot; x, \beta)) = \int_{\mathcal{S}} (s - x)^j K(s; x, \beta)ds \quad \text{(if it exists)}.$$

Note that $\mu_0(K(\cdot; x, \beta)) \equiv 1$, since the chosen kernel is a certain density with support $\mathcal{S}$. The results in this paper heavily depend on the moments up to the $2(p+1)$th order, for some $p \in \mathbb{N}$ (Assumption A4[p]), under which $\mu_j(K(\cdot; x, \beta))$, $x \in \mathcal{S}$, is expanded as a power of β. The regularity on the density f to be estimated (Assumption A5[p](i,ii) or A5$'$(i)) is standard in nonparametric density estimation. It should be remarked that Assumption A3, together with the latter part of A5[p](iii) (or A5$'$(ii)), is somewhat technical, but will be used only for the approximations of the integrated squared bias/variance when $\mathcal{S} = [0, \infty)$.

2.1. Assumptions

Throughout this paper, we denote by

$$\mathcal{S}_I = \begin{cases} (0, \infty), & \mathcal{S} = [0, \infty), \\ (0, 1), & \mathcal{S} = [0, 1] \end{cases} \quad \text{and} \quad \mathcal{S}_B = \begin{cases} \{0\}, & \mathcal{S} = [0, \infty), \\ \{0, 1\}, & \mathcal{S} = [0, 1] \end{cases}$$

the interior and boundary, respectively, of $\mathcal{S}$. We often distinguish between the two cases of a set of points far away from $\mathcal{S}_B$ and a set of points near $\mathcal{S}_B$, i.e.,

$$\mathcal{S}_{I,\beta} = \begin{cases} \left\{x \in \mathcal{S} \,\middle|\, \dfrac{x}{\beta} \to \infty\right\}, & \mathcal{S} = [0, \infty), \\ \left\{x \in \mathcal{S} \,\middle|\, \dfrac{x}{\beta} \to \infty, \dfrac{1-x}{\beta} \to \infty\right\}, & \mathcal{S} = [0, 1], \end{cases}$$

$$\mathcal{S}_{B,\beta,\kappa} = \begin{cases} \mathcal{S}_{0,\beta,\kappa}, & \mathcal{S} = [0, \infty), \\ \mathcal{S}_{0,\beta,\kappa} \cup \mathcal{S}_{1,\beta,\kappa}, & \mathcal{S} = [0, 1], \end{cases}$$

where $\mathcal{S}_{0,\beta,\kappa} = \{x \in \mathcal{S} \mid x/\beta \to \kappa\}$ and $\mathcal{S}_{1,\beta,\kappa} = \{x \in \mathcal{S} \mid (1-x)/\beta \to \kappa\}$ (here and subsequently, $\kappa \geq 0$ is a constant, unless otherwise stated). Also, we write

$$\psi(x) = \begin{cases} x, & \mathcal{S} = [0,\infty), \\ x(1-x), & \mathcal{S} = [0,1] \end{cases} \quad \text{and} \quad V(x;f) = \frac{f(x)}{2\sqrt{\pi\psi(x)}}.$$

Then, the following assumptions on $K(\cdot; x, \beta)$ and f are made for some $p \in \mathbb{N}$:

A1. (i) $\sup_{x\in\mathcal{S}} \sup_{s\in\mathcal{S}} K(s;x,\beta) \leq C_K \beta^{-1}$ for some constant $C_K > 0$, independent of β.

(ii) When $x \in \mathcal{S}_I$, $\sup_{s\in\mathcal{S}} K(s;x,\beta) \leq C'_K \{\beta\psi(x)\}^{-1/2}$ for some constant $C'_K > 0$, independent of β and x.

A2. For any constants $a_0, a'_0 > 0$,

$$\int_{\mathcal{S}} K(s;x,\beta/a_0)K(s;x,\beta/a'_0)ds = \begin{cases} \left(\dfrac{2a_0a'_0}{a_0+a'_0}\right)^{1/2} \dfrac{\beta^{-1/2}}{2\sqrt{\pi\psi(x)}} + O(\beta^{1/2}\{\psi(x)\}^{-3/2}), & x \in \mathcal{S}_{I,\beta}, \\ \beta^{-1}\varsigma_{a_0,a'_0}(\kappa)[1+\chi_{\{x\notin\mathcal{S}_B\}}o(1)], & x \in \mathcal{S}_{B,\beta,\kappa} \end{cases}$$

for some function ς_{a_0,a'_0}, independent of β.

A3. When $\mathcal{S} = [0,\infty)$, for any constants $k > 0$ and $\tau \in (0,1)$, and for all sufficiently small $\beta > 0$,

$$\int_{\beta^{-\tau}}^{\infty} K(s;x,\beta)dx = O(\beta^{\tau(k+1)}s^{k+1}), \quad s > 0.$$

A4[p]. The moments around $x \in \mathcal{S}$ admit asymptotic expansions, as follows: when $\mathcal{S} = [0,\infty)$,

$$\mu_j(K(\cdot;x,\beta)) = \begin{cases} \displaystyle\sum_{k=\lceil j/2\rceil}^{\min(j,p)} \zeta_{j,k}\, x^{j-k}\beta^k + \chi_{\{j>p\}}O(\beta^{p+1}(x+\beta)^{j-(p+1)}), & j = 1,\ldots,2p, \\ O(\beta^{p+1}(x+\beta)^{p+1}), & j = 2(p+1) \end{cases}$$

for some constants $\zeta_{j,k}$'s, independent of β and x, whereas, when $\mathcal{S} = [0,1]$, uniformly in $x \in [0,1]$,

$$\mu_j(K(\cdot;x,\beta)) = \begin{cases} \displaystyle\sum_{k=\lceil j/2\rceil}^{p} \zeta_{j,k}(x)\beta^k + O(\beta^{p+1}), & j = 1,\ldots,2p, \\ O(\beta^{p+1}), & j = 2(p+1) \end{cases}$$

for some polynomials in x; $\zeta_{j,k}(x)$'s, independent of β, where $\zeta_{2,1}(x) = \psi(x)$.

A5[p]. (i) f is $2p$ times continuously differentiable on $\mathcal{S}$, with $\sum_{j=0}^{2p} ||f^{(j)}||_{\mathcal{S}} < \infty$.

(ii) $f^{(2p)}$ is Hölder continuous on $\mathcal{S}$, i.e., there exist constants $\eta_{2p} \in (0,1]$ and $L_{2p} > 0$, such that $|f^{(2p)}(s) - f^{(2p)}(t)| \leq L_{2p}|s-t|^{\eta_{2p}}$ for any $s,t \in \mathcal{S}$.

(iii) When $\mathcal{S} = [0,\infty)$, $\sum_{j=p}^{2p}\int_0^\infty \{x^{j-p}f^{(j)}(x)\}^2dx < \infty$ and there exists a constant $k_{2p} > \{2p(2p+1)+(2p-1)\eta_{2p}\}/\eta_{2p}$ such that $\int_0^\infty x^{k_{2p}+1}f(x)dx < \infty$ (in this case, for any constant $\tau_{2p} \in (2p/(k_{2p}+1), \eta_{2p}/(2p+1+\eta_{2p}))$, $\int_0^{\beta^{-\tau_{2p}}} \beta^{\eta_{2p}}(1+x^{p+\eta_{2p}/2})^2dx = o(1)$).

Note that, in some cases, Assumptions A4[p] and A5[p] will be weakened, as follows:

A4$'$[J]. When $\mathcal{S} = [0, \infty)$,

$$\mu_j(K(\cdot; x, \beta)) = \sum_{k=\lceil j/2 \rceil}^{\min(j,J-1)} \zeta_{j,k}\, x^{j-k} \beta^k + \chi_{\{j>J-1\}} O(\beta^J (x+\beta)^{j-J}), \quad j = 1, \ldots, 2J$$

for some constants $\zeta_{j,k}$'s, independent of β and x, whereas, when $\mathcal{S} = [0,1]$, uniformly in $x \in [0,1]$,

$$\mu_j(K(\cdot; x, \beta)) = \sum_{k=\lceil j/2 \rceil}^{J-1} \zeta_{j,k}(x) \beta^k + O(\beta^J), \quad j = 1, \ldots, 2J$$

for some polynomials in x; $\zeta_{j,k}(x)$'s, independent of β, where $\zeta_{2,1}(x) = \psi(x)$.

A5$'$. (i) f is continuously differentiable on $\mathcal{S}$, with $||f||_{\mathcal{S}} + ||f^{(1)}||_{\mathcal{S}} < \infty$.

(ii) When $\mathcal{S} = [0, \infty)$, there exists a constant $k' > 0$, such that $\int_0^\infty x^{k'+1} f(x) dx < \infty$.

2.2. Asymptotic properties of asymmetric KDE (without bias corrections)

In this subsection, the asymptotic properties of the asymmetric KDE (1) are presented.

Theorem 1 (i) *Suppose that Assumptions* A4[p] *and* A5[p](i,ii) *hold for some* $p \in \mathbb{N}$. *Then,*

$$Bias[\widehat{f}_\beta(x)] = \sum_{k=1}^{p} \beta^k \gamma_k(x; f) + \mathcal{E}_{\beta,p}(x), \quad x \in \mathcal{S},$$

where $\gamma_k(x; f)$ *and* $\mathcal{E}_{\beta,p}(x)$ *are given, as follows: when* $\mathcal{S} = [0, \infty)$,

$$\gamma_k(x; f) = \sum_{j=k}^{2k} \zeta_{j,k}\, x^{j-k} \frac{f^{(j)}(x)}{j!}, \quad \mathcal{E}_{\beta,p}(x) = O(\beta^{p+\eta_{2p}/2}(1+x)^{p+\eta_{2p}/2}),$$

and, when $\mathcal{S} = [0,1]$,

$$\gamma_k(x; f) = \sum_{j=1}^{2k} \zeta_{j,k}(x) \frac{f^{(j)}(x)}{j!}, \quad \mathcal{E}_{\beta,p}(x) = O(\beta^{p+\eta_{2p}/2}) \text{ uniformly in } x \in [0,1].$$

(ii) *Suppose that Assumptions* A1, A2, A4$'$[1], *and* A5$'$(i) *hold. Then,*

$$V[\widehat{f}_\beta(x)] = \begin{cases} n^{-1}\beta^{-1/2} V(x; f)[1 + O(\beta\psi^{-1}(x))] + O(n^{-1}), & x \in \mathcal{S}_{I,\beta}, \\ n^{-1}\beta^{-1} f(x)[\varsigma_{1,1}(\kappa) + \chi_{\{x \notin \mathcal{S}_B\}} o(1)] + O(n^{-1}), & x \in \mathcal{S}_{B,\beta,\kappa}. \end{cases}$$

(iii) *Suppose that Assumption* A1(i) *holds. If* $n\beta/\log n \to \infty$, *then,* $\widehat{f}_\beta(x) - E[\widehat{f}_\beta(x)] \stackrel{a.s.}{\to} 0$, $x \in \mathcal{S}$.

Remark 1 Under Assumptions A4$'$[1] and A5[1](i), we have (see also Remark A.1)

$$\text{when } \mathcal{S} = [0, \infty), \ Bias[\widehat{f}_\beta(x)] = O(\beta(1+x)), \tag{2}$$

$$\text{when } \mathcal{S} = [0,1], \text{ uniformly in } x \in [0,1], \ Bias[\widehat{f}_\beta(x)] = O(\beta). \tag{2$'$}$$

Theorem 1(iii) immediately yields the strong consistency of the estimator (1);

$$\widehat{f}_\beta(x) \stackrel{a.s.}{\to} f(x) \quad \text{for fixed } x \in \mathcal{S}. \tag{3}$$

Theorem 2 *Suppose that Assumptions* A1*,* A2*,* A4′[1]*, and* A5′(i) *hold. Then,*

$$(n\beta^{1/2})^{1/2}\{\widehat{f}_\beta(x) - E[\widehat{f}_\beta(x)]\} \xrightarrow{d} N(0, V(x; f)) \quad \textit{for fixed } x \in \mathcal{S}_I,$$
$$(n\beta)^{1/2}\{\widehat{f}_\beta(x) - E[\widehat{f}_\beta(x)]\} \xrightarrow{d} N(0, \varsigma_{1,1}(0)f(x)) \quad \textit{for } x \in \mathcal{S}_B.$$

A replacement of $E[\widehat{f}_\beta(x)]$ by $f(x)$ (or $f(x) + \beta\gamma_1(x; f)$) is a routine problem in density estimation theory (use Slutsky's lemma; see Theorems 1(i) and 2)[1].

Theorem 1 shows that

$$MSE[\widehat{f}_\beta(x)] = \begin{cases} AMSE[\widehat{f}_\beta(x)] + o(\beta^2 + n^{-1}\beta^{-1/2}) & \text{for fixed } x \in \mathcal{S}_I, \\ AMSE[\widehat{f}_\beta(x)] + o(\beta^2 + n^{-1}\beta^{-1}) & \text{for } x \in \mathcal{S}_{B,\beta,\kappa}, \end{cases}$$

where

$$AMSE[\widehat{f}_\beta(x)] = \begin{cases} \beta^2\gamma_1^2(x; f) + n^{-1}\beta^{-1/2}V(x; f) & \text{for fixed } x \in \mathcal{S}_I, \\ \beta^2\gamma_1^2(x; f) + n^{-1}\beta^{-1}\varsigma_{1,1}(\kappa)f(x) & \text{for } x \in \mathcal{S}_{B,\beta,\kappa}. \end{cases}$$

Note that

$$\min_{\beta>0} AMSE[\widehat{f}_\beta(x)] = \begin{cases} \dfrac{5}{4}\left[4\gamma_1^2(x; f)\{V(x; f)n^{-1}\}^4\right]^{1/5} & \text{for fixed } x \in \mathcal{S}_I, \\ \dfrac{3}{2}\left[2\gamma_1^2(x; f)\{\varsigma_{1,1}(\kappa)f(x)n^{-1}\}^2\right]^{1/3} & \text{for } x \in \mathcal{S}_{B,\beta,\kappa} \end{cases}$$

(we assume $\gamma_1(x; f) \neq 0$). Although the estimator (1) has the slower convergence rate near the boundary $\mathcal{S}_B$, such a different rate is asymptotically negligible on the MISE.

Theorem 3 *Suppose that Assumptions* A1–A3*,* A4[1]*, and* A5[1] *hold. Then,*

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

where

$$AMISE[\widehat{f}_\beta] = \beta^2 \int_{\mathcal{S}} \gamma_1^2(x; f)dx + n^{-1}\beta^{-1/2} \int_{\mathcal{S}} V(x; f)dx$$

is minimized at

$$\beta = \left[\frac{\int_{\mathcal{S}} V(x; f)dx}{4\int_{\mathcal{S}} \gamma_1^2(x; f)dx} n^{-1}\right]^{2/5}$$

when $\gamma_1(x; f) \not\equiv 0$, that is,

$$\min_{\beta>0} AMISE[\widehat{f}_\beta] = \frac{5}{4}\left[4\int_{\mathcal{S}} \gamma_1^2(x; f)dx\left\{\int_{\mathcal{S}} V(x; f)dx\, n^{-1}\right\}^4\right]^{1/5}.$$

[1]Suppose that Assumptions A1, A2, A4[1], and A5[1](i,ii) hold.
(i). If $n\beta^{5/2+\eta_2} \to 0$, then, for fixed $x \in \mathcal{S}_I$,

$$(n\beta^{1/2})^{1/2}\{\widehat{f}_\beta(x) - f(x) - \beta\gamma_1(x; f)\} \xrightarrow{d} N(0, V(x; f)),$$

hence, if $n\beta^{5/2} \to 0$, then, $(n\beta^{1/2})^{1/2}\{\widehat{f}_\beta(x) - f(x)\} \xrightarrow{d} N(0, V(x; f))$.
(ii). If $n\beta^{3+\eta_2} \to 0$, then, for $x \in \mathcal{S}_B$,

$$(n\beta)^{1/2}\{\widehat{f}_\beta(x) - f(x) - \beta\gamma_1(x; f)\} \xrightarrow{d} N(0, \varsigma_{1,1}(0)f(x)),$$

hence, if $n\beta^3 \to 0$, then, $(n\beta)^{1/2}\{\widehat{f}_\beta(x) - f(x)\} \xrightarrow{d} N(0, \varsigma_{1,1}(0)f(x))$.

3. Additive, TS-type, and JF-type bias corrections

The main contribution of this paper is to study higher-order extensions of the previous works (e.g., Igarashi and Kakizawa (2015, 2018a) and Igarashi (2016a)). From now on, let $p \in \mathbb{N}\backslash\{1\}$, unless otherwise stated. Given a positive vector $\boldsymbol{a} = (a_1, \ldots, a_p)'$, such that the a_k's are distinct, the additive, TS-type, and JF-type bias-corrected KDEs of $f(x)$, $x \in \mathcal{S}$, are defined by

$$\widehat{f}_{\beta,ADD_{\boldsymbol{a}}^p}(x) = \sum_{k=1}^{p} c_k(\boldsymbol{a}) \widehat{f}_{\beta/a_k}(x) = \frac{1}{n}\sum_{i=1}^{n} K_{ADD_{\boldsymbol{a}}^p}(X_i; x, \beta) \quad \text{(say)}, \tag{4}$$

$$\widehat{f}_{\beta,TS_{\boldsymbol{a}}^p}(x) = \exp\left[\sum_{k=1}^{p} c_k(\boldsymbol{a}) \log\left\{\widehat{f}_{\beta/a_k}(x) + \frac{\epsilon}{a_k}\right\}\right] = \prod_{k=1}^{p}\left\{\widehat{f}_{\beta/a_k}(x) + \frac{\epsilon}{a_k}\right\}^{c_k(\boldsymbol{a})}, \tag{5}$$

$$\widehat{f}_{\beta,JF_{\boldsymbol{a}}^p}(x) = \{\widehat{f}_{\beta}(x) + \epsilon\} \exp\left[\sum_{j=1}^{p-1} \frac{(-1)^{j-1}}{j}\left\{\frac{\widehat{f}_{\beta,ADD_{\boldsymbol{a}}^p}(x)}{\widehat{f}_{\beta}(x)+\epsilon} - 1\right\}^j\right], \tag{6}$$

respectively, where $\epsilon = \epsilon_\beta \to 0$, specified later, is introduced to avoid $\log 0$ and the division by zero, and $\{c_1(\boldsymbol{a}), \ldots, c_p(\boldsymbol{a})\}$ is unique solution of

$$\sum_{k=1}^{p} c_k(\boldsymbol{a}) = 1, \quad \sum_{k=1}^{p} \frac{c_k(\boldsymbol{a})}{a_k^\ell} = 0, \quad \ell = 1, \ldots, p-1. \tag{7}$$

The following result (Lemma 4), independent of interest, enables us to see that

$$\sum_{k=1}^{p} \frac{c_k(\boldsymbol{a})}{a_k^p} = \frac{(-1)^{p-1}}{\prod_{k=1}^{p} a_k}. \tag{8}$$

Lemma 4 *For any $\boldsymbol{z} = (z_1, \ldots, z_p)' \in \mathbb{R}^p$, let $\mathcal{V}(\boldsymbol{z})$ be a Vandermonde matrix of $p \times p$, whose jth column is the vector $(1, z_j, \ldots, z_j^{p-1})'$ for $j = 1, \ldots, p$. If $\mathcal{V}(\boldsymbol{z})$ is invertible (i.e., the z_j's are assumed to be distinct), then, $\prod_{j=1}^p z_j = (-1)^{p-1} \sum_{j=1}^p z_j^p [\mathcal{V}^{-1}(\boldsymbol{z})]_{j1}$, where $[\mathcal{V}^{-1}(\boldsymbol{z})]_{jk}$ is the (j,k)th element of $\mathcal{V}^{-1}(\boldsymbol{z})$.*

Remark 2 The solution $\{c_k\}$ of (7) is computable for user's specified vector $\boldsymbol{a}$, i.e.,

$$c_k(\boldsymbol{a}) = \frac{a_k^{p-1}}{\prod_{j=1, j\neq k}^{p}(a_k - a_j)}, \quad k = 1, \ldots, p,$$

using the inversion of the Vandermonde matrix (e.g., (3.2) of Gautschi (1962)). For example,

- $\boldsymbol{a} = (1, 1/2, \ldots, 1/p)'$ yields $c_k(\boldsymbol{a}) = (-1)^{k-1}{}_pC_k$ for $k = 1, \ldots, p$, and
- $\boldsymbol{a} = (1, (p-1)/p, (p-2)/(p-1), \ldots, 1/2)'$ (i.e., $a_1 = 1$ and $a_k = (p-k+1)/(p-k+2)$ for $k = 2, \ldots, p$) yields $c_1(\boldsymbol{a}) = p!$ and $c_k(\boldsymbol{a}) = (-1)^{k-1}(p-k+1)^p{}_pC_{k-2}$ for $k = 2, \ldots, p$.

Practically, the selection of $\boldsymbol{a}$ is a difficult problem. In Section 5, numerical studies for $p = 2$ and $p = 3$ will be conducted by letting $\boldsymbol{a} = (1, a)$ and $\boldsymbol{a} = (1, a, 1/a)$, respectively, where $a \in (0, 1)$.

Before presenting the main results in this paper, we mention that, as an easy corollary of the strong consistency of the estimator (1), the estimators (4)–(6) are also strong consistent (for the estimators (5) and (6), we additionally assume that $\epsilon \to 0$), i.e., by virtue of Slutsky's lemma, (3) and (7) immediately yield, for fixed $x \in \mathcal{S}$,

- $\widehat{f}_{\beta,ADD^p_{\boldsymbol{a}}}(x) \stackrel{a.s.}{\to} \sum_{k=1}^{p} c_k(\boldsymbol{a}) f(x) = f(x)$, and
- if $f(x) > 0$, then, $\widehat{f}_{\beta,TS^p_{\boldsymbol{a}}}(x) \stackrel{a.s.}{\to} \exp\left[\sum_{k=1}^{p} c_k(\boldsymbol{a}) \log f(x)\right] = f(x)$ and $\widehat{f}_{\beta,JF^p_{\boldsymbol{a}}}(x) \stackrel{a.s.}{\to} f(x)$.

3.1. Asymptotic properties of additive estimator

To begin with, we consider the additive estimator (4). We write

$$B_{p,\boldsymbol{a}}(x;f) = \frac{(-1)^{p-1}}{\prod_{k=1}^{p} a_k}\gamma_p(x;f), \quad \lambda_{p,\boldsymbol{a}} = \sum_{j=1}^{p}\sum_{j'=1}^{p} c_j(\boldsymbol{a})c_{j'}(\boldsymbol{a})\Big(\frac{2a_j a_{j'}}{a_j + a_{j'}}\Big)^{1/2},$$
$$\upsilon_{p,\boldsymbol{a}}(\kappa) = \sum_{j=1}^{p}\sum_{j'=1}^{p} c_j(\boldsymbol{a})c_{j'}(\boldsymbol{a})\varsigma_{a_j,a_{j'}}(\kappa).$$

Theorem 5 (i) *Suppose that Assumptions* A4[p] *and* A5[p](i,ii) *hold for some* $p \in \mathbb{N}\backslash\{1\}$. *Then,*

$$Bias[\widehat{f}_{\beta,ADD^p_{\boldsymbol{a}}}(x)] = \beta^p B_{p,\boldsymbol{a}}(x;f) + \mathcal{E}_{\beta,ADD^p_{\boldsymbol{a}}}(x), \quad x \in \mathcal{S},$$

where $\mathcal{E}_{\beta,ADD^p_{\boldsymbol{a}}}(x) = \sum_{k=1}^{p} c_k(\boldsymbol{a})\mathcal{E}_{\beta/a_k,p}(x)$.
(ii) *Suppose that Assumptions* A1, A2, A4′[1], *and* A5′(i) *hold. Then,*

$$V[\widehat{f}_{\beta,ADD^p_{\boldsymbol{a}}}(x)] = \begin{cases} n^{-1}\beta^{-1/2}\lambda_{p,\boldsymbol{a}}V(x;f)[1 + O(\beta\psi^{-1}(x))] + O(n^{-1}), & x \in \mathcal{S}_{I,\beta}, \\ n^{-1}\beta^{-1}f(x)[\upsilon_{p,\boldsymbol{a}}(\kappa) + \chi_{\{x\notin\mathcal{S}_B\}}o(1)] + O(n^{-1}), & x \in \mathcal{S}_{B,\beta,\kappa}. \end{cases}$$

Theorem 6 *Suppose that Assumptions* A1, A2, A4′[1], *and* A5′(i) *hold. Then,*

$$(n\beta^{1/2})^{1/2}\{\widehat{f}_{\beta,ADD^p_{\boldsymbol{a}}}(x) - E[\widehat{f}_{\beta,ADD^p_{\boldsymbol{a}}}(x)]\} \stackrel{d}{\to} N(0, \lambda_{p,\boldsymbol{a}}V(x;f)) \quad \textit{for fixed } x \in \mathcal{S}_I,$$
$$(n\beta)^{1/2}\{\widehat{f}_{\beta,ADD^p_{\boldsymbol{a}}}(x) - E[\widehat{f}_{\beta,ADD^p_{\boldsymbol{a}}}(x)]\} \stackrel{d}{\to} N(0, \upsilon_{p,\boldsymbol{a}}(0)f(x)) \quad \textit{for } x \in \mathcal{S}_B.$$

A replacement of $E[\widehat{f}_{\beta,ADD^p_{\boldsymbol{a}}}(x)]$ by $f(x)$ (or $f(x) + \beta^p B_{p,\boldsymbol{a}}(x;f)$) is a routine problem in density estimation theory (use Slutsky's lemma; see Theorems 5(i) and 6)[2].

---

[2]Suppose that Assumptions A1, A2, A4[p], and A5[p](i,ii) hold for some $p \in \mathbb{N}\backslash\{1\}$.
(i). If $n\beta^{(4p+1)/2+\eta_{2p}} \to 0$, then, for fixed $x \in \mathcal{S}_I$,

$$(n\beta^{1/2})^{1/2}\{\widehat{f}_{\beta,ADD^p_{\boldsymbol{a}}}(x) - f(x) - \beta^p B_{p,\boldsymbol{a}}(x;f)\} \stackrel{d}{\to} N(0, \lambda_{p,\boldsymbol{a}}V(x;f)),$$

hence, if $n\beta^{(4p+1)/2} \to 0$, then, $(n\beta^{1/2})^{1/2}\{\widehat{f}_{\beta,ADD^p_{\boldsymbol{a}}}(x) - f(x)\} \stackrel{d}{\to} N(0, \lambda_{p,\boldsymbol{a}}V(x;f))$.
(ii). If $n\beta^{2p+1+\eta_{2p}} \to 0$, then, for $x \in \mathcal{S}_B$,

$$(n\beta)^{1/2}\{\widehat{f}_{\beta,ADD^p_{\boldsymbol{a}}}(x) - f(x) - \beta^p B_{p,\boldsymbol{a}}(x;f)\} \stackrel{d}{\to} N(0, \upsilon_{p,\boldsymbol{a}}(0)f(x)),$$

hence, if $n\beta^{2p+1} \to 0$, then, $(n\beta)^{1/2}\{\widehat{f}_{\beta,ADD^p_{\boldsymbol{a}}}(x) - f(x)\} \stackrel{d}{\to} N(0, \upsilon_{p,\boldsymbol{a}}(0)f(x))$.

Theorem 5 shows that

$$MSE[\widehat{f}_{\beta,ADD^p_{\boldsymbol{a}}}(x)] = \begin{cases} AMSE[\widehat{f}_{\beta,ADD^p_{\boldsymbol{a}}}(x)] + o(\beta^{2p} + n^{-1}\beta^{-1/2}) & \text{for fixed } x \in \mathcal{S}_I, \\ AMSE[\widehat{f}_{\beta,ADD^p_{\boldsymbol{a}}}(x)] + o(\beta^{2p} + n^{-1}\beta^{-1}) & \text{for } x \in \mathcal{S}_{B,\beta,\kappa}, \end{cases}$$

where

$$AMSE[\widehat{f}_{\beta,ADD^p_{\boldsymbol{a}}}(x)] = \begin{cases} \beta^{2p}B^2_{p,\boldsymbol{a}}(x;f) + n^{-1}\beta^{-1/2}\lambda_{p,\boldsymbol{a}}V(x;f) & \text{for fixed } x \in \mathcal{S}_I, \\ \beta^{2p}B^2_{p,\boldsymbol{a}}(x;f) + n^{-1}\beta^{-1}\upsilon_{p,\boldsymbol{a}}(\kappa)f(x) & \text{for } x \in \mathcal{S}_{B,\beta,\kappa}. \end{cases}$$

Note that

$$\min_{\beta>0} AMSE[\widehat{f}_{\beta,ADD^p_{\boldsymbol{a}}}(x)] = \begin{cases} \dfrac{4p+1}{4p}\left[4pB^2_{p,\boldsymbol{a}}(x;f)\left\{\lambda_{p,\boldsymbol{a}}V(x;f)n^{-1}\right\}^{4p}\right]^{1/(4p+1)} & \text{for fixed } x \in \mathcal{S}_I, \\ \dfrac{2p+1}{2p}\left[2pB^2_{p,\boldsymbol{a}}(x;f)\left\{\upsilon_{p,\boldsymbol{a}}(\kappa)f(x)n^{-1}\right\}^{2p}\right]^{1/(2p+1)} & \text{for } x \in \mathcal{S}_{B,\beta,\kappa} \end{cases}$$

(we assume $\gamma_p(x;f) \neq 0$). Although the additive estimator (4) has the slower convergence rate near the boundary $\mathcal{S}_B$, such a different rate is asymptotically negligible on the MISE.

Theorem 7 *Suppose that Assumptions* A1–A3*,* A4[p]*, and* A5[p] *hold for some* $p \in \mathbb{N}\backslash\{1\}$*. Then,*

$$MISE[\widehat{f}_{\beta,ADD^p_{\boldsymbol{a}}}] = AMISE[\widehat{f}_{\beta,ADD^p_{\boldsymbol{a}}}] + o(\beta^{2p} + n^{-1}\beta^{-1/2}),$$

where

$$AMISE[\widehat{f}_{\beta,ADD^p_{\boldsymbol{a}}}] = \beta^{2p}\int_{\mathcal{S}} B^2_{p,\boldsymbol{a}}(x;f)dx + n^{-1}\beta^{-1/2}\lambda_{p,\boldsymbol{a}}\int_{\mathcal{S}} V(x;f)dx$$

is minimized at

$$\beta = \left[\frac{\lambda_{p,\boldsymbol{a}}\int_{\mathcal{S}} V(x;f)dx}{4p\int_{\mathcal{S}} B^2_{p,\boldsymbol{a}}(x;f)dx}n^{-1}\right]^{2/(4p+1)}$$

when $\gamma_p(x;f) \not\equiv 0$*, that is,*

$$\min_{\beta>0} AMISE[\widehat{f}_{\beta,ADD^p_{\boldsymbol{a}}}] = \frac{4p+1}{4p}\left[4p\int_{\mathcal{S}} B^2_{p,\boldsymbol{a}}(x;f)dx\left\{\lambda_{p,\boldsymbol{a}}\int_{\mathcal{S}} V(x;f)dx\, n^{-1}\right\}^{4p}\right]^{1/(4p+1)}.$$

Remark 3 The additive estimator (4) loses the nonnegativity. However, it is easily remedied by considering the positive part $\widehat{f}^+_{\beta,ADD^p_{\boldsymbol{a}}}(x) = \max\{\widehat{f}_{\beta,ADD^p_{\boldsymbol{a}}}(x), 0\}$. Not surprisingly, $\widehat{f}^+_{\beta,ADD^p_{\boldsymbol{a}}}(x)$ is superior to $\widehat{f}_{\beta,ADD^p_{\boldsymbol{a}}}(x)$ in the (non-asymptotic) sense that, for any $x \in \mathcal{S}$,

$$\begin{aligned} &MSE[\widehat{f}_{\beta,ADD^p_{\boldsymbol{a}}}(x)] - MSE[\widehat{f}^+_{\beta,ADD^p_{\boldsymbol{a}}}(x)] \\ &= E[\widehat{f}^2_{\beta,ADD^p_{\boldsymbol{a}}}(x)\chi_{\{\widehat{f}_{\beta,ADD^p_{\boldsymbol{a}}}(x)<0\}}] - 2f(x)E[\widehat{f}_{\beta,ADD^p_{\boldsymbol{a}}}(x)\chi_{\{\widehat{f}_{\beta,ADD^p_{\boldsymbol{a}}}(x)<0\}}] \geq 0, \end{aligned}$$

hence, $MISE[\widehat{f}^+_{\beta,ADD^p_{\boldsymbol{a}}}] \leq MISE[\widehat{f}_{\beta,ADD^p_{\boldsymbol{a}}}]$.

3.2. Asymptotic properties of TS-type and JF-type estimators

We turn to the TS-type and JF-type estimators (5) and (6).

When $\mathcal{S} = [0, \infty)$, for rigorous asymptotic analyses as in Igarashi and Kakizawa (2018a), we pre-determine, for some constant $\eta \in (0, 1]$,

$$(\iota, \iota_0) \in \{(0,0)\} \bigcup \left\{ (\iota, \iota_0) \,\middle|\, 0 < \iota < \frac{\eta/2}{p + \eta/2} \text{ and } 0 < \iota_0 < \frac{1 - (p+1)\iota}{p} \right\} = \widetilde{I}_{p,\eta} \quad \text{(say)}, \qquad (9)$$

and consider a set of the points x, as follows:

$$\mathcal{I}_{\iota,\iota_0}[r_\beta] = \{x \in [0, r_\beta] \mid f(x) \geq \varrho \beta^{\iota_0}\} \quad \text{with } r_\beta = O(\beta^{-\iota})$$

for some $r_\beta \equiv r$ (fixed) or $r_\beta \to \infty$ (diverging slowly to infinity), according to $(\iota, \iota_0) = (0, 0)$ or $(\iota, \iota_0) \in \widetilde{I}_{p,\eta} \backslash \{(0, 0)\}$. Here and subsequently, $\varrho, r > 0$ are some constants. Note that, if $(\iota, \iota_0) \in \widetilde{I}_{p,\eta_{2p}} (\subset \widetilde{I}_{p,1})$ is pre-determined ($\eta_{2p} \in (0, 1]$ is given in Assumption A5[p](ii)), then,

$$r_\beta = O(\beta^{-\iota}) \quad \text{implies} \quad \beta^{\eta_{2p}/2}(1 + r_\beta)^{p+\eta_{2p}/2} + \beta^{1-\iota_0 p}(1 + r_\beta)^{p+1} = o(1).$$

For a technical reason, we use the weighted MISE criterion when $\mathcal{S} = [0, \infty)$, i.e.,

$$MISE[\widehat{f}; w] = \int_0^\infty w(x) MSE[\widehat{f}(x)] dx,$$

where the weight function w is nonnegative, bounded, and continuous except for a finite number of discontinuities (we assume $w(0) > 0$). On the other hand, when $\mathcal{S} = [0, 1]$, unlike the case $\mathcal{S} = [0, \infty)$, no technical difficulty is encountered in approximating the (unweighted) MISE.

In what follows, let $\# = TS, JF$, unless otherwise stated, and let

$$c_{p,TS} = 0 \quad \text{and} \quad c_{p,JF} = \begin{cases} 0, & p = 2 \text{ and } 0 < a_2 < a_1 = 1, \\ 1, & p = 2 \text{ and } 1 = a_1 < a_2, \\ p - 1, & p (> 2) \text{ is even}, \\ p - 2, & p (> 2) \text{ is odd}. \end{cases}$$

We write

$$B_{\#^p_{\boldsymbol{a}}}(x; f) = \begin{cases} B_{p,\boldsymbol{a}}(x; f) + \dfrac{(-1)^{p-1}}{\prod_{k=1}^p a_k} \displaystyle\sum_{j=2}^p \dfrac{(-1)^{j-1}}{j f^{j-1}(x)} \sum_{\mathcal{L}_{p,j}} \prod_{m=1}^j \gamma_{\ell_m}(x; f), & \# = TS, \\ B_{p,\boldsymbol{a}}(x; f) + \dfrac{1}{p f^{p-1}(x)} \gamma_1^p(x; f), & \# = JF, \end{cases}$$

where

$$\mathcal{L}_{p,j} = \left\{ \ell_1, \ldots, \ell_j \in \mathbb{N} \,\middle|\, \sum_{m=1}^j \ell_m = p \right\}.$$

We impose additional assumptions on β, ϵ, f, and w:

A6[p]$_{\iota_1,\iota_2}$. $\beta \propto n^{-\iota_1}$ and $\epsilon \propto \beta^{\iota_2}$ (independent of a_k, $k = 1, \ldots, p$) for some constants ι_1 and ι_2.

A7$[p]^{\#}_{\iota_0,\iota_2}$. When $\mathcal{S} = [0,\infty)$, given $r_\beta \equiv r$ or $r_\beta \to \infty$, f satisfies (i) $\min_{x\in[0,r_\beta]} f(x) \geq \varrho\beta^{\iota_0}$, and w is a weight function, independent of β, such that (ii) $\int_{r_\beta}^{\infty} w(x)dx \propto \exp(-\beta^{-A})$ for some constant $A > c_{p,\#}(1+\iota_2)$, independent of β, and that (iii) $w(x)B^2_{\#^p_{\boldsymbol{a}}}(x;f)$ is integrable, where ι_0 and ι_2 are some constants (when $r_\beta \equiv r$, the requirement (ii) holds iff w is a truncated weight function, with $w(y) = 0$ for any $y > r$).

A7$'$. When $\mathcal{S} = [0,1]$, f satisfies $\min_{x\in[0,1]} f(x) > 0$.

Below, given $p \in \mathbb{N}\backslash\{1\}$ and $(\iota,\iota_0) \in \widetilde{I}_{p,\eta_{2p}}$ ($\eta_{2p} \in (0,1]$ is given in Assumption A5$[p]$(ii)), we technically take $(\iota_1,\iota_2) \in I_{p,(\iota,\iota_0),\#}$, where

$$I_{p,(\iota,\iota_0),\#} = \left\{(\iota_1,\iota_2) \,\middle|\, 0 < \iota_1 < \frac{1}{1+2\iota_0+c_{p,\#}(1+\iota_2)}, \quad \iota_2 > 1+(\iota+\iota_0)(p-1)\right\}.$$

Remark 4 Note that $\beta \propto n^{-1/(2p+1/d)}$ for $d = 1,2$ (i.e., $\iota_1 = 2/(4p+1)$ and $\iota_1 = 1/(2p+1)$) are feasible for $\# = TS$ (the same remains valid for $\# = JF$ when $p = 2$ and $0 < a_2 < a_1 = 1$). In fact, with $c_{p,\#} = 0$, $\iota_0 \in [0,1/2)$ (see (9)) implies $1/(2p+1/2) < 1/2 < 1/(1+2\iota_0)$. On the other hand, for the JF-type (with $c_{p,JF} > 0$), as long as $(\iota,\iota_0) \in \widetilde{I}_{p,\eta_{2p}}$ satisfies

$$\iota_0 < \frac{2p-1/2-2c_{p,JF}-c_{p,JF}(p-1)\iota}{c_{p,JF}(p-1)+2},$$

we see that

$$1+(\iota+\iota_0)(p-1) < \iota_2 < \frac{2p-1/2-c_{p,JF}-2\iota_0}{c_{p,JF}}$$

implies that $1/(2p+1/2) < 1/\{1+2\iota_0+c_{p,JF}(1+\iota_2)\}$. Hence, to ensure the feasibility of $\beta \propto n^{-1/(2p+1/d)}$ for $d = 1,2$ and $\# = TS, JF$, we take $(\iota,\iota_0) \in \widetilde{I}_{p,\eta_{2p},\#}(\subset \widetilde{I}_{p,\eta_{2p}})$, where

$$\widetilde{I}_{p,\eta_{2p},\#} = \begin{cases} \widetilde{I}_{p,\eta_{2p}}, & c_{p,\#} = 0, \\ \widetilde{I}_{p,\eta_{2p}} \bigcap \left\{(\iota,\iota_0) \,\middle|\, \iota_0 < \dfrac{2p-1/2-2c_{p,\#}-c_{p,\#}(p-1)\iota}{c_{p,\#}(p-1)+2}\right\}, & c_{p,\#} > 0. \end{cases}$$

3.2.1. Case $\mathcal{S} = [0,\infty)$

We are ready to present the asymptotic properties of the TS-type and JF-type estimators (5) and (6) for the case $\mathcal{S} = [0,\infty)$.

Theorem 8 (i) *Given $p \in \mathbb{N}\backslash\{1\}$, suppose that Assumptions* A1(i), A4$[p]$, *and* A5$[p]$(i,ii) *hold. In addition, given $(\iota,\iota_0) \in \widetilde{I}_{p,\eta_{2p}}(\subset \widetilde{I}_{p,1})$, where $\eta_{2p} \in (0,1]$ is given in Assumption* A5$[p]$(ii), *suppose that Assumption* A6$[p]_{\iota_1,\iota_2}$ *holds for some constant $(\iota_1,\iota_2) \in I_{p,(\iota,\iota_0),\#}$, and define*

$$\omega_\beta(x) = \beta^{\eta_{2p}/2}(1+x)^{p+\eta_{2p}/2} + \beta^{1-\iota_0 p}(1+x)^{p+1} + \beta^{\iota_2-1-\iota_0(p-1)}(1+x)^{p-1}.$$

Then,

$$Bias[\widehat{f}_{\beta,\#^p_{\boldsymbol{a}}}(x)] = \beta^p B_{\#^p_{\boldsymbol{a}}}(x;f) + \mathcal{E}_{\beta,\#^p_{\boldsymbol{a}}}(x) \quad \textit{for } x \in \mathcal{I}_{\iota,\iota_0}[r_\beta],$$

where

$$\mathcal{E}_{\beta,\#^p_{\boldsymbol{a}}}(x) = O\bigg(\beta^p \omega_\beta(x) + \beta^{-\iota_0} \sum_{k=1}^{p} V[\widehat{f}_{\beta/a_k}(x)]\bigg).$$

(ii) *Given* $p \in \mathbb{N}\backslash\{1\}$ *and* $(\iota, \iota_0) \in \widetilde{I}_{p,1}$, *suppose that Assumptions* A1, A2, A4′[1], A5′(i), *and* A6$[p]_{\iota_1,\iota_2}$ *hold for some constant* $(\iota_1, \iota_2) \in I_{p,(\iota,\iota_0),\#}$. *Then,*

$$V[\widehat{f}_{\beta,\#^p_{\boldsymbol{a}}}(x)] = V[\widehat{f}_{\beta,ADD^p_{\boldsymbol{a}}}(x)] + \widetilde{\mathcal{E}}_{\beta,\#^p_{\boldsymbol{a}}}(x) \quad \textit{for } x \in \mathcal{I}_{\iota,\iota_0}[r_\beta],$$

where

$$\widetilde{\mathcal{E}}_{\beta,\#^p_{\boldsymbol{a}}}(x) = O\bigg(\beta^{2p+1-\iota_0 p}(1+x)^{p+1} + \{\beta^{1-\iota_0 p}(1+x)^{p+1} + n^{-1/2}\beta^{-(1/2+\iota_0)}\} \sum_{k=1}^{p} V[\widehat{f}_{\beta/a_k}(x)]\bigg).$$

Theorem 9 *Given* $p \in \mathbb{N}\backslash\{1\}$, *suppose that Assumptions* A1, A2, A4′[1], A5′(i), *and* A6$[p]_{\iota_1,\iota_2}$ *hold for some constant* $(\iota_1, \iota_2) \in I_{p,2,\#}$ *for* $x \in \mathcal{S}_I$ *or* $(\iota_1, \iota_2) \in I_{p,1,\#}$ *for* $x \in \mathcal{S}_B$ *(note that* $I_{p,2,\#} \subset I_{p,1,\#} \subset I_{p,(0,0),\#}$*), where*

$$I_{p,d,\#} = \left\{(\iota_1, \iota_2) \;\middle|\; \frac{1}{2p+2+1/d} < \iota_1 < \frac{1}{1+c_{p,\#}(1+\iota_2)}, \quad 1 < \iota_2 < \frac{2p+1+1/d-c_{p,\#}}{c_{p,\#}}\right\}$$

(exceptionally, when $c_{p,\#} = 0$*, the feasible range of* ι_2 *should read as "*$\iota_2 > 1$*"). Then,*

$$\begin{aligned}
(n\beta^{1/2})^{1/2}\{\widehat{f}_{\beta,\#^p_{\boldsymbol{a}}}(x) - E[\widehat{f}_{\beta,\#^p_{\boldsymbol{a}}}(x)]\} &\xrightarrow{d} N(0, \lambda_{p,\boldsymbol{a}} V(x;f)) \quad \textit{for fixed } x \in \mathcal{I}_{0,0}[r] \cap \mathcal{S}_I,\\
(n\beta)^{1/2}\{\widehat{f}_{\beta,\#^p_{\boldsymbol{a}}}(x) - E[\widehat{f}_{\beta,\#^p_{\boldsymbol{a}}}(x)]\} &\xrightarrow{d} N(0, v_{p,\boldsymbol{a}}(0) f(x)) \quad \textit{for } x \in \mathcal{I}_{0,0}[r] \cap \mathcal{S}_B.
\end{aligned}$$

A replacement of $E[\widehat{f}_{\beta,\#^p_{\boldsymbol{a}}}(x)]$ by $f(x)$ (or $f(x) + \beta^p B_{\#^p_{\boldsymbol{a}}}(x;f)$) is a routine problem in density estimation theory (use Slutsky's lemma; see Theorems 8(i) and 9)[3].

Theorem 8, together with Theorem 5(ii), shows that

$$MSE[\widehat{f}_{\beta,\#^p_{\boldsymbol{a}}}(x)] = \begin{cases} AMSE[\widehat{f}_{\beta,\#^p_{\boldsymbol{a}}}(x)] + o(\beta^{2p} + n^{-1}\beta^{-1/2}) & \text{for fixed } x \in \mathcal{I}_{0,0}[r] \cap \mathcal{S}_I, \\ AMSE[\widehat{f}_{\beta,\#^p_{\boldsymbol{a}}}(x)] + o(\beta^{2p} + n^{-1}\beta^{-1}) & \text{for } x \in \mathcal{I}_{0,0}[r] \cap \mathcal{S}_{B,\beta,\kappa}, \end{cases}$$

where

$$AMSE[\widehat{f}_{\beta,\#^p_{\boldsymbol{a}}}(x)] = \begin{cases} \beta^{2p} B^2_{\#^p_{\boldsymbol{a}}}(x;f) + n^{-1}\beta^{-1/2}\lambda_{p,\boldsymbol{a}} V(x;f) & \text{for fixed } x \in \mathcal{I}_{0,0}[r] \cap \mathcal{S}_I, \\ \beta^{2p} B^2_{\#^p_{\boldsymbol{a}}}(x;f) + n^{-1}\beta^{-1} v_{p,\boldsymbol{a}}(\kappa) f(x) & \text{for } x \in \mathcal{I}_{0,0}[r] \cap \mathcal{S}_{B,\beta,\kappa}. \end{cases}$$

[3]Given $p \in \mathbb{N}\backslash\{1\}$, suppose that Assumptions A1, A2, A4[p], A5[p](i,ii), and A6$[p]_{\iota_1,\iota_2}$ hold for some constant $(\iota_1, \iota_2) \in I_{p,2,\#}$ for $x \in \mathcal{S}_I$ or $(\iota_1, \iota_2) \in I_{p,1,\#}$ for $x \in \mathcal{S}_B$.
(i). If, in addition, $2/[4p+1+2\min\{\eta_{2p}, 2(\iota_2 - 1)\}] < \iota_1$, then, for fixed $x \in \mathcal{I}_{0,0}[r] \cap \mathcal{S}_I$,

$$(n\beta^{1/2})^{1/2}\{\widehat{f}_{\beta,\#^p_{\boldsymbol{a}}}(x) - f(x) - \beta^p B_{\#^p_{\boldsymbol{a}}}(x;f)\} \xrightarrow{d} N(0, \lambda_{p,\boldsymbol{a}} V(x;f)),$$

hence, if, in addition, $2/(4p+1) < \iota_1$, then, $(n\beta^{1/2})^{1/2}\{\widehat{f}_{\beta,\#^p_{\boldsymbol{a}}}(x) - f(x)\} \xrightarrow{d} N(0, \lambda_{p,\boldsymbol{a}} V(x;f))$.
(ii). If, in addition, $1/[2p+1+\min\{\eta_{2p}, 2(\iota_2 - 1)\}] < \iota_1$, then, for $x \in \mathcal{I}_{0,0}[r] \cap \mathcal{S}_B$,

$$(n\beta)^{1/2}\{\widehat{f}_{\beta,\#^p_{\boldsymbol{a}}}(x) - f(x) - \beta^p B_{\#^p_{\boldsymbol{a}}}(x;f)\} \xrightarrow{d} N(0, v_{p,\boldsymbol{a}}(0) f(x)),$$

hence, if, in addition, $1/(2p+1) < \iota_1$, then, $(n\beta)^{1/2}\{\widehat{f}_{\beta,\#^p_{\boldsymbol{a}}}(x) - f(x)\} \xrightarrow{d} N(0, v_{p,\boldsymbol{a}}(0) f(x))$.

Note that

$$\min_{\beta>0} AMSE[\widehat{f}_{\beta,\#_{\boldsymbol{a}}^p}(x)]$$
$$= \begin{cases} \dfrac{4p+1}{4p}\left[4pB^2_{\#_{\boldsymbol{a}}^p}(x;f)\{\lambda_{p,\boldsymbol{a}}V(x;f)n^{-1}\}^{4p}\right]^{1/(4p+1)} & \text{for fixed } x \in \mathcal{I}_{0,0}[r]\cap\mathcal{S}_I, \\ \dfrac{2p+1}{2p}\left[2pB^2_{\#_{\boldsymbol{a}}^p}(x;f)\{\upsilon_{p,\boldsymbol{a}}(\kappa)f(x)n^{-1}\}^{2p}\right]^{1/(2p+1)} & \text{for } x \in \mathcal{I}_{0,0}[r]\cap\mathcal{S}_{B,\beta,\kappa} \end{cases}$$

(see Remark 4, with $(\iota,\iota_0)=(0,0)$), provided that $B_{\#_{\boldsymbol{a}}^p}(x;f)\neq 0$. Although the TS/JF-type estimators (5) and (6) have the slower convergence rate near the boundary $\mathcal{S}_B$, such a different rate is asymptotically negligible on the weighted MISE.

Theorem 10 *Given $p\in\mathbb{N}\backslash\{1\}$, suppose that Assumptions* A1–A3, A4[p], *and* A5[p] *hold. In addition, given $(\iota,\iota_0)\in\widetilde{I}_{p,\eta_{2p}}$ ($\eta_{2p}\in(0,1]$ is given in Assumption* A5[p](ii)*), suppose that Assumptions* A6$[p]_{\iota_1,\iota_2}$ *and* A7$[p]^{\#}_{\iota_0,\iota_2}$ *hold for some constant $(\iota_1,\iota_2)\in I_{p,(\iota,\iota_0),\#}$. Then,*

$$MISE[\widehat{f}_{\beta,\#_{\boldsymbol{a}}^p};w] = AMISE[\widehat{f}_{\beta,\#_{\boldsymbol{a}}^p};w] + o(\beta^{2p}+n^{-1}\beta^{-1/2}),$$

where

$$AMISE[\widehat{f}_{\beta,\#_{\boldsymbol{a}}^p};w] = \beta^{2p}\int_0^\infty w(x)B^2_{\#_{\boldsymbol{a}}^p}(x;f)dx + n^{-1}\beta^{-1/2}\lambda_{p,\boldsymbol{a}}\int_0^\infty w(x)V(x;f)dx$$

is minimized at

$$\beta = \left[\frac{\lambda_{p,\boldsymbol{a}}\int_0^\infty w(x)V(x;f)dx}{4p\int_0^\infty w(x)B^2_{\#_{\boldsymbol{a}}^p}(x;f)dx}n^{-1}\right]^{2/(4p+1)}$$

(it is feasible for $(\iota,\iota_0)\in\widetilde{I}_{p,\eta_{2p},\#}(\subset\widetilde{I}_{p,\eta_{2p}})$; see Remark 4*) when $\sqrt{w(x)}B_{\#_{\boldsymbol{a}}^p}(x;f)\not\equiv 0$, that is,*

$$\min_{\beta>0} AMISE[\widehat{f}_{\beta,\#_{\boldsymbol{a}}^p};w]$$
$$= \frac{4p+1}{4p}\left[4p\int_0^\infty w(x)B^2_{\#_{\boldsymbol{a}}^p}(x;f)dx\left\{\lambda_{p,\boldsymbol{a}}\int_0^\infty w(x)V(x;f)dx\,n^{-1}\right\}^{4p}\right]^{1/(4p+1)}.$$

Remark 5 If possible, it will be better for us not to use the weighted MISE criterion. However, at present, we do not yet realize whether or not the valid asymptotic expansion

$$MISE[\widehat{f}_{\beta,\#_{\boldsymbol{a}}^p}] = \beta^{2p}\int_0^\infty B^2_{\#_{\boldsymbol{a}}^p}(x;f)dx + n^{-1}\beta^{-1/2}\lambda_{p,\boldsymbol{a}}\int_0^\infty V(x;f)dx + o(\beta^{2p}+n^{-1}\beta^{-1/2})$$

can be obtained for the case $w(x)\equiv 1$.

Here are some examples of (w,f) that we can apply Theorem 10.

(a) For a truncated weight function w, with $w(y)=0$ for any $y>r$, Theorem 10 is applicable, whenever $\min_{x\in[0,r]} f(x)>0$ (choose $(\iota,\iota_0)=(0,0)$ and $r_\beta\equiv r$).

(b) Let $w^\dagger(x)\propto x^{c_0-1}\exp\{x^{c_0}-\exp(x^{c_0})\}$ (say) for some constant $c_0>1$. Suppose that $w(x)\leq w^\dagger(x)$ for all sufficiently large x, and that there exists a constant $c_1>0$ such that

$\min_{x\geq 0}\{f(x)\exp(c_1 x)\} > 0$ (in this case, $w(x)B^2_{\#^p_{\boldsymbol{a}}}(x;f)$ is integrable). Then, given $p \in \mathbb{N}\backslash\{1\}$ and the pair $(\iota, \iota_0) \in \widetilde{I}_{p,\eta_{2p}}\backslash\{(0,0)\}$ ($\eta_{2p} \in (0,1]$ is given in Assumption A5[p](ii)), we choose $r_\beta = (\iota_0/c_1)\log(1/\beta)$ to verify that

- $\min_{x\in[0,r_\beta]} f(x) \geq \varrho\beta^{\iota_0}$, where $\varrho = \min_{x\geq 0}\{f(x)\exp(c_1 x)\}$,
- $\int_{r_\beta}^{\infty} x^{c_0-1}\exp\{x^{c_0} - \exp(x^{c_0})\}dx = \exp(-\beta^{-(\iota_0/c_1)^{c_0}\{\log(1/\beta)\}^{c_0-1}})$, where, for any constant $A > 0$ and all sufficiently small $\beta > 0$, $(\iota_0/c_1)^{c_0}\{\log(1/\beta)\}^{c_0-1} > A$.

(c) Let $w^\dagger(x) \propto e^{-x}$ or $w^\dagger(x) \propto \exp\{x - \exp(x)\}$ (say) according to $\# = TS$ or $\# = JF$. Suppose that $w(x) \leq w^\dagger(x)$ for all sufficiently large x, and that there exists a constant $c_1 > 1$ such that $\min_{x\geq 0}\{f(x)(1+x)^{c_1}\} > 0$ (in this case, $w(x)B^2_{\#^p_{\boldsymbol{a}}}(x;f)$ is integrable). We choose $r_\beta = \beta^{-\iota_0/c_1} - 1$ $(= O(\beta^{-\iota}))$, where, given $p \in \mathbb{N}\backslash\{1\}$, the pair $(\iota, \iota_0) \in \widetilde{I}_{p,\eta_{2p}}\backslash\{(0,0)\}$ ($\eta_{2p} \in (0,1]$ is given in Assumption A5[p](ii)) is pre-determined according to the inequalities $0 < \iota < \eta_{2p}/(2p + \eta_{2p})$, $0 < \iota_0 < \{1 - (p+1)\iota\}/p$, and $\iota_0 \leq c_1\iota$; more precisely,

- if $\eta_{2p} \in (0, 2/(1+c_1)]$, then, $(\iota, \iota_0) \in \widetilde{I}^{[1]}_{p,\eta_{2p}} \subset \widetilde{I}_{p,\eta_{2p}}\backslash\{(0,0)\}$, where

$$\widetilde{I}^{[1]}_{p,\eta_{2p}} = \left\{(\iota, \iota_0)\;\middle|\; 0 < \iota < \frac{\eta_{2p}/2}{p + \eta_{2p}/2} \text{ and } 0 < \iota_0 \leq c_1\iota\right\},$$

- if $\eta_{2p} \in (2/(1+c_1), 1]$, then, $(\iota, \iota_0) \in \widetilde{I}^{[2]}_{p} \cup \widetilde{I}^{[3]}_{p,\eta_{2p}} \subset \widetilde{I}_{p,\eta_{2p}}\backslash\{(0,0)\}$, where

$$\widetilde{I}^{[2]}_{p} = \left\{(\iota, \iota_0)\;\middle|\; 0 < \iota < \frac{1}{(1+c_1)p + 1} \text{ and } 0 < \iota_0 \leq c_1\iota\right\},$$
$$\widetilde{I}^{[3]}_{p,\eta_{2p}} = \left\{(\iota, \iota_0)\;\middle|\; \frac{1}{(1+c_1)p+1} \leq \iota < \frac{\eta_{2p}/2}{p + \eta_{2p}/2} \text{ and } 0 < \iota_0 < \frac{1-(p+1)\iota}{p}\right\}.$$

Then, we can verify that

- $\min_{x\in[0,r_\beta]} f(x) \geq \varrho\beta^{\iota_0}$, where $\varrho = \min_{x\geq 0}\{f(x)(1+x)^{c_1}\}$,
- $\int_{r_\beta}^{\infty} e^{-x}dx = \exp(-\beta^{-\iota_0/c_1} + 1)$,
- $\int_{r_\beta}^{\infty} \exp\{x - \exp(x)\}dx = \exp\{-\exp(\beta^{-\iota_0/c_1} - 1)\}$, where, for any constant $A > 0$ and all sufficiently small $\beta > 0$, $\beta^{-\iota_0/c_1} - 1 + A\log\beta > 0$.

3.2.2. Case $\mathcal{S} = [0,1]$

In line with Igarashi (2016a), let $\mathcal{I} = \{x \in [0,1] \mid f(x) \geq \varrho\}$ (note $\mathcal{I} = \mathcal{I}_{0,0}[1]$).

The following results for the case $\mathcal{S} = [0,1]$ are counterparts of Theorems 8–10 (here, we can handle the (unweighted) MISE without any difficulty).

Theorem 8′ (i) *Given* $p \in \mathbb{N}\backslash\{1\}$, *suppose that Assumptions* A1(i), A4[p], A5[p](i,ii), *and* A6$[p]_{\iota_1,\iota_2}$ *hold for some constant* $(\iota_1, \iota_2) \in I_{p,(0,0),\#}$. *Then,*

$$Bias[\widehat{f}_{\beta,\#^p_{\boldsymbol{a}}}(x)] = \beta^p B_{\#^p_{\boldsymbol{a}}}(x; f) + \mathcal{E}_{\beta,\#^p_{\boldsymbol{a}}}(x) \quad \text{for } x \in \mathcal{I},$$

where

$$\mathcal{E}_{\beta,\#^p_{\boldsymbol{a}}}(x) = O\bigg(\beta^{p+\min(\eta_{2p}/2,\iota_2-1)} + \sum_{k=1}^{p} V[\widehat{f}_{\beta/a_k}(x)]\bigg).$$

(ii) *Given* $p \in \mathbb{N}\backslash\{1\}$, *suppose that Assumptions* A1, A2, A4′[1], A5′(i), *and* A6$[p]_{\iota_1,\iota_2}$ *hold for some constant* $(\iota_1, \iota_2) \in I_{p,(0,0),\#}$. *Then,*

$$V[\widehat{f}_{\beta,\#^p_{\boldsymbol{a}}}(x)] = V[\widehat{f}_{\beta,ADD^p_{\boldsymbol{a}}}(x)] + \widetilde{\mathcal{E}}_{\beta,\#^p_{\boldsymbol{a}}}(x) \quad \text{for } x \in \mathcal{I},$$

where

$$\widetilde{\mathcal{E}}_{\beta,\#^p_{\boldsymbol{a}}}(x) = O\bigg(\beta^{2p+1} + (\beta + n^{-1/2}\beta^{-1/2})\sum_{k=1}^{p} V[\widehat{f}_{\beta/a_k}(x)]\bigg).$$

Theorem 9′ *Given* $p \in \mathbb{N}\backslash\{1\}$, *suppose that Assumptions* A1, A2, A4′[1], A5′(i), *and* A6$[p]_{\iota_1,\iota_2}$ *hold for some constant* $(\iota_1, \iota_2) \in I_{p,2,\#}$ *for* $x \in \mathcal{S}_I$ *or* $(\iota_1, \iota_2) \in I_{p,1,\#}$ *for* $x \in \mathcal{S}_B$ *(note that* $I_{p,2,\#} \subset I_{p,1,\#} \subset I_{p,(0,0),\#}$*). Then,*

$$(n\beta^{1/2})^{1/2}\{\widehat{f}_{\beta,\#^p_{\boldsymbol{a}}}(x) - E[\widehat{f}_{\beta,\#^p_{\boldsymbol{a}}}(x)]\} \xrightarrow{d} N(0, \lambda_{p,\boldsymbol{a}} V(x; f)) \quad \text{for fixed } x \in \mathcal{I} \cap \mathcal{S}_I,$$
$$(n\beta)^{1/2}\{\widehat{f}_{\beta,\#^p_{\boldsymbol{a}}}(x) - E[\widehat{f}_{\beta,\#^p_{\boldsymbol{a}}}(x)]\} \xrightarrow{d} N(0, v_{p,\boldsymbol{a}}(0) f(x)) \quad \text{for } x \in \mathcal{I} \cap \mathcal{S}_B.$$

A replacement of $E[\widehat{f}_{\beta,\#^p_{\boldsymbol{a}}}(x)]$ by $f(x)$ (or $f(x) + \beta^p B_{\#^p_{\boldsymbol{a}}}(x; f)$) is a routine problem in density estimation theory (use Slutsky's lemma; see Theorems 8′(i) and 9′)[4].

Theorem 8′, together with Theorem 5(ii), shows that

$$MSE[\widehat{f}_{\beta,\#^p_{\boldsymbol{a}}}(x)] = \begin{cases} AMSE[\widehat{f}_{\beta,\#^p_{\boldsymbol{a}}}(x)] + o(\beta^{2p} + n^{-1}\beta^{-1/2}) & \text{for fixed } x \in \mathcal{I} \cap \mathcal{S}_I, \\ AMSE[\widehat{f}_{\beta,\#^p_{\boldsymbol{a}}}(x)] + o(\beta^{2p} + n^{-1}\beta^{-1}) & \text{for } x \in \mathcal{I} \cap \mathcal{S}_{B,\beta,\kappa}, \end{cases}$$

where

$$AMSE[\widehat{f}_{\beta,\#^p_{\boldsymbol{a}}}(x)] = \begin{cases} \beta^{2p} B^2_{\#^p_{\boldsymbol{a}}}(x; f) + n^{-1}\beta^{-1/2}\lambda_{p,\boldsymbol{a}} V(x; f) & \text{for fixed } x \in \mathcal{I} \cap \mathcal{S}_I, \\ \beta^{2p} B^2_{\#^p_{\boldsymbol{a}}}(x; f) + n^{-1}\beta^{-1} v_{p,\boldsymbol{a}}(\kappa) f(x) & \text{for } x \in \mathcal{I} \cap \mathcal{S}_{B,\beta,\kappa}. \end{cases}$$

[4]Given $p \in \mathbb{N}\backslash\{1\}$, suppose that Assumptions A1, A2, A4[p], A5[p](i,ii), and A6$[p]_{\iota_1,\iota_2}$ hold for some constant $(\iota_1, \iota_2) \in I_{p,2,\#}$ for $x \in \mathcal{S}_I$ or $(\iota_1, \iota_2) \in I_{p,1,\#}$ for $x \in \mathcal{S}_B$.
(i). If, in addition, $2/[4p + 1 + 2\min\{\eta_{2p}, 2(\iota_2 - 1)\}] < \iota_1$, then, for fixed $x \in \mathcal{I} \cap \mathcal{S}_I$,

$$(n\beta^{1/2})^{1/2}\{\widehat{f}_{\beta,\#^p_{\boldsymbol{a}}}(x) - f(x) - \beta^p B_{\#^p_{\boldsymbol{a}}}(x; f)\} \xrightarrow{d} N(0, \lambda_{p,\boldsymbol{a}} V(x; f)),$$

hence, if, in addition, $2/(4p + 1) < \iota_1$, then, $(n\beta^{1/2})^{1/2}\{\widehat{f}_{\beta,\#^p_{\boldsymbol{a}}}(x) - f(x)\} \xrightarrow{d} N(0, \lambda_{p,\boldsymbol{a}} V(x; f))$.
(ii). If, in addition, $1/[2p + 1 + \min\{\eta_{2p}, 2(\iota_2 - 1)\}] < \iota_1$, then, for $x \in \mathcal{I} \cap \mathcal{S}_B$,

$$(n\beta)^{1/2}\{\widehat{f}_{\beta,\#^p_{\boldsymbol{a}}}(x) - f(x) - \beta^p B_{\#^p_{\boldsymbol{a}}}(x; f)\} \xrightarrow{d} N(0, v_{p,\boldsymbol{a}}(0) f(x)),$$

hence, if, in addition, $1/(2p + 1) < \iota_1$, then, $(n\beta)^{1/2}\{\widehat{f}_{\beta,\#^p_{\boldsymbol{a}}}(x) - f(x)\} \xrightarrow{d} N(0, v_{p,\boldsymbol{a}}(0) f(x))$.

Note that

$$\min_{\beta>0} AMSE[\widehat{f}_{\beta,\#_{\boldsymbol{a}}^p}(x)] = \begin{cases} \dfrac{4p+1}{4p}\left[4pB_{\#_{\boldsymbol{a}}^p}^2(x;f)\{\lambda_{p,\boldsymbol{a}}V(x;f)n^{-1}\}^{4p}\right]^{1/(4p+1)} & \text{for fixed } x \in \mathcal{I}\cap\mathcal{S}_I, \\ \dfrac{2p+1}{2p}\left[2pB_{\#_{\boldsymbol{a}}^p}^2(x;f)\{\upsilon_{p,\boldsymbol{a}}(\kappa)f(x)n^{-1}\}^{2p}\right]^{1/(2p+1)} & \text{for } x \in \mathcal{I}\cap\mathcal{S}_{B,\beta,\kappa} \end{cases}$$

(see Remark 4, with $(\iota,\iota_0) = (0,0)$), provided that $B_{\#_{\boldsymbol{a}}^p}(x;f) \neq 0$. Although the TS/JF-type estimators (5) and (6) have the slower convergence rate near the boundary $\mathcal{S}_B$, such a different rate is asymptotically negligible on the MISE.

Theorem 10′ *Given* $p \in \mathbb{N}\backslash\{1\}$, *suppose that Assumptions* A1, A2, A4[p], A5[p](i,ii), A6[p]$_{\iota_1,\iota_2}$, *and* A7′ *hold for some constant* $(\iota_1,\iota_2) \in I_{p,(0,0),\#}$. *Then,*

$$MISE[\widehat{f}_{\beta,\#_{\boldsymbol{a}}^p}] = AMISE[\widehat{f}_{\beta,\#_{\boldsymbol{a}}^p}] + o(\beta^{2p} + n^{-1}\beta^{-1/2}),$$

where

$$AMISE[\widehat{f}_{\beta,\#_{\boldsymbol{a}}^p}] = \beta^{2p}\int_0^1 B_{\#_{\boldsymbol{a}}^p}^2(x;f)dx + n^{-1}\beta^{-1/2}\lambda_{p,\boldsymbol{a}}\int_0^1 V(x;f)dx$$

is minimized at

$$\beta = \left[\frac{\lambda_{p,\boldsymbol{a}}\int_0^1 V(x;f)dx}{4p\int_0^1 B_{\#_{\boldsymbol{a}}^p}^2(x;f)dx}n^{-1}\right]^{2/(4p+1)}$$

(*it is feasible; see Remark* 4, *with* $(\iota,\iota_0) = (0,0)$), *when* $B_{\#_{\boldsymbol{a}}^p}(x;f) \not\equiv 0$, *that is,*

$$\min_{\beta>0} AMISE[\widehat{f}_{\beta,\#_{\boldsymbol{a}}^p}] = \frac{4p+1}{4p}\left[4p\int_0^1 B_{\#_{\boldsymbol{a}}^p}^2(x;f)dx\left\{\lambda_{p,\boldsymbol{a}}\int_0^1 V(x;f)dx\, n^{-1}\right\}^{4p}\right]^{1/(4p+1)}.$$

3.3. When Assumption A4[p] fails

A sufficient condition for Assumption A4[p] when $\mathcal{S} = [0,\infty)$ is that

$$\text{M.}\quad \text{for } j \in \mathbb{N},\ \mu_j(K(\cdot;x,\beta)) = \sum_{k=\lceil j/2\rceil}^{j} \zeta_{j,k}\, x^{j-k}\beta^k, \quad x \geq 0,$$

where $\zeta_{j,k}$'s are some constants, independent of β and x; see Examples 1 and 2 in Section 4. However, that is not always true and careful considerations are required, on a case-by-case basis.

Though there is a slight difference, the moments may be rational functions:

$$\text{M1}[p^*].\quad \text{for } j \in \mathbb{N},\ \mu_j(K(\cdot;x,\beta)) = \sum_{k=\lceil j/2\rceil}^{j} \zeta_{j,k}\, x^{j-k}\beta^k + \chi_{\{j\geq p^*\}}\beta^j Q_j\Big(\frac{x}{\beta}+1\Big), \quad x \geq 0,$$

where $Q_j(\cdot)$ is the rational function (independent of β and x) with $\rho|Q_j(\rho)| \leq \overline{Q}_j < \infty$, having the form of

$$Q_j(\rho) = \sum_{\ell_1=1}^{m_{1,j}}\sum_{\ell_2=1}^{m_{2,j}} \frac{\zeta_{j,\ell_1,\ell_2}}{(\rho+d_{\ell_2})^{\ell_1}} \text{ (type I) or } Q_j(\rho) = \frac{(\ell^*-1)\text{th polynomial in } \rho}{\ell^*\text{th polynomial in } \rho} \text{ (type II)}$$

for some natural numbers $p^*, m_{1,j}, m_{2,j}, \ell^*$ and real numbers $\zeta_{j,k}$'s, ζ_{j,ℓ_1,ℓ_2}'s, d_{ℓ_2}'s.

More generally, the moments are not be expressed in terms of elementary functions, i.e.,

$$\text{for } j \in \mathbb{N},\ \mu_j(K(\cdot; x, \beta)) = (-x)^j + \sum_{\ell=1}^{j} {}_jC_\ell (-x)^{j-\ell}(x + c\beta)^\ell g_\ell\Big(\frac{x}{\beta} + c\Big), \quad x \geq 0,$$

for some constant $c \geq 1$ and a set of continuous functions $\{g_\ell\}$ on $(0, \infty)$, with $g_\ell(\rho) \to 1$ as $\rho \to \infty$. Here, we focus on the particular cases that are appeared in Section 4 (Examples 3–5):

$$\text{(M2.1).} \quad g_\ell(\rho) = \Big(1 + \frac{1}{\rho}\Big)^{\ell(\nu+\ell/2)}, \qquad \text{(M2.2).} \quad g_\ell(\rho) = \frac{K_{\nu+\ell}(\rho)}{K_\nu(\rho)},$$

where $\nu \notin \{1/2 + M \mid M \in \mathbb{Z}\}$ (K_ν stands for the modified Bessel function of the third kind), or

$$\text{(M2.3).} \quad g_\ell(\rho) = \begin{cases} \Gamma^{\ell-1}(\rho/\gamma)\dfrac{\Gamma((\rho+\ell)/\gamma)}{\Gamma^\ell((\rho+1)/\gamma)}, & \gamma \in (0,\infty)\backslash\{1/M \mid M \in \mathbb{N}\}, \\ \Gamma^{\ell-1}((\rho+1)/|\gamma|)\dfrac{\Gamma((\rho+1-\ell)/|\gamma|)}{\Gamma^\ell(\rho/|\gamma|)}, & \gamma \in (-\infty, 0)\backslash\{-1/M \mid M \in \mathbb{N}\}. \end{cases}$$

It is worth noting that $g_1(\rho) \not\equiv 1$ for either of (M2.1) or (M2.2), while $g_1(\rho) \equiv 1$ for (M2.3). Thus, we are now interested in the following structure:

$$\text{for } j \in \mathbb{N},\ \mu_j(K(\cdot; x, \beta)) = c^j\beta^j + \sum_{\ell=p^*}^{j} {}_jC_\ell (-x)^{j-\ell}(x + c\beta)^\ell \Big\{g_\ell\Big(\frac{x}{\beta} + c\Big) - 1\Big\}, \quad x \geq 0$$

(we set $p^* = 1$ for either of (M2.1) or (M2.2) and set $p^* = 2$ for (M2.3)). Note that, for $j \in \mathbb{N}$,

$$\mu_j(K(\cdot; x, \beta)) = \beta^j\Big[c^j + \sum_{\ell=p^*}^{j} {}_jC_\ell(-\kappa)^{j-\ell}(\kappa + c)^\ell\{g_\ell(\kappa + c) - 1\} + \chi_{\{x=0\}}o(1)\Big] \quad \text{for } x \in \mathcal{S}_{B,\beta,\kappa}. \tag{10}$$

More importantly, using the large argument asymptotic expansion, we can verify that, for given $j \in \mathbb{N}$

$$\mu_j(K(\cdot; x, \beta)) = \sum_{k=\lceil j/2 \rceil}^{j} \zeta_{j,k}\, x^{j-k}\beta^k + \chi_{\{j \geq p^*\}}\beta^j R_j\Big(\frac{x}{\beta} + c\Big), \quad \frac{x}{\beta} + c \geq M_j, \tag{11}$$

for some constants $M_j (> c)$ and $\zeta_{j,k}$'s, independent of β and x, where a set of continuous functions on $(0, \infty)$; $\{R_j\}$ (independent of β and x) satisfies $\rho|R_j(\rho)| \leq \overline{R}_j < \infty$. Here, (10) and (11) are gathered as M2[p^*], which is an alternative to M1[p^*].

Now, we know (e.g., Igarashi and Kakizawa (2014b, 2018a) and Igarashi (2016b)) that the (bias-uncorrected) asymmetric KDEs even in either of M1[p^*] or M2[p^*] (at least, (M2.1)–(M2.3)) have the MISEs of order $O(n^{-4/5})$. Is it possible that using the bias correction methods reduces the convergence rate from $n^{-4/5}$, even if Assumption A4[p] is replaced by M1[p^*] or M2[p^*]? The answer is yes, but, the achievable rate when $f^{(p^*)}(0) \neq 0$ is shown to be $n^{-(4p^*+2)/(4p^*+3)}$, as follows:

Case M1[p^*]: M1[p^*] implies that Assumption A4[p] holds for any integer p; $1 \le p < p^*$; in this case, the results of the previous section are available. When $p \ge \max(2, p^*)$, we have

$$Bias[\widehat{f}_{\beta,ADD^p_{\boldsymbol{a}}}(x)] = \beta^p B_{p,\boldsymbol{a}}(x; f) + \sum_{j=p^*}^{p} \beta^j f^{(j)}(x)\widetilde{B}_{j,\boldsymbol{a}}(x,\beta) + \mathcal{E}_{\beta,ADD^p_{\boldsymbol{a}}}(x), \quad x \ge 0,$$

$$Bias[\widehat{f}_{\beta,\#^p_{\boldsymbol{a}}}(x)] = \beta^p B_{\#^p_{\boldsymbol{a}}}(x; f) + \sum_{j=p^*}^{p} \beta^j f^{(j)}(x)\widetilde{B}_{j,\boldsymbol{a}}(x,\beta) + \mathcal{E}^*_{\beta,\#^p_{\boldsymbol{a}}}(x) \quad \text{for } x \in \mathcal{I}_{\iota,\iota_0}[r_\beta],$$

where

$$\widetilde{B}_{j,\boldsymbol{a}}(x,\beta) = \frac{1}{j!}\sum_{k=1}^{p} \frac{c_k(\boldsymbol{a})}{a_k^j} Q_j\Big(\frac{a_k x}{\beta} + 1\Big), \quad p^* \le j \le p,$$

$$\mathcal{E}^*_{\beta,\#^p_{\boldsymbol{a}}}(x) = \begin{cases} \mathcal{E}_{\beta,TS^p_{\boldsymbol{a}}}(x) + \chi_{\{p>p^*\}} O(\beta^{p^*+1-\iota_0}(1+x)), & \# = TS, \\ \mathcal{E}_{\beta,JF^p_{\boldsymbol{a}}}(x) + \chi_{\{p^*=1\}} O(\beta^{2-\iota_0}(1+x)), & \# = JF. \end{cases}$$

Case M2[p^*] ($p^* = 1$ or $p^* = 2$): Let $p \in \mathbb{N}\backslash\{1\}$.

(i) When $p \ge \max(2, p^*)$, with $M \ge \max(M_1, \ldots, M_{2p}, M_{2(p+1)})$, we have

$$Bias[\widehat{f}_{\beta,ADD^p_{\boldsymbol{a}}}(x)] = \beta^p B_{p,\boldsymbol{a}}(x; f) + \sum_{j=p^*}^{p} \beta^j f^{(j)}(x)\check{B}_{j,\boldsymbol{a}}(x,\beta) + \mathcal{E}_{\beta,ADD^p_{\boldsymbol{a}}}(x), \quad \frac{x}{\beta} + c \ge M,$$

$$Bias[\widehat{f}_{\beta,\#^p_{\boldsymbol{a}}}(x)] = \beta^p B_{\#^p_{\boldsymbol{a}}}(x; f) + \sum_{j=p^*}^{p} \beta^j f^{(j)}(x)\check{B}_{j,\boldsymbol{a}}(x,\beta) + \mathcal{E}^*_{\beta,\#^p_{\boldsymbol{a}}}(x)$$

$$\text{for } x \in \mathcal{I}_{\iota,\iota_0}[r_\beta] \cap [(M-c)\beta, r_\beta],$$

where

$$\check{B}_{j,\boldsymbol{a}}(x,\beta) = \frac{1}{j!}\sum_{k=1}^{p} \frac{c_k(\boldsymbol{a})}{a_k^j} R_j\Big(\frac{a_k x}{\beta} + c\Big), \quad p^* \le j \le p.$$

(ii) When $p = p^* = 2$, we have

$$\begin{aligned} &Bias[\widehat{f}_{\beta,ADD^2_{(a_1,a_2)}}(x)] \\ &= \beta^2 \frac{f^{(2)}(0)}{2}\sum_{k=1}^{2} \frac{c_k(a_1,a_2)}{a_k^2}\big[c^2 + (a_k\kappa + c)^2\{g_2(a_k\kappa + c) - 1\}\big] + o(\beta^2) \quad \text{for } x \in \mathcal{S}_{B,\beta,\kappa}, \\ &Bias[\widehat{f}_{\beta,\#^2_{(a_1,a_2)}}(x)] \\ &= \beta^2\Big[\frac{f^{(2)}(0)}{2}\sum_{k=1}^{2} \frac{c_k(a_1,a_2)}{a_k^2}\big[c^2 + (a_k\kappa + c)^2\{g_2(a_k\kappa + c) - 1\}\big] + \frac{\{cf^{(1)}(0)\}^2}{2f(0)\big(\prod_{k=1}^2 a_k\big)^{\chi_{\{\#=TS\}}}}\Big] \\ &\quad + o(\beta^2) + O\Big(\beta^{-\iota_0}\sum_{k=1}^{2} V[\widehat{f}_{\beta/a_k}(x)]\Big) \quad \text{for } x \in \mathcal{I}_{\iota,\iota_0}[r_\beta] \cap \mathcal{S}_{B,\beta,\kappa}. \end{aligned}$$

On the other hand, when $p > p^*$, we have

$$Bias[\widehat{f}_{\beta,ADD_a^p}(x)] = \beta^{p^*}\frac{f^{(p^*)}(0)}{p^*!}\sum_{k=1}^{p}\frac{c_k(\boldsymbol{a})}{a_k^{p^*}}(a_k\kappa+c)^{p^*}\{g_{p^*}(a_k\kappa+c)-1\}+o(\beta^{p^*}) \quad \text{for } x\in\mathcal{S}_{B,\beta,\kappa},$$

$$Bias[\widehat{f}_{\beta,\#_a^p}(x)] = \beta^{p^*}\frac{f^{(p^*)}(0)}{p^*!}\sum_{k=1}^{p}\frac{c_k(\boldsymbol{a})}{a_k^{p^*}}(a_k\kappa+c)^{p^*}\{g_{p^*}(a_k\kappa+c)-1\}$$
$$+o(\beta^{p^*})+O\left(\beta^{-\iota_0}\sum_{k=1}^{p}V[\widehat{f}_{\beta/a_k}(x)]\right) \quad \text{for } x\in\mathcal{I}_{\iota,\iota_0}[r_\beta]\cap\mathcal{S}_{B,\beta,\kappa}.$$

Thus, unless $f^{(p^*)}(0)=0$, the biases of the additive, TS-type, and JF-type estimators when $x\in\mathcal{S}_{B,\beta,\kappa}$ ($\kappa>0$) are of order $O(\beta^{\min(p^*,p)})+\chi_{\{\#=TS,JF\}}O(\beta^{-\iota_0}\sum_{k=1}^{p}V[\widehat{f}_{\beta/a_k}(x)])$. Also, we can see that the resulting integrated squared biases, when $p=p^*(\geq 2)$, are of order $O(\beta^{2p})+\chi_{\{\#=TS,JF\}}o(n^{-1}\beta^{-1/2})$, and, when $p>p^*$, unless $f^{(p^*)}(0)=0$, they are of order $O(\beta^{\min(2p^*+1,2p)})+\chi_{\{\#=TS,JF\}}o(n^{-1}\beta^{-1/2}) = O(\beta^{2p^*+1})+\chi_{\{\#=TS,JF\}}o(n^{-1}\beta^{-1/2})$. Since the integrated variances are of order $O(n^{-1}\beta^{-1/2})$, the MISEs, when $p=p^*(\geq 2)$, achieves the order $O(n^{-4p/(4p+1)})$ by choosing $\beta\propto n^{-2/(4p+1)}$, and, when $p>p^*$, unless $f^{(p^*)}(0)=0$, they are of order $O(n^{-(4p^*+2)/(4p^*+3)})$ at most[5]. Therefore, attention should be paid to this phenomenon under M1[p^*] or M2[p^*].

4. Examples

Associated with the special functions $B(\theta_1,\theta_2)=\int_0^1 s^{\theta_1-1}(1-s)^{\theta_2-1}ds$, $\Gamma(\theta_1)=\int_0^\infty s^{\theta_1-1}e^{-s}ds$ (we have $B(\theta_1,\theta_2)=\Gamma(\theta_1)\Gamma(\theta_2)/\Gamma(\theta_1+\theta_2)$), and

$$K_\nu(\theta_1)=\int_0^\infty\frac{s^{\nu-1}}{2}\exp\left\{-\frac{\theta_1}{2}\left(s+\frac{1}{s}\right)\right\}ds \quad (\text{note that } K_{\pm 1/2}(\theta_1)=\{\pi/(2\theta_1)\}^{1/2}e^{-\theta_1}),$$

[5]Suppose that $p>p^*$.
The additive estimator has the MISE of order $O(n^{-(4p^*+2)/(4p^*+3)})$ by choosing $\beta\propto n^{-2/(4p^*+3)}$.
The TS/JF-type estimators have the MISEs of order $O(n^{-(4p^*+2)/(4p^*+3)})$ by choosing $\beta\propto n^{-2/(4p^*+3)}$ (if it is allowed). As in Remark 4, given $p\in\mathbb{N}\backslash\{1\}$ and $(\iota,\iota_0)\in\widetilde{I}_{p,\eta_{2p}}$, where $\eta_{2p}\in(0,1]$ is given in Assumption A5[p](ii), Assumption A6[p]$_{\iota_1,\iota_2}$ for some constant $(\iota_1,\iota_2)\in I_{p,(\iota,\iota_0),\#}$ implies that $\beta\propto n^{-1/(2p^*+3/2)}$ (i.e., $\iota_1=2/(4p^*+3)$) is feasible for $\#=TS$ (the same remains valid for $\#=JF$ when $p=2$ and $0<a_2<a_1=1$). In fact, with $c_{p,\#}=0$, $\iota_0\in[0,1/2)$ (see (9)) implies $1/(2p^*+3/2)<1/2<1/(1+2\iota_0)$. On the other hand, for the JF-type (with $c_{p,JF}>0$), consider the pair (p^*,p) with $p>p^*$, as follows: (i) if $p^*=1$, set $p=2$ (when $1=a_1<a_2$) or $p=3$, (ii) if p^* is even, set $p=p^*+1$, and (iii) if $p^*(>2)$ is odd, set $p=p^*+1,p^*+2$. Then, as long as $(\iota,\iota_0)\in\widetilde{I}_{p,\eta_{2p}}$ satisfies

$$\iota_0<\frac{2p^*+1/2-2c_{p,JF}-c_{p,JF}(p-1)\iota}{c_{p,JF}(p-1)+2},$$

we see that

$$1+(\iota+\iota_0)(p-1)<\iota_2<\frac{2p^*+1/2-c_{p,JF}-2\iota_0}{c_{p,JF}}$$

implies that $1/(2p^*+3/2)<1/\{1+2\iota_0+c_{p,JF}(1+\iota_2)\}$.

However, for other pair (p^*,p) with $p>p^*$ (i.e., $(2p^*+1/2-c_{p,JF}-2\iota_0)/c_{p,JF}-1<0$ for $(\iota,\iota_0)\in\widetilde{I}_{p,\eta_{2p}}$), the choice $\beta\propto n^{-1/(2p^*+3/2)}$ (i.e., $\iota_1=2/(4p^*+3)$) is not feasible, since, even if $1+(\iota+\iota_0)(p-1)<\iota_2$, it holds that, for $(\iota,\iota_0)\in\widetilde{I}_{p,\eta_{2p}}$,

$$\frac{2p^*+1/2-c_{p,JF}-2\iota_0}{c_{p,JF}}<1\leq 1+(\iota+\iota_0)(p-1)<\iota_2,$$

hence, $1/\{1+2\iota_0+c_{p,JF}(1+\iota_2)\}<1/(2p^*+3/2)$; in this case, the MISEs of the TS/JF-type estimators are worse than $O(n^{-(4p^*+2)/(4p^*+3)})$.

where $\theta_1, \theta_2 > 0$ and $\nu \in \mathbb{R}$, the following densities are well-defined:

$$
\begin{aligned}
K^{(B)}_{\theta_1,\theta_2}(s) &= \frac{s^{\theta_1-1}(1-s)^{\theta_2-1}}{B(\theta_1,\theta_2)}, \quad 0 \le s \le 1, \\
K^{(G)}_{\theta_1,\theta_2}(s) &= \frac{(s/\theta_2)^{\theta_1-1}e^{-s/\theta_2}}{\theta_2\Gamma(\theta_1)}, \quad s \ge 0, \\
K^{(MB)}_{\nu,\theta_1,\theta_2}(s) &= \frac{(s/\theta_2)^{\nu-1}}{2\theta_2 K_\nu(\theta_1)} \exp\biggl\{-\frac{\theta_1}{2}\biggl(\frac{s}{\theta_2}+\frac{\theta_2}{s}\biggr)\biggr\}, \quad s > 0.
\end{aligned}
$$

Although the last density is known as generalized inverse Gaussian (IG) density (Jørgensen (1982)), it was renamed as "modified Bessel (MB) density" (Igarashi and Kakizawa (2014b)), noting that, in analogy to the gamma and beta densities $K^{(G)}_{\theta_1,\theta_2}(s)$ and $K^{(B)}_{\theta_1,\theta_2}(s)$, the MB function of the third kind, K_ν, is the normalizing constant. Note that $K^{(MB)}_{-1/2,\theta_1,\theta_2}(s)$ and $K^{(MB)}_{1/2,\theta_1,\theta_2}(s)$ are IG and reciprocal IG (RIG) densities, respectively (Tweedie (1957)). A mixture of IG and RIG densities (MIG for short) is defined by

$$
K^{(MIG_\varepsilon)}_{\theta_1,\theta_2}(s) = \varepsilon K^{(MB)}_{1/2,\theta_1,\theta_2}(s) + (1-\varepsilon)K^{(MB)}_{-1/2,\theta_1,\theta_2}(s), \quad s > 0,\ \theta_1,\theta_2 > 0,\ 0 \le \varepsilon \le 1
$$

(Jørgensen et al. (1991)). Especially, $K^{(MIG_{1/2})}_{\theta_1,\theta_2}(s)$ is known as Birnbaum–Saunders (BS) density (Birnbaum and Saunders (1969)).

Besides, we list two densities: One is a weighted log-normal (LN_ν) density, defined by

$$
K^{(LN)}_{\nu,\theta_1,\theta_2}(s) = \frac{s^{\nu-1}}{\sqrt{2\pi\theta_2}} \exp\biggl\{-\frac{(\log s-\theta_1)^2}{2\theta_2} - \nu\theta_1 - \frac{\nu^2\theta_2}{2}\biggr\}, \quad s > 0,\ \nu \in \mathbb{R},\ \theta_1 \in \mathbb{R},\ \theta_2 > 0.
$$

Note that the LN_0 density is the ordinary LN density and that $K^{(LN)}_{\nu,\theta_1,\theta_2}(s) = K^{(LN)}_{0,\theta_1+\nu\theta_2,\theta_2}(s)$. The other is Amoroso (Stacy or generalized gamma) density, defined by

$$
K^{(A)}_{\theta_1,\theta_2,\gamma}(s) = \frac{|\gamma|(s/\theta_2)^{\theta_1\gamma-1}e^{-(s/\theta_2)^\gamma}}{\theta_2\Gamma(\theta_1)}, \quad s \ge 0,\ \theta_1,\theta_2 > 0,\ \gamma \ne 0
$$

(Amoroso (1925), Stacy (1962), and Stacy and Mihram (1965)). The gamma density $K^{(G)}_{\theta_1,\theta_2}(s)$ is a core member with $\gamma = 1$.

To build asymmetric KDEs from the above-mentioned densities, suitable parameterization is important, since, in principle, infinitely many parameterizations are possible[6]. In what follows, the parameter ν (or ε, γ) is chosen in advance, independent of β and x, unless otherwise stated. Then, we parameterize (θ_1, θ_2) as a function of β and x, in such a way that the resulting kernel concentrates around $s = x$ as $\beta \to 0$ (see the top of Section 2). By construction, the shape of such a kernel varies naturally according to the position $x \in \mathcal{S}$. This is a reason why the estimator is sometimes referred to as a varying asymmetric KDE.

[6]When $\mathcal{S} = [0,\infty)$, some existing estimators had the disadvantage that (i) $\widehat{f}_\beta(0) = 0$ even if $f(0) > 0$ (see Remark 7), and (ii) $\psi(x) = x^J$ for some $J \ge 2$ in Assumptions A1 and A2, hence, the asymptotic variance, being proportional to $f(x)/x^{J/2}$, is not integrable on $[0,\infty)$, unless $f(x) = O(x^\alpha)$ for $\alpha > J/2-1$ (in this case, $f(0) = 0$ must be imposed).

Example 1 The gamma KDE (Chen (2000))

$$\widehat{f}_{\beta}^{(G)}(x) = \frac{1}{n}\sum_{i=1}^{n} K_{x/\beta+1,\beta}^{(G)}(X_i), \quad x \geq 0,$$

satisfies Assumptions A1–A3 (see Igarashi and Kakizawa (2015)). We can prove, in supplemental issue (Supplemental appendix to "Higher-order bias corrections for kernel type density estimators on the unit or semi-infinite interval"), that, for any $j \in \mathbb{N}$,

$$\mu_j(K_{x/\beta+1,\beta}^{(G)}(\cdot)) = \sum_{k=\lceil j/2 \rceil}^{j} \zeta_{j,k}^{(G)} x^{j-k}\beta^k, \quad x \geq 0,$$

where $\zeta_{j,k}^{(G)} = \sum_{\ell=k}^{j}(-1)^{j+k-\ell}{}_jC_\ell\, s(\ell+1, \ell+1-k)$ with the Stirling number of the first kind $s(\cdot,\cdot)$. Hence, Assumption A4[p] holds for any $p \in \mathbb{N}$. For example,

$$\begin{aligned}
&\mu_1(K_{x/\beta+1,\beta}^{(G)}(\cdot)) = \beta, \quad \mu_2(K_{x/\beta+1,\beta}^{(G)}(\cdot)) = \beta x + 2\beta^2, \quad \mu_3(K_{x/\beta+1,\beta}^{(G)}(\cdot)) = 5\beta^2 x + 6\beta^3,\\
&\mu_4(K_{x/\beta+1,\beta}^{(G)}(\cdot)) = 3\beta^2x^2 + 26\beta^3 x + 24\beta^4, \quad \mu_5(K_{x/\beta+1,\beta}^{(G)}(\cdot)) = 35\beta^3x^2 + 154\beta^4 x + 120\beta^5,\\
&\mu_6(K_{x/\beta+1,\beta}^{(G)}(\cdot)) = 15\beta^3x^3 + 340\beta^4x^2 + 1044\beta^5 x + 720\beta^6, \quad \mu_8(K_{x/\beta+1,\beta}^{(G)}(\cdot)) = O(\beta^4(x+\beta)^4),
\end{aligned}$$

can be derived using a computer algebra system (e.g., Maple).

Example 2 For every $\varepsilon \in [0,1]$, the MIG$_\varepsilon$ KDE

$$\widehat{f}_{\beta}^{(MIG_\varepsilon)}(x) = \frac{1}{n}\sum_{i=1}^{n} K_{x/\beta+1,x+\beta}^{(MIG_\varepsilon)}(X_i), \quad x \geq 0$$

(the MIG$_\varepsilon$ KDEs when $\varepsilon = 0, 1/2, 1$ are referred to as the IG, BS, and RIG KDEs, respectively) satisfies Assumptions A1–A3 (see Igarashi and Kakizawa (2014b)). Given $j \in \mathbb{N}$,

$$\mu_j(K_{x/\beta+1,x+\beta}^{(MIG_\varepsilon)}(\cdot)) = \sum_{k=\lceil j/2 \rceil}^{j} \zeta_{j,k}^{(MIG_\varepsilon)} x^{j-k}\beta^k, \quad x \geq 0,$$

for some constants $\zeta_{j,k}^{(MIG_\varepsilon)}$'s, independent of β and x (we used a computer algebra system; Maple). For example,

$$\begin{aligned}
&\mu_1(K_{x/\beta+1,x+\beta}^{(MIG_\varepsilon)}(\cdot)) = (\varepsilon+1)\beta, \quad \mu_2(K_{x/\beta+1,x+\beta}^{(MIG_\varepsilon)}(\cdot)) = \beta x + (5\varepsilon+2)\beta^2,\\
&\mu_3(K_{x/\beta+1,x+\beta}^{(MIG_\varepsilon)}(\cdot)) = 3(\varepsilon+2)\beta^2 x + (30\varepsilon+7)\beta^3,\\
&\mu_4(K_{x/\beta+1,x+\beta}^{(MIG_\varepsilon)}(\cdot)) = 3\beta^2x^2 + 3(14\varepsilon+13)\beta^3 x + (229\varepsilon+37)\beta^4,\\
&\mu_5(K_{x/\beta+1,x+\beta}^{(MIG_\varepsilon)}(\cdot)) = 15(\varepsilon+3)\beta^3x^2 + 5(105\varepsilon+62)\beta^4 x + (2165\varepsilon+266)\beta^5,\\
&\mu_6(K_{x/\beta+1,x+\beta}^{(MIG_\varepsilon)}(\cdot)) = 15\beta^3x^3 + 45(9\varepsilon+13)\beta^4x^2 + 30(233\varepsilon+100)\beta^5 x + (24576\varepsilon+2431)\beta^6,\\
&\mu_8(K_{x/\beta+1,x+\beta}^{(MIG_\varepsilon)}(\cdot)) = O(\beta^4(x+\beta)^4).
\end{aligned}$$

Example 3 For every $\nu \in \mathbb{R}$, the LN_ν KDE

$$\widehat{f}_\beta^{(LN_\nu)}(x) = \frac{1}{n}\sum_{i=1}^n K^{(LN)}_{\nu,\mu_\beta(x),\sigma^2_\beta(x)}(X_i), \quad x \geq 0,$$

satisfies Assumptions A1–A3, where $\mu_\beta(x) = \log(x+\beta)$, $\sigma^2_\beta(x) = \log\{1+\beta/(x+\beta)\}$ (see Igarashi (2016b)). The LN_ν KDE, $\nu = -1/2$ or $1/2$, satisfies M1[p^*] of type I for $p^* = 4$ or $p^* = 2$, respectively (we used a computer algebra system; Maple). For example,

$$\begin{aligned}
&\mu_1(K^{(LN)}_{-1/2,\mu_\beta(x),\sigma^2_\beta(x)}(\cdot)) = \beta, \quad \mu_2(K^{(LN)}_{-1/2,\mu_\beta(x),\sigma^2_\beta(x)}(\cdot)) = \beta x + 2\beta^2,\\
&\mu_3(K^{(LN)}_{-1/2,\mu_\beta(x),\sigma^2_\beta(x)}(\cdot)) = 6\beta^2 x + 8\beta^3,\\
&\mu_4(K^{(LN)}_{-1/2,\mu_\beta(x),\sigma^2_\beta(x)}(\cdot)) = 3\beta^2x^2 + 40\beta^3 x + 57\beta^4 + \frac{6\beta^5}{x+\beta} + \frac{\beta^6}{(x+\beta)^2},\\
&\mu_1(K^{(LN)}_{1/2,\mu_\beta(x),\sigma^2_\beta(x)}(\cdot)) = 2\beta, \quad \mu_2(K^{(LN)}_{1/2,\mu_\beta(x),\sigma^2_\beta(x)}(\cdot)) = \beta x + 7\beta^2 + \frac{\beta^3}{x+\beta}
\end{aligned}$$

(we see $m_j = j(j-3)/2$ for $\nu = -1/2$ and $m_j = j(j-1)/2$ for $\nu = 1/2$). On the other hand, the $\mathrm{LN}_{J+1/2}$ KDE, where $J \in \mathbb{N}$, satisfies M1[p^*] of type I for $p^* = 1$ (with $m_{1,j} = j(j+2J-1)/2$ and $m_{2,j} = 1$), the $\mathrm{LN}_{-3/2}$ KDE satisfies M1[p^*] of type I for $p^* = 1$ (with $m_{2,j} = 2$), and the $\mathrm{LN}_{-(J+1/2)}$ KDE, where $J \in \mathbb{N}\backslash\{1\}$, satisfies M1[$p^*$] of type I for $p^* = 1$ (with $m_{2,j} = 1$). Also, the LN_ν KDE, when ν is not an half-integer, satisfies (M2.1) with $c = 1$.

Example 4 For every $\nu \in \mathbb{R}$, the MB_ν KDE

$$\widehat{f}_\beta^{(MB_\nu)}(x) = \frac{1}{n}\sum_{i=1}^n K^{(MB)}_{\nu,x/\beta+1,x+\beta}(X_i), \quad x \geq 0,$$

satisfies Assumptions A1–A3 (see Igarashi and Kakizawa (2014b)). The case $\nu = -1/2$ or $1/2$ corresponds to the MIG_ε KDE, with $\varepsilon = 0$ or 1, respectively. The $\mathrm{MB}_{\pm(J+1/2)}$ KDEs, where $J \in \mathbb{N}$, satisfy M1[p^*] of type II for $p^* = 1$ (with $\ell^* = J$). Also, the MB_ν KDE, when ν is not an half-integer, satisfies (M2.2) with $c = 1$.

Example 5 For every $\gamma \neq 0$, the Amoroso (A_γ) KDE

$$\widehat{f}_\beta^{(A_\gamma)}(x) = \frac{1}{n}\sum_{i=1}^n K^{(A)}_{\alpha_\gamma(x/\beta+c),\beta\theta_\gamma(x/\beta+c),\gamma}(X_i), \quad x \geq 0$$

(we choose $c = 1$ when $\gamma > 0$ or $c > 1$ when $\gamma < 0$[7]) satisfies Assumptions A1–A3 (see Igarashi and Kakizawa (2018a)), where α_γ and θ_γ are continuous functions on $(0,\infty)$, defined by

$$\alpha_\gamma(\rho) = \begin{cases} \dfrac{\rho}{\gamma}, & \gamma > 0, \\ \dfrac{\rho+1}{|\gamma|}, & \gamma < 0, \end{cases} \qquad \theta_\gamma(\rho) = \frac{\rho\Gamma(\alpha_\gamma(\rho))}{\Gamma(\alpha_\gamma(\rho)+1/\gamma)}.$$

[7]To ensure the existence of the higher-order moments of the resulting kernel, $c = 1$ is not allowed for $\gamma < 0$. More precisely, the jth moment is well-defined if $j < x/\beta + c + 1$ (see Igarashi and Kakizawa (2018a) for the details).

The case $\gamma = 1$ corresponds to the gamma KDE. The $\mathrm{A}_{1/J}$ KDE, where $J \in \mathbb{N}\backslash\{1\}$, satisfies M1[$p^*$] of type I for $p^* = 2$ (with $m_{1,j} = j-1$ and $m_{2,j} = J-1$), the $\mathrm{A}_{-1/J}$ KDE, where $J \in \mathbb{N}$, satisfies M1[p^*] of type I for $p^* = 2$ (with $m_{1,j} = 1$ and $m_{2,j} = J(j-1)$), and the A_γ KDE, except for $\gamma \in \{1/M \mid M \in \mathbb{Z}, M \neq 0\}$, satisfies (M2.3).

Example 6 The beta KDE (Chen (1999))

$$\widehat{f}_\beta^{(B)}(x) = \frac{1}{n}\sum_{i=1}^n K^{(B)}_{x/\beta+1,(1-x)/\beta+1}(X_i), \quad 0 \le x \le 1,$$

satisfies Assumptions A1 and A2 (see Igarashi (2016a)). Given $j \in \mathbb{N}$,

$$\mu_j(K^{(B)}_{x/\beta+1,(1-x)/\beta+1}(\cdot)) = \frac{\sum_{k=\lceil j/2\rceil}^{j} \widetilde{\zeta}^{(B)}_{j,k}(x)\beta^k}{\prod_{k=1}^{j}\{1+(k+1)\beta\}}, \quad 0 \le x \le 1,$$

for some polynomials $\widetilde{\zeta}^{(B)}_{j,k}(x)$'s, independent of β (we used a computer algebra system; Maple). Expanding the denominator, we have, for example, uniformly in $x \in [0,1]$,

$$\begin{aligned}
&\mu_1(K^{(B)}_{x/\beta+1,(1-x)/\beta+1}(\cdot)) = \beta(1-2x) + 2\beta^2\{-(1-2x)\} + 4\beta^3(1-2x) + O(\beta^4),\\
&\mu_2(K^{(B)}_{x/\beta+1,(1-x)/\beta+1}(\cdot)) = \beta x(1-x) + \beta^2\{2-11x(1-x)\} + \beta^3\{-10+49x(1-x)\} + O(\beta^4),\\
&\mu_3(K^{(B)}_{x/\beta+1,(1-x)/\beta+1}(\cdot)) = 5\beta^2(1-2x)x(1-x) + 3\beta^3(1-2x)\{2-19x(1-x)\} + O(\beta^4),\\
&\mu_4(K^{(B)}_{x/\beta+1,(1-x)/\beta+1}(\cdot)) = 3\beta^2x^2(1-x)^2 + 2\beta^3x(1-x)\{13-64x(1-x)\} + O(\beta^4),\\
&\mu_5(K^{(B)}_{x/\beta+1,(1-x)/\beta+1}(\cdot)) = 35\beta^3(1-2x)x^2(1-x)^2 + O(\beta^4),\\
&\mu_6(K^{(B)}_{x/\beta+1,(1-x)/\beta+1}(\cdot)) = 15\beta^3x^3(1-x)^3 + O(\beta^4), \quad \mu_8(K^{(B)}_{x/\beta+1,(1-x)/\beta+1}(\cdot)) = O(\beta^4).
\end{aligned}$$

Remark 6 (Two-regime type estimators) As variants of the gamma KDE, Chen (2000) and Igarashi and Kakizawa (2014b) introduced two-regime ρ-function

$$\rho_c(t) = \begin{cases} t+c, & t \ge 2,\\ (c+1)\left(\dfrac{t}{2}\right)^{2/(c+1)} + 1, & t \in [0,2) \end{cases} \quad \text{(say)},$$

and suggested a class of modified gamma KDEs $\widehat{f}_\beta^{(G2)}(x) = n^{-1}\sum_{i=1}^n K^{(G)}_{\rho_c(x/\beta),\beta}(X_i)$, $x \ge 0$ ($c = 1/4$ is the best choice in the sense of the $O(n^{-4/5})$-MISE). See also Igarashi and Kakizawa (2014b, 2018a) and Igarashi (2016b) for the related two-regime type MIG, LN_ν, and A_γ KDEs. However, it was revealed (see Igarashi and Kakizawa (2015, 2018a,c)) that the two-regime type KDEs ($\mathcal{S} = [0,\infty)$), after the bias corrections, have the MISEs of order $O(n^{-6/7})$ at most, unless f satisfies a shoulder condition $f^{(1)}(0) = 0$. Similarly, the two-regime type modified beta KDE $\widehat{f}_\beta^{(B2)}(x) = n^{-1}\sum_{i=1}^n K^{(B)}_{\rho_c(x/\beta),\rho_c((1-x)/\beta)}(X_i)$, $0 \le x \le 1$, after the bias corrections, has the MISE of order $O(n^{-6/7})$ at most, unless $f^{(1)}(0) = f^{(1)}(1) = 0$.

Remark 7 (Bad estimators with $\widehat{f}_\beta(0)=0$) The variants of the IG, BS, and LN KDEs due to Jin and Kawczak (2003) and Scaillet (2004) were based on other parameterization of the IG, BS, and LN_0 densities, i.e., $K^{(MIG_0)}_{1/(\beta x),x}$, $K^{(MIG_{1/2})}_{1/\beta,x}$, and $K^{(LN)}_{0,\log x,4\log(1+\beta)}$, respectively. It may be true that these estimators, with/without bias corrections, work when $x>0$ (the details are omitted here). But, their kernels converge to zero as $x\to 0$; consequently, their estimators yield $\widehat{f}_\beta(0)=0$ regardless of $f(0)=0$ or $f(0)>0$. Also, the variant of the RIG KDE due to Scaillet (2004) (his kernel is $K^{(MIG_1)}_{x/\beta-1,x-\beta}$) had the downward bias $\{e^{-2(1-x/\beta)}-1\}f(0)+O(\beta)$ for $x<\beta$ (see Igarashi and Kakizawa (2014b)). These problems were obviously caused by the bad parameterization; when $x=0$ ($x<\beta$), their parameters lie outside the parameter spaces of the IG, BS, and LN_0 (RIG) densities. Hence, these estimators are not appropriate for estimating a density with $f(0)>0$.

5. Simulation studies

We illustrate the finite sample performance of the bias-corrected estimators (4)–(6) for $p=2,3$ ($\boldsymbol{a}=(1,a)$ and $\boldsymbol{a}=(1,a,1/a)$ with $a=0.1,0.5,0.9$), through the simulations, using the Amoroso (A_γ, $\gamma=\pm0.5,\pm1,\pm1.5,\pm2$), IG, BS, RIG, and $\mathrm{LN}_{-1/2}$ kernels (Examples 1, 2, 3, and 5). Note that, when $p=3$, the use of the gamma (A_1), IG, BS, RIG, and $\mathrm{LN}_{-1/2}$ kernels enables us to attain the convergence rate $n^{-12/13}$ of the MISE; however, in general, the use of the A_γ kernel, where $\gamma\in\mathbb{R}\backslash\{0,1\}$, yields the convergence rate $n^{-10/11}$ of the MISE, though it is faster than $n^{-8/9}$ for the previous paper (e.g., Igarashi and Kakizawa (2018a)).

We generated 1000 samples of $n=100,200$ from four densities:

$$\text{A. } f(x)=\frac{1}{2}\Big(\frac{e^{-x/3}}{3}+\frac{xe^{-x/3}}{9}\Big),\quad \text{B. } f(x)=\frac{e^{-x/3}}{3},\quad \text{C. } f(x)=\frac{1}{2}\Big(\frac{e^{-x/10}}{10}+xe^{-x}\Big),$$
$$\text{D. } f(x)=\frac{1}{2}\bigg[\frac{1}{\sqrt{2\pi}0.8x}\exp\Big\{-\frac{(\log x-1)^2}{2(0.8)^2}\Big\}+\frac{1}{\sqrt{2\pi}0.4x}\exp\Big\{-\frac{(\log x-2)^2}{2(0.4)^2}\Big\}\bigg],$$

and then calculated the integrated squared error (ISE) $ISE^{[k]}=\int_{\mathcal{S}}\{\widehat{f}^{[k]}_\beta(x)-f(x)\}^2dx$ for the kth sample. Each smoothing parameter β was chosen using the least squared cross-validation method. We chose $\epsilon=0.000001\times\beta^2$.

From Tables 1–4, we observe that the average ISEs, $\sum_{k=1}^{1000}ISE^{[k]}/1000$, decreased, as the sample size n increased. For the cases A and B, the average ISE of $\widehat{f}_{\beta,\#^3_{(1,a,1/a)}}$ decreased, as a was close to one. Among the bias corrections using the IG, BS, and RIG kernels, the IG kernel had the best performance and the BS kernel was the second best. The IG kernel was inferior to the best implemented A_γ kernel. The $\mathrm{LN}_{-1/2}$ kernel had the similar performance to the IG kernel. Hereafter, we pay attention to the A_γ kernel ($n=200$ and $a=0.9$).

- For the case A, $\widehat{f}^{(A_\gamma)}_{\beta,\#^3_{(1,a,1/a)}}$, $\#=ADD,TS,JF$, outperformed $\widehat{f}^{(A_\gamma)}_{\beta,\#^2_{(1,a)}}$, except for some kernels (overall, $\gamma<0$ was not good). Here, $\widehat{f}^{(A_1)}_{\beta,TS^3_{(1,a,1/a)}}$ had the best performance.

- For the case B, $\widehat{f}^{(A_\gamma)}_{\beta,\#^3_{(1,a,1/a)}}$, $\# = ADD, TS$, outperformed $\widehat{f}^{(A_\gamma)}_{\beta,\#^2_{(1,a)}}$, except for some kernels (overall, $\gamma < 0$ was not good), and $\widehat{f}^{(A_{\pm 0.5})}_{\beta,JF^3_{(1,a,1/a)}}$ outperformed $\widehat{f}^{(A_{\pm 0.5})}_{\beta,JF^2_{(1,a)}}$. Here, $\widehat{f}^{(A_2)}_{\beta,TS^3_{(1,a,1/a)}}$ had the best performance.

- For the case C, $\widehat{f}^{(A_\gamma)}_{\beta,\#^3_{(1,a,1/a)}}$, $\# = ADD, TS$, outperformed $\widehat{f}^{(A_\gamma)}_{\beta,\#^2_{(1,a)}}$ for $\gamma \leq 0.5$, and $\widehat{f}^{(A_\gamma)}_{\beta,JF^3_{(1,a,1/a)}}$ worked well. Here, $\widehat{f}^{(A_{0.5})}_{\beta,TS^3_{(1,a,1/a)}}$ had the best performance, and, among the A_γ's except $\gamma = 0.5$, $\widehat{f}^{(A_{-0.5})}_{\beta,\#^3_{(1,a,1/a)}}$, $\# = ADD, TS, JF$, worked well.

- For the case D (in this case, $f(0) = 0$), $\widehat{f}^{(A_{-0.5})}_{\beta,\#^2_{(1,a)}}$ and $\widehat{f}^{(A_{-0.5})}_{\beta,\#^3_{(1,a,1/a)}}$, $\# = ADD, TS, JF$, worked well. Here, $\widehat{f}^{(A_{-0.5})}_{\beta,ADD^2_{(1,a)}}$ had the best performance and $\widehat{f}^{(A_{-0.5})}_{\beta,ADD^3_{(1,a,1/a)}}$ was the second best. Note that a better performance of the exponent $\gamma < 0$ is also found in other contexts (Kakizawa and Igarashi (2017) and Igarashi and Kakizawa (2018a,b,c)).

6. Concluding remarks

6.1. Limiting estimator

Given a positive vector $\boldsymbol{a} = (1, H_2(a), \ldots, H_p(a))'$, such that $\lim_{a\to 1} H_k(a) = 1$ for $k = 2, \ldots, p$, we guess that $\lim_{a\to 1} \widehat{f}_{\beta,\#^p_{\boldsymbol{a}}}(x)$, $\# = ADD, TS, JF$, are also the bias-corrected asymmetric KDEs, provided that, for the TS-type and JF-type, $\epsilon > 0$ is independent of $a \in (0, 1)$. In fact, extending Igarashi and Kakizawa (2015, 2018b) and Igarashi (2016a) for $p = 2$, we can construct the limiting estimators $\lim_{a\to 1} \widehat{f}_{\beta,\#^3_{(1,a,1/a)}}(x) = \widehat{f}_{\beta,\#^3_{(1,1,1)}}(x)$ (say), where

$$
\begin{aligned}
\widehat{f}_{\beta,ADD^3_{(1,1,1)}}(x) &= \widehat{f}_\beta(x) - \beta \frac{\partial}{\partial \beta} \widehat{f}_\beta(x) + \frac{1}{2}\beta^2 \frac{\partial^2}{\partial \beta^2} \widehat{f}_\beta(x), \\
\widehat{f}_{\beta,TS^3_{(1,1,1)}}(x) &= \{\widehat{f}_\beta(x) + \epsilon\} \exp\left[\frac{\widehat{f}_{\beta,ADD^3_{(1,1,1)}}(x)}{\widehat{f}_\beta(x) + \epsilon} - 1 - \frac{\{\beta \frac{\partial}{\partial \beta} \widehat{f}_\beta(x) + \epsilon\}^2}{2\{\widehat{f}_\beta(x) + \epsilon\}^2} \right], \\
\widehat{f}_{\beta,JF^3_{(1,1,1)}}(x) &= \{\widehat{f}_\beta(x) + \epsilon\} \exp\left[\sum_{j=1}^{2} \frac{(-1)^{j-1}}{j} \left\{ \frac{\widehat{f}_{\beta,ADD^3_{(1,1,1)}}(x)}{\widehat{f}_\beta(x) + \epsilon} - 1 \right\}^j \right].
\end{aligned}
$$

6.2. Case $\mathcal{S} = \mathbb{R}$

Suppose that $supp(f) = \mathbb{R}$. If f is $2p$ times continuously differentiable for some $p \in \mathbb{N}\backslash\{1\}$, the classical Rosenblatt–Parzen KDE $\widehat{f}^{(K_{[2]})}_h(x) = (nh)^{-1} \sum_{i=1}^n K_{[2]}((x - X_i)/h)$, using a symmetric second-order kernel $K_{[2]}$, yields

$$
E[\widehat{f}^{(K_{[2]})}_h(x)] = f(x) + \sum_{k=1}^{p} h^{2k} \frac{f^{(2k)}(x)}{(2k)!} \int_{-\infty}^{\infty} z^{2k} K_{[2]}(z) dz + o(h^{2p})
$$

(in this case, $\beta = h^2$). For a given positive vector $\boldsymbol{a} = (a_1, \ldots, a_p)'$, such that the a_k's are distinct, let $\boldsymbol{a}^2 = (a_1^2, \ldots, a_p^2)'$. The bias-corrected KDE, $(nh)^{-1} \sum_{i=1}^n K_{[2p],\boldsymbol{a}}((x - X_i)/h)$, can

Table 1: Case A. The average ISEs$\times 10^6$ of estimators with/without bias corrections, where A_γ, IG, BS, RIG, and $LN_{-1/2}$ stand for the asymmetric KDEs (see Examples 1, 2, 3, and 5). The number in the parentheses stands for the standard deviation$\times 10^6$ of the ISEs.

$n = 100$

	p	a	A_2	$A_{1.5}$	A_1	$A_{0.5}$	$A_{-0.5}$	A_{-1}	$A_{-1.5}$	A_{-2}	IG	BS	RIG	$LN_{-1/2}$
original			3389	3263	3006	2914	3201	3392	3735	4135	3316	3539	3751	3288
			(2946)	(3216)	(3144)	(3206)	(3235)	(2898)	(2674)	(2941)	(3346)	(3411)	(3468)	(3317)
ADD	2	0.1	2715	2607	2498	2595	2782	2750	3099	3419	2627	2944	3071	2564
			(3023)	(2994)	(3129)	(3061)	(2931)	(2606)	(2595)	(2666)	(2843)	(3396)	(3312)	(2756)
		0.5	2369	2356	2255	2382	2531	2519	2905	3344	2352	2574	2738	2328
			(3005)	(3185)	(3023)	(2826)	(2405)	(2279)	(2362)	(2572)	(2615)	(2992)	(3271)	(2649)
		0.9	2268	2326	2207	2372	2511	2565	2956	3476	2336	2564	2713	2327
			(2627)	(3185)	(2856)	(2781)	(2262)	(2432)	(2442)	(2700)	(2587)	(3042)	(3265)	(2682)
ADD	3	0.1	3533	3182	2866	2773	3193	3675	4237	4863	3336	2975	2954	3332
			(3758)	(3619)	(3584)	(3630)	(3633)	(3771)	(4052)	(4380)	(3809)	(3585)	(3533)	(3841)
		0.5	2221	2162	2000	2298	2502	2626	3311	4118	2281	2413	2513	2241
			(3074)	(3374)	(2786)	(2878)	(2710)	(2794)	(3047)	(3622)	(3047)	(3052)	(3258)	(2994)
		0.9	2090	2034	2004	2314	2446	2568	3455	4465	2176	2299	2392	2138
			(2814)	(2971)	(2784)	(2890)	(2411)	(2450)	(2996)	(3494)	(2661)	(2858)	(3142)	(2752)
TS	2	0.1	2434	2396	2334	2600	2653	2329	2500	2752	2314	2600	2717	2295
			(3047)	(3023)	(3016)	(3101)	(2465)	(2270)	(2542)	(2945)	(2616)	(2988)	(3162)	(2695)
		0.5	2061	2164	2071	2407	2557	2298	2767	3376	2202	2386	2477	2121
			(2459)	(3016)	(2690)	(2808)	(2242)	(2327)	(2630)	(2963)	(2478)	(2912)	(3096)	(2549)
		0.9	2026	2091	2066	2411	2606	2390	2967	3665	2165	2369	2470	2117
			(2383)	(2789)	(2702)	(2882)	(2303)	(2462)	(2782)	(3082)	(2581)	(2934)	(3163)	(2602)
TS	3	0.1	3864	3426	2997	2832	3307	3905	4687	5457	3545	3143	2990	3551
			(3911)	(3769)	(3637)	(3658)	(3696)	(3826)	(4282)	(4742)	(3929)	(3825)	(3672)	(3946)
		0.5	2593	2303	$\underline{\underline{1989}}$	2321	2631	2898	3777	4862	2456	2310	2355	2459
			(3315)	(3409)	(2969)	(2862)	(2739)	(2879)	(3218)	(3788)	(3040)	(3218)	(3172)	(3105)
		0.9	2574	2146	$\underline{1933}$	2342	2614	2820	3941	5242	2380	2259	2313	2375
			(3246)	(2979)	(2919)	(2862)	(2690)	(2639)	(3251)	(3842)	(2979)	(3106)	(2998)	(3041)
JF	2	0.1	2692	2600	2483	2591	2753	2738	3088	3364	2615	2949	3044	2556
			(3010)	(3009)	(3054)	(3063)	(2823)	(2617)	(2610)	(2585)	(2842)	(3413)	(3291)	(2766)
		0.5	2226	2272	2150	2388	2531	2464	2838	3279	2266	2501	2643	2243
			(2795)	(3141)	(2849)	(2819)	(2272)	(2317)	(2436)	(2629)	(2492)	(2986)	(3182)	(2618)
		0.9	2029	2140	2069	2401	2581	2384	2898	3571	2198	2413	2533	2164
			(2377)	(3009)	(2680)	(2839)	(2249)	(2440)	(2624)	(3034)	(2574)	(2989)	(3202)	(2619)
JF	3	0.1	3435	3147	2807	2634	2748	3215	3805	4262	3013	3400	3659	3045
			(2975)	(2966)	(3219)	(3446)	(3380)	(2995)	(2942)	(2912)	(3228)	(3149)	(3232)	(3182)
		0.5	2368	2328	2154	2343	2455	2663	3136	3617	2396	2537	2694	2401
			(2960)	(3325)	(2877)	(2856)	(2520)	(2590)	(2600)	(2751)	(2833)	(3082)	(3205)	(2849)
		0.9	2139	2174	2112	2332	2464	2592	3124	3667	2297	2445	2622	2325
			(2615)	(3166)	(2933)	(2876)	(2467)	(2411)	(2527)	(2774)	(2527)	(3033)	(3412)	(2754)

The underlined or double-underlined number indicates the smallest or second smallest average ISE, respectively.

Table 1: (continued).

$n = 200$

	p	a	A_2	$A_{1.5}$	A_1	$A_{0.5}$	$A_{-0.5}$	A_{-1}	$A_{-1.5}$	A_{-2}	IG	BS	RIG	$LN_{-1/2}$
original			2107	2005	1802	1690	1831	2048	2283	2489	1971	2157	2329	1942
			(2285)	(2353)	(2171)	(2077)	(2098)	(2044)	(2018)	(1913)	(2229)	(2372)	(2475)	(2166)
ADD	2	0.1	1579	1532	1458	1421	1607	1650	1779	1988	1566	1724	1763	1538
			(1884)	(2155)	(2185)	(1739)	(2011)	(1932)	(1508)	(1902)	(2153)	(2337)	(2235)	(2134)
		0.5	1323	1282	1256	1333	1473	1465	1624	1850	1361	1516	1599	1363
			(1799)	(1793)	(1805)	(1724)	(1787)	(1580)	(1259)	(1320)	(1884)	(2257)	(2315)	(1991)
		0.9	1301	1265	1232	1316	1468	1437	1639	1886	1336	1505	1589	1337
			(1717)	(1747)	(1734)	(1728)	(1664)	(1300)	(1252)	(1321)	(1848)	(2273)	(2330)	(1950)
ADD	3	0.1	1879	1723	1562	1460	1663	1924	2204	2538	1767	1662	1733	1768
			(2531)	(2418)	(2350)	(2194)	(2136)	(2370)	(2330)	(2617)	(2397)	(2406)	(2346)	(2428)
		0.5	1255	1176	1165	1267	1466	1421	1744	2083	1244	1355	1401	1239
			(1904)	(1798)	(1796)	(1622)	(1750)	(1442)	(1629)	(1856)	(1689)	(2061)	(2010)	(1769)
		0.9	1192	1122	1153	1232	1417	1401	1770	2282	1220	1296	1372	1220
			(1429)	(1733)	(1819)	(1455)	(1420)	(1380)	(1546)	(1869)	(1443)	(1812)	(1996)	(1745)
TS	2	0.1	1381	1384	1352	1432	1577	1404	1361	1445	1376	1560	1571	1331
			(1807)	(2093)	(2119)	(1798)	(1950)	(1813)	(1324)	(1382)	(1995)	(2280)	(2189)	(1923)
		0.5	1239	1217	1198	1333	1475	1296	1417	1678	1245	1378	1435	1221
			(1730)	(1760)	(1749)	(1711)	(1529)	(1238)	(1286)	(1470)	(1619)	(2006)	(2134)	(1629)
		0.9	1202	1182	1180	1313	1475	1320	1512	1826	1238	1374	1421	1238
			(1714)	(1717)	(1729)	(1555)	(1534)	(1318)	(1401)	(1533)	(1579)	(2025)	(2131)	(1683)
TS	3	0.1	2029	1814	1616	1495	1692	2005	2388	2749	1844	1673	1614	1853
			(2580)	(2422)	(2366)	(2216)	(2106)	(2326)	(2513)	(2579)	(2405)	(2437)	(2396)	(2434)
		0.5	1365	1219	1106	1304	1514	1490	1908	2403	1262	1299	1364	1264
			(1682)	(1904)	(1791)	(1612)	(1836)	(1473)	(1571)	(1884)	(1719)	(1858)	(2121)	(1713)
		0.9	1331	1173	1080	1257	1480	1478	1983	2659	1232	1272	1310	1229
			(1518)	(1823)	(1750)	(1413)	(1423)	(1367)	(1616)	(1996)	(1634)	(1811)	(1895)	(1619)
JF	2	0.1	1570	1512	1456	1419	1600	1641	1767	1963	1559	1728	1775	1533
			(1907)	(2131)	(2180)	(1739)	(1991)	(1931)	(1505)	(1873)	(2146)	(2357)	(2353)	(2128)
		0.5	1280	1244	1226	1330	1480	1382	1551	1796	1303	1465	1537	1302
			(1779)	(1757)	(1771)	(1715)	(1760)	(1285)	(1224)	(1398)	(1657)	(2198)	(2243)	(1751)
		0.9	1208	1185	1194	1311	1466	1343	1516	1802	1248	1389	1441	1247
			(1713)	(1704)	(1772)	(1556)	(1526)	(1312)	(1363)	(1509)	(1585)	(2036)	(2149)	(1687)
JF	3	0.1	1926	1855	1656	1484	1587	1891	2135	2351	1808	1993	2085	1793
			(2021)	(2420)	(2211)	(2139)	(2106)	(2064)	(1881)	(1792)	(2332)	(2464)	(2383)	(2201)
		0.5	1287	1225	1205	1291	1447	1489	1694	1951	1341	1462	1493	1338
			(1819)	(1789)	(1809)	(1773)	(1693)	(1590)	(1298)	(1521)	(1921)	(2292)	(2044)	(1978)
		0.9	1226	1157	1182	1238	1397	1466	1705	1997	1245	1400	1416	1266
			(1798)	(1726)	(1827)	(1457)	(1387)	(1552)	(1253)	(1499)	(1379)	(2168)	(1983)	(1716)

Table 2: Case B. The average ISEs$\times 10^6$ of estimators with/without bias corrections, where A_γ, IG, BS, RIG, and $LN_{-1/2}$ stand for the asymmetric KDEs (see Examples 1, 2, 3, and 5). The number in the parentheses stands for the standard deviation$\times 10^6$ of the ISEs.

$n = 100$

	p	a	A_2	$A_{1.5}$	A_1	$A_{0.5}$	$A_{-0.5}$	A_{-1}	$A_{-1.5}$	A_{-2}	IG	BS	RIG	$LN_{-1/2}$
original			6452	6480	6049	5578	5924	6652	7505	7983	6565	7184	7466	6478
			(5871)	(7440)	(7483)	(7131)	(6366)	(6154)	(6561)	(6147)	(7350)	(8495)	(7694)	(7071)
ADD	2	0.1	5247	5315	4989	5022	5215	5615	6042	6552	5220	5827	5942	5239
			(5829)	(7027)	(6508)	(7021)	(6105)	(5991)	(5760)	(5512)	(6176)	(7052)	(6380)	(6481)
		0.5	4142	4365	4490	4686	4684	4707	5340	5912	4537	5046	5400	4419
			(4704)	(5369)	(5935)	(6591)	(5231)	(4215)	(4569)	(4675)	(4995)	(5758)	(6548)	(5051)
		0.9	4069	4312	4368	4659	4693	4662	5337	5940	4505	5013	5208	4337
			(4699)	(5469)	(5789)	(6653)	(5309)	(4154)	(4556)	(4668)	(5644)	(5811)	(6185)	(4881)
ADD	3	0.1	5417	5211	4748	4816	5292	5966	6757	7402	5431	5794	5874	5393
			(6988)	(7563)	(6885)	(7329)	(6916)	(7284)	(7439)	(7606)	(7545)	(7618)	(6431)	(7726)
		0.5	3697	3819	4109	4532	4613	4450	5132	5926	4129	4543	4495	4039
			(5053)	(5180)	(5736)	(6727)	(5319)	(4371)	(5057)	(5721)	(4401)	(5325)	(5297)	(4667)
		0.9	3234	3627	3893	4369	4589	4258	5012	6105	4007	4447	4324	3805
			(3697)	(5093)	(5439)	(6268)	(5429)	(4240)	(4902)	(5671)	(4310)	(5362)	(5313)	(4336)
TS	2	0.1	4577	4593	4681	4980	4983	4496	4619	4870	4627	5312	5344	4638
			(5600)	(6078)	(6526)	(7108)	(5656)	(4617)	(4379)	(4458)	(5460)	(6822)	(6464)	(6346)
		0.5	3583	3869	4081	4529	4674	4069	4447	5130	3926	4588	4490	3753
			(4768)	(5368)	(5697)	(6471)	(5223)	(4352)	(3999)	(4223)	(4352)	(5425)	(5480)	(4394)
		0.9	3498	3785	4013	4503	4714	4124	4765	5547	3906	4501	4432	3730
			(4686)	(5342)	(5594)	(6483)	(5304)	(4172)	(4793)	(5000)	(4330)	(5404)	(5530)	(4375)
TS	3	0.1	5308	4990	4584	4747	5362	6117	7263	8087	5488	5298	5144	5548
			(7213)	(7508)	(7008)	(7387)	(6899)	(7029)	(7961)	(8237)	(7766)	(7936)	(6567)	(8052)
		0.5	3057	3326	3732	4434	4785	4618	5877	7406	3855	4157	4036	3708
			(4216)	(5215)	(5687)	(6652)	(5835)	(4666)	(5427)	(6395)	(4611)	(5016)	(5222)	(4714)
		0.9	$\underline{2874}$	$\underline{\underline{2987}}$	3583	4353	4738	4560	6292	8495	3813	4056	3931	3570
			(3934)	(4640)	(5510)	(6700)	(5585)	(4569)	(5859)	(6950)	(4414)	(4929)	(5309)	(4658)
JF	2	0.1	5175	5307	4974	5014	5221	5587	6058	6421	5224	5804	5926	5267
			(5740)	(7057)	(6524)	(7018)	(6112)	(6018)	(5957)	(4998)	(6320)	(7074)	(6431)	(6608)
		0.5	3961	4195	4416	4610	4682	4554	5065	5614	4265	4964	5038	4150
			(4743)	(5412)	(6560)	(6549)	(5257)	(4441)	(4511)	(4498)	(4603)	(5891)	(6210)	(4706)
		0.9	3548	3887	4070	4506	4748	4155	4777	5468	3980	4656	4554	3817
			(4640)	(5438)	(5609)	(6477)	(5474)	(4146)	(4683)	(4878)	(4351)	(5579)	(5625)	(4403)
JF	3	0.1	6808	6348	6139	5570	5760	6905	7568	8463	6715	7146	7315	6572
			(6137)	(5891)	(6817)	(7097)	(5966)	(6381)	(5599)	(6085)	(7364)	(8813)	(6553)	(6831)
		0.5	4303	4329	4393	4627	4554	4711	5400	6201	4540	4919	5137	4511
			(5092)	(5433)	(5794)	(6618)	(4591)	(4243)	(4661)	(5190)	(4965)	(5559)	(6004)	(5007)
		0.9	3877	4000	4195	4439	4520	4579	5265	6262	4301	4689	4793	4257
			(4514)	(5052)	(5623)	(6062)	(5266)	(4248)	(4574)	(5216)	(4557)	(5280)	(5434)	(4767)

The underlined or double-underlined number indicates the smallest or second smallest average ISE, respectively.

Table 2: (continued).

$n = 200$

	p	a	A_2	$A_{1.5}$	A_1	$A_{0.5}$	$A_{-0.5}$	A_{-1}	$A_{-1.5}$	A_{-2}	IG	BS	RIG	$LN_{-1/2}$
original			3726	3531	3222	2882	3084	3701	4098	4515	3470	3749	4063	3456
			(2917)	(3335)	(3304)	(2891)	(2702)	(3209)	(2766)	(2800)	(3329)	(3342)	(3441)	(3299)
ADD	2	0.1	2917	2798	2620	2502	2656	3020	3342	3711	2839	3083	3248	2792
			(2771)	(2849)	(2843)	(2772)	(2581)	(3253)	(2720)	(2883)	(3334)	(3403)	(3163)	(3223)
		0.5	2601	2515	2342	2357	2482	2655	3028	3463	2476	2772	2948	2502
			(2983)	(2955)	(2724)	(2720)	(2472)	(2291)	(2244)	(2400)	(2676)	(2931)	(3172)	(2786)
		0.9	2464	2403	2321	2336	2477	2627	3010	3557	2430	2753	2893	2420
			(2704)	(2782)	(2757)	(2730)	(2500)	(2259)	(2208)	(2492)	(2627)	(3084)	(3116)	(2682)
ADD	3	0.1	2859	2676	2488	2433	2591	3024	3448	3886	2750	2990	3161	2721
			(2899)	(2870)	(2855)	(2804)	(2673)	(3426)	(3219)	(3471)	(3431)	(3034)	(3043)	(3438)
		0.5	2155	2195	2122	2240	2388	2398	2806	3350	2226	2501	2592	2209
			(2666)	(2975)	(2729)	(2663)	(2476)	(2083)	(2245)	(2633)	(2530)	(2955)	(3032)	(2584)
		0.9	2015	2075	2074	2206	2322	2318	2863	3425	2138	2414	2532	2125
			(2362)	(2889)	(2786)	(2633)	(2416)	(2079)	(2311)	(2560)	(2483)	(2902)	(3113)	(2509)
TS	2	0.1	2566	2473	2398	2437	2553	2627	2742	2920	2544	2823	2880	2503
			(2827)	(2779)	(2799)	(2720)	(2414)	(3028)	(2415)	(2331)	(3289)	(3418)	(3107)	(3209)
		0.5	2095	2164	2137	2297	2445	2302	2600	2934	2199	2512	2624	2173
			(2532)	(2817)	(2717)	(2669)	(2382)	(2090)	(2273)	(2304)	(2511)	(2907)	(3070)	(2612)
		0.9	2015	2131	2103	2289	2443	2279	2602	3083	2191	2467	2592	2101
			(2276)	(2834)	(2722)	(2676)	(2387)	(2035)	(2153)	(2376)	(2548)	(2862)	(3097)	(2479)
TS	3	0.1	2612	2436	2268	2332	2559	2979	3472	3954	2612	2666	2860	2621
			(3095)	(3379)	(2936)	(2800)	(2593)	(3508)	(3455)	(3781)	(3435)	(3288)	(3531)	(3514)
		0.5	1740	1824	1943	2212	2430	2328	3022	3909	2057	2299	2371	1965
			(2207)	(2660)	(2761)	(2664)	(2382)	(2113)	(2558)	(3165)	(2451)	(2841)	(3233)	(2380)
		0.9	1580	1633	1901	2193	2363	2320	3321	4510	2023	2191	2259	1862
			(1841)	(2131)	(2772)	(2667)	(2263)	(2034)	(2739)	(3380)	(2408)	(2727)	(3040)	(2061)
JF	2	0.1	2896	2778	2604	2493	2655	3022	3326	3680	2827	3046	3240	2773
			(2777)	(2850)	(2829)	(2767)	(2582)	(3286)	(2738)	(2862)	(3338)	(3389)	(3201)	(3219)
		0.5	2404	2355	2290	2318	2453	2541	2954	3344	2384	2694	2819	2341
			(2710)	(2878)	(2822)	(2717)	(2432)	(2270)	(2335)	(2453)	(2656)	(2944)	(3134)	(2684)
		0.9	2078	2171	2126	2291	2445	2321	2699	3148	2213	2527	2632	2156
			(2337)	(2827)	(2730)	(2678)	(2397)	(2027)	(2162)	(2351)	(2502)	(2862)	(3075)	(2514)
JF	3	0.1	3492	3294	3035	2768	2937	3438	3869	4252	3252	3445	3750	3236
			(2868)	(2850)	(2771)	(2940)	(2696)	(2779)	(2915)	(2872)	(2887)	(2804)	(3116)	(2824)
		0.5	2468	2360	2274	2289	2410	2584	2925	3356	2408	2654	2840	2420
			(2790)	(2856)	(2785)	(2742)	(2548)	(2237)	(2286)	(2446)	(2675)	(2973)	(3197)	(2727)
		0.9	2240	2263	2187	2230	2346	2454	2925	3439	2295	2555	2699	2282
			(2357)	(2939)	(2782)	(2637)	(2537)	(2017)	(2306)	(2536)	(2499)	(2987)	(3044)	(2526)

Table 3: Case C. The average ISEs$\times 10^6$ of estimators with/without bias corrections, where A_γ, IG, BS, RIG, and $LN_{-1/2}$ stand for the asymmetric KDEs (see Examples 1, 2, 3, and 5). The number in the parentheses stands for the standard deviation$\times 10^6$ of the ISEs.

$n = 100$

	p	a	A_2	$A_{1.5}$	A_1	$A_{0.5}$	$A_{-0.5}$	A_{-1}	$A_{-1.5}$	A_{-2}	IG	BS	RIG	$LN_{-1/2}$
original			7178	6656	6045	5411	5763	6688	7454	8060	6462	6926	7545	6482
			(4902)	(4722)	(4604)	(4325)	(4565)	(4967)	(4944)	(4936)	(4970)	(5060)	(5310)	(4912)
ADD	2	0.1	7477	6978	6194	5339	5869	6845	7774	8532	6627	7148	7936	6707
			(4909)	(4915)	(4651)	(4249)	(4682)	(4964)	(4932)	(4979)	(4884)	(5261)	(5564)	(4897)
		0.5	7818	7199	6286	5335	5818	7126	8381	9492	6744	7294	8162	6814
			(5031)	(4894)	(4526)	(4440)	(4394)	(4465)	(4841)	(5337)	(4767)	(5267)	(5476)	(4653)
		0.9	7788	7200	6309	5299	5871	7269	8489	9430	6769	7310	8144	6814
			(4900)	(4911)	(4600)	(4395)	(4562)	(4511)	(4977)	(5501)	(4667)	(5238)	(5321)	(4677)
ADD	3	0.1	7358	6711	5920	5164	5561	6482	7569	8409	6311	6995	7948	6369
			(5172)	(4992)	(4618)	(4128)	(4685)	(5134)	(5715)	(5877)	(5081)	(5189)	(5554)	(5091)
		0.5	8374	7636	6468	5301	5588	6191	7015	8125	6483	7503	8523	6620
			(5196)	(5102)	(4652)	(4386)	(4253)	(4423)	(4727)	(5347)	(4598)	(5307)	(5645)	(4709)
		0.9	8534	7766	6495	5237	5600	6133	6967	7704	6527	7606	8770	6583
			(4637)	(5094)	(4742)	(4285)	(4238)	(4388)	(4840)	(5199)	(4582)	(5288)	(5678)	(4587)
TS	2	0.1	7898	7258	6314	5331	5891	7612	9214	10476	6882	7345	8292	6992
			(5059)	(4934)	(4687)	(4287)	(4530)	(4595)	(4846)	(5234)	(4659)	(5119)	(5531)	(4737)
		0.5	8570	7694	6543	5303	5961	7884	9762	11405	7341	7537	8715	7681
			(5066)	(5186)	(4768)	(4403)	(4506)	(4681)	(5929)	(7180)	(4621)	(5017)	(5451)	(4491)
		0.9	8522	7698	6582	5274	5797	6898	8032	9209	6972	7577	8757	7179
			(4612)	(5161)	(4980)	(4410)	(4402)	(4628)	(5375)	(6409)	(4654)	(5075)	(5430)	(4728)
TS	3	0.1	7808	6964	6009	5117	5524	6566	7881	8975	6408	7103	8213	6491
			(5548)	(5183)	(4798)	(4154)	(4735)	(5264)	(5942)	(6215)	(5214)	(5329)	(5694)	(5265)
		0.5	8803	7906	6719	5274	5527	5981	6672	7750	6328	7667	9289	6484
			(5285)	(5011)	(4807)	(4447)	(4431)	(4881)	(5048)	(5534)	(4828)	(5196)	(5335)	(5092)
		0.9	8880	8132	6809	5212	5430	5725	6468	7659	6291	7761	9706	6474
			(4900)	(5046)	(5038)	(4319)	(4166)	(4412)	(5121)	(5914)	(4647)	(5221)	(5382)	(5156)
JF	2	0.1	7493	6994	6202	5332	5872	6873	7797	8585	6645	7165	7966	6747
			(4922)	(4913)	(4652)	(4254)	(4727)	(4899)	(4931)	(4990)	(4860)	(5262)	(5583)	(5073)
		0.5	8130	7402	6439	5313	5903	7698	9224	10554	6942	7437	8329	6986
			(5080)	(5080)	(4748)	(4396)	(4431)	(4377)	(4948)	(5869)	(4652)	(5257)	(5394)	(4634)
		0.9	8416	7628	6521	5283	5837	7051	8131	9201	7006	7525	8678	7182
			(4791)	(5136)	(4763)	(4421)	(4440)	(4656)	(5287)	(6011)	(4645)	(5060)	(5417)	(4661)
JF	3	0.1	7367	6837	6222	5370	5685	6874	7657	8392	6591	6950	7519	6633
			(4911)	(4754)	(4617)	(4227)	(4240)	(5031)	(5011)	(5431)	(4805)	(4819)	(4898)	(4808)
		0.5	7739	7266	6365	5264	5770	6866	7692	8399	6741	7389	8264	6829
			(4941)	(4791)	(4594)	(4383)	(4389)	(4362)	(4609)	(4672)	(4499)	(5161)	(5440)	(4560)
		0.9	7797	7347	6425	5223	5792	6842	7576	8122	6845	7500	8471	6936
			(4804)	(5022)	(4746)	(4288)	(4305)	(4264)	(4179)	(4494)	(4614)	(5296)	(5572)	(4503)

The underlined or double-underlined number indicates the smallest or second smallest average ISE, respectively.

Table 3: (continued).

$n = 200$

	p	a	A_2	$A_{1.5}$	A_1	$A_{0.5}$	$A_{-0.5}$	A_{-1}	$A_{-1.5}$	A_{-2}	IG	BS	RIG	$LN_{-1/2}$
original			4241	3911	3492	3048	3224	3774	4248	4696	3646	3977	4359	3681
			(2773)	(2770)	(2659)	(2492)	(2634)	(2799)	(2842)	(2942)	(2777)	(2813)	(2895)	(2802)
ADD	2	0.1	4455	4042	3544	2933	3136	3779	4435	5108	3670	4055	4549	3701
			(2927)	(2823)	(2740)	(2504)	(2432)	(2612)	(3050)	(3308)	(2657)	(2887)	(3032)	(2689)
		0.5	4505	4049	3426	2835	3038	4127	5337	6288	3667	4058	4554	3695
			(2852)	(2712)	(2322)	(2041)	(2202)	(3139)	(3757)	(4285)	(2777)	(2893)	(3076)	(2816)
		0.9	4472	3979	3424	2818	3023	4313	5776	6805	3651	4057	4526	3666
			(2741)	(2572)	(2305)	(2018)	(2200)	(3153)	(3886)	(4392)	(3011)	(2917)	(3064)	(2819)
ADD	3	0.1	4466	4048	3542	2900	3083	3662	4271	4836	3599	4081	4578	3614
			(2965)	(2846)	(2749)	(2268)	(2371)	(2571)	(2847)	(3189)	(2587)	(2902)	(3048)	(2595)
		0.5	4676	4095	3484	2816	3013	3594	4338	5097	3513	4077	4657	3591
			(3053)	(2760)	(2436)	(2032)	(2162)	(2621)	(3058)	(3460)	(2557)	(2901)	(3272)	(2739)
		0.9	4877	4137	3433	2804	3035	3764	4633	5523	3593	4104	4760	3626
			(3125)	(2908)	(2415)	(2021)	(2194)	(2788)	(3332)	(4130)	(2697)	(2964)	(3400)	(2866)
TS	2	0.1	4691	4169	3588	2894	3115	4307	5705	6884	3773	4168	4706	3805
			(3013)	(2873)	(2584)	(2275)	(2435)	(3167)	(3759)	(4343)	(2745)	(2988)	(3120)	(2779)
		0.5	4851	4234	3560	2816	2983	5476	7348	8759	4146	4187	4780	4273
			(3053)	(2842)	(2471)	(2019)	(2127)	(3967)	(5067)	(5958)	(3164)	(2836)	(3179)	(3257)
		0.9	4791	4194	3546	2806	2995	4477	5837	6713	3942	4188	4825	4076
			(2936)	(2741)	(2474)	(2032)	(2183)	(3344)	(4336)	(4960)	(3068)	(2853)	(3263)	(3196)
TS	3	0.1	4651	4142	3522	2873	3052	3634	4308	4945	3586	4116	4734	3605
			(3083)	(2902)	(2471)	(2253)	(2358)	(2579)	(2884)	(3258)	(2601)	(2756)	(3154)	(2616)
		0.5	4948	4297	3560	2803	2946	3459	4065	4667	3541	4234	5141	3579
			(3240)	(2845)	(2409)	(2026)	(2074)	(2499)	(2923)	(3188)	(2591)	(2938)	(3497)	(2780)
		0.9	5367	4335	3548	$\underline{2792}$	2979	3499	4109	4943	3509	4346	5557	3520
			(3543)	(2918)	(2440)	(2003)	(2131)	(2480)	(3061)	(4306)	(2511)	(3068)	(3774)	(2551)
JF	2	0.1	4471	4049	3550	2932	3132	3794	4449	5139	3677	4073	4561	3694
			(2923)	(2827)	(2748)	(2512)	(2433)	(2634)	(3048)	(3338)	(2662)	(2920)	(3060)	(2670)
		0.5	4688	4147	3450	2821	3036	4601	6325	7562	3744	4126	4635	3773
			(3021)	(3003)	(2278)	(2031)	(2237)	(3395)	(4163)	(4860)	(2844)	(2917)	(3086)	(2892)
		0.9	4730	4156	3525	$\underline{\underline{2802}}$	3006	4596	6092	6985	3948	4179	4759	4063
			(2821)	(2642)	(2382)	(2000)	(2194)	(3391)	(4346)	(4864)	(3075)	(2854)	(3207)	(3169)
JF	3	0.1	4214	3921	3511	2959	3158	3781	4290	4676	3649	3955	4323	3664
			(2922)	(2919)	(2812)	(2592)	(2704)	(2709)	(2954)	(2881)	(2680)	(2914)	(3036)	(2685)
		0.5	4431	4030	3444	2807	3031	3908	4639	5127	3577	4074	4561	3625
			(2859)	(2741)	(2364)	(2037)	(2214)	(2910)	(3279)	(3372)	(2565)	(2968)	(3134)	(2681)
		0.9	4414	3994	3472	2809	3064	4211	4984	5425	3656	4077	4656	3774
			(2877)	(2692)	(2665)	(2030)	(2284)	(3016)	(3257)	(3365)	(2687)	(3013)	(3301)	(3000)

Table 4: Case D. The average ISEs$\times 10^6$ of estimators with/without bias corrections, where A_γ, IG, BS, RIG, and $LN_{-1/2}$ stand for the asymmetric KDEs (see Examples 1, 2, 3, and 5). The number in the parentheses stands for the standard deviation$\times 10^6$ of the ISEs.

$n = 100$

	p	a	A_2	$A_{1.5}$	A_1	$A_{0.5}$	$A_{-0.5}$	A_{-1}	$A_{-1.5}$	A_{-2}	IG	BS	RIG	$LN_{-1/2}$
original			4665	4230	3669	3168	3155	3800	4600	5371	3708	4170	4705	3721
			(2851)	(2744)	(2491)	(2295)	(2316)	(2659)	(3011)	(3326)	(2599)	(2736)	(2907)	(2586)
ADD	2	0.1	5011	4399	3736	3190	3120	4095	5443	6363	3699	4223	4890	3786
			(3046)	(2861)	(2592)	(2436)	(2405)	(3053)	(3537)	(3660)	(2681)	(2826)	(3113)	(2742)
		0.5	5191	4602	3818	3159	3046	4362	5559	6291	3763	4257	5065	3940
			(3076)	(3013)	(2689)	(2400)	(2372)	(3204)	(3318)	(3280)	(2895)	(2967)	(3359)	(3065)
		0.9	5209	4627	3848	3151	3071	3881	4744	5299	3662	4247	5103	3783
			(2817)	(3021)	(2751)	(2350)	(2460)	(2977)	(3083)	(3130)	(2795)	(2970)	(3430)	(2898)
ADD	3	0.1	4861	4301	3625	3129	3099	3702	4466	5138	3644	4180	4847	3681
			(3227)	(2952)	(2612)	(2403)	(2414)	(2698)	(2997)	(3230)	(2679)	(2827)	(3109)	(2691)
		0.5	5342	4708	3879	3146	3023	3442	4189	4861	3522	4217	5174	3585
			(3145)	(3087)	(2796)	(2267)	(2397)	(2648)	(2996)	(3208)	(2679)	(2984)	(3530)	(2744)
		0.9	5254	4766	3864	3155	2984	3422	4111	4800	3497	4243	5289	3561
			(2751)	(3094)	(2729)	(2294)	(2276)	(2684)	(3016)	(3238)	(2694)	(3021)	(3650)	(2722)
TS	2	0.1	5265	4583	3834	3233	3171	4589	5628	6328	4008	4368	5168	4211
			(3144)	(2986)	(2613)	(2392)	(2409)	(3346)	(3790)	(4061)	(2972)	(2901)	(3277)	(3157)
		0.5	5446	4875	4000	3240	3177	4241	5285	6080	3906	4572	5616	4052
			(2984)	(3064)	(2711)	(2371)	(2464)	(2984)	(3158)	(3268)	(2839)	(3079)	(3649)	(2988)
		0.9	5424	4854	4014	3236	3152	4004	4794	5312	3822	4541	5552	3922
			(2876)	(2987)	(2702)	(2341)	(2382)	(2969)	(3013)	(3071)	(2718)	(3019)	(3559)	(2783)
TS	3	0.1	5108	4493	3740	3183	3153	3804	4624	5372	3753	4313	5020	3795
			(3315)	(3073)	(2689)	(2422)	(2414)	(2724)	(3035)	(3292)	(2711)	(2887)	(3149)	(2722)
		0.5	5641	5008	4079	3284	3167	3692	4381	5003	3817	4537	5444	3853
			(3108)	(3114)	(2751)	(2394)	(2420)	(2666)	(2997)	(3180)	(2728)	(3017)	(3423)	(2727)
		0.9	5489	5033	4134	3273	3128	3675	4278	4838	3816	4587	5584	3882
			(2642)	(3024)	(2755)	(2353)	(2279)	(2689)	(3004)	(3161)	(2716)	(3034)	(3547)	(2782)
JF	2	0.1	5046	4415	3747	3196	3126	4139	5510	6479	3718	4240	4919	3804
			(3086)	(2864)	(2593)	(2442)	(2405)	(3048)	(3571)	(3699)	(2695)	(2831)	(3124)	(2747)
		0.5	5328	4751	3910	3200	3103	4353	5429	6107	3860	4403	5263	4047
			(2960)	(3006)	(2678)	(2395)	(2386)	(3103)	(3208)	(3170)	(2902)	(2979)	(3394)	(3060)
		0.9	5402	4836	3997	3227	3144	3979	4774	5282	3805	4502	5534	3907
			(2852)	(2987)	(2701)	(2340)	(2381)	(2946)	(3007)	(3054)	(2718)	(2992)	(3561)	(2791)
JF	3	0.1	4554	4102	3588	3155	3222	3852	4414	4855	3730	4141	4621	3757
			(2874)	(2727)	(2548)	(2343)	(2374)	(2693)	(2837)	(3027)	(2640)	(2748)	(2893)	(2644)
		0.5	5366	4751	3938	3151	3082	3590	4253	4887	3665	4347	5235	3738
			(3034)	(3007)	(2739)	(2235)	(2394)	(2716)	(2961)	(3183)	(2704)	(2999)	(3402)	(2758)
		0.9	5326	4819	3933	3153	3034	3521	4121	4706	3622	4371	5370	3698
			(2683)	(2972)	(2731)	(2209)	(2266)	(2710)	(2974)	(3088)	(2706)	(3040)	(3553)	(2794)

The underlined or double-underlined number indicates the smallest or second smallest average ISE, respectively.

Table 4: (continued).

$n = 200$

	p	a	A_2	$A_{1.5}$	A_1	$A_{0.5}$	$A_{-0.5}$	A_{-1}	$A_{-1.5}$	A_{-2}	IG	BS	RIG	$LN_{-1/2}$
original			2896	2615	2286	1988	1976	2311	2683	3055	2287	2579	2874	2293
			(1845)	(1780)	(1677)	(1564)	(1548)	(1724)	(1828)	(2007)	(1714)	(1775)	(1851)	(1717)
ADD	2	0.1	3016	2677	2273	1975	1924	2240	2870	3772	2195	2568	2913	2222
			(1957)	(1908)	(1762)	(1633)	(1593)	(1792)	(2245)	(2902)	(1726)	(1862)	(1963)	(1757)
		0.5	3155	2711	2273	1963	1891	2424	3998	5128	2136	2509	2971	2159
			(2089)	(1988)	(1775)	(1624)	(1593)	(2281)	(2988)	(3031)	(1732)	(1869)	(2127)	(1811)
		0.9	3175	2715	2257	1952	$\underline{1877}$	2155	3056	3774	2145	2521	2948	2140
			(1950)	(1986)	(1754)	(1610)	(1593)	(1861)	(2492)	(2542)	(1755)	(1890)	(2103)	(1754)
ADD	3	0.1	3031	2693	2269	1968	1924	2222	2645	3034	2207	2567	2927	2226
			(2028)	(1930)	(1759)	(1613)	(1685)	(1771)	(1895)	(1998)	(1745)	(1868)	(1983)	(1766)
		0.5	3319	2834	2318	1944	1888	2038	2450	2908	2107	2552	2981	2133
			(2063)	(2067)	(1860)	(1389)	(1631)	(1746)	(1941)	(2109)	(1751)	(1982)	(2167)	(1802)
		0.9	3604	2876	2320	1945	$\underline{\underline{1883}}$	2023	2439	3022	2109	2573	3014	2098
			(1851)	(2135)	(1856)	(1388)	(1626)	(1740)	(1795)	(2351)	(1785)	(2013)	(2233)	(1764)
TS	2	0.1	3131	2771	2352	2012	1964	2590	3413	4009	2295	2667	3036	2329
			(2028)	(1919)	(1777)	(1638)	(1598)	(2297)	(2900)	(3286)	(1779)	(1886)	(1985)	(1814)
		0.5	3376	2912	2404	2020	1945	2536	3863	4928	2281	2722	3191	2308
			(2081)	(2017)	(1810)	(1635)	(1596)	(2173)	(2745)	(2872)	(1771)	(1937)	(2152)	(1842)
		0.9	3385	2914	2416	2010	1935	2281	3079	3780	2289	2729	3205	2285
			(1919)	(2014)	(1809)	(1610)	(1587)	(1812)	(2319)	(2411)	(1790)	(1919)	(2162)	(1793)
TS	3	0.1	3156	2795	2350	2002	1962	2314	2758	3164	2295	2681	3046	2313
			(2048)	(1940)	(1769)	(1603)	(1601)	(1793)	(1911)	(2025)	(1761)	(1904)	(1994)	(1782)
		0.5	3565	3042	2471	2029	1953	2229	2643	3060	2299	2777	3281	2309
			(1989)	(2068)	(1851)	(1636)	(1636)	(1809)	(1972)	(2100)	(1793)	(1998)	(2197)	(1848)
		0.9	3790	3103	2505	2017	1946	2216	2677	3140	2317	2807	3303	2305
			(1717)	(2048)	(1854)	(1441)	(1628)	(1801)	(2005)	(2266)	(1805)	(1992)	(2200)	(1821)
JF	2	0.1	3035	2693	2282	1979	1930	2266	2926	3860	2219	2582	2929	2238
			(2010)	(1910)	(1762)	(1634)	(1592)	(1810)	(2297)	(2950)	(1741)	(1863)	(1962)	(1756)
		0.5	3282	2833	2347	1997	1917	2499	3942	5001	2223	2636	3091	2229
			(2094)	(2009)	(1805)	(1637)	(1594)	(2229)	(2815)	(2852)	(1771)	(1914)	(2106)	(1805)
		0.9	3359	2900	2394	2003	1937	2269	3100	3762	2279	2718	3175	2283
			(1905)	(2012)	(1784)	(1608)	(1604)	(1814)	(2357)	(2399)	(1790)	(1920)	(2134)	(1814)
JF	3	0.1	2957	2693	2347	2009	2029	2433	2817	3101	2389	2666	2955	2402
			(1869)	(1796)	(1682)	(1521)	(1525)	(1730)	(1842)	(1907)	(1703)	(1803)	(1881)	(1706)
		0.5	3338	2889	2357	1997	1922	2147	2514	2960	2208	2633	3088	2218
			(1959)	(2037)	(1788)	(1640)	(1615)	(1805)	(1922)	(2146)	(1752)	(1934)	(2132)	(1793)
		0.9	3658	2930	2394	1960	1919	2085	2503	3031	2189	2659	3105	2178
			(1806)	(2096)	(1858)	(1390)	(1653)	(1725)	(1862)	(2286)	(1767)	(1988)	(2193)	(1783)

be constructed, where $K_{[2p],\boldsymbol{a}}(\cdot) = \sum_{j=1}^{p} a_j c_j(\boldsymbol{a}^2) K_{[2]}(a_j \cdot)$ is a $2p$th-order kernel, that is an extension of Schucany and Sommers' fourth-order kernel $K_{[4],(1,a)}$ ($a \neq 1$), as mentioned in Introduction. Such kernels (independent of interest) form a class of $2p$th-order kernels, whose limiting version may be also considered. For example, we produce a class of 6th-order kernels from a given $K_{[2]}$, as follows:

$$K_{[6],(1,a,1/a)}(s) = \begin{cases} \dfrac{1}{(a^2+1)(a^2-1)^2}\left\{-a^2(a^2+1)K_{[2]}(s) + a^7 K_{[2]}(as) + \dfrac{1}{a}K_{[2]}\left(\dfrac{s}{a}\right)\right\}, & a \neq 1, \\ \dfrac{1}{8}\left\{15K_{[2]}(s) + 9sK'_{[2]}(s) + s^2 K''_{[2]}(s)\right\}, & a = 1. \end{cases}$$

Setting $K_{[2]}(s) = e^{-s^2/2}/\sqrt{2\pi} = \phi(s)$ (say), we obtain the (Gaussian-based) 6th-order kernel $\phi(s)(15 - 10s^2 + s^4)/8$, which is found in Nadaraya (1974) and Wand and Schucany (1990). As alternatives to the additive-type bias-corrected estimator; $\sum_{j=1}^{p} c_j(a_1^2, \ldots, a_p^2)\widehat{f}_{h/a_j}^{(K_{[2]})}(x)$, the TS/JF-type bias-corrected estimators can be further proposed (the details are omitted).

Remark 8 In Terrell and Scott (1980), a linear combination of the Rosenblatt–Parzen KDEs; $\sum_{j=1}^{p} c_j(1, 1/2^2, \ldots, 1/p^2)\widehat{f}_{jh}^{(K_{[2]})}(x)$, as well as a multiplicative analogue (they are what we call the additive-type and TS-type bias correction methods), were already mentioned to obtain the faster convergence rate of the MISE; $n^{-4p/(4p+1)}$, where

$$c_j(1, 1/2^2, \ldots, 1/p^2) = 2(-1)^{j-1}\frac{\prod_{i=1}^{j}(p-i+1)}{\prod_{i=1}^{j}(p+i)} \quad \text{for } j = 1, \ldots, p.$$

6.3. Multivariate density estimation

We briefly discuss the extension of the bias correction methods to the multivariate setting with $supp(f) = [0,\infty)^m$ (the case $[0,1]^m$ is similar)[8]. Let $\widehat{f}_\beta(\boldsymbol{x})$, $\boldsymbol{x} = (x_1, \ldots, x_m)' \in [0,\infty)^m$, be a density estimator, such that $E[\widehat{f}_\beta(\boldsymbol{x})] = f(\boldsymbol{x}) + \sum_{j=1}^{p} \beta^j \gamma_{j,m}(\boldsymbol{x}; f) + o(\beta^p)$ for some $p \in \mathbb{N}\backslash\{1\}$ and functions $\gamma_{j,m}(\cdot; f)$, $j = 1, \ldots, p$, independent of β. Such an estimator can be constructed using the product kernel, as follows: for a random sample $\{\boldsymbol{X}_i = (X_{i1}, \ldots, X_{im})', i = 1, \ldots, n\}$ of size n, $\widehat{f}_\beta(\boldsymbol{x}) = n^{-1}\sum_{i=1}^{n}\prod_{j=1}^{m} K(X_{ij}, x_j, c_j\beta)$ is a product-type asymmetric KDE, where c_j's are positive constants, independent of β and $\boldsymbol{x}$. Given a positive vector $\boldsymbol{a} = (a_1, \ldots, a_p)'$, such that the a_k's are distinct, the bias-corrected estimators are defined by

$$\widehat{f}_{\beta, ADD_{\boldsymbol{a}}^p}(\boldsymbol{x}) = \sum_{k=1}^{p} c_k(\boldsymbol{a})\widehat{f}_{\beta/a_k}(\boldsymbol{x}), \quad \widehat{f}_{\beta, TS_{\boldsymbol{a}}^p}(\boldsymbol{x}) = \prod_{k=1}^{p}\left\{\widehat{f}_{\beta/a_k}(\boldsymbol{x}) + \frac{\epsilon}{a_k}\right\}^{c_k(\boldsymbol{a})},$$

$$\widehat{f}_{\beta, JF_{\boldsymbol{a}}^p}(\boldsymbol{x}) = \{\widehat{f}_\beta(\boldsymbol{x}) + \epsilon\}\exp\left[\sum_{j=1}^{p-1}\frac{(-1)^{j-1}}{j}\left\{\frac{\widehat{f}_{\beta, ADD_{\boldsymbol{a}}^p}(\boldsymbol{x})}{\widehat{f}_\beta(\boldsymbol{x}) + \epsilon} - 1\right\}^j\right].$$

The details are omitted here.

[8]We do not pursue the density estimation with mixed support, such as $supp(f) = [0,\infty) \times \mathbb{R}$.

Appendix A.1. Preliminary results

We write

$$\overline{\Delta}_\beta(x) = \widehat{f}_\beta(x) - E[\widehat{f}_\beta(x)],$$

which is the average of zero-mean independent random variables

$$\Delta(X_i; x, \beta) = K(X_i; x, \beta) - E[K(X_i; x, \beta)], \quad i = 1, \ldots, n.$$

Note that $V[\widehat{f}_\beta(x)] = V[\overline{\Delta}_\beta(x)] = n^{-1}E[\Delta^2(X_1; x, \beta)]$. Also, we write

$$\overline{\Delta}_{\beta,ADD^p_{\boldsymbol{a}}}(x) = \sum_{k=1}^{p} c_k(\boldsymbol{a})\overline{\Delta}_{\beta/a_k}(x) = \widehat{f}_{\beta,ADD^p_{\boldsymbol{a}}}(x) - E[\widehat{f}_{\beta,ADD^p_{\boldsymbol{a}}}(x)],$$

which is the average of zero-mean independent random variables

$$\Delta_{ADD^p_{\boldsymbol{a}}}(X_i; x, \beta) = \sum_{k=1}^{p} c_k(\boldsymbol{a})\Delta(X_i; x, \beta/a_k), \quad i = 1, \ldots, n.$$

Lemma A.1 *Suppose that Assumption* A1(i) *holds. For any $n \in \mathbb{N}$, $\beta, t > 0$, $x \in \mathcal{S}$, and $j \geq 2$, we have*

$$\begin{aligned}
\text{(i)} \quad E\big[|\overline{\Delta}_\beta(x)|^j\big] &\leq C(j)\bigg\{\frac{1}{n^{j-1}}E\big[|\Delta(X_1; x, \beta)|^j\big] + \Big(\frac{1}{n}E\big[\Delta^2(X_1; x, \beta)\big]\Big)^{j/2}\bigg\} \\
&\quad (\textit{the constant } C(j) > 0 \textit{ depends only on } j) \\
&\leq C(j)\bigg\{\Big(\frac{C_K\beta^{-1}}{n}\Big)^{j-2} + \Big(\frac{C_K\beta^{-1}}{n}\Big)^{(j-2)/2}\bigg\}\frac{1}{n}E\big[\Delta^2(X_1; x, \beta)\big] \\
&= O((n\beta)^{-(j-2)/2}V[\widehat{f}_\beta(x)]) \quad (\textit{if } j > 2, \textit{ assume } n\beta \to \infty),
\end{aligned}$$

$$\text{(ii)} \quad P\big[|\overline{\Delta}_\beta(x)| \geq t\big] \leq 2\exp\bigg\{-\frac{n\beta t^2}{C_K(2||f||_{\mathcal{S}} + t)}\bigg\},$$

$$\text{(iii)} \quad P[|\overline{\Delta}_{\beta,ADD^p_{\boldsymbol{a}}}(x)| \geq t] \leq 2\exp\bigg[-\frac{n\beta t^2}{C_K\{2p\sum_{k=1}^p c_k^2(\boldsymbol{a})a_k||f||_{\mathcal{S}} + t\sum_{k=1}^p |c_k(\boldsymbol{a})|a_k\}}\bigg].$$

Proof Assumption A1(i) enables us to see that, for $i = 1, \ldots, n$,

$$\begin{aligned}
|\Delta(X_i; x, \beta)| &\leq \sup_{s\in\mathcal{S}} K(s; x, \beta) \leq C_K\beta^{-1}, \\
V[\Delta(X_i; x, \beta)] &\leq \int_{\mathcal{S}} K^2(s; x, \beta)f(s)ds \leq C_K\beta^{-1}||f||_{\mathcal{S}}.
\end{aligned}$$

Hence, Rosenthal's inequality and Bennett's inequality yield the results (i) and (ii). Similarly, we have the result (iii), noting that

$$\begin{aligned}
|\Delta_{ADD^p_{\boldsymbol{a}}}(X_i; x, \beta)| &\leq \sum_{k=1}^{p} |c_k(\boldsymbol{a})||\Delta(X_i; x, \beta/a_k)| \leq \sum_{k=1}^{p} |c_k(\boldsymbol{a})|C_K a_k\beta^{-1}, \\
V[\Delta_{ADD^p_{\boldsymbol{a}}}(X_i; x, \beta)] &\leq p\sum_{k=1}^{p} c_k^2(\boldsymbol{a})V[\Delta(X_i; x, \beta/a_k)] \leq p\sum_{k=1}^{p} c_k^2(\boldsymbol{a})C_K a_k\beta^{-1}||f||_{\mathcal{S}}. \quad \square
\end{aligned}$$

Lemma A.2 *Let $a_0, a_0' > 0$ be arbitrary constants.*
(i) *Suppose that Assumptions* A1*,* A2*,* A4$'$[1]*, and* A5$'$(i) *hold. Then,*

$$Cov[\widehat{f}_{\beta/a_0}(x), \widehat{f}_{\beta/a_0'}(x)] = \begin{cases} n^{-1}\beta^{-1/2}\left(\dfrac{2a_0a_0'}{a_0+a_0'}\right)^{1/2} V(x;f)[1+O(\beta\psi^{-1}(x))] + O(n^{-1}), & x \in \mathcal{S}_{I,\beta}, \\ n^{-1}\beta^{-1}f(x)[\varsigma_{a_0,a_0'}(\kappa) + \chi_{\{x\notin\mathcal{S}_B\}}o(1)] + O(n^{-1}), & x \in \mathcal{S}_{B,\beta,\kappa}. \end{cases}$$

(ii) *Suppose that Assumptions* A1–A3*,* A4$'$[1]*, and* A5$'$ *hold. Then,*

$$\int_{\mathcal{S}} Cov[\widehat{f}_{\beta/a_0}(x), \widehat{f}_{\beta/a_0'}(x)]dx = n^{-1}\beta^{-1/2}\left(\frac{2a_0a_0'}{a_0+a_0'}\right)^{1/2}\int_{\mathcal{S}} V(x;f)dx + o(n^{-1}\beta^{-1/2}).$$

Proof (i) Assumptions A1, A4$'$[1], and A5$'$(i) yield $0 < \int_{\mathcal{S}} K(s;x,\beta)f(s)ds \le ||f||_{\mathcal{S}}$ and

$$\begin{aligned} &\left|\int_{\mathcal{S}} K(s;x,\beta/a_0)K(s;x,\beta/a_0')(s-x)\int_0^1 f'(x+\theta(s-x))d\theta ds\right| \\ &\le ||f'||_{\mathcal{S}}\int_{\mathcal{S}} |s-x|K(s;x,\beta/a_0)K(s;x,\beta/a_0')ds = O(1) \quad \text{for } x \in \mathcal{S}_{I,\beta}\cup\mathcal{S}_{B,\beta,\kappa}, \end{aligned}$$

since

$$\int_{\mathcal{S}} |s-x|K(s;x,\beta/a_0)K(s;x,\beta/a_0')ds \le \begin{cases} C_K' a_0^{1/2}\{\beta\psi(x)\}^{-1/2}\mu_2^{1/2}(K(\cdot;x,\beta/a_0')), & x \in \mathcal{S}_{I,\beta}, \\ C_K a_0\beta^{-1}\mu_2^{1/2}(K(\cdot;x,\beta/a_0')), & x \in \mathcal{S}_{B,\beta,\kappa}. \end{cases}$$

Hence, under Assumption A2, we have

$$\begin{aligned} &Cov[\widehat{f}_{\beta/a_0}(x), \widehat{f}_{\beta/a_0'}(x)] \\ &= n^{-1}\left\{\int_{\mathcal{S}} K(s;x,\beta/a_0)K(s;x,\beta/a_0')f(s)ds - \int_{\mathcal{S}} K(s;x,\beta/a_0)f(s)ds\int_{\mathcal{S}} K(s;x,\beta/a_0')f(s)ds\right\} \\ &= n^{-1}\int_{\mathcal{S}} K(s;x,\beta/a_0)K(s;x,\beta/a_0')\left\{f(x) + (s-x)\int_0^1 f'(x+\theta(s-x))d\theta\right\}ds + O(n^{-1}) \\ &= \begin{cases} n^{-1}\beta^{-1/2}\left(\dfrac{2a_0a_0'}{a_0+a_0'}\right)^{1/2} V(x;f)[1+O(\beta\psi^{-1}(x))] + O(n^{-1}), & x \in \mathcal{S}_{I,\beta}, \\ n^{-1}\beta^{-1}f(x)[\varsigma_{a_0,a_0'}(\kappa) + \chi_{\{x\notin\mathcal{S}_B\}}o(1)] + O(n^{-1}), & x \in \mathcal{S}_{B,\beta,\kappa}. \end{cases} \end{aligned}$$

(ii) Note that, under Assumption A1(i), we have, for any interval $I(\subset \mathcal{S})$,

$$\begin{aligned} &\left|\int_I Cov[\widehat{f}_{\beta/a_0}(x), \widehat{f}_{\beta/a_0'}(x)]dx\right| \\ &\le \left\{\int_I V[\widehat{f}_{\beta/a_0}(x)]dx\int_I V[\widehat{f}_{\beta/a_0'}(x)]dx\right\}^{1/2} \\ &\le n^{-1}\left\{\int_I\int_{\mathcal{S}} K^2(s;x,\beta/a_0)f(s)dsdx\int_I\int_{\mathcal{S}} K^2(s;x,\beta/a_0')f(s)dsdx\right\}^{1/2} \\ &\le n^{-1}\beta^{-1}C_K\left\{a_0a_0'\int_I\int_{\mathcal{S}} K(s;x,\beta/a_0)f(s)dsdx\int_I\int_{\mathcal{S}} K(s;x,\beta/a_0')f(s)dsdx\right\}^{1/2}. \end{aligned}$$

The case $\mathcal{S} = [0,\infty)$: Under Assumption A1(i) and the boundedness of f, we can see that, choosing $\tau \in (1/2, 1)$,

$$\left|\int_0^{\beta^\tau} Cov[\widehat{f}_{\beta/a_0}(x), \widehat{f}_{\beta/a_0'}(x)]dx\right| \le n^{-1}\beta^{\tau-1}C_K(a_0a_0')^{1/2}||f||_{[0,\infty)} = o(n^{-1}\beta^{-1/2}).$$

Under Assumptions A1(i), A3, and A5$'$(ii), the choice of $\tau' \in (1/\{2(k'+1)\}, 1/2)$ yields

$$\begin{aligned}
&\left|\int_{\beta^{-\tau'}}^{\infty} Cov[\widehat{f}_{\beta/a_0}(x), \widehat{f}_{\beta/a_0'}(x)]dx\right| \\
&\leq n^{-1}\beta^{-1}C_K\left\{a_0a_0'\int_0^{\infty}\int_{\beta^{-\tau'}}^{\infty} K(s;x,\beta/a_0)dxf(s)ds\int_0^{\infty}\int_{\beta^{-\tau'}}^{\infty} K(s;x,\beta/a_0')dxf(s)ds\right\}^{1/2} \\
&= O(n^{-1}\beta^{\tau'(k'+1)-1}) = o(n^{-1}\beta^{-1/2}).
\end{aligned}$$

Also,

$$\begin{aligned}
n^{-1}\int_{\beta^{\tau}}^{\beta^{-\tau'}}\left\{\beta^{1/2}\frac{f(x)}{\sqrt{\psi^3(x)}}+1\right\}dx &\leq n^{-1}\left\{\beta^{1/2-\tau}\int_0^{\infty}\frac{f(x)}{\sqrt{\psi(x)}}dx+\beta^{-\tau'}\right\} = o(n^{-1}\beta^{-1/2}), \\
n^{-1}\beta^{-1/2}\left(\int_0^{\beta^{\tau}}+\int_{\beta^{-\tau'}}^{\infty}\right)\frac{f(x)}{\sqrt{\psi(x)}}dx &\leq n^{-1}\beta^{-1/2}\left\{||f||_{[0,\infty)}\int_0^{\beta^{\tau}}\frac{1}{\sqrt{\psi(x)}}dx+\beta^{\tau'/2}\int_{\beta^{-\tau'}}^{\infty}f(x)dx\right\} \\
&= o(n^{-1}\beta^{-1/2}).
\end{aligned}$$

Combining them with the result (i) yields

$$\begin{aligned}
&\left|\int_0^{\infty} Cov[\widehat{f}_{\beta/a_0}(x), \widehat{f}_{\beta/a_0'}(x)]dx - n^{-1}\beta^{-1/2}\left(\frac{2a_0a_0'}{a_0+a_0'}\right)^{1/2}\int_0^{\infty}V(x;f)dx\right| \\
&\leq \int_{\beta^{\tau}}^{\beta^{-\tau'}}\left|Cov[\widehat{f}_{\beta/a_0}(x), \widehat{f}_{\beta/a_0'}(x)] - n^{-1}\beta^{-1/2}\left(\frac{2a_0a_0'}{a_0+a_0'}\right)^{1/2}V(x;f)\right|dx + o(n^{-1}\beta^{-1/2}) \\
&= o(n^{-1}\beta^{-1/2}).
\end{aligned}$$

The case $\mathcal{S} = [0,1]$: Under Assumption A1(i) and the boundedness of f, we can see that, choosing $\tau \in (1/2, 1)$,

$$\left|\left(\int_0^{\beta^{\tau}}+\int_{1-\beta^{\tau}}^{1}\right)Cov[\widehat{f}_{\beta/a_0}(x), \widehat{f}_{\beta/a_0'}(x)]dx\right| \leq 2n^{-1}\beta^{\tau-1}C_K(a_0a_0')^{1/2}||f||_{[0,1]} = o(n^{-1}\beta^{-1/2}).$$

Also,

$$\begin{aligned}
n^{-1}\int_{\beta^{\tau}}^{1-\beta^{\tau}}\left\{\frac{\beta^{1/2}}{\sqrt{\psi^3(x)}}+1\right\}dx &\leq n^{-1}\left\{\frac{\beta^{1/2}}{\beta^{\tau}(1-\beta^{\tau})}\int_0^{1}\frac{1}{\sqrt{\psi(x)}}dx+1\right\} = o(n^{-1}\beta^{-1/2}), \\
n^{-1}\beta^{-1/2}\left(\int_0^{\beta^{\tau}}+\int_{1-\beta^{\tau}}^{1}\right)\frac{1}{\sqrt{\psi(x)}}dx &= o(n^{-1}\beta^{-1/2}).
\end{aligned}$$

Combining them with the result (i) yields

$$\begin{aligned}
&\left|\int_0^{1} Cov[\widehat{f}_{\beta/a_0}(x), \widehat{f}_{\beta/a_0'}(x)]dx - n^{-1}\beta^{-1/2}\left(\frac{2a_0a_0'}{a_0+a_0'}\right)^{1/2}\int_0^{1}V(x;f)dx\right| \\
&\leq \int_{\beta^{\tau}}^{1-\beta^{\tau}}\left|Cov[\widehat{f}_{\beta/a_0}(x), \widehat{f}_{\beta/a_0'}(x)] - n^{-1}\beta^{-1/2}\left(\frac{2a_0a_0'}{a_0+a_0'}\right)^{1/2}V(x;f)\right|dx + o(n^{-1}\beta^{-1/2}) \\
&= o(n^{-1}\beta^{-1/2}). \quad \square
\end{aligned}$$

Appendix A.2. Original estimator (without bias corrections)

In this section, we prove Theorems 1–3.

Proof of Theorem 1 (i) Under Assumption A5[p](i,ii), the $2p$-term Taylor expansion of f around $s = x$ yields

$$
\begin{aligned}
E[\widehat{f}_\beta(x)] &= \int_{\mathcal{S}} K(s; x, \beta) f(s) ds \\
&= \int_{\mathcal{S}} K(s; x, \beta) \left\{ f(x) + \sum_{j=1}^{2p} \frac{1}{j!} (s - x)^j f^{(j)}(x) \right\} ds + \mathcal{R}_\beta(x) \\
&= f(x) + \sum_{j=1}^{2p} \frac{1}{j!} \mu_j(K(\cdot; x, \beta)) f^{(j)}(x) + \mathcal{R}_\beta(x),
\end{aligned}
$$

where

$$
\mathcal{R}_\beta(x) = \frac{1}{(2p-1)!} \int_{\mathcal{S}} K(s; x, \beta)(s - x)^{2p} \int_0^1 \{f^{(2p)}(x + \theta(s - x)) - f^{(2p)}(x)\}(1 - \theta)^{2p-1} d\theta ds
$$

satisfies

$$
|\mathcal{R}_\beta(x)| \le \frac{L_{2p}}{(2p)!} \int_{\mathcal{S}} |s - x|^{2p+\eta_{2p}} K(s; x, \beta) ds \le \frac{L_{2p}}{(2p)!} \mu_{2(p+1)}^{(2p+\eta_{2p})/\{2(p+1)\}}(K(\cdot; x, \beta)).
$$

The result follows from Assumption A4[p], i.e., when $\mathcal{S} = [0, \infty)$,

$$
\mu_j(K(\cdot; x, \beta)) = \begin{cases} \displaystyle\sum_{k=\lceil j/2 \rceil}^{p} \chi_{\{k \le j\}} \zeta_{j,k}\, x^{j-k} \beta^k + O(\beta^{p+1}(1 + x)^{p-1}), & j = 1, \dots, 2p, \\ O(\beta^{p+1}(1 + x)^{p+1}), & j = 2(p + 1), \end{cases}
$$

hence,

$$
\begin{aligned}
&\sum_{j=1}^{2p} \mu_j(K(\cdot; x, \beta)) \frac{f^{(j)}(x)}{j!} + \mathcal{R}_\beta(x) \\
&= \sum_{m=1}^{p} \sum_{k=m}^{p} \left\{ \chi_{\{k \le 2m-1\}} \zeta_{2m-1,k}\, x^{2m-1-k} \beta^k \frac{f^{(2m-1)}(x)}{(2m-1)!} + \chi_{\{k \le 2m\}} \zeta_{2m,k}\, x^{2m-k} \beta^k \frac{f^{(2m)}(x)}{(2m)!} \right\} \\
&\quad + O(\beta^{p+1}(1 + x)^{p-1}) + O(\beta^{p+\eta_{2p}/2}(1 + x)^{p+\eta_{2p}/2}) \\
&= \sum_{k=1}^{p} \beta^k \sum_{m=1}^{k} \left\{ \chi_{\{k \le 2m-1\}} \zeta_{2m-1,k}\, x^{2m-1-k} \frac{f^{(2m-1)}(x)}{(2m-1)!} + \chi_{\{k \le 2m\}} \zeta_{2m,k}\, x^{2m-k} \frac{f^{(2m)}(x)}{(2m)!} \right\} \\
&\quad + O(\beta^{p+\eta_{2p}/2}(1 + x)^{p+\eta_{2p}/2}) \\
&= \sum_{k=1}^{p} \beta^k \gamma_k(x; f) + O(\beta^{p+\eta_{2p}/2}(1 + x)^{p+\eta_{2p}/2}),
\end{aligned}
$$

with $\gamma_k(x;f) = \sum_{j=1}^{2k} \chi_{\{k\le j\}}\zeta_{j,k}\, x^{j-k} f^{(j)}(x)/j!$, whereas, when $\mathcal{S}=[0,1]$, uniformly in $x\in[0,1]$,

$$\begin{aligned}
&\sum_{j=1}^{2p} \mu_j(K(\cdot;x,\beta))\frac{f^{(j)}(x)}{j!} + \mathcal{R}_\beta(x)\\
&= \sum_{m=1}^{p}\left\{\sum_{k=m}^{p} \zeta_{2m-1,k}(x)\beta^k \frac{f^{(2m-1)}(x)}{(2m-1)!} + \sum_{k=m}^{p} \zeta_{2m,k}(x)\beta^k \frac{f^{(2m)}(x)}{(2m)!}\right\} + O(\beta^{p+1}) + O(\beta^{p+\eta_{2p}/2})\\
&= \sum_{k=1}^{p} \beta^k \left\{\sum_{m=1}^{k} \zeta_{2m-1,k}(x) \frac{f^{(2m-1)}(x)}{(2m-1)!} + \sum_{m=1}^{k} \zeta_{2m,k}(x) \frac{f^{(2m)}(x)}{(2m)!}\right\} + O(\beta^{p+\eta_{2p}/2})\\
&= \sum_{k=1}^{p} \beta^k \gamma_k(x;f) + O(\beta^{p+\eta_{2p}/2}),
\end{aligned}$$

with $\gamma_k(x;f) = \sum_{j=1}^{2k} \zeta_{j,k}(x) f^{(j)}(x)/j!$.

(ii) Use Lemma A.2(i) (set $a_0 = a_0' = 1$).

(iii) Use Lemma A.1(ii) and the Borel–Cantelli lemma. $\square$

Remark A.1 Assumption A4[p] for some $p \in \mathbb{N}\backslash\{1\}$ implies that[9] Assumption A4$'$[J] holds for $J = 2, \dots, p$. Then, under Assumption A5[J](i) (which is, of course, implied by A5[p](i)), we have

$$\text{when } \mathcal{S}=[0,\infty),\ Bias[\widehat{f}_\beta(x)] = \sum_{k=1}^{J-1} \beta^k \gamma_k(x;f) + O(\beta^J(1+x)^J), \tag{A.1}$$

$$\text{when } \mathcal{S}=[0,1], \text{ uniformly in } x\in[0,1],\ Bias[\widehat{f}_\beta(x)] = \sum_{k=1}^{J-1} \beta^k \gamma_k(x;f) + O(\beta^J). \tag{A.1$'$}$$

The proofs of (A.1) and (A.1$'$) for the case $J = 2, \dots, p$ (also (2) and (2$'$) for the case $J = 1$) are easy, as follows: we have, as in Proof of Theorem 1(i),

$$E[\widehat{f}_\beta(x)] = f(x) + \sum_{j=1}^{2J-1} \frac{1}{j!}\mu_j(K(\cdot;x,\beta)) f^{(j)}(x) + \mathcal{R}^\dagger_\beta(x),$$

[9] The case $\mathcal{S}=[0,1]$ is trivial. When $\mathcal{S}=[0,\infty)$, it holds that, for $j = 1, \dots, 2J-2$ (note $J = 2, \dots, p$),

$$\begin{aligned}
\mu_j(K(\cdot;x,\beta)) &= \sum_{k=\lceil j/2\rceil}^{J-1} \chi_{\{k\le j\}}\zeta_{j,k}\, x^{j-k}\beta^k + \sum_{k=J}^{p} \chi_{\{k\le j\}}\zeta_{j,k}\, x^{j-k}\beta^k + \chi_{\{j>p\}} O(\beta^J(1+x)^{j-(p+1)})\\
&= \sum_{k=\lceil j/2\rceil}^{J-1} \chi_{\{k\le j\}}\zeta_{j,k}\, x^{j-k}\beta^k + \chi_{\{j>J-1\}} O(\beta^J(1+x)^{j-J}),
\end{aligned}$$

and that

$$\mu_{2J-1}(K(\cdot;x,\beta)) = \sum_{k=J}^{p} \chi_{\{k\le 2J-1\}}\zeta_{2J-1,k}\, x^{2J-1-k}\beta^k + \chi_{\{2J-1>p\}} O(\beta^J(1+x)^{J-2}) = O(\beta^J(1+x)^{J-1}),$$

$$\mu_{2J}(K(\cdot;x,\beta)) = \sum_{k=J}^{p} \chi_{\{k\le 2J\}}\zeta_{2J-1,k}\, x^{2J-k}\beta^k + \chi_{\{2J>p\}} O(\beta^J(1+x)^{J-1}) = O(\beta^J(1+x)^J).$$

where

$$\mathcal{R}_\beta^\dagger(x) = \frac{1}{(2J-1)!}\int_{\mathcal{S}} K(s;x,\beta)(s-x)^{2J}\int_0^1 f^{(2J)}(x+\theta(s-x))(1-\theta)^{2J-1}d\theta ds$$

satisfies

$$|\mathcal{R}_\beta^\dagger(x)| \le \frac{||f^{(2J)}||_{\mathcal{S}}}{(2J)!}\int_{\mathcal{S}}(s-x)^{2J}K(s;x,\beta)ds = \frac{||f^{(2J)}||_{\mathcal{S}}}{(2J)!}\mu_{2J}(K(\cdot;x,\beta)).$$

It follows that, when $\mathcal{S} = [0,\infty)$,

$$\begin{aligned}
&\sum_{j=1}^{2J-1}\mu_j(K(\cdot;x,\beta))\frac{f^{(j)}(x)}{j!} + \mathcal{R}_\beta^\dagger(x) \\
&= \sum_{m=1}^{J-1}\left\{\sum_{k=m}^{J-1}\chi_{\{k\le 2m-1\}}\zeta_{2m-1,k}\, x^{2m-1-k}\beta^k\frac{f^{(2m-1)}(x)}{(2m-1)!} + \sum_{k=m}^{J-1}\chi_{\{k\le 2m\}}\zeta_{2m,k}\, x^{2m-k}\beta^k\frac{f^{(2m)}(x)}{(2m)!}\right\} \\
&\quad + O(\beta^J(1+x)^{J-1}) + O(\beta^J(1+x)^J) \\
&= \sum_{k=1}^{J-1}\beta^k\left\{\sum_{m=1}^{k}\chi_{\{k\le 2m-1\}}\zeta_{2m-1,k}\, x^{2m-1-k}\frac{f^{(2m-1)}(x)}{(2m-1)!} + \sum_{m=1}^{k}\chi_{\{k\le 2m\}}\zeta_{2m,k}\, x^{2m-k}\frac{f^{(2m)}(x)}{(2m)!}\right\} \\
&\quad + O(\beta^J(1+x)^J) \\
&= \sum_{k=1}^{J-1}\beta^k\gamma_k(x;f) + O(\beta^J(1+x)^J),
\end{aligned}$$

whereas, when $\mathcal{S} = [0,1]$, uniformly in $x \in [0,1]$,

$$\begin{aligned}
&\sum_{j=1}^{2J-1}\mu_j(K(\cdot;x,\beta))\frac{f^{(j)}(x)}{j!} + \mathcal{R}_\beta^\dagger(x) \\
&= \sum_{m=1}^{J-1}\left\{\sum_{k=m}^{J-1}\zeta_{2m-1,k}(x)\beta^k\frac{f^{(2m-1)}(x)}{(2m-1)!} + \sum_{k=m}^{J-1}\zeta_{2m,k}(x)\beta^k\frac{f^{(2m)}(x)}{(2m)!}\right\} + O(\beta^J) \\
&= \sum_{k=1}^{J-1}\beta^k\left\{\sum_{m=1}^{k}\zeta_{2m-1,k}(x)\frac{f^{(2m-1)}(x)}{(2m-1)!} + \sum_{m=1}^{k}\zeta_{2m,k}(x)\frac{f^{(2m)}(x)}{(2m)!}\right\} + O(\beta^J) \\
&= \sum_{k=1}^{J-1}\beta^k\gamma_k(x;f) + O(\beta^J).
\end{aligned}$$

Proof of Theorem 2 Under Assumption A1, we have

$$\sup_{s\in\mathcal{S}}|\Delta(s;x,\beta)| \le \begin{cases} C_K'\{\beta\psi(x)\}^{-1/2} & \text{for fixed } x\in\mathcal{S}_I, \\ C_K\beta^{-1} & \text{for } x\in\mathcal{S}_B. \end{cases} \tag{A.2}$$

Also, from Theorem 1(ii),

$$\lim_{n\to\infty} n\beta^{1/2}V[\widehat{f}_\beta(x)] = V(x;f) \quad \text{for fixed } x\in\mathcal{S}_I,$$
$$\lim_{n\to\infty} n\beta V[\widehat{f}_\beta(x)] = \varsigma_{1,1}(0)f(x) \quad \text{for } x\in\mathcal{S}_B.$$

Noting that $V[\widehat{f}_\beta(x)] = V[\sum_{i=1}^n n^{-1}\Delta(X_i;x,\beta)] = \sum_{i=1}^n n^{-2}E[\Delta^2(X_i;x,\beta)]$, we have

$$\frac{\sum_{i=1}^n E\left[|n^{-1}\Delta(X_i;x,\beta)|^{2+\delta}\right]}{\left\{\sum_{i=1}^n V[n^{-1}\Delta(X_i;x,\beta)]\right\}^{1+\delta/2}} = \begin{cases} O((n\beta^{1/2})^{-\delta/2}) & \text{for fixed } x\in\mathcal{S}_I, \\ O((n\beta)^{-\delta/2}) & \text{for } x\in\mathcal{S}_B, \end{cases}$$

where $\delta > 0$ is arbitrary. Hence, Lyapunov's central limit theorem enables us to see that

$$\frac{\widehat{f}_\beta(x) - E[\widehat{f}_\beta(x)]}{\{V[\widehat{f}_\beta(x)]\}^{1/2}} \xrightarrow{d} N(0,1) \quad \text{for fixed } x \in \mathcal{S}.$$

The results follow from Slutsky's lemma. □

Proof of Theorem 3 The integrated variance approximation easily follows from Lemma A.2(ii) (set $a_0 = a_0' = 1$). It suffices to approximate the integrated squared bias.

The case $\mathcal{S} = [0,1]$: Theorem 1(i) immediately yields

$$\int_0^1 \{Bias[\widehat{f}_\beta(x)]\}^2 dx = \beta^2 \int_0^1 \gamma_1^2(x;f)dx + o(\beta^2).$$

The case $\mathcal{S} = [0,\infty)$: Under Assumptions A3 and A5[1](iii) and the boundedness of f, we have

$$\begin{aligned}
\int_{\beta^{-\tau_2}}^\infty \{Bias[\widehat{f}_\beta(x)]\}^2 dx &\le 2\int_{\beta^{-\tau_2}}^\infty \left[\left\{\int_0^\infty K(s;x,\beta)f(s)ds\right\}^2 + f^2(x)\right]dx \\
&\le 2||f||_{[0,\infty)}\left\{\int_0^\infty \int_{\beta^{-\tau_2}}^\infty K(s;x,\beta)dxf(s)ds + \int_{\beta^{-\tau_2}}^\infty f(x)dx\right\} \\
&= O(\beta^{\tau_2(k_2+1)}) = o(\beta^2).
\end{aligned}$$

Theorem 1(i) and Assumption A5[1](iii) (note that $\int_0^{\beta^{-\tau_2}} \mathcal{E}_{\beta,1}^2(x)dx = o(\beta^2)$) yield

$$\begin{aligned}
&\left|\int_0^{\beta^{-\tau_2}} \{Bias[\widehat{f}_\beta(x)]\}^2 dx - \beta^2 \int_0^{\beta^{-\tau_2}} \gamma_1^2(x;f)dx\right| \\
&\le 2\beta\left\{\int_0^\infty \gamma_1^2(x;f)dx \int_0^{\beta^{-\tau_2}} \mathcal{E}_{\beta,1}^2(x)dx\right\}^{1/2} + \int_0^{\beta^{-\tau_2}} \mathcal{E}_{\beta,1}^2(x)dx = o(\beta^2). \quad \square
\end{aligned}$$

Appendix A.3. Proof of Lemma 4

Proof of Lemma 4 Let $Z = \mathrm{diag}(z_1,\ldots,z_p)$. Then, $\prod_{j=1}^p z_j = |Z| = |\mathcal{V}(\boldsymbol{z})||Z||\mathcal{V}^{-1}(\boldsymbol{z})| = |\mathcal{V}(\boldsymbol{z})Z\mathcal{V}^{-1}(\boldsymbol{z})|$, where $|\cdot|$ denotes the determinant. Also, it is not difficult to see that

$$\begin{aligned}
\mathcal{V}(\boldsymbol{z})Z\mathcal{V}^{-1}(\boldsymbol{z}) &= \begin{pmatrix} z_1 & z_2 & \cdots & z_p \\ z_1^2 & z_2^2 & \cdots & z_p^2 \\ \vdots & \vdots & \ddots & \vdots \\ z_1^p & z_2^p & \cdots & z_p^p \end{pmatrix} \begin{pmatrix} [\mathcal{V}^{-1}(\boldsymbol{z})]_{11} & [\mathcal{V}^{-1}(\boldsymbol{z})]_{12} & \cdots & [\mathcal{V}^{-1}(\boldsymbol{z})]_{1p} \\ [\mathcal{V}^{-1}(\boldsymbol{z})]_{21} & [\mathcal{V}^{-1}(\boldsymbol{z})]_{22} & \cdots & [\mathcal{V}^{-1}(\boldsymbol{z})]_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ [\mathcal{V}^{-1}(\boldsymbol{z})]_{p1} & [\mathcal{V}^{-1}(\boldsymbol{z})]_{p2} & \cdots & [\mathcal{V}^{-1}(\boldsymbol{z})]_{pp} \end{pmatrix} \\
&= \begin{pmatrix} \sum_{j=1}^p z_j[\mathcal{V}^{-1}(\boldsymbol{z})]_{j1} & \sum_{j=1}^p z_j[\mathcal{V}^{-1}(\boldsymbol{z})]_{j2} & \cdots & \sum_{j=1}^p z_j[\mathcal{V}^{-1}(\boldsymbol{z})]_{jp} \\ \sum_{j=1}^p z_j^2[\mathcal{V}^{-1}(\boldsymbol{z})]_{j1} & \sum_{j=1}^p z_j^2[\mathcal{V}^{-1}(\boldsymbol{z})]_{j2} & \cdots & \sum_{j=1}^p z_j^2[\mathcal{V}^{-1}(\boldsymbol{z})]_{jp} \\ \vdots & \vdots & \ddots & \vdots \\ \sum_{j=1}^p z_j^p[\mathcal{V}^{-1}(\boldsymbol{z})]_{j1} & \sum_{j=1}^p z_j^p[\mathcal{V}^{-1}(\boldsymbol{z})]_{j2} & \cdots & \sum_{j=1}^p z_j^p[\mathcal{V}^{-1}(\boldsymbol{z})]_{jp} \end{pmatrix} \\
&= \left(\begin{array}{c:ccc} 0_{p-1} & & \mathrm{I}_{p-1} & \\ \hdashline \sum_{j=1}^p z_j^p[\mathcal{V}^{-1}(\boldsymbol{z})]_{j1} & \sum_{j=1}^p z_j^p[\mathcal{V}^{-1}(\boldsymbol{z})]_{j2} & \cdots & \sum_{j=1}^p z_j^p[\mathcal{V}^{-1}(\boldsymbol{z})]_{jp} \end{array}\right)
\end{aligned}$$

(the last equality is a direct consequence of $\mathcal{V}(\boldsymbol{z})\mathcal{V}^{-1}(\boldsymbol{z}) = I_p$, i.e., $\sum_{j=1}^{p} z_j^{k-1}[\mathcal{V}^{-1}(\boldsymbol{z})]_{jk} = 1$ and $\sum_{j=1}^{p} z_j^{\ell-1}[\mathcal{V}^{-1}(\boldsymbol{z})]_{jk} = 0$ for $k = 1, \ldots, p$; $\ell \in \{1, \ldots, p\}\backslash\{k\}$), hence,

$$|\mathcal{V}(\boldsymbol{z})Z\mathcal{V}^{-1}(\boldsymbol{z})| = (-1)^{p-1}\sum_{j=1}^{p} z_j^p[\mathcal{V}^{-1}(\boldsymbol{z})]_{j1}. \quad \square$$

Appendix A.4. Additive estimator

In this section, we prove Theorems 5–7, with a slight modification of Proofs of Theorems 1–3.

Proof of Theorem 5 Theorem 1(i), together with (7) and (8), yields the result (i). Using

$$V[\widehat{f}_{\beta,ADD_{\boldsymbol{a}}^p}(x)] = \sum_{j=1}^{p}\sum_{j'=1}^{p} c_j(\boldsymbol{a})c_{j'}(\boldsymbol{a})Cov[\widehat{f}_{\beta/a_j}(x), \widehat{f}_{\beta/a_{j'}}(x)], \tag{A.3}$$

Lemma A.2(i) yields the result (ii). $\square$

Proof of Theorem 6 Under Assumption A1, we have

$$\sup_{s\in\mathcal{S}}|\Delta_{ADD_{\boldsymbol{a}}^p}(s;x,\beta)| \leq \begin{cases} \sum_{k=1}^{p}|c_k(\boldsymbol{a})|C_K' a_k^{1/2}\{\beta\psi(x)\}^{-1/2} & \text{for fixed } x \in \mathcal{S}_I, \\ \sum_{k=1}^{p}|c_k(\boldsymbol{a})|C_K a_k\beta^{-1} & \text{for } x \in \mathcal{S}_B, \end{cases}$$

since $\Delta_{ADD_{\boldsymbol{a}}^p}(s;x,\beta) = \sum_{k=1}^{p} c_k(\boldsymbol{a})\Delta(s;x,\beta/a_k)$ (we used (A.2)). Also, from Theorem 5(ii),

$$\lim_{n\to\infty} n\beta^{1/2}V[\widehat{f}_{\beta,ADD_{\boldsymbol{a}}^p}(x)] = \lambda_{p,\boldsymbol{a}}V(x;f) \quad \text{for fixed } x \in \mathcal{S}_I,$$
$$\lim_{n\to\infty} n\beta V[\widehat{f}_{\beta,ADD_{\boldsymbol{a}}^p}(x)] = v_{p,\boldsymbol{a}}(0)f(x) \quad \text{for } x \in \mathcal{S}_B.$$

Noting that $V[\widehat{f}_{\beta,ADD_{\boldsymbol{a}}^p}(x)] = V[\sum_{i=1}^{n} n^{-1}\Delta_{ADD_{\boldsymbol{a}}^p}(X_i;x,\beta)] = \sum_{i=1}^{n} n^{-2}E[\Delta^2_{ADD_{\boldsymbol{a}}^p}(X_i;x,\beta)]$, we have

$$\frac{\sum_{i=1}^{n} E\big[|n^{-1}\Delta_{ADD_{\boldsymbol{a}}^p}(X_i;x,\beta)|^{2+\delta}\big]}{\big\{\sum_{i=1}^{n} V[n^{-1}\Delta_{ADD_{\boldsymbol{a}}^p}(X_i;x,\beta)]\big\}^{1+\delta/2}} = \begin{cases} O((n\beta^{1/2})^{-\delta/2}) & \text{for fixed } x \in \mathcal{S}_I, \\ O((n\beta)^{-\delta/2}) & \text{for } x \in \mathcal{S}_B, \end{cases}$$

where $\delta > 0$ is arbitrary. Hence, Lyapunov's central limit theorem enables us to see that

$$\frac{\widehat{f}_{\beta,ADD_{\boldsymbol{a}}^p}(x) - E[\widehat{f}_{\beta,ADD_{\boldsymbol{a}}^p}(x)]}{\big\{V[\widehat{f}_{\beta,ADD_{\boldsymbol{a}}^p}(x)]\big\}^{1/2}} \xrightarrow{d} N(0,1) \quad \text{for fixed } x \in \mathcal{S}.$$

The results follow from Slutsky's lemma. $\square$

Proof of Theorem 7 The integrated variance approximation easily follows from (A.3) and Lemma A.2(ii). It suffices to approximate the integrated squared bias.

The case $\mathcal{S} = [0,1]$: Theorem 5(i) immediately yields

$$\int_0^1\{Bias[\widehat{f}_{\beta,ADD_{\boldsymbol{a}}^p}(x)]\}^2dx = \beta^{2p}\int_0^1 B_{p,\boldsymbol{a}}^2(x;f)dx + o(\beta^{2p}).$$

The case $\mathcal{S} = [0,\infty)$: Under Assumptions A3 and A5[p](iii) and the boundedness of f, we have

$$
\begin{aligned}
&\int_{\beta^{-\tau_{2p}}}^{\infty} \{Bias[\widehat{f}_{\beta,ADD^p_{\boldsymbol{a}}}(x)]\}^2 dx \\
&\le 2\int_{\beta^{-\tau_{2p}}}^{\infty} \left[\left\{\int_0^{\infty} K_{ADD^p_{\boldsymbol{a}}}(s;x,\beta) f(s)ds\right\}^2 + f^2(x)\right] dx \\
&\le 2||f||_{[0,\infty)} \left[\left\{\sum_{k=1}^{p} |c_k(\boldsymbol{a})|\right\} \sum_{k=1}^{p} |c_k(\boldsymbol{a})| \int_0^{\infty}\int_{\beta^{-\tau_{2p}}}^{\infty} K(s;x,\beta/a_k)dx f(s)ds + \int_{\beta^{-\tau_{2p}}}^{\infty} f(x)dx\right] \\
&= O(\beta^{\tau_{2p}(k_{2p}+1)}) = o(\beta^{2p}).
\end{aligned}
$$

Theorem 5(i) and Assumption A5[p](iii) (note that $\int_0^{\beta^{-\tau_{2p}}} \mathcal{E}^2_{\beta,ADD^p_{\boldsymbol{a}}}(x)dx = o(\beta^{2p})$) yield

$$
\begin{aligned}
&\left|\int_0^{\beta^{-\tau_{2p}}} \{Bias[\widehat{f}_{\beta,ADD^p_{\boldsymbol{a}}}(x)]\}^2 dx - \beta^{2p}\int_0^{\beta^{-\tau_{2p}}} B^2_{p,\boldsymbol{a}}(x;f)dx\right| \\
&\le 2\beta^p\left\{\int_0^{\infty} B^2_{p,\boldsymbol{a}}(x;f)dx \int_0^{\beta^{-\tau_{2p}}} \mathcal{E}^2_{\beta,ADD^p_{\boldsymbol{a}}}(x)dx\right\}^{1/2} + \int_0^{\beta^{-\tau_{2p}}} \mathcal{E}^2_{\beta,ADD^p_{\boldsymbol{a}}}(x)dx = o(\beta^{2p}). \quad \square
\end{aligned}
$$

Appendix A.5. TS-type and JF-type estimators

In this section, we will prove Theorems 8–10 and 8′–10′. For this, we prepare the stochastic expansions of the TS-type and JF-type estimators (5) and (6), together with technical lemmas.

A.5.1. Stochastic expansion of TS-type estimator and auxiliary lemmas

Write $\overline{D}_{\beta/a_k}(x) = \widehat{f}_{\beta/a_k}(x) + \epsilon/a_k - f(x)$. Whenever $f(x) > 0$, we have

$$
\begin{aligned}
\widehat{f}_{\beta,TS^p_{\boldsymbol{a}}}(x) &= f(x)\exp\left[\sum_{k=1}^{p} c_k(\boldsymbol{a})\log\left\{1 + \frac{\overline{D}_{\beta/a_k}(x)}{f(x)}\right\}\right] \\
&= f(x)\exp\left\{\mathcal{Q}_{\beta,TS^p_{\boldsymbol{a}}}(x) + \mathcal{R}_{\beta,p+1}(x)\right\} \\
&= f(x) + \sum_{i=1}^{p} \frac{f(x)}{i!}\mathcal{Q}^i_{\beta,TS^p_{\boldsymbol{a}}}(x) + \mathcal{R}_{\beta,TS^p_{\boldsymbol{a}}}(x),
\end{aligned}
$$

where

$$
\begin{aligned}
\mathcal{Q}_{\beta,TS^p_{\boldsymbol{a}}}(x) &= \sum_{j=1}^{p} \frac{(-1)^{j-1}}{jf^j(x)} \sum_{k=1}^{p} c_k(\boldsymbol{a})\overline{D}^j_{\beta/a_k}(x), \\
\mathcal{R}_{\beta,p+1}(x) &= \frac{(-1)^p}{f^{p+1}(x)} \sum_{k=1}^{p} c_k(\boldsymbol{a})\overline{D}^{p+1}_{\beta/a_k}(x) \int_0^1 \frac{(1-\theta)^p}{(1+\theta\overline{D}_{\beta/a_k}(x)/f(x))^{p+1}} d\theta, \\
\mathcal{R}_{\beta,TS^p_{\boldsymbol{a}}}(x) &= f(x)\mathcal{R}_{\beta,p+1}(x)\exp\left\{\mathcal{Q}_{\beta,TS^p_{\boldsymbol{a}}}(x)\right\}\int_0^1 \exp\{\theta\mathcal{R}_{\beta,p+1}(x)\}d\theta \\
&\quad + \frac{f(x)}{p!}\mathcal{Q}^{p+1}_{\beta,TS^p_{\boldsymbol{a}}}(x)\int_0^1 \exp\{\theta\mathcal{Q}_{\beta,TS^p_{\boldsymbol{a}}}(x)\}(1-\theta)^p d\theta.
\end{aligned}
$$

For simplicity, we use the notation

$$\mathcal{I}_{\beta,TS_{\boldsymbol{a}}^p}(x) = \sum_{J=1}^{2} \mathcal{I}_{\beta,TS_{\boldsymbol{a}}^p}^{[J]}(x)$$

with

$$\mathcal{I}_{\beta,TS_{\boldsymbol{a}}^p}^{[1]}(x) = \sum_{j=2}^{p} \frac{(-1)^{j-1}}{j f^{j-1}(x)} \sum_{k=1}^{p} c_k(\boldsymbol{a}) \overline{D}_{\beta/a_k}^j(x),$$

$$\mathcal{I}_{\beta,TS_{\boldsymbol{a}}^p}^{[2]}(x) = \sum_{i=2}^{p} \frac{f(x)}{i!} \Bigg[\sum_{j=1}^{p} \frac{(-1)^{j-1}}{j f^j(x)} \sum_{k=1}^{p} c_k(\boldsymbol{a}) \overline{D}_{\beta/a_k}^j(x) \Bigg]^i.$$

Note that $\sum_{k=1}^{p} c_k(\boldsymbol{a}) \overline{D}_{\beta/a_k}(x) = \widehat{f}_{\beta,ADD_{\boldsymbol{a}}^p}(x) - f(x)$, using (7). In summary, we have:

Lemma A.3 *When $f(x) > 0$, the stochastic expansion of $\widehat{f}_{\beta,TS_{\boldsymbol{a}}^p}(x)$ is given by*

$$\widehat{f}_{\beta,TS_{\boldsymbol{a}}^p}(x) = \widehat{f}_{\beta,ADD_{\boldsymbol{a}}^p}(x) + \mathcal{I}_{\beta,TS_{\boldsymbol{a}}^p}(x) + \mathcal{R}_{\beta,TS_{\boldsymbol{a}}^p}(x).$$

We prepare the following lemmas to prove Theorems 8–10 and 8′–10′ for the TS-type.

Lemma A.4 *Under Assumption* A1(i)*, we have*

$$0 \le \widehat{f}_{\beta,TS_{\boldsymbol{a}}^p}(x) \le \prod_{k=1}^{p} \Big\{ \Big(C_K \frac{a_k}{\beta} + \frac{\epsilon}{a_k} \Big)^{\chi_{\{c_k(\boldsymbol{a})>0\}} c_k(\boldsymbol{a})} \Big(\frac{\epsilon}{a_k} \Big)^{\chi_{\{c_k(\boldsymbol{a})<0\}} c_k(\boldsymbol{a})} \Big\} = M_{\beta,TS_{\boldsymbol{a}}^p} \quad (say).$$

Proof Use $0 \le \widehat{f}_\beta(x) \le C_K \beta^{-1}$ (see Assumption A1(i)). □

Lemma A.5 **($\mathcal{S} = [0,\infty)$)** *Given $p \in \mathbb{N}\backslash\{1\}$ and $(\iota, \iota_0) \in \widetilde{I}_{p,1}$, suppose that Assumptions* A1(i) *and* A6$[p]_{\iota_1,\iota_2}$ *hold for some constant $(\iota_1, \iota_2) \in I_{p,(\iota,\iota_0),TS}$, i.e.,*

$$0 < \iota_1 < \frac{1}{1 + 2\iota_0}, \quad \iota_2 > 1 + (\iota + \iota_0)(p-1),$$

and let $x \in \mathcal{I}_{\iota,\iota_0}[r_\beta]$.

(i) *In addition, suppose that Assumptions* A4′$[p]$ *and* A5$[p]$(i) *hold. Then,*

$$E[\mathcal{I}_{\beta,TS_{\boldsymbol{a}}^p}^{[1]}(x)] = \frac{(-1)^{p-1}\beta^p}{\prod_{k=1}^p a_k} \sum_{j=2}^{p} \frac{(-1)^{j-1}}{j f^{j-1}(x)} \sum_{\mathcal{L}_{p,j}} \prod_{m=1}^{j} \gamma_{\ell_m}(x; f)$$
$$+ O\Big(\beta^{p+1-\iota_0(p-1)}(1+x)^{p+1} + \beta^{p+\iota_2-1-\iota_0(p-1)}(1+x)^{p-1} + \beta^{-\iota_0} \sum_{k=1}^{p} V[\widehat{f}_{\beta/a_k}(x)] \Big), \tag{A.4}$$

$$E[\mathcal{I}_{\beta,TS_{\boldsymbol{a}}^p}^{[2]}(x)] = O\Big(\beta^{p+1-\iota_0 p}(1+x)^{p+1} + \beta^{-\iota_0} \sum_{k=1}^{p} V[\widehat{f}_{\beta/a_k}(x)] \Big), \tag{A.5}$$

hence,

$$E[\mathcal{I}_{\beta,TS_{\boldsymbol{a}}^p}(x)] = \frac{(-1)^{p-1}\beta^p}{\prod_{k=1}^p a_k} \sum_{j=2}^{p} \frac{(-1)^{j-1}}{j f^{j-1}(x)} \sum_{\mathcal{L}_{p,j}} \prod_{m=1}^{j} \gamma_{\ell_m}(x; f)$$
$$+ O\Big(\beta^{p+1-\iota_0 p}(1+x)^{p+1} + \beta^{p+\iota_2-1-\iota_0(p-1)}(1+x)^{p-1} + \beta^{-\iota_0} \sum_{k=1}^{p} V[\widehat{f}_{\beta/a_k}(x)] \Big).$$

(ii) *On the other hand, suppose that Assumptions* A4′[1] *and* A5′(i) *hold. Then,*

$$V[\mathcal{I}^{[1]}_{\beta,TS^p_{\boldsymbol{a}}}(x)] = O\biggl(\{\beta^{2(1-\iota_0)}(1+x)^2 + n^{-1}\beta^{-(1+2\iota_0)}\}\sum_{k=1}^{p} V[\widehat{f}_{\beta/a_k}(x)]\biggr), \tag{A.6}$$

$$V[\mathcal{I}^{[2]}_{\beta,TS^p_{\boldsymbol{a}}}(x)] = O\biggl(\{\beta^{2(1-\iota_0)}(1+x)^2 + n^{-1}\beta^{-(1+2\iota_0)}\}\sum_{k=1}^{p} V[\widehat{f}_{\beta/a_k}(x)]\biggr), \tag{A.7}$$

hence,

$$V[\mathcal{I}_{\beta,TS^p_{\boldsymbol{a}}}(x)] \le 2\sum_{J=1}^{2} V[\mathcal{I}^{[J]}_{\beta,TS^p_{\boldsymbol{a}}}(x)] = O\biggl(\{\beta^{2(1-\iota_0)}(1+x)^2 + n^{-1}\beta^{-(1+2\iota_0)}\}\sum_{k=1}^{p} V[\widehat{f}_{\beta/a_k}(x)]\biggr).$$

Also, for any $u \ge 1$,

$$E[|\mathcal{R}_{\beta,TS^p_{\boldsymbol{a}}}(x)|^u] = O\biggl(\beta^{u(p+1-\iota_0 p)}(1+x)^{u(p+1)} + \beta^{-\iota_0(2-u)}(n\beta^{1+2\iota_0})^{-\{u(p+1)-2\}/2}\sum_{k=1}^{p} V[\widehat{f}_{\beta/a_k}(x)]\biggr). \tag{A.8}$$

Lemma A.5′ ($\mathcal{S} = [0,1]$) *Given* $p \in \mathbb{N}\backslash\{1\}$, *suppose that Assumptions* A1(i) *and* A6$[p]_{\iota_1,\iota_2}$ *hold for some constant* $(\iota_1,\iota_2) \in I_{p,(0,0),TS}$, *i.e.,* $0 < \iota_1 < 1$ *and* $\iota_2 > 1$, *and let* $x \in \mathcal{I}$.

(i) *In addition, suppose that Assumptions* A4′[p] *and* A5[p](i) *hold. Then,*

$$E[\mathcal{I}^{[1]}_{\beta,TS^p_{\boldsymbol{a}}}(x)] = \frac{(-1)^{p-1}\beta^p}{\prod_{k=1}^{p} a_k}\sum_{j=2}^{p}\frac{(-1)^{j-1}}{jf^{j-1}(x)}\sum_{\mathcal{L}_{p,j}}\prod_{m=1}^{j}\gamma_{\ell_m}(x;f) + O\biggl(\beta^{p+\min(1,\iota_2-1)} + \sum_{k=1}^{p} V[\widehat{f}_{\beta/a_k}(x)]\biggr), \tag{A.4′}$$

$$E[\mathcal{I}^{[2]}_{\beta,TS^p_{\boldsymbol{a}}}(x)] = O\biggl(\beta^{p+1} + \sum_{k=1}^{p} V[\widehat{f}_{\beta/a_k}(x)]\biggr), \tag{A.5′}$$

hence,

$$E[\mathcal{I}_{\beta,TS^p_{\boldsymbol{a}}}(x)] = \frac{(-1)^{p-1}\beta^p}{\prod_{k=1}^{p} a_k}\sum_{j=2}^{p}\frac{(-1)^{j-1}}{jf^{j-1}(x)}\sum_{\mathcal{L}_{p,j}}\prod_{m=1}^{j}\gamma_{\ell_m}(x;f) + O\biggl(\beta^{p+\min(1,\iota_2-1)} + \sum_{k=1}^{p} V[\widehat{f}_{\beta/a_k}(x)]\biggr).$$

(ii) *On the other hand, suppose that Assumptions* A4′[1] *and* A5′(i) *hold. Then,*

$$V[\mathcal{I}^{[1]}_{\beta,TS^p_{\boldsymbol{a}}}(x)] = O\biggl((\beta^2 + n^{-1}\beta^{-1})\sum_{k=1}^{p} V[\widehat{f}_{\beta/a_k}(x)]\biggr), \tag{A.6′}$$

$$V[\mathcal{I}^{[2]}_{\beta,TS^p_{\boldsymbol{a}}}(x)] = O\biggl((\beta^2 + n^{-1}\beta^{-1})\sum_{k=1}^{p} V[\widehat{f}_{\beta/a_k}(x)]\biggr), \tag{A.7′}$$

hence,

$$V[\mathcal{I}_{\beta,TS^p_{\boldsymbol{a}}}(x)] \le 2\sum_{J=1}^{2} V[\mathcal{I}^{[J]}_{\beta,TS^p_{\boldsymbol{a}}}(x)] = O\biggl((\beta^2 + n^{-1}\beta^{-1})\sum_{k=1}^{p} V[\widehat{f}_{\beta/a_k}(x)]\biggr).$$

Also, for any $u \ge 1$,

$$E[|\mathcal{R}_{\beta,TS^p_{\boldsymbol{a}}}(x)|^u] = O\biggl(\beta^{u(p+1)} + (n\beta)^{-\{u(p+1)-2\}/2}\sum_{k=1}^{p} V[\widehat{f}_{\beta/a_k}(x)]\biggr). \tag{A.8′}$$

The proofs of Lemmas A.5 and A.5$'$ are in supplemental issue (Supplemental appendix to "Higher-order bias corrections for kernel type density estimators on the unit or semi-infinite interval").

A.5.2. Stochastic expansion of JF-type estimator and auxiliary lemmas

Write

$$\overline{D}^{\dagger}_{\beta}(x) = \widehat{f}_{\beta}(x) + \epsilon - f(x) \quad \text{and} \quad \overline{D}_{\beta,ADD^p_{\mathbf{a}}}(x) = \widehat{f}_{\beta,ADD^p_{\mathbf{a}}}(x) - f(x).$$

Noting that, on the event $\{\widehat{f}_{\beta,ADD^p_{\mathbf{a}}}(x) > 0\}$,

$$\log\left\{\frac{\widehat{f}_{\beta,ADD^p_{\mathbf{a}}}(x)}{\widehat{f}_{\beta}(x)+\epsilon}\right\} = \sum_{j=1}^{p} \frac{(-1)^{j-1}}{j}\left\{\frac{\widehat{f}_{\beta,ADD^p_{\mathbf{a}}}(x)}{\widehat{f}_{\beta}(x)+\epsilon} - 1\right\}^{j} - \mathcal{R}_{\beta,\mathrm{i}}(x),$$

we have

$$\begin{aligned}
\widehat{f}_{\beta,JF^p_{\mathbf{a}}}(x) &= \widehat{f}_{\beta,JF^p_{\mathbf{a}}}(x)\chi_{\{\widehat{f}_{\beta,ADD^p_{\mathbf{a}}}(x)>0\}} + \widehat{f}_{\beta,JF^p_{\mathbf{a}}}(x)\chi_{\{\widehat{f}_{\beta,ADD^p_{\mathbf{a}}}(x)\le 0\}} \\
&= \widehat{f}_{\beta,ADD^p_{\mathbf{a}}}(x)\exp\left[\frac{(-1)^p}{p}\left\{\frac{\widehat{f}_{\beta,ADD^p_{\mathbf{a}}}(x)}{\widehat{f}_{\beta}(x)+\epsilon} - 1\right\}^{p}\right]\{1+\mathcal{R}_{\beta,\mathrm{ii}}(x)\}\chi_{\{\widehat{f}_{\beta,ADD^p_{\mathbf{a}}}(x)>0\}} \\
&\quad + \widehat{f}_{\beta,JF^p_{\mathbf{a}}}(x)\chi_{\{\widehat{f}_{\beta,ADD^p_{\mathbf{a}}}(x)\le 0\}} \\
&= \widehat{f}_{\beta,ADD^p_{\mathbf{a}}}(x)\Bigg[1 + \frac{(-1)^p}{p}\left\{\frac{\widehat{f}_{\beta,ADD^p_{\mathbf{a}}}(x)}{\widehat{f}_{\beta}(x)+\epsilon} - 1\right\}^{p} + \mathcal{R}_{\beta,\mathrm{iii}}(x) \\
&\qquad + \exp\left[\frac{(-1)^p}{p}\left\{\frac{\widehat{f}_{\beta,ADD^p_{\mathbf{a}}}(x)}{\widehat{f}_{\beta}(x)+\epsilon} - 1\right\}^{p}\right]\mathcal{R}_{\beta,\mathrm{ii}}(x)\Bigg]\chi_{\{\widehat{f}_{\beta,ADD^p_{\mathbf{a}}}(x)>0\}} \\
&\quad + \widehat{f}_{\beta,JF^p_{\mathbf{a}}}(x)\chi_{\{\widehat{f}_{\beta,ADD^p_{\mathbf{a}}}(x)\le 0\}} \\
&= \widehat{f}_{\beta,ADD^p_{\mathbf{a}}}(x) + \frac{(-1)^p}{p}f(x)\left\{\frac{\widehat{f}_{\beta,ADD^p_{\mathbf{a}}}(x)}{\widehat{f}_{\beta}(x)+\epsilon} - 1\right\}^{p} + \sum_{j=1}^{3}\mathcal{R}^{[j]}_{\beta,JF^p_{\mathbf{a}}}(x) \\
&= \widehat{f}_{\beta,ADD^p_{\mathbf{a}}}(x) + \frac{(-1)^p}{pf^{p-1}(x)}\{\overline{D}_{\beta,ADD^p_{\mathbf{a}}}(x) - \overline{D}^{\dagger}_{\beta}(x)\}^p + \sum_{j=1}^{4}\mathcal{R}^{[j]}_{\beta,JF^p_{\mathbf{a}}}(x)
\end{aligned}$$

(for the last equality, we assumed $f(x) > 0$), where

$$\mathcal{R}_{\beta,\mathrm{i}}(x) = (-1)^{p+1}\left\{\frac{\widehat{f}_{\beta,ADD^p_{\mathbf{a}}}(x)}{\widehat{f}_{\beta}(x)+\epsilon} - 1\right\}^{p+1}\int_0^1\left\{1-\theta+\frac{\theta\widehat{f}_{\beta,ADD^p_{\mathbf{a}}}(x)}{\widehat{f}_{\beta}(x)+\epsilon}\right\}^{-(p+1)}(1-\theta)^p d\theta,$$

$$\mathcal{R}_{\beta,\mathrm{ii}}(x) = \mathcal{R}_{\beta,\mathrm{i}}(x)\int_0^1 \exp\{\theta\mathcal{R}_{\beta,\mathrm{i}}(x)\}d\theta,$$

$$\mathcal{R}_{\beta,\mathrm{iii}}(x) = \frac{1}{p^2}\left\{\frac{\widehat{f}_{\beta,ADD^p_{\mathbf{a}}}(x)}{\widehat{f}_{\beta}(x)+\epsilon} - 1\right\}^{2p}\int_0^1 \exp\left[\theta\frac{(-1)^p}{p}\left\{\frac{\widehat{f}_{\beta,ADD^p_{\mathbf{a}}}(x)}{\widehat{f}_{\beta}(x)+\epsilon} - 1\right\}^{p}\right](1-\theta)d\theta,$$

$$\mathcal{R}^{[1]}_{\beta,JF^p_{\mathbf{a}}}(x) = \left[\widehat{f}_{\beta,JF^p_{\mathbf{a}}}(x) - \widehat{f}_{\beta,ADD^p_{\mathbf{a}}}(x) - \frac{(-1)^p}{p}\widehat{f}_{\beta,ADD^p_{\mathbf{a}}}(x)\left\{\frac{\widehat{f}_{\beta,ADD^p_{\mathbf{a}}}(x)}{\widehat{f}_{\beta}(x)+\epsilon} - 1\right\}^{p}\right]\chi_{\{\widehat{f}_{\beta,ADD^p_{\mathbf{a}}}(x)\le 0\}},$$

$$\mathcal{R}^{[2]}_{\beta,JF^p_{\mathbf{a}}}(x) = \widehat{f}_{\beta,ADD^p_{\mathbf{a}}}(x)\left[\exp\left[\frac{(-1)^p}{p}\left\{\frac{\widehat{f}_{\beta,ADD^p_{\mathbf{a}}}(x)}{\widehat{f}_{\beta}(x)+\epsilon} - 1\right\}^{p}\right]\mathcal{R}_{\beta,\mathrm{ii}}(x) + \mathcal{R}_{\beta,\mathrm{iii}}(x)\right]\chi_{\{\widehat{f}_{\beta,ADD^p_{\mathbf{a}}}(x)>0\}},$$

$$\mathcal{R}^{[3]}_{\beta,JF^p_{\boldsymbol{a}}}(x) = \frac{(-1)^p}{p}\overline{D}_{\beta,ADD^p_{\boldsymbol{a}}}(x)\left\{\frac{\widehat{f}_{\beta,ADD^p_{\boldsymbol{a}}}(x)}{\widehat{f}_\beta(x)+\epsilon}-1\right\}^p,$$

$$\begin{aligned}\mathcal{R}^{[4]}_{\beta,JF^p_{\boldsymbol{a}}}(x) &= \frac{(-1)^p}{p}f(x)\left\{\frac{\widehat{f}_{\beta,ADD^p_{\boldsymbol{a}}}(x)}{\widehat{f}_\beta(x)+\epsilon}-1\right\}^p\left[1-\left\{1+\frac{\overline{D}^\dagger_\beta(x)}{f(x)}\right\}^p\right]\\ &= \frac{(-1)^{p-1}}{p}f(x)\left\{\frac{\widehat{f}_{\beta,ADD^p_{\boldsymbol{a}}}(x)}{\widehat{f}_\beta(x)+\epsilon}-1\right\}^p\sum_{j=1}^p {}_pC_j\left\{\frac{\overline{D}^\dagger_\beta(x)}{f(x)}\right\}^j.\end{aligned}$$

For simplicity, we use the notations

$$\mathcal{I}_{\beta,JF^p_{\boldsymbol{a}}}(x) = \frac{(-1)^p}{pf^{p-1}(x)}\{\overline{D}_{\beta,ADD^p_{\boldsymbol{a}}}(x)-\overline{D}^\dagger_\beta(x)\}^p \quad\text{and}\quad \mathcal{R}_{\beta,JF^p_{\boldsymbol{a}}}(x) = \sum_{j=1}^4 \mathcal{R}^{[j]}_{\beta,JF^p_{\boldsymbol{a}}}(x).$$

In summary, we have:

Lemma A.6 *When $f(x)>0$, the stochastic expansion of $\widehat{f}_{\beta,JF^p_{\boldsymbol{a}}}(x)$ is given by*

$$\widehat{f}_{\beta,JF^p_{\boldsymbol{a}}}(x) = \widehat{f}_{\beta,ADD^p_{\boldsymbol{a}}}(x)+\mathcal{I}_{\beta,JF^p_{\boldsymbol{a}}}(x)+\mathcal{R}_{\beta,JF^p_{\boldsymbol{a}}}(x).$$

We prepare the following lemmas to prove Theorems 8–10 and $8'$–$10'$ for the JF-type.

Lemma A.7 *Under Assumption* A1(i), *we have*

$$0 \le \widehat{f}_{\beta,JF^p_{\boldsymbol{a}}}(x) \le M_{\beta,JF^p_{\boldsymbol{a}}},$$

where

$$M_{\beta,JF^p_{\boldsymbol{a}}} = \begin{cases}(C_K\beta^{-1}+\epsilon)\exp\{c_1(\boldsymbol{a})\}, & p=2 \text{ and } 0<a_2<a_1=1,\\ (C_K\beta^{-1}+\epsilon)\exp\left\{\sum_{k=1}^p |c_k(\boldsymbol{a})|C_K a_k(\beta\epsilon)^{-1}+1\right\}, & p=2 \text{ and } 1=a_1<a_2,\\ (C_K\beta^{-1}+\epsilon)\exp\left[\frac{c_p+1}{2}\left\{\sum_{k=1}^p |c_k(\boldsymbol{a})|C_K a_k(\beta\epsilon)^{-1}+1\right\}^{c_p}\right], & p>2,\end{cases}$$

with

$$c_p = \begin{cases}p-1, & p(>2) \text{ is even},\\ p-2, & p(>2) \text{ is odd}.\end{cases}$$

Proof Use $0\le\widehat{f}_\beta(x)\le C_K\beta^{-1}$ (see Assumption A1(i)) to bound

$$0 \le \widehat{f}_{\beta,JF^p_{\boldsymbol{a}}}(x) \le \{\widehat{f}_\beta(x)+\epsilon\}\exp\left[\sum_{j=1}^{p-1}\chi_{\{j \text{ is odd}\}}\frac{1}{j}\left\{\frac{|\widehat{f}_{\beta,ADD^p_{\boldsymbol{a}}}(x)|}{\widehat{f}_\beta(x)+\epsilon}+1\right\}^j\right].$$

Exceptionally, if $p=2$ and $0<a_2<a_1=1$, then, $c_1(\boldsymbol{a})>0$ and $c_2(\boldsymbol{a})<0$, hence,

$$0 \le \widehat{f}_{\beta,JF^p_{\boldsymbol{a}}}(x) \le \{\widehat{f}_\beta(x)+\epsilon\}\exp\{c_1(\boldsymbol{a})\}. \quad \square$$

Lemma A.8 **($\mathcal{S}=[0,\infty)$)** *Given $p\in\mathbb{N}\backslash\{1\}$ and $(\iota,\iota_0)\in\widetilde{I}_{p,1}$, suppose that Assumptions* A1(i) *and* A6$[p]_{\iota_1,\iota_2}$ *hold for some constant $(\iota_1,\iota_2)\in I_{p,(\iota,\iota_0),JF}$, i.e.,*

$$0<\iota_1<\frac{1}{1+2\iota_0+c_{p,JF}(1+\iota_2)},\quad \iota_2>1+(\iota+\iota_0)(p-1),$$

and let $x\in\mathcal{I}_{\iota,\iota_0}[r_\beta]$.

(i) *In addition, suppose that Assumptions* A4$'$[2] *and* A5[2](i) *hold. Then,*

$$\begin{aligned}&E[\mathcal{I}_{\beta,JF^p_{\boldsymbol{a}}}(x)]\\&=\beta^p\frac{\gamma_1^p(x;f)}{pf^{p-1}(x)}+O\bigg(\beta^{p+1-\iota_0(p-1)}(1+x)^{p+1}+\beta^{p+\iota_2-1-\iota_0(p-1)}(1+x)^{p-1}+\beta^{-\iota_0}\sum_{k=1}^{p}V[\widehat{f}_{\beta/a_k}(x)]\bigg).\end{aligned}\tag{A.9}$$

(ii) *On the other hand, suppose that Assumptions* A4$'$[1] *and* A5$'$(i) *hold. Then,*

$$V\big[\mathcal{I}_{\beta,JF^p_{\boldsymbol{a}}}(x)\big]=O\bigg(\big\{\beta^{2(1-\iota_0)}(1+x)^2+n^{-1}\beta^{-(1+2\iota_0)}\big\}\sum_{k=1}^{p}V[\widehat{f}_{\beta/a_k}(x)]\bigg).\tag{A.10}$$

Also, for any $u\geq 1$,

$$\begin{aligned}&E\big[\big|\mathcal{R}_{\beta,JF^p_{\boldsymbol{a}}}(x)\big|^u\\&=O\bigg(\beta^{u(p+1-\iota_0 p)}(1+x)^{u(p+1)}+\beta^{-\iota_0(2-u)}\big(n\beta^{1+2\iota_0}\big)^{-\{u(p+1)-2\}/2}\sum_{k=1}^{p}V[\widehat{f}_{\beta/a_k}(x)]\bigg).\end{aligned}\tag{A.11}$$

Lemma A.8$'$ **($\mathcal{S}=[0,1]$)** *Given $p\in\mathbb{N}\backslash\{1\}$, suppose that Assumptions* A1(i) *and* A6$[p]_{\iota_1,\iota_2}$ *hold for some constant $(\iota_1,\iota_2)\in I_{p,(0,0),JF}$, i.e.,*

$$0<\iota_1<\frac{1}{1+c_{p,JF}(1+\iota_2)},\quad \iota_2>1,$$

and let $x\in\mathcal{I}$.

(i) *In addition, suppose that Assumptions* A4$'$[2] *and* A5[2](i) *hold. Then,*

$$E[\mathcal{I}_{\beta,JF^p_{\boldsymbol{a}}}(x)]=\beta^p\frac{\gamma_1^p(x;f)}{pf^{p-1}(x)}+O\bigg(\beta^{p+\min(1,\iota_2-1)}+\sum_{k=1}^{p}V[\widehat{f}_{\beta/a_k}(x)]\bigg).\tag{A.9$'$}$$

(ii) *On the other hand, suppose that Assumptions* A4$'$[1] *and* A5$'$(i) *hold. Then,*

$$V\big[\mathcal{I}_{\beta,JF^p_{\boldsymbol{a}}}(x)\big]=O\bigg((\beta^2+n^{-1}\beta^{-1})\sum_{k=1}^{p}V[\widehat{f}_{\beta/a_k}(x)]\bigg).\tag{A.10$'$}$$

Also, for any $u\geq 1$,

$$E\big[\big|\mathcal{R}_{\beta,JF^p_{\boldsymbol{a}}}(x)\big|^u\big]=O\bigg(\beta^{u(p+1)}+(n\beta)^{-\{u(p+1)-2\}/2}\sum_{k=1}^{p}V[\widehat{f}_{\beta/a_k}(x)]\bigg).\tag{A.11$'$}$$

The proofs of Lemmas A.8 and A.8$'$ are in supplemental issue (Supplemental appendix to "Higher-order bias corrections for kernel type density estimators on the unit or semi-infinite interval").

A.5.3. Proofs of Theorems 8–10 and 8′–10′

Assuming $f(x) > 0$, Lemma A.3 (or A.6) yields

$$
\begin{aligned}
E\big[\widehat{f}_{\beta,\#^p_{\boldsymbol{a}}}(x)\big] &= E\big[\widehat{f}_{\beta,ADD^p_{\boldsymbol{a}}}(x)\big] + E\big[\mathcal{I}_{\beta,\#^p_{\boldsymbol{a}}}(x)\big] + E\big[\mathcal{R}_{\beta,\#^p_{\boldsymbol{a}}}(x)\big], && \text{(A.12)}\\
V\big[\widehat{f}_{\beta,\#^p_{\boldsymbol{a}}}(x)\big] &= V\big[\widehat{f}_{\beta,ADD^p_{\boldsymbol{a}}}(x)\big] + V\big[\mathcal{I}_{\beta,\#^p_{\boldsymbol{a}}}(x) + \mathcal{R}_{\beta,\#^p_{\boldsymbol{a}}}(x)\big]\\
&\qquad +2Cov\big[\widehat{f}_{\beta,ADD^p_{\boldsymbol{a}}}(x), \mathcal{I}_{\beta,\#^p_{\boldsymbol{a}}}(x) + \mathcal{R}_{\beta,\#^p_{\boldsymbol{a}}}(x)\big], && \text{(A.13)}
\end{aligned}
$$

where

$$
\begin{aligned}
V\big[\mathcal{I}_{\beta,\#^p_{\boldsymbol{a}}}(x) + \mathcal{R}_{\beta,\#^p_{\boldsymbol{a}}}(x)\big] &\le 2\big\{V\big[\mathcal{I}_{\beta,\#^p_{\boldsymbol{a}}}(x)\big] + V\big[\mathcal{R}_{\beta,\#^p_{\boldsymbol{a}}}(x)\big]\big\}\\
&\le 2\big\{V\big[\mathcal{I}_{\beta,\#^p_{\boldsymbol{a}}}(x)\big] + E\big[\mathcal{R}^2_{\beta,\#^p_{\boldsymbol{a}}}(x)\big]\big\},\\
\big|Cov\big[\widehat{f}_{\beta,ADD^p_{\boldsymbol{a}}}(x), \mathcal{I}_{\beta,\#^p_{\boldsymbol{a}}}(x) + \mathcal{R}_{\beta,\#^p_{\boldsymbol{a}}}(x)\big]\big| &\le \big\{V\big[\widehat{f}_{\beta,ADD^p_{\boldsymbol{a}}}(x)\big]V\big[\mathcal{I}_{\beta,\#^p_{\boldsymbol{a}}}(x) + \mathcal{R}_{\beta,\#^p_{\boldsymbol{a}}}(x)\big]\big\}^{1/2}
\end{aligned}
$$

with $V\big[\widehat{f}_{\beta,ADD^p_{\boldsymbol{a}}}(x)\big] = O\big(\sum_{k=1}^p V[\widehat{f}_{\beta/a_k}(x)]\big)$.

We are ready to prove Theorems 8–10 and 8′–10′.

Proof of Theorem 8 (i) Given $p \in \mathbb{N}\backslash\{1\}$ and $(\iota, \iota_0) \in \widetilde{I}_{p,\eta_{2p}} (\subset \widetilde{I}_{p,1})$, where $\eta_{2p} \in (0, 1]$ is given in Assumption A5[p](ii), suppose that Assumption A6$[p]_{\iota_1,\iota_2}$ holds for some constant $(\iota_1, \iota_2) \in I_{p,(\iota,\iota_0),\#}$, and let $x \in \mathcal{I}_{\iota,\iota_0}[r_\beta]$. The bias follows from (A.12), Theorem 5(i), and Lemmas A.5 (or A.8).
(ii) Given $p \in \mathbb{N}\backslash\{1\}$ and $(\iota, \iota_0) \in \widetilde{I}_{p,1}$, suppose that Assumption A6$[p]_{\iota_1,\iota_2}$ holds for some constant $(\iota_1, \iota_2) \in I_{p,(\iota,\iota_0),\#}$, and let $x \in \mathcal{I}_{\iota,\iota_0}[r_\beta]$; note that $n^{-1}\beta^{-(1+2\iota_0)} \propto n^{-1+\iota_1(1+2\iota_0)} = o(1)$, and that $r_\beta = O(\beta^{-\iota})$ implies $\beta^{1-\iota_0 p}(1+r_\beta)^{p+1} = O(\beta^{1-\iota_0 p-(p+1)\iota}) = o(1)$. The variance follows from (A.13) and Lemma A.5(ii) (or A.8(ii)), i.e.,

$$
\begin{aligned}
&V\big[\mathcal{I}_{\beta,\#^p_{\boldsymbol{a}}}(x) + \mathcal{R}_{\beta,\#^p_{\boldsymbol{a}}}(x)\big]\\
&= O\bigg(\beta^{2(p+1-\iota_0 p)}(1+x)^{2(p+1)} + \{\beta^{2(1-\iota_0)}(1+x)^2 + n^{-1}\beta^{-(1+2\iota_0)}\}\sum_{k=1}^p V[\widehat{f}_{\beta/a_k}(x)]\bigg) \qquad \text{(A.14)}
\end{aligned}
$$

and

$$
\begin{aligned}
&\big|Cov\big[\widehat{f}_{\beta,ADD^p_{\boldsymbol{a}}}(x), \mathcal{I}_{\beta,\#^p_{\boldsymbol{a}}}(x) + \mathcal{R}_{\beta,\#^p_{\boldsymbol{a}}}(x)\big]\big|\\
&= O\bigg(\beta^{p+1-\iota_0 p}(1+x)^{p+1}\bigg(\sum_{k=1}^p V[\widehat{f}_{\beta/a_k}(x)]\bigg)^{1/2} + \{\beta^{1-\iota_0}(1+x) + n^{-1/2}\beta^{-(1/2+\iota_0)}\}\sum_{k=1}^p V[\widehat{f}_{\beta/a_k}(x)]\bigg)\\
&= O\bigg(\beta^{2p+1-\iota_0 p}(1+x)^{p+1} + \{\beta^{1-\iota_0 p}(1+x)^{p+1} + n^{-1/2}\beta^{-(1/2+\iota_0)}\}\sum_{k=1}^p V[\widehat{f}_{\beta/a_k}(x)]\bigg). \quad \square
\end{aligned}
$$

Proof of Theorem 9 Recall Lemma A.3 (or A.6). Under Assumption A6$[p]_{\iota_1,\iota_2}$ for some constant $(\iota_1, \iota_2) \in I_{p,2,\#}$ for $x \in \mathcal{S}_I$ or $(\iota_1, \iota_2) \in I_{p,1,\#}$ for $x \in \mathcal{S}_B$, we have

$$\begin{aligned}
&(n\beta^{1/2})^{1/2}\{\widehat{f}_{\beta,\#^p_{\boldsymbol{a}}}(x) - E[\widehat{f}_{\beta,\#^p_{\boldsymbol{a}}}(x)]\} \\
&= (n\beta^{1/2})^{1/2}\{\widehat{f}_{\beta,ADD^p_{\boldsymbol{a}}}(x) - E[\widehat{f}_{\beta,ADD^p_{\boldsymbol{a}}}(x)]\} + o_p(1) \quad \text{for fixed } x \in \mathcal{I}_{0,0}[r] \cap \mathcal{S}_I, \\
&(n\beta)^{1/2}\{\widehat{f}_{\beta,\#^p_{\boldsymbol{a}}}(x) - E[\widehat{f}_{\beta,\#^p_{\boldsymbol{a}}}(x)]\} \\
&= (n\beta)^{1/2}\{\widehat{f}_{\beta,ADD^p_{\boldsymbol{a}}}(x) - E[\widehat{f}_{\beta,ADD^p_{\boldsymbol{a}}}(x)]\} + o_p(1) \quad \text{for } x \in \mathcal{I}_{0,0}[r] \cap \mathcal{S}_B,
\end{aligned}$$

since, from (A.14) (set $(\iota, \iota_0) = (0, 0)$ and $r_\beta \equiv r$),

$$\begin{aligned}
n\beta^{1/2}V[\mathcal{I}_{\beta,\#^p_{\boldsymbol{a}}}(x) + \mathcal{R}_{\beta,\#^p_{\boldsymbol{a}}}(x)] &= o(1) \quad \text{for fixed } x \in \mathcal{I}_{0,0}[r] \cap \mathcal{S}_I, \\
n\beta V[\mathcal{I}_{\beta,\#^p_{\boldsymbol{a}}}(x) + \mathcal{R}_{\beta,\#^p_{\boldsymbol{a}}}(x)] &= o(1) \quad \text{for } x \in \mathcal{I}_{0,0}[r] \cap \mathcal{S}_B.
\end{aligned}$$

This, together with Theorem 6, yields the results. □

Proof of Theorem 10 Using Lemma A.4 (or A.7), we have

$$\int_{r_\beta}^\infty w(x)E[\{\widehat{f}_{\beta,\#^p_{\boldsymbol{a}}}(x) - f(x)\}^2]dx \leq (M_{\beta,\#^p_{\boldsymbol{a}}} + ||f||_{[0,\infty)})^2 \int_{r_\beta}^\infty w(x)dx = o(\beta^{2p}),$$

hence,

$$MISE[\widehat{f}_{\beta,\#^p_{\boldsymbol{a}}}; w] = \int_0^{r_\beta} w(x)\big[\{Bias[\widehat{f}_{\beta,\#^p_{\boldsymbol{a}}}(x)]\}^2 + V[\widehat{f}_{\beta,\#^p_{\boldsymbol{a}}}(x)]\big]dx + o(\beta^{2p}).$$

Given $p \in \mathbb{N}\backslash\{1\}$ and $(\iota, \iota_0) \in \widetilde{I}_{p,\eta_{2p}} (\subset \widetilde{I}_{p,1})$, where $\eta_{2p} \in (0, 1]$ is given in Assumption A5[p](ii), suppose that Assumption A6[p]$_{\iota_1,\iota_2}$ holds for some constant $(\iota_1, \iota_2) \in I_{p,(\iota,\iota_0),\#}$. It is easy to see that $\int_0^{r_\beta} w(x)\mathcal{E}^2_{\beta,\#^p_{\boldsymbol{a}}}(x)dx = o(\beta^{2p} + n^{-1}\beta^{-1/2})$, since, for $x \in [0, r_\beta]$,

$$w(x)\mathcal{E}^2_{\beta,\#^p_{\boldsymbol{a}}}(x) = O\left(\beta^{2p}\omega^2_\beta(r_\beta)w(x) + n^{-1}\beta^{-(1+2\iota_0)}\sum_{k=1}^p V[\widehat{f}_{\beta/a_k}(x)]\right)$$

(we used $\omega_\beta(r_\beta) = o(1)$ and $n^{-1}\beta^{-(1+2\iota_0)} \propto n^{-1+\iota_1(1+2\iota_0)} = o(1)$). Then, Theorem 8(i) yields

$$\begin{aligned}
&\left|\int_0^{r_\beta} w(x)\{Bias[\widehat{f}_{\beta,\#^p_{\boldsymbol{a}}}(x)]\}^2dx - \beta^{2p}\int_0^\infty w(x)B^2_{\#^p_{\boldsymbol{a}}}(x; f)dx\right| \\
&\leq 2\beta^p\left\{\int_0^\infty w(x)B^2_{\#^p_{\boldsymbol{a}}}(x; f)dx\int_0^{r_\beta} w(x)\mathcal{E}^2_{\beta,\#^p_{\boldsymbol{a}}}(x)dx\right\}^{1/2} + \int_0^{r_\beta} w(x)\mathcal{E}^2_{\beta,\#^p_{\boldsymbol{a}}}(x)dx \\
&\quad + \beta^{2p}\int_{r_\beta}^\infty w(x)B^2_{\#^p_{\boldsymbol{a}}}(x; f)dx \\
&= o(\beta^{2p} + n^{-1}\beta^{-1/2}),
\end{aligned}$$

whereas, Theorem 8(ii) yields

$$\begin{aligned}
&\int_0^{r_\beta} w(x)V[\widehat{f}_{\beta,\#^p_{\boldsymbol{a}}}(x)]dx \\
&= \int_0^{r_\beta} w(x)V[\widehat{f}_{\beta,ADD^p_{\boldsymbol{a}}}(x)]dx + o(\beta^{2p})\int_0^\infty w(x)dx + o(1)\int_0^\infty \sum_{k=1}^p V[\widehat{f}_{\beta/a_k}(x)]dx \\
&= \int_0^{r_\beta} w(x)V[\widehat{f}_{\beta,ADD^p_{\boldsymbol{a}}}(x)]dx + o(\beta^{2p} + n^{-1}\beta^{-1/2}).
\end{aligned}$$

With a slight modification of Proof of Lemma A.2(ii) (recall (A.3)), we can show that, choosing $\tau \in (1/2, 1)$,

$$\begin{aligned}
&\left|\int_0^{r_\beta} w(x)V[\widehat{f}_{\beta,ADD^p_{\boldsymbol{a}}}(x)]dx - n^{-1}\beta^{-1/2}\lambda_{p,\boldsymbol{a}}\int_0^\infty w(x)V(x;f)dx\right| \\
&\le ||w||_{[0,\infty)}\sum_{j=1}^p\sum_{j'=1}^p |c_j(\boldsymbol{a})||c_{j'}(\boldsymbol{a})| \\
&\qquad\qquad\times\left[\left|\int_0^{\beta^\tau} Cov[\widehat{f}_{\beta/a_j}(x), \widehat{f}_{\beta/a'_j}(x)]dx\right|\right. \\
&\qquad\qquad\left. + \int_{\beta^\tau}^{r_\beta}\left|Cov[\widehat{f}_{\beta/a_j}(x), \widehat{f}_{\beta/a'_j}(x)] - n^{-1}\beta^{-1/2}\Big(\frac{2a_ja_{j'}}{a_j+a_{j'}}\Big)^{1/2}V(x;f)\right|dx\right] \\
&\quad + n^{-1}\beta^{-1/2}\lambda_{p,\boldsymbol{a}}\frac{||f||_{[0,\infty)}}{2\sqrt{\pi}}\left\{||w||_{[0,\infty)}\int_0^{\beta^\tau}\frac{1}{\sqrt{\psi(x)}}dx + \frac{1}{\sqrt{\psi(r_\beta)}}\int_{r_\beta}^\infty w(x)dx\right\} \\
&= o(n^{-1}\beta^{-1/2}),
\end{aligned}$$

using

$$\begin{aligned}
n^{-1}\int_{\beta^\tau}^{r_\beta} w(x)\left\{\beta^{1/2}\frac{f(x)}{\sqrt{\psi^3(x)}}+1\right\}dx &\le n^{-1}\left\{||w||_{[0,\infty)}\,\beta^{1/2-\tau}\int_0^\infty\frac{f(x)}{\sqrt{\psi(x)}}dx + \int_0^\infty w(x)dx\right\} \\
&= o(n^{-1}\beta^{-1/2}). \quad \square
\end{aligned}$$

Proof of Theorem 8′ Given $p \in \mathbb{N}\backslash\{1\}$, suppose that Assumption A6$[p]_{\iota_1,\iota_2}$ holds for some constant $(\iota_1, \iota_2) \in I_{p,(0,0),\#}$, and let $x \in \mathcal{I}$.
(i) The bias follows from (A.12), Theorem 5(i), and Lemmas A.5′ (or A.8′).
(ii) Noting that $n^{-1}\beta^{-1} \propto n^{-1+\iota_1} = o(1)$, the variance follows from (A.13) and Lemma A.5′(ii) (or A.8′(ii)), i.e.,

$$V\big[\mathcal{I}_{\beta,\#^p_{\boldsymbol{a}}}(x) + \mathcal{R}_{\beta,\#^p_{\boldsymbol{a}}}(x)\big] = O\bigg(\beta^{2(p+1)} + (\beta^2 + n^{-1}\beta^{-1})\sum_{k=1}^p V[\widehat{f}_{\beta/a_k}(x)]\bigg) \qquad \text{(A.14′)}$$

and

$$\begin{aligned}
&\big|Cov\big[\widehat{f}_{\beta,ADD^p_{\boldsymbol{a}}}(x), \mathcal{I}_{\beta,\#^p_{\boldsymbol{a}}}(x) + \mathcal{R}_{\beta,\#^p_{\boldsymbol{a}}}(x)\big]\big| \\
&= O\bigg(\beta^{p+1}\Big(\sum_{k=1}^p V[\widehat{f}_{\beta/a_k}(x)]\Big)^{1/2} + (\beta + n^{-1/2}\beta^{-1/2})\sum_{k=1}^p V[\widehat{f}_{\beta/a_k}(x)]\bigg) \\
&= O\bigg(\beta^{2p+1} + (\beta + n^{-1/2}\beta^{-1/2})\sum_{k=1}^p V[\widehat{f}_{\beta/a_k}(x)]\bigg). \quad \square
\end{aligned}$$

Proof of Theorem 9′ Recall Lemma A.3 (or A.6). Under Assumption A6$[p]_{\iota_1,\iota_2}$ for some constant $(\iota_1, \iota_2) \in I_{p,2,\#}$ for $x \in \mathcal{S}_I$ or $(\iota_1, \iota_2) \in I_{p,1,\#}$ for $x \in \mathcal{S}_B$, we have

$$\begin{aligned}
&(n\beta^{1/2})^{1/2}\{\widehat{f}_{\beta,\#^p_{\boldsymbol{a}}}(x) - E[\widehat{f}_{\beta,\#^p_{\boldsymbol{a}}}(x)]\} \\
&= (n\beta^{1/2})^{1/2}\{\widehat{f}_{\beta,ADD^p_{\boldsymbol{a}}}(x) - E[\widehat{f}_{\beta,ADD^p_{\boldsymbol{a}}}(x)]\} + o_p(1) \quad \text{for fixed } x \in \mathcal{I}\cap\mathcal{S}_I,
\end{aligned}$$

$$
\begin{aligned}
&(n\beta)^{1/2}\{\widehat{f}_{\beta,\#^p_{\boldsymbol{a}}}(x) - E[\widehat{f}_{\beta,\#^p_{\boldsymbol{a}}}(x)]\} \\
&= (n\beta)^{1/2}\{\widehat{f}_{\beta,ADD^p_{\boldsymbol{a}}}(x) - E[\widehat{f}_{\beta,ADD^p_{\boldsymbol{a}}}(x)]\} + o_p(1) \quad \text{for } x \in \mathcal{I}\cap\mathcal{S}_B,
\end{aligned}
$$

since, from (A.14′),

$$
\begin{aligned}
n\beta^{1/2}V[\mathcal{I}_{\beta,\#^p_{\boldsymbol{a}}}(x) + \mathcal{R}_{\beta,\#^p_{\boldsymbol{a}}}(x)] &= o(1) \quad \text{for fixed } x \in \mathcal{I}\cap\mathcal{S}_I, \\
n\beta V[\mathcal{I}_{\beta,\#^p_{\boldsymbol{a}}}(x) + \mathcal{R}_{\beta,\#^p_{\boldsymbol{a}}}(x)] &= o(1) \quad \text{for } x \in \mathcal{I}\cap\mathcal{S}_B.
\end{aligned}
$$

This, together with Theorem 6, yields the results. □

Proof of Theorem 10′ Given $p \in \mathbb{N}\backslash\{1\}$, suppose that Assumption A6$[p]_{\iota_1,\iota_2}$ holds for some constant $(\iota_1, \iota_2) \in I_{p,(0,0),\#}$. It is easy see that $\int_0^1 \mathcal{E}^2_{\beta,\#^p_{\boldsymbol{a}}}(x)dx = o(\beta^{2p} + n^{-1}\beta^{-1/2})$, since, for $x \in [0,1]$,

$$
\mathcal{E}^2_{\beta,\#^p_{\boldsymbol{a}}}(x) = O\Bigg(\beta^{2p+\min\{\eta_{2p},2(\iota_2-1)\}} + n^{-1}\beta^{-1}\sum_{k=1}^{p} V[\widehat{f}_{\beta/a_k}(x)]\Bigg)
$$

(we used $n^{-1}\beta^{-1} \propto n^{-1+\iota_1} = o(1)$). Then, Theorem 8′(i) yields

$$
\begin{aligned}
&\left|\int_0^1 \{Bias[\widehat{f}_{\beta,\#^p_{\boldsymbol{a}}}(x)]\}^2 dx - \beta^{2p}\int_0^1 B^2_{\#^p_{\boldsymbol{a}}}(x;f)dx\right| \\
&\le 2\beta^p\left\{\int_0^1 B^2_{\#^p_{\boldsymbol{a}}}(x;f)dx \int_0^1 \mathcal{E}^2_{\beta,\#^p_{\boldsymbol{a}}}(x)dx\right\}^{1/2} + \int_0^1 \mathcal{E}^2_{\beta,\#^p_{\boldsymbol{a}}}(x)dx = o(\beta^{2p} + n^{-1}\beta^{-1/2}),
\end{aligned}
$$

whereas, Theorem 8′(ii) yields

$$
\begin{aligned}
\int_0^1 V[\widehat{f}_{\beta,\#^p_{\boldsymbol{a}}}(x)]dx &= \int_0^1 V[\widehat{f}_{\beta,ADD^p_{\boldsymbol{a}}}(x)]dx + o(\beta^{2p}) + o(1)\int_0^1 \sum_{k=1}^{p} V[\widehat{f}_{\beta/a_k}(x)]dx \\
&= \int_0^1 V[\widehat{f}_{\beta,ADD^p_{\boldsymbol{a}}}(x)]dx + o(\beta^{2p} + n^{-1}\beta^{-1/2}).
\end{aligned}
$$

With a slight modification of Proof of Lemma A.2(ii) (recall (A.3)), we can show that, choosing $\tau \in (1/2, 1)$,

$$
\begin{aligned}
&\left|\int_0^1 V[\widehat{f}_{\beta,ADD^p_{\boldsymbol{a}}}(x)]dx - n^{-1}\beta^{-1/2}\lambda_{p,\boldsymbol{a}}\int_0^1 V(x;f)dx\right| \\
&\le \sum_{j=1}^{p}\sum_{j'=1}^{p} |c_j(\boldsymbol{a})||c_{j'}(\boldsymbol{a})|\Bigg[\left|\left(\int_0^{\beta^\tau} + \int_{1-\beta^\tau}^1\right) Cov[\widehat{f}_{\beta/a_j}(x), \widehat{f}_{\beta/a'_j}(x)]dx\right| \\
&\qquad + \int_{\beta^\tau}^{1-\beta^\tau}\left|Cov[\widehat{f}_{\beta/a_j}(x), \widehat{f}_{\beta/a'_j}(x)] - n^{-1}\beta^{-1/2}\Big(\frac{2a_j a_{j'}}{a_j + a_{j'}}\Big)^{1/2} V(x;f)\right| dx\Bigg] \\
&\quad + n^{-1}\beta^{-1/2}\lambda_{p,\boldsymbol{a}}\frac{||f||_{[0,1]}}{2\sqrt{\pi}}\left(\int_0^{\beta^\tau} + \int_{1-\beta^\tau}^1\right)\frac{1}{\sqrt{\psi(x)}}dx \\
&= o(n^{-1}\beta^{-1/2}),
\end{aligned}
$$

using

$$
n^{-1}\int_{\beta^\tau}^{1-\beta^\tau}\left\{\frac{\beta^{1/2}}{\sqrt{\psi^3(x)}} + 1\right\}dx \le n^{-1}\left\{\frac{\beta^{1/2}}{\beta^\tau(1-\beta^\tau)}\int_0^1 \frac{1}{\sqrt{\psi(x)}}dx + 1\right\} = o(n^{-1}\beta^{-1/2}). \quad \square
$$

References

Amoroso, L. (1925) "Ricerche intorno alla curva dei redditi", *Annali di Matematica Pura ed Applicata*, 2, 123–159.

Birnbaum, Z. W. and Saunders, S. C. (1969) "A new family of life distributions", *Journal of Applied Probability*, 6, 319–327.

Chen, S. X. (1999) "Beta kernel estimators for density functions", *Computational Statistics and Data Analysis*, 31, 131–145.

Chen, S. X. (2000) "Probability density function estimation using gamma kernels", *Annals of the Institute of Statistical Mathematics*, 52, 471–480.

Gautschi, W. (1962) "On inverses of Vandermonde and confluent Vandermonde matrices", *Numerische Mathematik*, 4, 117–123.

Hirukawa, M. (2010) "Nonparametric multiplicative bias correction for kernel-type density estimation on the unit interval", *Computational Statistics and Data Analysis*, 54, 473–495. Correction: (2016), 95, 240–242.

Hirukawa, M. and Sakudo, M. (2014) "Nonnegative bias reduction methods for density estimation using asymmetric kernels", *Computational Statistics and Data Analysis*, 75, 112–123.

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.

Igarashi, G. (2016a) "Bias reductions for beta kernel estimation", *Journal of Nonparametric Statistics*, 28, 1–30.

Igarashi, G. (2016b) "Weighted log-normal kernel density estimation", *Communications in Statistics - Theory and Methods*, 45, 6670–6687.

Igarashi, G. and Kakizawa, Y. (2014a) "On improving convergence rate of Bernstein polynomial density estimator", *Journal of Nonparametric Statistics*, 26, 61–84.

Igarashi, G. and Kakizawa, Y. (2014b) "Re-formulation of inverse Gaussian, reciprocal inverse Gaussian, and Birnbaum–Saunders kernel estimators", *Statistics and Probability Letters*, 84, 235–246.

Igarashi, G. and Kakizawa, Y. (2015) "Bias corrections for some asymmetric kernel estimators", *Journal of Statistical Planning and Inference*, 159, 37–63.

Igarashi, G. and Kakizawa, Y. (2018a) "Generalised gamma kernel density estimation for nonnegative data and its bias reduction", *Journal of Nonparametric Statistics*, 30, 598–639.

Igarashi, G. and Kakizawa, Y. (2018b) "Limiting bias-reduced Amoroso kernel density estimators for nonnegative data", *Communications in Statistics - Theory and Methods*, 47, 4905–4937.

Igarashi, G. and Kakizawa, Y. (2018c) "Multiplicative bias correction for asymmetric kernel density estimators revisited", Department of Policy and Planning Sciences Discussion Paper Series No. 1355, University of Tsukuba.

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.

Jones, M. C. (1993) "Simple boundary correction for kernel density estimation", *Statistics and Computing*, 3, 135–146.

Jones, M. C. and Foster, P. J. (1993) "Generalized jackknifing and higher order kernels", *Journal of Nonparametric Statistics*, 3, 81–94.

Jones, M. C., Linton, O. and Nielsen, J. P. (1995) "A simple bias reduction method for density estimation", *Biometrika*, 82, 327–338.

Jørgensen, B. (1982) *Statistical Properties of the Generalized Inverse Gaussian Distribution*, New York: Springer-Verlag.

Jørgensen, B., Seshadri, V. and Whitmore, G. A. (1991) "On the mixture of the inverse Gaussian distribution with its complementary reciprocal", *Scandinavian Journal of Statistics*, 18, 77–89.

Kakizawa, Y. (2018) "Nonparametric density estimation for nonnegative data, using symmetrical-based inverse and reciprocal inverse Gaussian kernels through dual transformation", *Journal of Statistical Planning and Inference*, 193, 117–135.

Kakizawa, Y. and Igarashi, G. (2017) "Inverse gamma kernel density estimation for nonnegative data", *Journal of the Korean Statistical Society*, 46, 194–207.

Leblanc, A. (2010) "A bias-reduced approach to density estimation using Bernstein polynomials", *Journal of Nonparametric Statistics*, 22, 459–475.

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.

Marron, J. S. and Ruppert, D. (1994) "Transformations to reduce boundary bias in kernel density estimation", *Journal of the Royal Statistical Society, Series B*, 56, 653–671.

Nadaraya, E. A. (1974) "On the integral mean square error of some nonparametric estimates for the density function", *Theory of Probability and Its Applications*, 19, 133–141.

Parzen, E. (1962) "On estimation of a probability density function and mode", *The Annals of Mathematical Statistics*, 33, 1065–1076.

Rosenblatt, M. (1956) "Remarks on some nonparametric estimates of a density function", *The Annals of Mathematical Statistics*, 27, 832–837.

Saulo, H., Leiva, V., Ziegelmann, F. A. and Marchant, C. (2013) "A nonparametric method for estimating asymmetric densities based on skewed Birnbaum–Saunders distributions applied to environmental data", *Stochastic Environmental Research and Risk Assessment*, 27, 1479–1491.

Scaillet, O. (2004) "Density estimation using inverse and reciprocal inverse Gaussian kernels", *Journal of Nonparametric Statistics*, 16, 217–226.

Schucany, W. R. and Sommers, J. P. (1977) “Improvement of kernel type density estimators”, *Journal of the American Statistical Association*, 72, 420–423.

Stacy, E. W. (1962) “A generalization of the gamma distribution”, *The Annals of Mathematical Statistics*, 33, 1187–1192.

Stacy, E. W. and Mihram, G. A. (1965) “Parameter estimation for a generalized gamma distribution”, *Technometrics*, 7, 349–358.

Terrell, G. R. and Scott, D. W. (1980) “On improving convergence rates for nonnegative kernel density estimators”, *The Annals of Statistics*, 8, 1160–1163.

Tweedie, M. C. K. (1957) “Statistical properties of inverse Gaussian distributions. I.”, *The Annals of Mathematical Statistics*, 28, 362–377.

Wand, M. P. and Schucany, W. R. (1990) “Gaussian-based kernels”, *Canadian Journal of Statistics*, 18, 197–204.

Zhang, S., Karunamuni, R. J. and Jones, M. C. (1999) “An improved estimator of the density function at the boundary”, *Journal of the American Statistical Association*, 94, 1231–1241.

Zougab, N. and Adjabi, S. (2016) “Multiplicative bias correction for generalized Birnbaum–Saunders kernel density estimators and application to nonnegative heavy tailed data”, *Journal of the Korean Statistical Society*, 45, 51–63.