Multiplicative bias correction for asymmetric kernel density estimators revisited

著者	IGARASHI Gaku, KAKIZAWA Yoshihide
year	2018-08
シリーズ	Department of Policy and Planning Sciences Discussion Paper Series;no.1355
URL	http://hdl.handle.net/2241/00153036

Department of Policy and Planning Sciences

Discussion Paper Series

No.1355

Multiplicative bias correction for asymmetric kernel density estimators revisited

by

Gaku IGARASHI and Yoshihide KAKIZAWA

Aug 2018

UNIVERSITY OF TSUKUBA
Tsukuba, Ibaraki 305-8573
JAPAN

Multiplicative bias correction for asymmetric kernel density estimators revisited

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

This paper revisits multiplicative bias correction for some asymmetric kernel density estimators (KDEs) when the data is supported on $[0, \infty)^d$ or $[0, 1]^d$. The original method was introduced by Jones et al. (1995) for the standard KDE with symmetric kernel. After Hirukawa (2010) for beta KDE, there have been renewed interests for applications to the asymmetric KDEs. We stress that the variance manipulation must be performed by looking at four terms from the law of total variance/covariance, in which only one term is negligible, while other three terms contribute to the variance formula. It turns out that, even for recently developed asymmetric KDEs, the achievement of the reduced bias is available, at the expense of the constant-factor inflation of the variance. Interestingly, the same factor appears in other bias correction methods.

Keywords: multiplicative bias correction; nonparametric density estimation; boundary bias problem; asymmetric kernel.

1. Introduction

The kernel density estimator (KDE), introduced by Rosenblatt (1956), is perhaps the most popular in the context of nonparametric density estimation. Several asymptotic results using the location-scale form $K_h(\cdot - x)$, with $K_h(\cdot) = K(\cdot/h)/h$, where K is a kernel and $h > 0$ is a bandwidth, have been well established when the support $\mathcal{S}$ of the underlying density is $\mathbb{R}$. See, e.g., Silverman (1986) and Wand and Jones (1995). However, if $\mathcal{S} \neq \mathbb{R}$, the standard KDE is, in general, inconsistent, due to the bias that is $O(1)$ near the boundary. For this boundary bias problem, various remedies have been discussed in the literature; for example, renormalization, reflection, and generalized jackknifing (Jones (1993)), transformation (Marron and Ruppert (1994)), and advanced reflection (Zhang et al. (1999)).

During recent years, when $\mathcal{S} \neq \mathbb{R}$, there has been a growing interest of the development of the nonparametric density estimation using a certain asymmetric kernel. To the best of our knowledge, Silverman (1986, page 28) first mentioned the possibility of using gamma or log-normal (LN) kernel for the nonnegative data, and Chen (1999) did pioneering studies on beta KDE using a beta kernel for the data from the unit interval. Note that Chen's beta KDE is boundary bias free and nonnegative. Since then, there have been many attempts to suggest a suitable asymmetric kernel $K(\cdot; x, \beta)$, where

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

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].

x is the location where the density estimation is made, and $\beta > 0$ is a smoothing parameter. Given an iid sample $\{X_1, \ldots, X_n\}$ from the density f with the support $\mathcal{S}$, we construct an average estimator

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

which is customarily called an asymmetric KDE when $K(\cdot; x, \beta)$ is chosen to be supported on $\mathcal{S} \neq \mathbb{R}$. The following points should be distinguished between the classical KDE and the recent asymmetric KDE. Whenever $\mathcal{S} \neq \mathbb{R}$, the support of the asymmetric kernel $K(\cdot; x, \beta)$ at the location $x \in \mathcal{S}$ under consideration matchs the support $\mathcal{S}$ of the underlying density, whereas any location-scale kernel $K_h(\cdot - x)$, at the location x near the boundary, necessarily has a mass outside the support $\mathcal{S}$. Such an asymmetric KDE has often been built within a certain parametric form like (generalized) gamma density, inverse Gaussian (IG) density, reciprocal IG (RIG) density, Birnbaum–Saunders (BS) density, LN density, and so on. See, e.g., Igarashi and Kakizawa (2018a) and references cited therein.

Typically, when the support $\mathcal{S}$ of the density f to be estimated has only one boundary point $\{0\}$ (for the case $\mathcal{S} = [0, 1]$, $\{0\}$ should read $\{0, 1\}$, of course), there exist constants[1] $q, r, r' > 0$; $q \geq r \geq r'$ and functions $\gamma_1(\cdot; f)$ and $\delta(\cdot)$, independent of β, such that, for all sufficiently small $\beta > 0$,

$$Bias[\widetilde{f}_\beta(x)] = \beta^q \gamma_1(x; f) + o(\beta^q), \tag{2}$$

$$Var[\widetilde{f}_\beta(x)] = \begin{cases} n^{-1}\beta^{-r'}\delta(x)f(x) + o(n^{-1}\beta^{-r'}) & \text{for } x \in \mathcal{S}\backslash\{0\}, \\ n^{-1}\beta^{-r}\delta(0)f(0) + o(n^{-1}\beta^{-r}) & \text{for } x = 0. \end{cases} \tag{3}$$

Then, assuming that $\beta = \beta_n \searrow 0$ and $n\beta^r \to \infty$ (in this case, $n\beta^{r'} \to \infty$) as $n \to \infty$, the leading term of the mean squared error $MSE[\widetilde{f}_\beta(x)]$ (in short, asymptotic MSE (AMSE)) is given by

$$AMSE[\widetilde{f}_\beta(x)] = \begin{cases} \beta^{2q}\gamma_1^2(x; f) + n^{-1}\beta^{-r'}\delta(x)f(x) & \text{for } x \in \mathcal{S}\backslash\{0\}, \\ \beta^{2q}\gamma_1^2(0; f) + n^{-1}\beta^{-r}\delta(0)f(0) & \text{for } x = 0 \end{cases}$$

(i.e., the asymmetric KDE is pointwise consistent). Choosing

$$\beta^{opt}(x) = \begin{cases} \left[\dfrac{r'\delta(x)f(x)}{2q\gamma_1^2(x; f)} n^{-1}\right]^{1/(2q+r')} & \text{for } x \in \mathcal{S}\backslash\{0\}, \\ \left[\dfrac{r\delta(0)f(0)}{2q\gamma_1^2(0; f)} n^{-1}\right]^{1/(2q+r)} & \text{for } x = 0, \end{cases}$$

we have the optimal AMSE;

$$AMSE^{opt}[\widetilde{f}_\beta(x)] = \begin{cases} \dfrac{2q + r'}{r'}\{\gamma_1^2(x; f)\}^{r'/(2q+r')}\left[\dfrac{r'}{2q}\delta(x)f(x)n^{-1}\right]^{2q/(2q+r')} & \text{for } x \in \mathcal{S}\backslash\{0\}, \\ \dfrac{2q + r}{r}\{\gamma_1^2(0; f)\}^{r/(2q+r)}\left[\dfrac{r}{2q}\delta(0)f(0)n^{-1}\right]^{2q/(2q+r)} & \text{for } x = 0. \end{cases} \tag{4}$$

[1]The particular case of $q = r = 1$ and $r' = 1/2$ is important, in accordance with the pioneering work on the gamma KDE (Chen (2000)). However, it would not be guaranteed generally, except that one index "r" (say) could be assumed to be 1, without loss of generality.

1.1. Additive bias reduction

Improving the performance (4) is an important topic even when $\mathcal{S} \neq \mathbb{R}$. There have been recent renewed interests on bias reductions for the asymmetric KDE (1). Suppose that, in addition to (2),

$$Bias[\widetilde{f}_\beta(x)] = \sum_{j=1}^{2} \beta^{jq}\gamma_j(x; f) + o(\beta^{2q}), \tag{5}$$

for some function $\gamma_2(\cdot; f)$, independent of β. In this case, the easiest way to remove the $O(\beta^q)$ bias is perhaps an additive bias reduction due to Schucany and Sommers (1977). That is, for each constant $a \in (0,1)$, the estimator, defined by

$$\widetilde{f}_{\beta,SS_a}(x) = \frac{1}{1-a^q}\widetilde{f}_\beta(x) - \frac{a^q}{1-a^q}\widetilde{f}_{\beta/a}(x), \quad x \in \mathcal{S} \quad \text{(we call the SS-type)},$$

has the asymptotic bias, $Bias[\widetilde{f}_{\beta,SS_a}(x)] = -(\beta^{2q}/a^q)\gamma_2(x; f) + o(\beta^{2q})$. This technique was originally applied to the standard KDE, and recently to the asymmetric KDEs. See Leblanc (2010), Igarashi and Kakizawa (2014a, 2015, 2018a), and Igarashi (2016a) for the data supported on $\mathcal{S} = [0,\infty)$ or $[0,1]$. Furthermore, as discussed in Jones and Foster (1993, Example 2.1) for the standard KDE, and Igarashi and Kakizawa (2015, 2018b) and Igarashi (2016a) for the asymmetric KDEs, if the estimator (1) is differentiable with respect to β, the limiting estimator, defined by

$$\lim_{a\to 1}\widetilde{f}_{\beta,SS_a}(x) = \widetilde{f}_\beta(x) - \frac{\beta}{q}\frac{\partial}{\partial\beta}\widetilde{f}_\beta(x) = \widetilde{f}_{\beta,SS_1}(x) \quad \text{(say)}, \quad x \in \mathcal{S},$$

will have the asymptotic bias, $Bias[\widetilde{f}_{\beta,SS_1}(x)] = -\beta^{2q}\gamma_2(x; f) + o(\beta^{2q})$. Here, the SS-type estimator is written as the form (1), i.e., $\widetilde{f}_{\beta,SS_a}(x) = n^{-1}\sum_{i=1}^{n} K_{SS_a}(X_i; x, \beta)$, $x \in \mathcal{S}$, where

$$K_{SS_a}(s; x, \beta) = \begin{cases} \dfrac{1}{1-a^q}K(s; x, \beta) - \dfrac{a^q}{1-a^q}K(s; x, \beta/a), & a \in (0,1), \\ K(s; x, \beta) - \dfrac{\beta}{q}\dfrac{\partial}{\partial\beta}K(s; x, \beta), & a = 1, \end{cases}$$

so that the resulting asymmetric kernel will be interpreted as "a higher order asymmetric kernel" derived from a given asymmetric kernel $K(\cdot; x, \beta)$ only[2]. Its asymptotic variance and AMSE are found in Igarashi and Kakizawa (2015, 2018a, 2018b) and Igarashi (2016a) for the data supported on $\mathcal{S} = [0,\infty)$ or $[0,1]$ (if $a \in (0,1)$, an additional task is to approximate $Cov[\widetilde{f}_\beta(x), \widetilde{f}_{\beta/a}(x)]$). Compared to the (uncorrected) estimator $\widetilde{f}_\beta(x)$, the order of $V[\widetilde{f}_{\beta,SS_a}(x)]$ remains unchanged, but the coefficient of the leading term of $V[\widetilde{f}_{\beta,SS_a}(x)]$ increases with the factor

$$\lambda(a) = \frac{1}{(1-a)^2}\left\{1 - 2a\left(\frac{2a}{a+1}\right)^{1/2} + a^{5/2}\right\},$$

where λ is increasing for $a \in (0,1)$; $\lim_{a\to 0}\lambda(a) = 1$ and $\lim_{a\to 1}\lambda(a) = 27/16$. Interestingly, the factor $\lambda(a)$ appeared in Wand and Schucany (1990) for the (standard) Gaussian KDE ($\mathcal{S} = \mathbb{R}$). It should be remarked that, although the SS_a-type, for each $a \in (0,1]$, loses the nonnegativity by construction, the positive part estimator $\widetilde{f}^{+}_{\beta,SS_a}(x) = \max\{0, \widetilde{f}_{\beta,SS_a}(x)\}$ not only keeps the nonnegativity, but also improves the performance in the sense that $MSE[\widetilde{f}^{+}_{\beta,SS_a}(x)] \leq MSE[\widetilde{f}_{\beta,SS_a}(x)]$.

[2] For the standard KDE using the location-scale form $K_{[2]}((\cdot - x)/h)/h$ as the kernel, Jones and Foster (1993) considered an enormous variety of higher order kernels from a given 2nd order kernel $K_{[2]}$ only, by generalized jackknifing.

1.2. Nonnegative bias reductions

There have been other bias reduction methods due to Terrell and Scott (1980) and Jones and Foster (1993), that are guaranteed to be nonnegative. Technically, introducing a parameter $\epsilon \searrow 0$ to avoid the division by zero, two nonnegative bias-corrected estimators (we call the TS-type and JF-type) are, respectively, defined by, for each constant $a \in (0,1)$,

$$\widetilde{f}_{\beta,TS_a}(x) = \frac{\{\widetilde{f}_{\beta}(x)+\epsilon\}^{1/(1-a^q)}}{\{\widetilde{f}_{\beta/a}(x)+\epsilon/a^q\}^{a^q/(1-a^q)}}, \quad \widetilde{f}_{\beta,JF_a}(x) = \{\widetilde{f}_{\beta}(x)+\epsilon\}\exp\left\{\frac{\widetilde{f}_{\beta,SS_a}(x)}{\widetilde{f}_{\beta}(x)+\epsilon}-1\right\}, \quad x \in \mathcal{S}$$

(Terrell and Scott (1980) and Jones and Foster (1993) originally used $\epsilon = 0$ for the standard KDE ($\mathcal{S} = \mathbb{R}$)). By construction, we can see that, if ϵ is independent of a, then,

$$\lim_{a\to 1}\widetilde{f}_{\beta,TS_a}(x) = \lim_{a\to 1}\widetilde{f}_{\beta,JF_a}(x) = \{\widetilde{f}_{\beta}(x)+\epsilon\}\exp\left\{\frac{\widetilde{f}_{\beta,SS_1}(x)}{\widetilde{f}_{\beta}(x)+\epsilon}-1\right\}, \quad x \in \mathcal{S}$$

(hence, the TS-type is linked with the JF-type; $\widetilde{f}_{\beta,TS_1}(x) = \widetilde{f}_{\beta,JF_1}(x)$). As demonstrated in Igarashi (2016a) and Igarashi and Kakizawa (2018a, 2018b) (note that a careful analysis of the remainder term was carried out there), the TS-type and JF-type have the stochastic expansions

$$\widetilde{f}_{\beta,\#_a}(x) \approx \widetilde{f}_{\beta,SS_a}(x) + \frac{\{\widetilde{f}_{\beta}(x)-\widetilde{f}_{\beta,SS_a}(x)+\epsilon\}^2}{2a^{\chi\{\#=TS\}}f(x)}, \quad \# = TS, JF, \quad \text{if } f(x) > 0,$$

which help understanding of asymptotic properties of the TS_a-type and JF_a-type similar to those of the SS_a-type, except for an additional bias term coming from $E[\{\widetilde{f}_{\beta}(x)-\widetilde{f}_{\beta,SS_a}(x)+\epsilon\}^2/f(x)]/(2a^{\chi\{\#=TS\}})$. Consequently, the TS_a-type and JF_a-type bias reductions work even for the asymmetric KDEs, at expense of the constant-factor inflation; $\lambda(a)$ of the variance, as in the SS_a-type bias reduction.

On the other hand, in the seminal paper, Jones et al. (1995) developed another multiplicative (hence, nonnegative) bias reduction method (we call the JLN-type), focusing on the standard KDE ($\mathcal{S} = \mathbb{R}$). A possible application to asymmetric KDEs was mentioned in Hagmann and Scaillet (2007). Hirukawa (2010, correction 2016) showed that the asymptotic variance of the JLN-type bias-corrected beta KDE is equivalent to that of the (uncorrected) beta KDE, despite the variance inflation $\lambda(a)$ for the TS_a-type bias-corrected beta KDE (the TS_1-type or JF_a-type bias-corrected beta KDE was additionally studied by Igarashi (2016a)). However[3], as will be revisited in this paper, it turns out that Hirukawa's asymptotic variance miss two terms, and, to make matters worse, Hirukawa's original incorrect proof may lead to similar incorrect conclusions in his companion papers (Hirukawa and Sakudo (2014, 2015)) and subsequent papers (Funke and Kawka (2015), Zougab and Adjabi (2016), and Zougab et al. (2018)). To be exact, as Jones et al. (1995) did for the standard KDE ($\mathcal{S} = \mathbb{R}$), we need, for the variance manipulation, to look at four terms from the law of total variance/covariance, in which only one term is negligible, while other three terms contribute to the final result (i.e., the aforementioned authors's asymptotic variances would be incorrectly asserted in common).

[3]The first author, in his master thesis (March, 2012; in Japanese), realized that the asymptotic variances of the JLN-type bias-corrected beta/gamma/Bernstein KDEs have the inflation factor $\lambda(1/2)$. However, his proof at that time was formal without a rigorous analysis of the remainder term of the stochastic expansion.

1.3. Overview of the paper

The contribution of this paper is three fold. First, we are changing the original JLN-type definition involved the division by $\widetilde{f}_\beta(X_i)$, $i = 1, \ldots, n$, as follows: In line with Jones et al. (1995) (they originally used $\epsilon = 0$ for the standard KDE ($\mathcal{S} = \mathbb{R}$)), we define the multiplicative bias-corrected estimator by

$$\widetilde{f}_{\beta,JLN}(x) = \{\widetilde{f}_\beta(x) + \epsilon\}\frac{1}{n}\sum_{i=1}^{n}\frac{K(X_i; x, \beta)}{\widetilde{f}_\beta(X_i) + \epsilon}, \quad x \in \mathcal{S}, \tag{6}$$

where the introduction of $\epsilon \searrow 0$ avoids dividing by zero. By construction, this estimator keeps the nonnegativity. Second, we are carefully examining the asymptotic negligibility of the remainder term of the stochastic expansion, in the spirit of Chen et al. (2009) (see Igarashi and Kakizawa (2018a)). The basic tools are the Rosenthal and Bennett inequalities of the absolute moment and tail probability of the sum of zero-mean independent random variables (and their conditional variants). Third, revisiting the variance manipulation, it is shown that the JLN-type bias reduction, even when it is applied to the asymmetric KDEs, works, at expense of the constant-factor inflation of the variance, whose factor is given by $\lambda(1/2) = 4 - 4\sqrt{2}/\sqrt{3} + 1/\sqrt{2} \approx 1.441$, at least, for three specific asymmetric KDEs suggested in Igarashi and Kakizawa (2015, 2018a) and Igarashi (2016b) ($\mathcal{S} = [0, \infty)$), or the beta KDE ($\mathcal{S} = [0, 1]$). Interestingly, the same factor $\lambda(1/2)$ appeared in the $\mathrm{SS}_{1/2}$-type, $\mathrm{TS}_{1/2}$-type, and $\mathrm{JF}_{1/2}$-type. It is worth noting that the ratio of the integral $\int_{-\infty}^{\infty}\{2\phi(u) - (\phi * \phi)(u)\}^2 du$ relative to the integral $\int_{-\infty}^{\infty}\phi^2(u)du$ is equal to $\lambda(1/2)$, where ϕ is Gaussian density and $*$ is a convolution operator.

The rest of the paper is organized as follows. After presenting our assumptions, Section 2 gives the main results for the JLN-type bias-corrected estimator (6). Section 3 is devoted to the special cases of three different families of asymmetric KDEs for the nonnegative data. Section 4 describes extension to d-variate density estimation by using the product kernel method. Section 5 discusses the product-type beta KDE for the data supported on $[0, 1]^d$. In Section 6, we present results from simulation studies. The proofs are given in the Appendix; some technical details are deferred without further reference to an supplemental issue: Supplemental appendix to "Multiplicative bias correction for asymmetric kernel density estimators revisited", Faculty of Economics, Hokkaido University, Discussion Paper Series A: No. 2018–328.

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 use the notation $||h||_{\mathcal{S}} = \sup_{x\in\mathcal{S}}|h(x)|$ for any bounded function h on $\mathcal{S}$. As usual, we also denote by $h^{(j)}(x) = (d/dx)^j h(x)$ the jth derivative of h (if it exists), and write $h^{(0)}(x) = h(x)$.

2. JLN-type bias correction

In what follows, we assume that $\mathcal{S} = [0, \infty)$. Let $\mathcal{X}_n = \{X_1, \ldots, X_n\}$ be a random sample of size n, drawn from a univariate unknown density f with the support $\mathcal{S}$. The goal of this section is to study the JLN-type bias-corrected KDE (6). We formulate our asymptotic results at given point $x_0 \in [0, c_U]$, where $c_U > 0$ is fixed but arbitrary.

For simplicity, the jth moment around x ($\in \mathcal{S}$) of $K(\cdot; x, \beta)$ is denoted by

$$\widetilde{\mu}_j(K(\cdot; x, \beta)) = \int_{\mathcal{S}} (t-x)^j K(t; x, \beta) dt \quad \text{(if it exists)}.$$

2.1. Assumptions

We make the following assumptions (although, as in Introduction, we can formulate the assumptions indexed by $q, r, r' > 0$; $q \geq r \geq r'$, we here focus on the particular case $q = r = 1$ and $r' = 1/2$, in accordance with the gamma KDE (Chen (2000))):

A1. There exists a density $p_K(\cdot; \cdot)$, such that

1. $K(t; x, \beta) = p_K(t/\beta; x/\beta)/\beta$ for any $t, x \in \mathcal{S}$, where the functional form of p_K is independent of β and x (this implies that $K(\cdot; x, \beta)$ is nonnegative and satisfies $\widetilde{\mu}_0(K(\cdot; x, \beta)) \equiv 1$, where $\widetilde{\mu}_j(K(\cdot; x, \beta)) = \beta^j \int_{\mathcal{S}} (u - x/\beta)^j p_K(u; x/\beta) du$ for $j \in \mathbb{N}$ (if it exists)),
2. the jth raw moment $\mu'_j(y) = \int_{\mathcal{S}} u^j p_K(u; y) du$ exists for any $y \in \mathcal{S}$, having the polynomial grawth of degree j, i.e., $\sup_{y \in \mathcal{S}} \{\mu'_j(y)/(1+y)^j\} < \infty$, for $j = 1, 2, 3, 4, 6$, and
3. for any $y \in \mathcal{S}$, $\mu'_1(y) = y + \zeta_{1,1}$ for some constant $\zeta_{1,1}$, independent of y (in other words, for any $x \in \mathcal{S}$, $K(\cdot; x, \beta)$ satisfies $\widetilde{\mu}_1(K(\cdot; x, \beta)) = \beta\zeta_{1,1}$); in this case, the constant $\zeta_{1,1}$ is necessarily equal to $\mu'_1(0) > 0$.

A2. (i) When $x/\beta \to \infty$ (it holds at least for any fixed bounded $x > 0$), the jth moments around x; $\widetilde{\mu}_j(K(\cdot; x, \beta))$, $j = 2, 3, 4, 6$, admit the asymptotic expansions

$$\widetilde{\mu}_j(K(\cdot; x, \beta)) = \begin{cases} \beta\zeta_{1,2}\, x + \beta^2\zeta_{2,2} + O(\beta^3 x^{-1}), & j = 2, \\ \beta^2\zeta_{2,3}\, x + O(\beta^3), & j = 3, \\ \beta^2\zeta_{2,4}\, x^2 + O(\beta^3 x), & j = 4, \\ O(\beta^3 x^3), & j = 6 \end{cases}$$

for some constants $\zeta_{1,2}, \zeta_{2,2}, \zeta_{2,3}, \zeta_{2,4}$ ($\zeta_{1,2}, \zeta_{2,4} > 0$), independent of β and x. More precisely, the remainder terms $r_{2,\beta}(x) = O(\beta^3 x^{-1})$, $r_{3,\beta}(x) = O(\beta^3)$, $r_{4,\beta}(x) = O(\beta^3 x)$, and $r_{6,\beta}(x) = O(\beta^3 x^3)$ for $j = 2, 3, 4, 6$ are estimated, as follows: Given constants $\eta \in [0, 1)$ and $c_L > 0$, for all sufficiently small $\beta > 0$, $x \geq c_L \beta^\eta$ implies that $|r_{2,\beta}(x)| \leq M_2\beta^3/(x+\beta)$, $|r_{3,\beta}(x)| \leq M_3\beta^3$, $|r_{4,\beta}(x)| \leq M_4\beta^3(x+\beta)$, and $|r_{6,\beta}(x)| \leq M_6\beta^3(x+\beta)^3$ for some constants $M_2, M_3, M_4, M_6 > 0$, independent of β and x.
(ii) The following uniform/nonuniform bounds hold:

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

(iii) Given constants $\eta \in [0,1)$ and $c_L > 0$, for all sufficiently small $\beta > 0$, $x \geq c_L \beta^\eta$ implies that

$$\Big|\int_{\mathcal{S}} K^2(s;x,\beta)ds - \frac{1}{2\sqrt{\pi\beta x}}\Big| \leq \frac{M}{\sqrt{\beta x}}\Big(\frac{\beta}{x+\beta}\Big)$$

for some constant $M > 0$, independent of β and x.

(iv) Given constants $\eta \in [0,1/4)$ and $0 < c_L < c_U$, for all sufficiently small $\beta > 0$, the following approximations hold for $x \in [c_L\beta^\eta, c_U]$:

$$\int_{\mathcal{S}} K(t;x,\beta)\int_{\mathcal{S}} K(s;x,\beta)K(s;t,\beta)dsdt = \frac{1}{\sqrt{6\pi\beta x}} + o((\beta x)^{-1/2}),$$

$$\int_{\mathcal{S}}\int_{\mathcal{S}}\Big\{\prod_{j=1}^{2} K(t_j;x,\beta)\Big\}\int_{\mathcal{S}}\Big\{\prod_{j=1}^{2} K(s;t_j,\beta)\Big\}dsdt_1dt_2 = \frac{1}{2\sqrt{2\pi\beta x}} + o((\beta x)^{-1/2}).$$

A3. (i) f is four times continuously differentiable on $\mathcal{S}$, with $\sum_{j=0}^{4}||f^{(j)}||_{\mathcal{S}} < \infty$.

(ii) There exist constants $\eta_4 \in (0,1]$ and $L_4 > 0$, such that $|f^{(4)}(y) - f^{(4)}(z)| \leq L_4|y-z|^{\eta_4}$ for any $y, z \in \mathcal{S}$.

A4. (i) There exists a constant $R(> c_U)$, such that, for some $\overline{\ell} \in \mathbb{N}$,

1. $\inf_{0\leq x\leq R} f(x) > 0$, and
2. for all sufficiently small $\beta > 0$, $\sup_{x\in[0,c_U]}\int_R^\infty (1+s^2)K(s;x,\beta)/f^{\overline{\ell}}(s)ds$ is smaller than any positive power of β.

(ii) There exists a constant $\delta \in (0,1)$, such that, given constants $\eta \in [0,1)$ and $0 < c_L < c_U$, for all sufficiently small $\beta > 0$, $\sup_{x\in[c_L\beta^\eta,c_U]}\int_0^{\delta x} K(s;x,\beta)ds$ is smaller than any positive power of β.

A5. $\beta \propto n^{-\iota_1}$ for some constant $\iota_1 \in (0,1)$.

Assumptions A3 and A5 are standard in nonparametric density estimation. To make the statements of other assumptions clearer, we will give, in the next section, some examples which are covered (or not covered) in our framework of Assumption A1.1. Define

$$\gamma_1(x;f) = \zeta_{1,1}f^{(1)}(x) + \frac{\zeta_{1,2}}{2}xf^{(2)}(x), \quad \gamma_2(x;f) = \frac{\zeta_{2,2}}{2}f^{(2)}(x) + \frac{\zeta_{2,3}}{6}xf^{(3)}(x) + \frac{\zeta_{2,4}}{24}x^2f^{(4)}(x).$$

Before presenting our main results on the JLN-type (Subsection 2.2), let us now illustrate the key quantities in our analysis (compared to (3) and (5)). For simplicity, we write

$$B_\beta^{(K),f}(s) = \int_{\mathcal{S}} K(t;s,\beta)f(t)dt - f(s), \quad s \in \mathcal{S}.$$

As usual, Taylor's expansion around $t = s$ yields

$$\Big|\int_{\mathcal{S}} K(t;s,\beta)f(t)dt - f(s) - \sum_{j=1}^{4}\frac{f^{(j)}(s)}{j!}\widetilde{\mu}_j(K(\cdot;s,\beta))\Big| \leq \frac{L_4}{24}\widetilde{\mu}_6^{(4+\eta_4)/6}(K(\cdot;s,\beta)),$$

hence, for all sufficiently small $\beta > 0$, $s \geq c_L \beta^\eta$ ($\eta \in [0, 1)$ and $c_L > 0$ are constants) implies that

$$\Big|B_\beta^{(K),f}(s) - \sum_{j=1}^{2} \beta^j \gamma_j(s; f)\Big| \leq \sum_{j=2}^{4} \frac{||f^{(j)}||_{\mathcal{S}}}{j!}|r_{j,\beta}(s)| + \frac{L_4}{24}|r_{6,\beta}(s)|^{(4+\eta_4)/6}$$
$$\leq M'\Big[\beta^{3-\eta} + \beta^{2+\eta_4/2}(1 + s^{2+\eta_4/2})\Big] \quad (7)$$

for some constant $M' > 0$, independent of β and s (see Assumptions A1.3 and A2(i)). Our framework of Assumption A1.1–2 facilitates the boundary analysis when $s = \beta\kappa$ ($\kappa \geq 0$ is a constant), i.e.,

$$B_\beta^{(K),f}(\beta\kappa) = \beta\{\mu_1'(\kappa) - \kappa\}f^{(1)}(0) + \frac{\beta^2}{2}\{\mu_2'(\kappa) - \kappa^2\}f^{(2)}(0) + O(\beta^3) \quad (8)$$

(for the expression (8), we can use $\mu_1'(\kappa) - \kappa = \mu_1'(0)$ under Assumption A1.3), since

$$\Big|\int_{\mathcal{S}} K(t; \beta\kappa, \beta)f(t)dt - f(\beta\kappa) - \sum_{j=1}^{2} \frac{f^{(j)}(\beta\kappa)}{j!}\widetilde{\mu}_j(K(\cdot; \beta\kappa, \beta))\Big| \leq \frac{||f^{(3)}||_{\mathcal{S}}}{6}\widetilde{\mu}_4^{3/4}(K(\cdot; \beta\kappa, \beta)),$$

with $\widetilde{\mu}_j(K(\cdot; \beta\kappa, \beta)) = \beta^j \int_{\mathcal{S}}(u - \kappa)^j p_K(u; \kappa)du$, $j = 1, 2, 4$ (note that $|f^{(2)}(\beta\kappa) - f^{(2)}(0)| \leq \beta\kappa||f^{(3)}||_{\mathcal{S}}$ and $|f^{(1)}(\beta\kappa) - f^{(1)}(0) - \beta\kappa f^{(2)}(0)| \leq (1/2)(\beta\kappa)^2||f^{(3)}||_{\mathcal{S}}$).

Remark 1 (i) $\sup_{x\in\mathcal{S}} |B_\beta^{(K),f}(x)| \leq 2||f||_{\mathcal{S}}$ under the boundedness of f and Assumption A1.1.
(ii) Under Assumptions A1, A2(i), and A3(i), we have, for all sufficiently small $\beta > 0$,

$$\sup_{x\in[0,\widetilde{R}]} |B_\beta^{(K),f}(x)| = O(\beta) \quad \text{for any constant } \widetilde{R} > 0. \quad (9)$$

Also,

$$\sup_{x\in[0,\beta^\tau]} |B_\beta^{(K),f}(x) - \beta\xi_{1,1}f^{(1)}(0)| = O(\beta^{2\tau}) \quad \text{for any constant } \tau \in (0, 1).$$

Proof of Remark 1(ii) We see that, for $x \geq 0$,

$$\Big|\int_{\mathcal{S}} K(t; x, \beta)f(t)dt - f(x) - \beta\xi_{1,1}f^{(1)}(x)\Big| \leq \frac{||f^{(2)}||_{\mathcal{S}}}{2}\widetilde{\mu}_2(K(\cdot; x, \beta)),$$

hence, $|B_\beta^{(K),f}(x)| \leq \beta\xi_{1,1}||f^{(1)}||_{\mathcal{S}} + (1/2)||f^{(2)}||_{\mathcal{S}}\,\widetilde{\mu}_2(K(\cdot; x, \beta))$. For any $\widetilde{R} > 0$, Assumption A2(i) yields, for all sufficiently small $\beta > 0$, $\sup_{x\in[\beta^{1/2},\widetilde{R}]} \widetilde{\mu}_2(K(\cdot; x, \beta)) \leq \beta\zeta_{1,2}\widetilde{R} + \beta^2(|\zeta_{2,2}| + M_2)$, whereas Assumption A1.2 implies that

$$\widetilde{\mu}_2(K(\cdot; x, \beta)) \leq 2\Big[\int_{\mathcal{S}} t^2 \frac{1}{\beta}p_K\Big(\frac{t}{\beta}; \frac{x}{\beta}\Big)dt + x^2\Big] \leq 2\Big[\beta^2\Big\{1 + \Big(\frac{x}{\beta}\Big)^2\Big\} \sup_{y\in\mathcal{S}} \frac{\mu_2'(y)}{1 + y^2} + x^2\Big],$$

hence, $\sup_{x\in[0,\beta^{\widetilde{\tau}}]} \widetilde{\mu}_2(K(\cdot; x, \beta)) = O(\beta^{\min(2,2\widetilde{\tau})})$ for any constant $\widetilde{\tau} > 0$. The latter result then follows from $\sup_{x\in[0,\beta^{\widetilde{\tau}}]} |f^{(1)}(x) - f^{(1)}(0)| \leq \beta^{\widetilde{\tau}}||f^{(2)}||_{\mathcal{S}}$. □

Remark 2 The error bound (7) should read

$$\Big|B_\beta^{(K),f}(s) - \sum_{j=1}^{2} \beta^j \gamma_j(s; f)\Big| \leq M'\beta^{2+\eta_4/2}(1 + s^{2+\eta_4/2}) \quad \text{for } s \geq 0,$$

if Assumption A2(i) can be replaced by the following stronger version[4]:

A2. ($\mathrm{i}^\sharp$) The jth moments around $x \geq 0$; $\widetilde{\mu}_j(K(\cdot;x,\beta))$, $j = 2,3,4,6$, are given by

$$\widetilde{\mu}_j(K(\cdot;x,\beta)) = \begin{cases} \beta\zeta_{1,2}\,x + \beta^2\zeta_{2,2} & j=2, \\ \beta^2\zeta_{2,3}\,x \ + r_{3,\beta}(x), & j=3, \\ \beta^2\zeta_{2,4}\,x^2 + r_{4,\beta}(x), & j=4, \\ r_{6,\beta}(x), & j=6 \end{cases}$$

with $|r_{3,\beta}(x)| \leq M_3\beta^3$, $|r_{4,\beta}(x)| \leq M_4\beta^3(x+\beta)$, and $|r_{6,\beta}(x)| \leq M_6\beta^3(x+\beta)^3$ for some constants $\zeta_{1,2}, \zeta_{2,2}, \zeta_{2,3}, \zeta_{2,4}$ ($\zeta_{1,2}, \zeta_{2,2}, \zeta_{2,4} > 0$) and $M_3, M_4, M_6 > 0$, independent of β and x.

Also, if g is continuously differentiable g on $\mathcal{S}$, with $\sum_{i=0}^{1} ||g^{(i)}||_{\mathcal{S}} < \infty$, Assumptions A1.1–2, A2(i,ii.2,iii), and A3(i) yield, for all sufficiently small $\beta > 0$,

$$\int_{\mathcal{S}} K^2(t;x,\beta)g(t)dt = \begin{cases} \beta^{-1/2}\dfrac{g(x)}{2\sqrt{\pi x}} + o((\beta x)^{-1/2}) + O(1) & \text{for } x \in [c_L\beta^\eta, c_U], \\ \beta^{-1}g(0)\displaystyle\int_{\mathcal{S}} p_K^2(u;\kappa)du + O(1) & \text{for } x = \beta\kappa, \end{cases} \tag{10}$$

where $\eta \in [0,1)$, $0 < c_L < c_U$, and $\kappa \geq 0$ are constants. The proof is easy, as follows: For $x \in [c_L\beta^\eta, c_U]$, Taylor's expansion around $t = x$ yields

$$\begin{aligned} \Big|\int_{\mathcal{S}} K^2(t;x,\beta)g(t)dt - g(x)\int_{\mathcal{S}} K^2(t;x,\beta)dt\Big| &\leq ||g^{(1)}||_{\mathcal{S}} \int_{\mathcal{S}} |t-x| K^2(t;x,\beta)dt \\ &\leq ||g^{(1)}||_{\mathcal{S}} \frac{C_K'}{\sqrt{\beta x}} \widetilde{\mu}_2^{1/2}(K(\cdot;x,\beta)) \\ &\leq ||g^{(1)}||_{\mathcal{S}} C_K' \Big\{\zeta_{1,2} + \frac{\beta^{1-\eta}}{c_L}(|\zeta_{2,2}| + M_2)\Big\}^{1/2}, \end{aligned}$$

whereas, for $x = \beta\kappa$, Taylor's expansion around $t = 0$ yields

$$\Big|\int_{\mathcal{S}} K^2(t;\beta\kappa,\beta)g(t)dt - g(0)\int_{\mathcal{S}} K^2(t;\beta\kappa,\beta)dt\Big| \leq ||g^{(1)}||_{\mathcal{S}} \int_{\mathcal{S}} tK^2(t;\beta\kappa,\beta)dt,$$

with $\int_{\mathcal{S}} t^j K^2(t;\beta\kappa,\beta)dt = \beta^{j-1}\int_{\mathcal{S}} u^j p_K^2(u;\kappa)du$, $j = 0,1$.

2.2. Main results

2.2.1. Bias and variance approximations

We are ready to present the bias and variance of the JLN-type bias-corrected KDE (6).

Proposition 1 *Suppose that Assumptions* A1, A2(i,ii.1), *and* A3–A5 *hold (we set* $\overline{\ell} \geq 3$ *for* A4(i))[5], *and that* $\epsilon \propto \beta^{\iota_2}$ *for some constant* $\iota_2 > 1$. *Then, for all sufficiently small* $\beta > 0$,

$$Bias[\widetilde{f}_{\beta,JLN}(x)] = \begin{cases} \beta^2 B_{JLN}(x;f) + o(\beta^2) + O(n^{-1}(\beta x)^{-1/2}) & \textit{for } x \in [c_L\beta^\eta, c_U], \\ \beta^2 I(\kappa;f) + o(\beta^2) + O(n^{-1}\beta^{-1}) & \textit{for } x = \beta\kappa, \end{cases} \tag{11}$$

[4]The requirement of $\widetilde{\mu}_2(K(\cdot;x,\beta)) = \beta\zeta_{1,2}\,x + \beta^2\zeta_{2,2}$ is satisfied, if, in addition to Assumption A1.3, the 2nd raw moment $\mu_2'(y)$, for $y \geq 0$, is a quadratic polynomial in y; $\mu_2'(y) = y^2 + (2\zeta_{1,1} + \zeta_{1,2})y + \zeta_{2,2}$; in this case, we have $\zeta_{j,j} = \mu_j'(0) > 0$, $j = 1,2$.

[5]Hirukawa and Sakudo (2014) (see also the subsequent papers; Funke and Kawka (2015), Hirukawa and Sakudo (2015), Zougab and Adjabi (2016), and Zougab et al. (2018)) assumed $n\beta^3 \to \infty$ to control the remainder term of the bias, i.e., $O(n^{-1}\beta^{-1}) = o(b^2)$. Their stronger assumption is, however, redundant for analyzing the M(I)SE.

where $\eta \in [0,1)$, $0 < c_L < c_U$, *and* $\kappa \geq 0$ *are constants,*

$$B_{JLN}(x;f) = -f(x)\gamma_1\Big(x; \frac{\gamma_1(\cdot;f)}{f(\cdot)}\Big),$$
$$I(\kappa;f) = \{\mu_1'(0)\}^2\frac{\{f^{(1)}(0)\}^2}{f(0)} + \Big\{-\frac{1}{2}\int_{\mathcal{S}}\mu_2'(y)p_K(y;\kappa)dy + \mu_2'(\kappa) - \frac{\kappa^2}{2}\Big\}f^{(2)}(0).$$

Note that, under Assumption A2($\mathrm{i}^\sharp$), *we can use*

$$I(\kappa;f) = \zeta_{1,1}^2\frac{\{f^{(1)}(0)\}^2}{f(0)} - \zeta_{1,1}\Big(\zeta_{1,1} + \frac{\zeta_{1,2}}{2}\Big)f^{(2)}(0) = B_{JLN}(\beta\kappa;f) + O(\beta).$$

Proposition 2 *Suppose that Assumptions* A1–A5 *hold (we set* $\bar{\ell} \geq 5$ *for* A4(i)), *and that* $\epsilon \propto \beta^{\iota_2}$ *for some constant* $\iota_2 \geq 1$. *Then, for all sufficiently small* $\beta > 0$,

$$V[\widetilde{f}_{\beta,JLN}(x)] = \begin{cases} n^{-1}\beta^{-1/2}\lambda(1/2)\dfrac{f(x)}{2\sqrt{\pi x}} + o(n^{-1}(\beta x)^{-1/2}) + O(\beta^5) & \text{for } x \in [c_L\beta^\eta, c_U], \\ n^{-1}\beta^{-1}J(\kappa)f(0) + o(n^{-1}\beta^{-1}) + O(\beta^5) & \text{for } x = \beta\kappa, \end{cases} \tag{12}$$

where $\eta \in [0,1/4)$, $0 < c_L < c_U$, *and* $\kappa \geq 0$ *are constants, and* $J(\kappa) = 4J_1(\kappa) - 4J_2(\kappa) + J_3(\kappa)$ *with*

$$J_1(\kappa) = \int_{\mathcal{S}} p_K^2(u;\kappa)du, \quad J_2(\kappa) = \int_{\mathcal{S}} p_K(y;\kappa)\int_{\mathcal{S}} p_K(u;\kappa)p_K(u;y)dudy,$$
$$J_3(\kappa) = \int_{\mathcal{S}}\int_{\mathcal{S}}\Big[\prod_{j=1}^2 p_K(y_j;\kappa)\Big]\int_{\mathcal{S}}\Big[\prod_{j=1}^2 p_K(u;y_j)\Big]dudy_1dy_2.$$

Note that $V[\widetilde{f}_{\beta,JLN}(x)] = O(n^{-1}(\beta x)^{-1/2} + \beta^5)$ *for* $x \in [c_L\beta^\eta, c_U]$, *if* $\eta \in [0,1)$.

Remark 3 A careful analysis of Proofs of Propositions 1 and 2 shows that, for all sufficiently small $\beta > 0$,

$$\sup_{x\in[0,c_U]} V[\widetilde{f}_{\beta,JLN}(x)] = O(n^{-1}\beta^{-1} + \beta^5),$$
$$\sup_{x\in[0,\beta^\tau]} |Bias[\widetilde{f}_{\beta,JLN}(x)]| = O(\beta^{2\tau} + n^{-1}\beta^{-1}) \quad \text{for any constant } \tau \in (1/2,1).$$

These bounds will be required to show that the different rates in the variance and the remainder term of the bias has negligible impact on the truncated mean integrated squared error (MISE).

2.2.2. MSE

Propositions 1 and 2 (i.e., the bias (11) and variance (12)) immediately yield

$$MSE[\widetilde{f}_{\beta,JLN}(x_0)] = \begin{cases} \beta^4B_{JLN}^2(x_0;f) + n^{-1}\beta^{-1/2}\lambda(1/2)\dfrac{f(x_0)}{2\sqrt{\pi x_0}} + o(\beta^4 + n^{-1}\beta^{-1/2}) & \text{for fixed } x_0 \in (0,c_U], \\ \beta^4I^2(0;f) + n^{-1}\beta^{-1}J(0)f(0) + o(\beta^4 + n^{-1}\beta^{-1}) & \text{for } x_0 = 0, \end{cases}$$

whose leading term is minimized by choosing

$$\beta_{JLN}^{opt}(x_0) = \begin{cases} \left[\dfrac{\lambda(1/2)\frac{f(x_0)}{2\sqrt{\pi x_0}}}{8B_{JLN}^2(x_0; f)} n^{-1} \right]^{2/9} & \text{for fixed } x_0 \in (0, c_U], \\ \left[\dfrac{J(0)f(0)}{4I^2(0; f)} n^{-1} \right]^{1/5} & \text{for } x_0 = 0, \end{cases}$$

and the optimal AMSE is then given by

$$AMSE^{opt}[\widetilde{f}_{\beta,JLN}(x_0)] = \begin{cases} 9\{B_{JLN}^2(x_0; f)\}^{1/9} \left[\dfrac{1}{8}\lambda(1/2)\dfrac{f(x_0)}{2\sqrt{\pi x_0}} n^{-1} \right]^{8/9} & \text{for fixed } x_0 \in (0, c_U], \\ 5\{I^2(0; f)\}^{1/5} \left[\dfrac{1}{4} J(0)f(0)n^{-1} \right]^{4/5} & \text{for } x_0 = 0. \end{cases}$$

This order is faster, compared with (4) when $q = r = 1$ and $r' = 1/2$.

2.2.3. MISE

To measure a global performance of the density estimator, we now technically use the truncated MISE, defined by $MISE_w[\widehat{f}] = \int_{[0,w]} MSE[\widehat{f}(x)]dx$, where $w \in (0, c_U]$ is a constant. Propositions 1 and 2, together with Remark 3, yield $MISE_w[\widetilde{f}_{\beta,JLN}] = AMISE_w[\widetilde{f}_{\beta,JLN}] + o(\beta^4 + n^{-1}\beta^{-1/2})$, where

$$AMISE_w[\widetilde{f}_{\beta,JLN}] = \beta^4 \int_0^w B_{JLN}^2(x; f)dx + n^{-1}\beta^{-1/2}\lambda(1/2) \int_0^w \frac{f(x)}{2\sqrt{\pi x}} dx$$

(with $\beta \propto n^{-2/9}$, the convergence rate $n^{-8/9}$ is achieved, which is faster than that of the uncorrected case, $n^{-4/5}$).

Proof Choosing constants $\tau_1 \in (4/5, 1)$ and $\tau_2 \in (0, 1/4)$, we have, for all sufficiently small $\beta > 0$,

$$\begin{aligned} \int_{\beta^{\tau_1}}^w Bias^2[\widetilde{f}_{\beta,JLN}(x)]dx &= \beta^4 \int_{\beta^{\tau_1}}^w B_{JLN}^2(x; f)dx + o(n^{-1}\beta^{-1/2} + \beta^4) + O((n^{-1}\beta^{-1/2})^2 \log(1/\beta)) \\ &= \beta^4 \int_0^w B_{JLN}^2(x; f)dx + o(n^{-1}\beta^{-1/2} + \beta^4), \\ \int_{\beta^{\tau_2}}^w V[\widetilde{f}_{\beta,JLN}(x)]dx &= n^{-1}\beta^{-1/2}\lambda(1/2) \int_{\beta^{\tau_2}}^w \frac{f(x)}{2\sqrt{\pi x}} dx + o(n^{-1}\beta^{-1/2} + \beta^4) \\ &= n^{-1}\beta^{-1/2}\lambda(1/2) \int_0^w \frac{f(x)}{2\sqrt{\pi x}} dx + o(n^{-1}\beta^{-1/2} + \beta^4), \\ \int_0^{\beta^{\tau_1}} MSE[\widetilde{f}_{\beta,JLN}(x)]dx &= O(\{\beta^{4\tau_1} + (n^{-1}b^{-1})^2 + n^{-1}\beta^{-1} + \beta^5\}\beta^{\tau_1}) = o(n^{-1}\beta^{-1/2} + \beta^4), \\ \int_{\beta^{\tau_1}}^{\beta^{\tau_2}} V[\widetilde{f}_{\beta,JLN}(x)]dx &= O(n^{-1}\beta^{-1/2+\tau_1/2} + \beta^{5+\tau_2}) = o(n^{-1}\beta^{-1/2} + \beta^4). \quad \square \end{aligned}$$

3. Special cases and discussion

To build the asymmetric kernel $K(\cdot; x, \beta)$ supported on $[0, \infty)$, we focus on the application of three different families:

- The modified Bessel (MB) density, with the parameter $\omega, \sigma > 0$ and $\lambda \in \mathbb{R}$, is defined by

$$K^{(MB)}_{\omega,\sigma}(s;\lambda) = \frac{s^{\lambda-1}}{2\sigma^\lambda K_\lambda(\omega)} \exp\Big\{-\frac{\omega}{2}\Big(\frac{s}{\sigma}+\frac{\sigma}{s}\Big)\Big\},$$

where K_ν is the modified Bessel function of the third kind with index $\nu \in \mathbb{R}$, i.e.,

$$K_\nu(\omega) = \int_0^\infty \frac{s^{\nu-1}}{2} \exp\Big\{-\frac{\omega}{2}(s+s^{-1})\Big\} ds = \int_0^\infty \cosh(\nu t)\exp\{-\omega\cosh(t)\}dt$$

(note that $K_\nu(\omega) = K_{-\nu}(\omega)$ and $K_{1/2}(\omega) = \{\pi/(2\omega)\}^{1/2}e^{-\omega}$).

- The weighted LN[λ] density, with the parameter $\mu, \lambda \in \mathbb{R}$ and $\sigma^2 > 0$, is defined by

$$K^{(LN)}_{\mu,\sigma^2}(s;\lambda) = \frac{s^{\lambda-1}}{\sqrt{2\pi}\sigma} \exp\Big\{-\frac{(\log s-\mu)^2}{2\sigma^2} - \lambda\mu - \frac{\lambda^2\sigma^2}{2}\Big\}$$

(by definition, $K^{(LN)}_{\mu,\sigma^2}(\cdot;\lambda) = K^{(LN)}_{\mu+\lambda\sigma^2,\sigma^2}(\cdot;0)$ for $\lambda \in \mathbb{R}$).

- The Amoroso density, with the parameter $\alpha, \beta > 0$ and $\gamma \neq 0$, is defined by

$$K^{(A)}_{\alpha,\beta,\gamma}(s) = \frac{|\gamma| s^{\alpha\gamma-1}}{\beta^{\alpha\gamma}\Gamma(\alpha)} \exp\Big\{-\Big(\frac{s}{\beta}\Big)^\gamma\Big\}.$$

3.1. Application to MIG/MLN/Amoroso KDEs

Example 1 (MIG KDE) Reformulating the IG/BS/RIG KDEs due to Jin and Kawczak (2003) and Scaillet (2004), Igarashi and Kakizawa (2014b, 2015) suggested the mixture of IG and RIG kernels (in short, the MIG kernel)

$$K^{(MIG_p)}(\cdot;x,b) = (1-p)K^{(MB)}_{x/b+1,x+b}(\cdot;-1/2) + pK^{(MB)}_{x/b+1,x+b}(\cdot;1/2), \quad x \geq 0$$

to construct the (uncorrected) MIG KDE defined by $\widehat{f}^{(MIG_p)}_b(x) = n^{-1}\sum_{i=1}^n K^{(MIG_p)}(X_i;x,b)$, where $p \in [0,1]$ is a mixing proportion, independent of β and x. This class of estimators contains the (reformulated) IG/BS/RIG KDEs as special cases $p = 0, 1/2, 1$, respectively.

Example 2 (MLN KDE) Reformulating the LN KDE due to Jin and Kawczak (2003), Igarashi (2016b) (see also Igarashi and Kakizawa (2015)) used the LN[$\pm 1/2$] kernel $K^{(LN)}_{\mu_b(x),\sigma_b^2(x)}(\cdot;\pm 1/2)$, $x \geq 0$, where

$$\mu_b(x) = \log(x+b) = \log b\rho(x/b), \quad \sigma_b^2(x) = \log\Big(1+\frac{b}{x+b}\Big) = \log\Big\{1+\frac{1}{\rho(x/b)}\Big\} \leq \log 2$$

(we write $\rho(t) = t+1$). Parallel to the MIG kernel, we may choose a mixture of the LN[$\pm 1/2$] kernels (in short, the MLN kernel)

$$K^{(MLN_p)}(\cdot;x,b) = (1-p)K^{(LN)}_{\mu_b(x),\sigma_b^2(x)}(\cdot;-1/2) + pK^{(LN)}_{\mu_b(x),\sigma_b^2(x)}(\cdot;1/2), \quad x \geq 0$$

to construct the (uncorrected) MLN KDE defined by $\widehat{f}^{(MLN_p)}_b(x) = n^{-1}\sum_{i=1}^n K^{(MLN_p)}(X_i;x,b)$.

Example 3 (Amoroso (or generalized gamma) KDE) For every constant $\gamma > 0$, Igarashi and Kakizawa (2018a) suggested the Amoroso kernel

$$K^{(A_\gamma)}(s; x, b) = K^{(A)}_{\alpha_b(x), b\beta_b(x), \gamma}(s), \quad x \geq 0,$$

where

$$\alpha_b(x) = \frac{x/b + 1}{\gamma} = \frac{\rho(x/b)}{\gamma}, \quad \beta_b(x) = (x/b + 1)\frac{\Gamma\left(\frac{x/b+1}{\gamma}\right)}{\Gamma\left(\frac{x/b+2}{\gamma}\right)} = \rho(x/b)\frac{\Gamma\left(\frac{\rho(x/b)}{\gamma}\right)}{\Gamma\left(\frac{\rho(x/b)+1}{\gamma}\right)}.$$

In this paper, we use the (uncorrected) Amoroso KDE defined by $\widehat{f}^{(A_\gamma)}_b(x) = n^{-1}\sum_{i=1}^n K^{(A_\gamma)}(X_i; x, b\gamma)$, for every constant $\gamma > 0$[6]. The gamma KDE due to Chen (2000), $\widehat{f}^{(G)}_b(x) = n^{-1}\sum_{i=1}^n K^{(A)}_{x/b+1,b,1}(s)$ is a core member with $\gamma = 1$.

Now, to ensure Propositions 1 and 2 for the bias-corrected MIG/MLN/Amoroso KDEs, defined by

$$\widehat{f}^{(\#)}_{b,JLN}(x) = \{\widehat{f}^{(\#)}_b(x) + \epsilon\}\frac{1}{n}\sum_{i=1}^n \frac{K^{(\#)}(X_i; x, b)}{\widehat{f}^{(\#)}_b(X_i) + \epsilon}, \quad x \geq 0, \quad \# = MIG_p, MLN_p, A_\gamma,$$

we have only to verify Assumptions on the respective kernel $K^{(\#)}(\cdot;\cdot,b)$. The most of them is not restrictive or available from the existing literature. By construction, Assumptions A1 and A2(i,iii) can be readily verified for three examples. In addition to $q = r = 1$ and $r' = 1/2$, from Lemma 1 of Igarashi and Kakizawa (2015) and Lemma A.1 of Igarashi and Kakizawa (2018a), we obtain

- $(\zeta^{(\#_p)}_{1,1}, \zeta^{(\#_p)}_{1,2}) = (p+1, 1)$, $(\zeta^{(\#_p)}_{2,2}, \zeta^{(\#_p)}_{2,3}, \zeta^{(\#_p)}_{2,4}) = (5p+2, 3(p+2), 3)$, where $\# = MIG, MLN$,
- $(\zeta^{(A_\gamma)}_{1,1}, \zeta^{(A_\gamma)}_{1,2}) = (\gamma, 1)$, $(\zeta^{(A_\gamma)}_{2,2}, \zeta^{(A_\gamma)}_{2,3}, \zeta^{(A_\gamma)}_{2,4}) = \left(\frac{1}{2}(3\gamma^2+1), 2\gamma+3, 3\right)$.

Assumption A2(i$^\sharp$) holds for the MIG/LN[−1/2]/gamma kernels[7]. The uniform/nonuniform bounds in Assumption A2(ii) for the MB kernel $K^{(MB)}_{x/b+1,x+b}(\cdot;\lambda)$ and the weighted LN kernel $K^{(LN)}_{\mu_b(x),\sigma_b^2(x)}(\cdot;\lambda)$, where $\lambda \in \mathbb{R}$, are found in Lemma A.2 of Igarashi and Kakizawa (2014b), Lemma 4 of Igarashi (2016b), and Remark A.1(i) of Igarashi and Kakizawa (2015). See also Lemma A.3 of Igarashi and Kakizawa (2018a) for the Amoroso kernel $K^{(A_\gamma)}(\cdot;\cdot,b\gamma)$.

On the other hand, after some algebra, it is straightforward to see that

---

[6] Igarashi and Kakizawa's original definition should read $n^{-1}\sum_{i=1}^n K^{(A_\gamma)}(X_i; x, b)$.

In their paper (Igarashi and Kakizawa (2018a)), the negative exponent $\gamma < 0$ has been allowed, for which the definition of the parameter $(\alpha_b(x), \beta_b(x))$ should read, as follows: For every constant $\gamma < 0$,

$$\alpha_b(x) = \frac{\rho^*(x/b) + 1}{|\gamma|}, \quad \beta_b(x) = \rho^*(x/b)\frac{\Gamma\left(\frac{\rho^*(x/b)+1}{|\gamma|}\right)}{\Gamma\left(\frac{\rho^*(x/b)}{|\gamma|}\right)}, \quad \text{where } \rho^*(x) = x + c^*, \text{ with } c^* > 1 \text{ (rather than } c^* = 1).$$

For the simulation studies in Section 6, $c^* = 1.1$ was chosen.

[7] Let $x \geq 0$. We can see that $\widetilde{\mu}_j(K^{(\#)}(\cdot; x, b))$'s, where $\# = G, MIG_p$, are the polynomials in x;

$$\widetilde{\mu}_j(K^{(G)}(\cdot;x,b)) = \begin{cases} b, & j=1, \\ bx + 2b^2, & j=2, \\ 5b^2x + O(b^3), & j=3, \\ 3b^2x^2 + O(b^3(x+b)), & j=4, \\ O(b^3(x+b)^3), & j=6 \end{cases} \quad \text{and} \quad \widetilde{\mu}_j(K^{(MIG_p)}(\cdot;x,b)) = \begin{cases} (p+1)b, & j=1, \\ bx + (5p+2)b^2, & j=2, \\ 3(p+2)b^2x + O(b^3), & j=3, \\ 3b^2x^2 + O(b^3(x+b)), & j=4, \\ O(b^3(x+b)^3), & j=6, \end{cases}$$

- the kernels $K^{(MIG_p)}(\cdot;\cdot,b)$, $K^{(MLN_p)}(\cdot;\cdot,b)$, and $K^{(A_\gamma)}(\cdot;\cdot,b\gamma)$ can be well approximated by a Gaussian, i.e., choosing constants $\eta \in [0,1/4)$ and $\tau \in ((2\eta+1)/3,1/2)$, we have, for all sufficiently small $\beta > 0$,

$$\sqrt{\beta x}K(x+\sqrt{\beta x}z_1; x+\sqrt{\beta x}z_2, \beta) = \phi(z_1 - z_2)[1+o(1)]$$

  uniformly in $z_1, z_2 \in [-\beta^\tau/\sqrt{\beta x}, \beta^\tau/\sqrt{\beta x}\,]$ and $x \in [c_L\beta^\eta, c_U]$.

This Gaussian approximation is a sufficient condition for Assumption A2(iv); note that, in this case, two coefficients $\zeta_{1,2}$ and $\zeta_{2,4}$ are equal to 1 and 3, respectively.

It remains to verify Assumption A4. For this, we now consider the two situations (we do not pursue more general conditions) where either of the following assumptions holds for some constant $\iota > 0$:

f1[ι]. $\inf_{x\in\mathcal{S}}\{f(x)\exp(C_2x^\iota)\} \geq C_1$ for some constants $C_1, C_2 > 0$.

f2. $\inf_{x\in\mathcal{S}}\{f(x)(1+x)^{C_2}\} \geq C_1$ for some constants $C_1 > 0$ and $C_2 > 1$.

In either case, $\inf_{0\leq x\leq R} f(x) > 0$ for any $R > 0$ (Assumption A4(i.1) is automatically satisfied).

Lemma 3 (i) *If either of Assumption* f1[ι] *or* f2 *holds for some* $\iota \in (0,1]$*, Assumption* A4(i.2,ii) *holds for any* $\overline{\ell} \in \mathbb{N}$ *when* K *is the MIG kernel* $K^{(MIG_p)}(\cdot;\cdot,b)$.
(ii) *If Assumption* f2 *holds, Assumption* A4(i.2,ii) *holds for any* $\overline{\ell} \in \mathbb{N}$ *when* K *is the MLN kernel* $K^{(MLN_p)}(\cdot;\cdot,b)$.
(iii) *Given* $\gamma > 0$*, if either of Assumption* f1[ι] *or* f2 *holds for some* $\iota \in (0,\gamma]$*, Assumption* A4(i.2,ii) *holds for any* $\overline{\ell} \in \mathbb{N}$ *when* K *is the Amoroso kernel* $K^{(A_\gamma)}(\cdot;\cdot,b\gamma)$.

3.2. Discussion

Examples 1–3 have in common that the linear function $\rho(t) = t+1$ (one can use $\rho(t) = t+c$, where $c \geq 1$ is a constant) is adopted. As demonstrated in Chen (2000) and Igarashi and Kakizawa (2014b, 2015, 2018a) (see also Hirukawa and Sakudo (2015) for a subfamily of the generalized gamma KDE under an additional restriction $\gamma \geq 1$, focusing on the Nakagami case with $\gamma = 2$), it may be true that

---

whereas $\widetilde{\mu}_j(K^{(MLN_p)}(\cdot;x,b))$'s are, in general, the rational functions in x;

$$\widetilde{\mu}_j(K^{(MLN_p)}(\cdot;x,b)) = \begin{cases} (p+1)b, & j=1, \\ bx + (5p+2)b^2 + \dfrac{pb^3}{x+b}, & j=2, \\ 3(p+2)b^2x + r_{3,b}^{(MLN_p)}(x), & j=3, \\ 3b^2x^2 + r_{4,b}^{(MLN_p)}(x), & j=4, \\ r_{6,b}^{(MLN_p)}(x) & j=6 \end{cases}$$

with $|r_{3,b}^{(MLN_p)}(x)| \leq M_3b^3$, $|r_{4,b}^{(MLN_p)}(x)| \leq M_4b^3(x+b)$, and $|r_{6,b}^{(MLN_p)}(x)| \leq M_6b^3(x+b)^3$ for some constants $M_3, M_4, M_6 > 0$, independent of b and x.

The MLN$_p$ kernel for $p \in (0,1]$ does not satisfy Assumption A2(i$^\sharp$).

the use of two-regime ρ-function, having the form

$$\rho_c(t) = \begin{cases} c+t, & t>2, \\ r_c(t), & t \in [0,2] \end{cases} \quad \text{(we assume } c+2 = r_c(2) \geq r_c(0) \geq 1 \text{ for some constant } c \in \mathbb{R}\text{)},$$

where the function ρ_c is continuous and non-decreasing on $[0,\infty)$, has the following advantages for the resulting estimator $\widetilde{f}_b^\star(\cdot;\rho_c)$ (we call the two-regime type): (i) choosing $c=0$, $Bias[\widetilde{f}_b^\star(x;\rho_0)]$ does not involve f' in the leading $O(b)$-term, when $x \in [2b,\infty)$, or (ii) one can minimize the $O(n^{-4/5})$-MISE of $\widetilde{f}_b^\star(\cdot;\rho_c)$ with respect to the additional parameter c; see Igarashi and Kakizawa (2014b). However, Assumption A1.3 is violated, due to the introduction of the non-linear ρ-function.

Remark 4 Even if Assumption A1.3 is replaced by the following assumption, (9) remains valid:

A1. $3^\sharp$. $\mu_1'(y)$ has the form[8] of

$$\mu_1'(y) = \begin{cases} \zeta_{1,1}+y, & y>c', \\ \zeta(y), & y \in [0,c'] \end{cases} \quad \text{for some constants } c'>0 \text{ and } \zeta_{1,1} \in \mathbb{R},$$

where ζ is a continuous and non-decreasing function on $[0,\infty)$, with $\zeta(c') = \zeta_{1,1}+c' \geq \zeta(0) > 0$.

The proof is easy: We see that, for $x \geq 0$,

$$\Big|\int_{\mathcal{S}} K(t;x,\beta)f(t)dt - f(x) - \beta\{\mu_1'(x/\beta) - x/\beta\}f^{(1)}(x)\Big| \leq \frac{||f^{(2)}||_{\mathcal{S}}}{2}\widetilde{\mu}_2(K(\cdot;x,\beta)),$$

where $\sup_{x\in[0,\widetilde{R}]}\widetilde{\mu}_2(K(\cdot;x,\beta)) = O(\beta)$ for any constant $\widetilde{R}>0$ (see Proof of (9)). By definition, $x \in [c'\beta,\widetilde{R}]$ implies $\mu_1'(x/\beta) - x/\beta \equiv \zeta_{1,1}$. Also, $\sup_{x\in[0,c'\beta]}|\mu_1'(x/\beta)-x/\beta| \leq \zeta(c')+c'$.

The following result reveals the weakness of the two-regime type.

Proposition 4 (Violation of Assumption A1.3) *In Propositions* 1 *and* 2*, if Assumption* A1.3 *is replaced by* A1.$3^\sharp$*, then, in general, the order of the bias can not be improved near the boundary, i.e., the asymptotic bias* (11) *when* $x=\beta\kappa$ *should read as*

$$Bias[\widetilde{f}_{\beta,JLN}(\beta\kappa)] = \beta\Big[\{\mu_1'(\kappa)-\kappa\} - \int_{\mathcal{S}}\{\mu_1'(y)-y\}p_K(y;\kappa)dy\Big]f^{(1)}(0) + O(\beta^2+n^{-1}\beta^{-1})$$

(*without a shoulder condition* $f^{(1)}(0)=0$*, Assumption* A1.3[9] *plays a crucial role of the bias reduction*).

[8]Using the non-linear ρ-function ρ_c (rather than $\rho(t)=t+1$), the two-regime $\mathrm{MIG}_p/\mathrm{MLN}_p$ kernels, for $p\in[0,1]$, yield $\mu_1'(y) = p+\rho_c(y)$, whereas the two-regime $\mathrm{Amoroso}_\gamma$ kernel yields $\mu_1'(y)=\gamma\rho_c(y/\gamma)$.

[9]Generally speaking, it is difficult to solve the integral equation $\mathcal{M}(\kappa) - \int_{\mathcal{S}}\mathcal{M}(y)p_K(y;\kappa)dy = 0$ with respect to the function $\mathcal{M}$; of course, $\mathcal{M}(\cdot)\equiv$ constant (hence, the case $\mu_1'(y)-y\equiv\zeta_{1,1}$) is an exceptional solution.

Anyway, letting $r_c(t) = (c+1)(t/2)^{2/(c+1)}+1$, where $c>-1$, we numerically verify that $\mu_1'(y)-y = p+\rho_c(y)-y$ for the two-regime $\mathrm{MIG}_p/\mathrm{MLN}_p$ kernels is not the solution of the integral equation; in this case, the JLN-type bias correction does not work, unless $f^{(1)}(0)=0$. The same argument is valid for the two-regime $\mathrm{Amoroso}_\gamma$ kernel by considering the case $\mu_1'(y)=\gamma\rho_c(y/\gamma)$.

We close this section by pointing out that, in the literature, the variants of the IG/BS/LN KDEs using $K^{(IG)_S}(s;x,b)=K^{(MB)}_{1/(bx),x}(s;-1/2)$, $K^{(BS)_{JK}}(s;x,b)=(1/2)\{K^{(MB)}_{1/b,x}(s;-1/2)+K^{(MB)}_{1/b,x}(s;1/2)\}$, and $K^{(LN)_{JK}}(s;x,b)=K^{(LN)}_{\log x,\log(1+b)}(s;0)$, due to Jin and Kawczak (2003)[10] and Scaillet (2004), have been discussed, but the resulting estimators yield $\widehat{f}_b(0)=0$ by construction. Clearly, such an unrealistic constraint is not suitable for estimating the density $f(0)>0$. See also Koul and Song (2013), Marchant et al. (2013), and Saulo et al. (2013). To make matters worse, a variant of the RIG KDE using $K^{(RIG)_S}(s;x,b)=K^{(MB)}_{(x-b)/b,x-b}(s;1/2)$, due to Scaillet (2004), had the downward bias $(e^{-2}-1)f(0)$ at $x=0$; see Igarashi and Kakizawa (2014b). Also, the $\mathrm{IG}_S/\mathrm{BS}_{JK}/\mathrm{LN}_{JK}$ KDEs had the asymptotic variance $n^{-1}b^{-1/2}(2\sqrt{\pi}x^J)^{-1}f(x)$, $J=1,3/2$ (rather than $n^{-1}b^{-1/2}(2\sqrt{\pi x})^{-1}f(x)$) under $\int_0^\infty x^{-J}f(x)dx<\infty$ (any bounded continuous density f on $[0,\infty)$, with $f(0)>0$, was implicitly excluded). These problems were apparently caused by the bad parameterization; when $x=0$, the parameter $(1/(bx),x)$, $(1/b,x)$, $(\log x,\log(1+b))$ lies outside the parameter space of the IG/BS/LN density, respectively, and, when $x\in[0,b]$, $K^{(RIG)_S}(\cdot;x,b)$ is not the density. That is the reason why the authors have so far suggested a suitable parameterization for a certain parametric family $K_\theta(\cdot)$; see Examples 1–3, in such a way that, choosing a subcomponent θ_1 to be a function $\theta_1(x,b)$ of the location $x\in\mathcal{S}$ and a smoothing parameter b, the resulting density estimator $n^{-1}\sum_{i=1}^n K_{\theta_1(x,b),\theta_2}(X_i)$ shares common properties to the gamma KDE (Chen (2000)). See also Kakizawa (2018).

Remark 5 The above-mentioned "bad" asymmetric KDEs may be applied, if $f(0)=0$ is known in advance; in this case, the corresponding "bad" kernels ($\mathrm{IG}_S/\mathrm{RIG}_S/\mathrm{BS}_{JK}/\mathrm{LN}_{JK}$) share similar properties to $K(\cdot;x,\beta)$ for fixed $x\in(0,c_U]$ (the details are omitted here). We stress that, after the JLN-type bias correction, the asymptotic variances at $x\in(0,c_U]$ increase with the factor $\lambda(1/2)$, hence, incorrect asymptotic variances of Theorem 2 in Hirukawa and Sakudo (2014) (see also Funke and Kawka (2015)) should be corrected as $n^{-1}b^{-1/2}\lambda(1/2)(2\sqrt{\pi}x^J)^{-1}f(x)$.

4. The case $[0,\infty]^d$: product-type asymmetric KDE

Once the univariate case is studied in detail, the product kernel method is available for estimating the multivariate density. To illustrate it, we focus on the situation where the data $\mathbf{X}_i=(X_{i1},\ldots,X_{id})'$ is supported on $\mathcal{S}_d=[0,\infty)^d$, and construct d-variate product-type asymmetric KDE, defined by[11]

$$\widetilde{f}^{\Pi}_\beta(\mathbf{x})=\frac{1}{n}\sum_{i=1}^n\prod_{j=1}^d K(X_{ij};x_j,\beta),\quad \mathbf{x}=(x_1,\ldots,x_d)'\in\mathcal{S}_d.$$

Similar theoretical results as in one dimension can be easily derived under the following assumptions:

- (i_d) A random sample $\{\mathbf{X}_1,\ldots,\mathbf{X}_n\}$ is drawn from an unknown density f with support $\mathcal{S}_d$.

[10] Jin and Kawczak (2003) originally considered $K^{(LN)}_{\log x,4\log(1+b)}(s;0)$; but the "4" in their definition of $4\log(1+b)$ seemed to be not important. Their estimator should read as $K^{(LN)}_{\log x,\log(1+b)}(s;0)$ (of course, $K^{(LN)}_{\log x,b}(s;0)$ can be used).

[11] Although, in this paper, we adopt the single smoothing parameter β for simplicity, one can use vectors of the smoothing parameter proportional to a given vector $(\widetilde{c}_1,\ldots,\widetilde{c}_d)'$, i.e., $(\beta_1,\ldots,\beta_d)'=\beta(\widetilde{c}_1,\ldots,\widetilde{c}_d)'$, where $\widetilde{c}_j>0$ is a constant.

(ii$_d$) There exists a function φ, being four times continuously differentiable on $(\underline{c}, \infty)^d$ ($\supset \mathcal{S}_d$), such that the target density f is the restriction of φ on $\mathcal{S}_d$, and that f and $\partial_{i_1} \cdots \partial_{i_q} f$ are all bounded for $q = 1, 2, 3, 4$; $i_1, \ldots, i_q \in \{1, \ldots, d\}$, where $\partial_j = \partial/\partial x_j$. In addition, there exist constants $\eta_4 \in (0, 1]$ and $L_4 > 0$, such that $|\partial_{i_1}\partial_{i_2}\partial_{i_3}\partial_{i_4} f(\mathbf{u}) - \partial_{i_1}\partial_{i_2}\partial_{i_3}\partial_{i_4} f(\mathbf{v})| \le L_4 ||\mathbf{u} - \mathbf{v}||^{\eta_4}$ for any $\mathbf{u}, \mathbf{v} \in \mathcal{S}_d$, where $||\mathbf{u}||$ is an Euculidian norm $(\sum_{j=1}^d u_j^2)^{1/2}$ of $\mathbf{u} = (u_1, \ldots, u_d)'$.

(iii$_d$) $\beta \propto n^{-\iota_1}$ for some constant $\iota_1 \in (0, 1/d)$.

Note that one technical assumption, analogous to Assumption A4 (see Lemma 3 in the previous section), can be verified under either of the following assumptions for some constant $\iota > 0$:

f1$_d[\iota]$. $\inf_{\mathrm{x} \in \mathcal{S}_d} \{f(\mathbf{x}) \exp(C_2 ||\mathbf{x}||^\iota)\} \ge C_1$ for some constants $C_1, C_2 > 0$.

f2$_d$. $\inf_{\mathrm{x} \in \mathcal{S}_d} \{f(\mathbf{x})(1 + ||\mathbf{x}||)^{C_2}\} \ge C_1$ for some constants $C_1 > 0$ and $C_2 > 1$.

In either case, $\inf_{\mathrm{x} \in [0,R]^d} f(\mathbf{x}) > 0$ for any $R > 0$. Other assumptions related to the product-type asymmetric kernel $K^\Pi(\mathbf{s}; \mathbf{x}, \beta) = \prod_{j=1}^d K(s_j; x_j, \beta)$ are met, provided that the selected asymmetric kernel $K(\cdot; \cdot, \beta)$ in one dimensional case satisfies Assumptions A1 and A2. Here, the $j_1, \ldots, j_d$-th moment around $\mathbf{x}$ ($\in \mathcal{S}_d$) of $K^\Pi(\cdot; \mathbf{x}, \beta)$, denoted by

$$\widetilde{\mu}_{j_1,\ldots,j_d}(K^\Pi(\cdot; \mathbf{x}, \beta)) = \int_{\mathcal{S}_d} (t_1 - x_1)^{j_1} \cdots (t_d - x_d)^{j_d} K^\Pi(\mathbf{t}; \mathbf{x}, \beta) d\mathbf{t} \quad \text{(if it exists)},$$

can be written as

$$\widetilde{\mu}_{j_1,\ldots,j_d}(K^\Pi(\cdot; \mathbf{x}, \beta)) = \prod_{i=1}^d \widetilde{\mu}_{j_i}(K(\cdot; x_i, \beta)) \quad \text{due to the independence},$$

so that the cross-moments up to the fourth-order, except for the marginal moments, are given by

$$\widetilde{\mu}_{\underbrace{1,\ldots,1}_{j \text{ times}},0,\ldots,0}(K^\Pi(\cdot; \mathbf{x}, \beta)) = (\beta\zeta_{1,1})^j, \quad j = 2, 3, 4,$$

$$\widetilde{\mu}_{2,1,0,\ldots,0}(K^\Pi(\cdot; \mathbf{x}, \beta)) = \beta^2 \zeta_{1,1}\zeta_{1,2}\, x_1 x_2 + O(\beta^3) \quad \text{when } x_1/\beta \to \infty,$$

$$\widetilde{\mu}_{3,1,0,\ldots,0}(K^\Pi(\cdot; \mathbf{x}, \beta)) = \widetilde{\mu}_{2,1,1,0,\ldots,0}(K^\Pi(\cdot; \mathbf{x}, \beta)) = O(\beta^3 x_1) \quad \text{when } x_1/\beta \to \infty,$$

$$\widetilde{\mu}_{2,2,0,\ldots,0}(K^\Pi(\cdot; \mathbf{x}, \beta)) = \beta^2 \zeta_{1,2}^2\, x_1 x_2 + O(\beta^3 (x_1 + x_2)) \quad \text{when } x_1/\beta \to \infty \text{ and } x_2/\beta \to \infty$$

(there are, of course, the permutation variants, being omitted here). Similarly, the uniform/nonuniform bounds and the approximations of certain integrals (see Assumption A2(ii–iv)) are readily extended:

- $\sup_{\mathrm{x} \in \mathcal{S}_d} \sup_{\mathrm{s} \in \mathcal{S}_d} K^\Pi(\mathbf{s}; \mathbf{x}, \beta) \le (C_K \beta^{-1})^d$,
- whenever $x_1, \ldots, x_{d'} > 0$, $\sup_{x_{d'+1}, \ldots, x_d \in \mathcal{S}} \sup_{\mathrm{s} \in \mathcal{S}_d} K^\Pi(\mathbf{s}; \mathbf{x}, \beta) \le (C_K \beta^{-1})^{d-d'} \prod_{j=1}^{d'} (C'_K \beta x_j)^{-1/2}$, where $d' = 1, \ldots, d$ (the permutation variants for $(x_1, \ldots, x_d)$ are omitted here),
- given constants $\eta \in [0, 1)$ and $c_L > 0$, for all sufficiently small $\beta > 0$, $x_1, \ldots, x_d \ge c_L \beta^\eta$ imply that

$$\left| \int_{\mathcal{S}_d} \{K^\Pi(\mathbf{s}; \mathbf{x}, \beta)\}^2 d\mathbf{s} - \frac{\beta^{-d/2}}{\prod_{j=1}^d (2\sqrt{\pi x_j})} \right| \le \frac{\beta^{-d/2} M_d}{\prod_{j=1}^d \sqrt{x_j}} \sum_{j=1}^d \frac{\beta}{x_j + \beta}$$

  for some constant $M_d > 0$, independent of β and $\mathbf{x}$, and

- given constants $\eta \in [0, 1/4)$ and $0 < c_L < c_U$, for all sufficiently small $\beta > 0$, the following approximations hold for $x_1, \ldots, x_d \in [c_L \beta^\eta, c_U]$:

$$\int_{\mathcal{S}_d} K^\Pi(\mathbf{t}; \mathbf{x}, \beta) \int_{\mathcal{S}_d} K^\Pi(\mathbf{s}; \mathbf{x}, \beta) K^\Pi(\mathbf{s}; \mathbf{t}, \beta) d\mathbf{s} d\mathbf{t} = \frac{\beta^{-d/2}}{\prod_{j=1}^d (\sqrt{6\pi x_j})} + o(\beta^{-d/2} \prod_{j=1}^d x_j^{-1/2}),$$

$$\int_{\mathcal{S}_d} \int_{\mathcal{S}_d} \Big\{ \prod_{j=1}^2 K^\Pi(\mathbf{t}_j; \mathbf{x}, \beta) \Big\} \int_{\mathcal{S}_d} \Big\{ \prod_{j=1}^2 K^\Pi(\mathbf{s}; \mathbf{t}_j, \beta) \Big\} d\mathbf{s} d\mathbf{t}_1 d\mathbf{t}_2 = \frac{\beta^{-d/2}}{\prod_{j=1}^d (2\sqrt{2\pi x_j})} + o(\beta^{-d/2} \prod_{j=1}^d x_j^{-1/2}).$$

It turns out that the JLN-type bias-corrected estimator

$$\widetilde{f}_{\beta,JLN}^\Pi(\mathbf{x}) = \{\widetilde{f}_\beta^\Pi(\mathbf{x}) + \epsilon\} \frac{1}{n} \sum_{i=1}^n \frac{K^\Pi(\mathbf{X}_i; \mathbf{x}, \beta)}{\widetilde{f}_\beta^\Pi(\mathbf{X}_i) + \epsilon}, \quad \mathbf{x} \in \mathcal{S}_d$$

(we assume $\epsilon \propto \beta^{\iota_2}$ for some $\iota_2 > 1$) has the asymptotic bias and variance at $\mathbf{x} = (x_1, \ldots, x_d)'$, as follows: For all sufficiently small $\beta > 0$, we have, for $x_1, \ldots, x_d \in [c_L \beta^\eta, c_U]$,

$$Bias[\widetilde{f}_{\beta,JLN}^\Pi(\mathbf{x})] = -\beta^2 f(\mathbf{x}) \gamma_{1,d}\Big(\mathbf{x}; \frac{\gamma_{1,d}(\cdot; f)}{f(\cdot)}\Big) + o(\beta^2) + O(n^{-1}\beta^{-d/2} \prod_{j=1}^d x_j^{-1/2}), \quad \text{if } \eta \in [0,1),$$

$$V[\widetilde{f}_{\beta,JLN}^\Pi(\mathbf{x})] = n^{-1}\beta^{-d/2} \frac{\lambda_d f(\mathbf{x})}{\prod_{j=1}^d (2\sqrt{\pi x_j})} + o(n^{-1}\beta^{-d/2} \prod_{j=1}^d x_j^{-1/2}) + O(\beta^5), \quad \text{if } \eta \in [0, 1/4)$$

(note that, if $\eta \in [0,1)$, then, $V[\widetilde{f}_{\beta,JLN}^\Pi(\mathbf{x})] = O(n^{-1}\beta^{-d/2} \prod_{j=1}^d x_j^{-1/2} + \beta^5))$, where

$$\gamma_{1,d}(\mathbf{x}; f) = \sum_{j=1}^d \Big\{ \zeta_{1,1}\, \partial_j f(\mathbf{x}) + \frac{\zeta_{1,2}}{2} x_j\, \partial_j^2 f(\mathbf{x}) \Big\} \quad \text{and} \quad \lambda_d = 4 - 4\Big(\frac{\sqrt{2}}{\sqrt{3}}\Big)^d + \Big(\frac{1}{\sqrt{2}}\Big)^d.$$

Remark 6 The corresponding results when some components are near the boundary $\{0\}$ can be also obtained. For examples, when $\mathbf{x}_0 = (y_1, \ldots, y_{d'}, \beta\kappa_{d'+1}, \ldots, \beta\kappa_d)'$ (its permutation variants are omitted), where $y_1, \ldots, y_{d'} \in (0, c_U]$ and $\kappa_{d'+1}, \ldots, \kappa_d \geq 0$ are fixed,

$$V[\widetilde{f}_{\beta,JLN}^\Pi(\mathbf{x}_0)] = n^{-1}\beta^{-(d-d')-d'/2} J(\kappa_{d'+1}, \ldots, \kappa_d) \frac{f(y_1, \ldots, y_{d'}, 0, \ldots, 0)}{\prod_{j=1}^{d'} (2\sqrt{\pi y_j})}$$
$$+ o(n^{-1}\beta^{-(d-d')-d'/2}) + O(\beta^5), \quad d' = 0, 1, \ldots, d-1,$$

where

$$J(\kappa_{d'+1}, \ldots, \kappa_d) = 4 \prod_{j=d'+1}^d J_1(\kappa_j) - 4\Big(\frac{\sqrt{2}}{\sqrt{3}}\Big)^{d'} \prod_{j=d'+1}^d J_2(\kappa_j) + \Big(\frac{1}{\sqrt{2}}\Big)^{d'} \prod_{j=d'+1}^d J_3(\kappa_j).$$

The leading term of the variance of $\widetilde{f}_{\beta,JLN}^\Pi$ is different from that of the (uncorrected) estimator $\widetilde{f}_\beta^\Pi$ (actually, for the fixed interior case with $x_j > 0$, $j = 1, \ldots, d$, the asymptotic variance increases with the factor $\lambda_d > 1$); an incorrect asymptotic variance of Theorem 2.2 in Funke and Kawka (2015) (see also Zougab et al. (2018)) should be corrected so.

Remark 7 Introduce the subset $\mathcal{D}^{(d',\tau)}_{[0,c_U]^d} = \{\mathbf{x} \in [0, c_U]^d \mid$ the only d' components belong to $[0, \beta^\tau]\}$ for $d' = 1, \ldots, d$. Then, for all sufficiently small $\beta > 0$, $\mathbf{x} \in \mathcal{D}^{(d',\tau)}_{[0,c_U]^d}$ implies that

$$Bias[\widetilde{f}^{\Pi}_{\beta,JLN}(\mathbf{x})] = O(\beta^{2\tau} + n^{-1}\beta^{-d'-(d-d')/2} \prod_{x_j \in [\beta^\tau, c_U]} x_j^{-1/2}), \quad \text{if } \tau \in (1/2, 1),$$
$$V[\widetilde{f}^{\Pi}_{\beta,JLN}(\mathbf{x})] = O(n^{-1}\beta^{-d'-(d-d')/2} \prod_{x_j \in [\beta^\tau, c_U]} x_j^{-1/2} + \beta^5), \quad \text{if } \tau \in [0, 1).$$

As in the univariate case (Subsection 2.2.3), using these bounds, the different rates of the variance and the remainder term of the bias has negligible impact on the MISE (here, the integration is performed in $[0, w]^d$). That is, $MISE_w[\widetilde{f}^{\Pi}_{\beta,JLN}] = AMISE_w[\widetilde{f}^{\Pi}_{\beta,JLN}] + o(\beta^4 + n^{-1}\beta^{-1/2})$, where

$$AMISE_w[\widetilde{f}^{\Pi}_{\beta,JLN}] = \beta^4 \int_{[0,w]^d} B^2_{JLN,d}(\mathbf{x}; f)d\mathbf{x} + n^{-1}\beta^{-d/2}\lambda_d \int_{[0,w]^d} \frac{f(\mathbf{x})}{\prod_{j=1}^d (2\sqrt{\pi x_j})} d\mathbf{x}$$

with

$$B_{JLN,d}(\mathbf{x}; f) = -f(\mathbf{x})\gamma_{1,d}\Big(\mathbf{x}; \frac{\gamma_{1,d}(\cdot; f)}{f(\cdot)}\Big).$$

Using $\beta \propto n^{-2/(d+8)}$, which is feasible if $d < 8$ (see Assumption (iii$_d$)), the convergence rate $n^{-8/(d+8)}$ is achieved.

Proof Choosing constants $\tau_1 \in (4/5, 1)$ and $\tau_2 \in (0, 1/4)$, we have, for all sufficiently small $\beta > 0$,

$$\begin{aligned}
\int_{[\beta^{\tau_1},w]^d} Bias^2[\widetilde{f}^{\Pi}_{\beta,JLN}(\mathbf{x})]d\mathbf{x} &= \beta^4 \int_{[\beta^{\tau_1},w]^d} B^2_{JLN,d}(\mathbf{x}; f)d\mathbf{x} + o(n^{-1}\beta^{-d/2} + \beta^4) \\
&\quad + O((n^{-1}\beta^{-d/2})^2\{\log(1/\beta)\}^d) \\
&= \beta^4 \int_{[0,w]^d} B^2_{JLN,d}(\mathbf{x}; f)d\mathbf{x} + o(n^{-1}\beta^{-d/2} + \beta^4), \\
\int_{[\beta^{\tau_2},w]^d} V[\widetilde{f}^{\Pi}_{\beta,JLN}(\mathbf{x})]d\mathbf{x} &= n^{-1}\beta^{-d/2}\lambda_d \int_{[\beta^{\tau_2},w]^d} \frac{f(\mathbf{x})}{\prod_{j=1}^d (2\sqrt{\pi x_j})} d\mathbf{x} + o(n^{-1}\beta^{-d/2} + \beta^4) \\
&= n^{-1}\beta^{-d/2}\lambda_d \int_{[0,w]^d} \frac{f(\mathbf{x})}{\prod_{j=1}^d (2\sqrt{\pi x_j})} d\mathbf{x} + o(n^{-1}\beta^{-d/2} + \beta^4),
\end{aligned}$$

and[12]

$$\begin{aligned}
\int_{[0,w]^d \setminus [\beta^{\tau_1},w]^d} Bias^2[\widetilde{f}^{\Pi}_{\beta,JLN}(\mathbf{x})]d\mathbf{x} &= \sum_{d_L=1}^{d} O(\beta^{4\tau_1 + d_L\tau_1} + (n^{-1}\beta^{-d_L-(d-d_L)/2})^2\beta^{d_L\tau_1}\{\log(1/\beta)\}^{d-d_L}) \\
&= o(n^{-1}\beta^{-d/2} + \beta^4),
\end{aligned}$$

[12]The subset $[0, w]^d \setminus [\beta^{\tau_2}, w]^d$ consists of the following two patterns:

(I) For $d_L = 1, \ldots, d$, the d_L components belong to $[0, \beta^{\tau_1}]$, and the remaining $d - d_L$ components belong to $[\beta^{\tau_1}, w]$, and

(II) for $d_M = 1, \ldots, d$, the d_M components belong to $[\beta^{\tau_1}, \beta^{\tau_2}]$, and the remaining $d - d_M$ components belong to $[\beta^{\tau_2}, w]$.

Note that the subset $[0, w]^d \setminus [\beta^{\tau_1}, w]^d$ consists of the pattern (I) only.

$$\int_{[0,w]^d \setminus [\beta^{\tau_2},w]^d} V[\widetilde{f}^{\Pi}_{\beta,JLN}(\mathbf{x})]d\mathbf{x} = \sum_{d_L=1}^{d} O(n^{-1}\beta^{-d_L-(d-d_L)/2+d_L\tau_1+(d-d_L)\tau_1/2} + \beta^{5+d_L\tau_1})$$
$$+ \sum_{d_M=1}^{d} O(n^{-1}\beta^{-d/2+d_M\tau_1/2+(d-d_M)\tau_2/2} + \beta^{5+d_M\tau_2})$$
$$= o(n^{-1}\beta^{-d/2} + \beta^4). \quad \square$$

5. The case $[0,1]^d$: product-type beta KDE

To estimate a density f with support $[0,1]^d$, we now consider the product-type beta KDE, defined by

$$\widehat{f}^{(B)\Pi}_b(x) = \frac{1}{n}\sum_{i=1}^{n}\prod_{j=1}^{d} K^{(B)}(X_{ij}; x_j, b), \quad \mathbf{x} \in [0,1]^d$$

(we write $K^{(B)\Pi}(\mathbf{s};\mathbf{x},b) = \prod_{j=1}^d K^{(B)}(s_j; x_j, b)$), where

$$K^{(B)}(s; x, b) = \frac{s^{x/b}(1-s)^{(1-x)/b}}{B(x/b+1, (1-x)/b+1)}$$

is the beta kernel due to Chen (1999). Then, the JLN-type bias-corrected beta KDE is constructed as

$$\widehat{f}^{(B)\Pi}_{b,JLN}(\mathbf{x}) = \{\widehat{f}^{(B)\Pi}_b(\mathbf{x}) + \epsilon\}\frac{1}{n}\sum_{i=1}^{n}\frac{K^{(B)\Pi}(\mathbf{X}_i;\mathbf{x},b)}{\widehat{f}^{(B)\Pi}_b(\mathbf{X}_i)}, \quad \mathbf{x} \in [0,1]^d.$$

Let $\psi(x) = x(1-x)$. Our results in Section 2–4 can be extended, with minor modifications, as follows:

- In addition to (iii$_d$), we make the following assumptions:

  (i$^\sharp_d$) A random sample $\{\mathbf{X}_1, \ldots, \mathbf{X}_n\}$ is drawn from an unknown density f with support $[0,1]^d$.

  (ii$^\sharp_d$) There exists a function φ, being four times continuously differentiable on $(\underline{c}, \overline{c})^d$ ($\supset [0,1]^d$), such that the target density f is the restriction of φ on $[0,1]^d$, and there exist constants $\eta_4 \in (0,1]$ and $L_4 > 0$, such that $|\partial_{i_1}\partial_{i_2}\partial_{i_3}\partial_{i_4} f(\mathbf{u}) - \partial_{i_1}\partial_{i_2}\partial_{i_3}\partial_{i_4} f(\mathbf{v})| \le L_4||\mathbf{u}-\mathbf{v}||^{\eta_4}$ for any $\mathbf{u}, \mathbf{v} \in [0,1]^d$.

  (f$^\sharp_d$)[13] $\inf_{\mathbf{x}\in[0,1]^d} f(\mathbf{x}) > 0$.

- The beta kernel $K^{(B)}(s; x, b)$ in one dimensional case has no scale parameter, unlike the examples in Section 3; hence, some arguments, relying on Assumption A1, must be re-considered. Also, as variants of the properties in Assumption A2, the beta kernel $K^{(B)}(s; x, b)$ satisfies:

  (i) The jth moments around x; $\widetilde{\mu}_j(K^{(B)}(\cdot; x, b))$, $j = 1, 2, 3, 4, 6$, are given by

$$\widetilde{\mu}_j(K^{(B)}(\cdot; x, b)) = \begin{cases} b(1-2x) - 2b^2(1-2x) + O(b^3), & j = 1, \\ bx(1-x) + b^2\{2 - 11x(1-x)\} + O(b^3), & j = 2, \\ 5b^2(1-2x)x(1-x) + O(b^3), & j = 3, \\ 3b^2x^2(1-x)^2 + O(b^3), & j = 4, \\ O(b^3), & j = 6 \end{cases}$$

[13]This assumption is natural for the compact support case, in which a rigorous treatment of the integrals involving the powers of $1/f$ is a fairly easy task. On the other hand, for the previous section (the unbounded support case $[0,\infty)^d$), we believe that a rather technical assumption (e.g., Assumption A4) is indispensable, due to the unboundedness of $1/f$.

uniformly in $x \in [0,1]$ (Lemma A.1 in Igarashi (2016a)),

(ii) the mode of the density $K^{(B)}(\cdot; x, b)$ is given by x, where the uniform/nonuniform bounds for $\sup_{s\in[0,1]} K^{(B)}(s; x, b) = K^{(B)}(x; x, b)$ are available (Lemma A.3 in Igarashi (2016a))[14]:

$$1. \sup_{x\in[0,1]} K^{(B)}(x; x, b) \leq b^{-1}(1 + b), \quad 2. \text{ whenever } x \in (0,1),\ K^{(B)}(x; x, b) \leq \frac{b^{-1/2}(1 + b)}{\sqrt{2\pi\psi(x)}},$$

(iii) given constants $\eta \in [0, 1)$ and $c_L > 0$, for all sufficiently small $b > 0$, $x \in [c_L b^\eta, 1 - c_L b^\eta]$ implies that

$$\Big|\int_{\mathcal{S}} \{K^{(B)}(s; x, b)\}^2 ds - \frac{1}{2\sqrt{\pi b\psi(x)}}\Big| \leq \frac{M}{\sqrt{b\psi(x)}}\Big\{\frac{b}{\psi(x) + b}\Big\}$$

for some constant $M > 0$, independent of b and x,

(iv) given constants $\eta \in [0, 1/4)$ and $c_L > 0$, for all sufficiently small $b > 0$, the following approximations hold for $x \in [c_L b^\eta, 1 - c_L b^\eta]$:

$$\int_0^1 K^{(B)}(t; x, b) \int_0^1 K^{(B)}(s; x, b)K^{(B)}(s; t, b)dsdt = \frac{1}{\sqrt{6\pi b\psi(x)}} + o(\{b\psi(x)\}^{-1/2}),$$

$$\int_0^1 \int_0^1 \Big\{\prod_{j=1}^{2} K^{(B)}(t_j; x, b)\Big\} \int_0^1 \Big\{\prod_{j=1}^{2} K^{(B)}(s; t_j, b)\Big\}dsdt_1dt_2 = \frac{1}{2\sqrt{2\pi b\psi(x)}} + o(\{b\psi(x)\}^{-1/2})$$

(as mentioned in Section 3, these calculations can be verified by a Gaussian approximation

$$\sqrt{b\psi(x)}K^{(B)}(x + \sqrt{b\psi(x)}z_1; x + \sqrt{b\psi(x)}z_2, b) = \phi(z_1 - z_2)[1 + o(1)]$$

uniformly in $z_1, z_2 \in [-b^\tau/\sqrt{b\psi(x)}, b^\tau/\sqrt{b\psi(x)}\,]$ and $x \in [c_L b^\eta, 1 - c_L b^\tau]$; the detail is omitted), and

(iv′) for all sufficiently small $b > 0$, the following approximations hold for $x = b\kappa, 1 - b\kappa$:

$$\int_0^1 K^{(B)}(t; x, b) \int_0^1 K^{(B)}(s; x, b)K^{(B)}(s; t, b)dsdt = b^{-1}J_2^{(B)}(\kappa) + o(b^{-1}),$$

$$\int_0^1 \int_0^1 \Big\{\prod_{j=1}^{2} K^{(B)}(t_j; x, b)\Big\} \int_0^1 \Big\{\prod_{j=1}^{2} K^{(B)}(s; t_j, b)\Big\}dsdt_1dt_2 = b^{-1}J_3^{(B)}(\kappa) + o(b^{-1}),$$

where

$$J_2^{(B)}(\kappa) = \frac{1}{2^{\kappa+1}\Gamma^2(\kappa + 1)} \int_0^\infty y^\kappa \frac{e^{-y}\Gamma(y + \kappa + 1)}{2^y\Gamma(y + 1)} dy,$$

$$J_3^{(B)}(\kappa) = \frac{1}{2\Gamma^2(\kappa + 1)} \int_0^\infty \int_0^\infty y_1^\kappa y_2^\kappa \frac{e^{-(y_1+y_2)}\Gamma(y_1 + y_2 + 1)}{2^{y_1+y_2}\Gamma(y_1 + 1)\Gamma(y_2 + 1)} dy_1dy_2.$$

[14] Whenever $x \in (0, 1)$, the density $K^{(B)}(\cdot; x, b)$ has an exponential small tail as $b \searrow 0$, in the sense that

3-1. for $s_0 \in (0, x]$, $\sup_{0\leq s\leq s_0} K^{(B)}(s; x, b) \leq b^{-1}(1 + b) \exp\Big[\frac{1}{b}\Big(x \log \frac{s_0}{x} + x - s_0\Big)\Big]$,

3-2. for $s_0' \in [x, 1)$, $\sup_{s_0'\leq s\leq 1} K^{(B)}(s; x, b) \leq b^{-1}(1 + b) \exp\Big[\frac{1}{b}\Big\{(1 - x) \log \frac{1 - s_0'}{1 - x} + (1 - x) - (1 - s_0')\Big\}\Big]$.

The proof is easy, as follows: $K^{(B)}(s; x, b)/K^{(B)}(x; x, b) = (s/x)^{x/b}\{(1 - s)/(1 - x)\}^{(1-x)/b}$ is strictly increasing on $[0, x]$ (strictly decreasing on $[x, 1]$); note $\log z - z + 1 \leq 0$.

Note that the above mentioned properties have the d-variate counterparts as in Section 4 (the details are omitted to save space).

Now, we define

$$\gamma_{1,d}^{(B)}(\mathbf{x};f) = \sum_{j=1}^{d}\Big\{(1-2x_j)\partial_j f(\mathbf{x}) + \frac{1}{2}x_j(1-x_j)\partial_j^2 f(\mathbf{x})\Big\}.$$

Provided that $\epsilon \propto b^{\iota_2}$ for some constant $\iota_2 > 1$, the asymptotic bias and variance at $\mathbf{x} = (x_1, \ldots, x_d)'$ are given, as follows: For all sufficiently small $b > 0$, we have, for $x_1, \ldots, x_d \in [c_L b^{\eta}, 1 - c_L b^{\eta}]$,

$$Bias[\widehat{f}_{b,JLN}^{(B)\Pi}(\mathbf{x})] = b^2 B_{JLN,d}^{(B)}(\mathbf{x};f) + o(b^2) + O(n^{-1}b^{-d/2}\prod_{j=1}^{d}\{\psi(x_j)\}^{-1/2}), \quad \text{if } \eta \in [0,1),$$

$$V[\widehat{f}_{b,JLN}^{(B)\Pi}(\mathbf{x})] = n^{-1}b^{-d/2}\frac{\lambda_d f(\mathbf{x})}{\prod_{j=1}^{d} 2\sqrt{\pi\psi(x_j)}} + o(n^{-1}b^{-d/2}\prod_{j=1}^{d}\{\psi(x_j)\}^{-1/2}) + O(b^5), \quad \text{if } \eta \in [0,1/4)$$

(note that, if $\eta \in [0,1)$, then, $V[\widehat{f}_{b,JLN}^{(B)\Pi}(\mathbf{x})] = O(n^{-1}b^{-d/2}\prod_{j=1}^{d}\{\psi(x_j)\}^{-1/2} + b^5))$, where

$$B_{JLN,d}^{(B)}(\mathbf{x};f) = -f(\mathbf{x})\gamma_{1,d}^{(B)}\Big(\mathbf{x}; \frac{\gamma_{1,d}^{(B)}(\cdot;f)}{f(\cdot)}\Big).$$

Remark 8 The corresponding results when some components are near the boundary $\{0,1\}$ can be also obtained. For examples, when $\mathbf{x}_0 = (y_1, \ldots, y_{d'}, b\kappa_{d'+1}, \ldots, b\kappa_d)'$ (its permutation variants are omitted), where, $y_1, \ldots, y_{d'} \in (0,1)$ and $\kappa_{d'+1}, \ldots, \kappa_d \geq 0$ are fixed,

$$V[\widehat{f}_{b,JLN}^{(B)\Pi}(\mathbf{x}_0)] = n^{-1}b^{-d'/2-(d-d')}J^{(B)}(\kappa_{d'+1}, \ldots, \kappa_d)\frac{f(y_1, \ldots, y_{d'}, b\kappa_{d'+1}, \ldots, b\kappa_d)}{\prod_{j=1}^{d'}(2\sqrt{\pi\psi(y_j)})}$$
$$+o(n^{-1}b^{-d'/2-(d-d')}) + O(b^5), \quad d' = 0, 1, \ldots, d-1,$$

where

$$J^{(B)}(\kappa_{d'+1}, \ldots, \kappa_d) = 4\prod_{j=d'+1}^{d} J_1^{(B)}(\kappa_j) - 4\Big(\frac{\sqrt{2}}{\sqrt{3}}\Big)^{d'}\prod_{j=d'+1}^{d} J_2^{(B)}(\kappa_j) + \Big(\frac{1}{\sqrt{2}}\Big)^{d'}\prod_{j=d'+1}^{d} J_3^{(B)}(\kappa_j).$$

Similar results remain valid for $\mathbf{x}_0 = (y_1, \ldots, y_{d'}, b\kappa_{d'+1}, \ldots, b\kappa_{d''}, 1 - b\kappa_{d''+1}, \ldots, 1 - b\kappa_d)'$ (and its permutation variants), except that f should be evaluated at $(y_1, \ldots, y_{d'}, 0, \ldots, 0, 1, \ldots, 1)$.

The leading term of the variance of $\widehat{f}_{b,JLN}^{(B)\Pi}$ is different from that of the (uncorrected) beta KDE $\widehat{f}_{b}^{(B)\Pi}$ (actually, for the fixed interior case with $x_j \in (0,1)$, $j = 1, \ldots, d$, the asymptotic variance increases with the factor $\lambda_d > 1$); an incorrect asymptotic variance of Theorem 2.2 in Funke and Kawka (2015) (see also Hirukawa (2010)) should be corrected so. Furthermore, the MISE should read as $\int_{[0,1]^d} MSE[\widehat{f}_{b,JLN}^{(B)\Pi}(\mathbf{x})]d\mathbf{x} = AMISE[\widehat{f}_{b,JLN}^{(B)\Pi}] + o(n^{-1}b^{-d/2} + b^4)$, where

$$AMISE[\widehat{f}_{b,JLN}^{(B)\Pi}] = b^4\int_{[0,1]^d}\{B_{JLN,d}^{(B)}(\mathbf{x};f)\}^2 d\mathbf{x} + n^{-1}b^{-d/2}\lambda_d\int_{[0,1]^d}\frac{f(\mathbf{x})}{\prod_{j=1}^{d}(2\sqrt{\pi\psi(x_j)})}d\mathbf{x}$$

(the formula given by Funke and Kawka (2015) (see also Hirukawa (2010)) miss the factor λ_d). Using $b \propto n^{-2/(d+8)}$, which is feasible if $d < 8$ (see Assumption (iii$_d$)), the convergence rate $n^{-8/(d+8)}$ is achieved.

Proof Introducing $\mathcal{D}_d^{(d',\eta)} = \{\mathbf{x} \in [0,1]^d \mid \text{the only } d' \text{ components belong to } [0,b^\eta]\bigcup[1-b^\eta,1]\}$ for $d' = 1,\ldots,d$, where $\eta \in [0,1)$ is a constant, we can see that, for all sufficiently small $\beta > 0$,

$$Bias[\widehat{f}_{b,JLN}^{(B)\Pi}(\mathbf{x})] = O(b^2 + n^{-1}b^{-d'-(d-d')/2}\prod_{x_j\in[b^\eta,1-b^\eta]}\{\psi(x_j)\}^{-1/2}),$$
$$V[\widehat{f}_{b,JLN}^{(B)\Pi}(\mathbf{x})] = O(n^{-1}b^{-d'-(d-d')/2}\prod_{x_j\in[b^\eta,1-b^\eta]}\{\psi(x_j)\}^{-1/2} + b^5)$$

for $\mathbf{x} \in \mathcal{D}_d^{(d',\eta)}$. Then, choosing constants $\tau_1 \in (1/2,1)$ and $\tau_2 \in (0,1/4)$, we have, for all sufficiently small $b > 0$,

$$\begin{aligned}\int_{[b^{\tau_1},1-b^{\tau_1}]^d} Bias^2[\widehat{f}_{b,JLN}^{(B)\Pi}(\mathbf{x})]d\mathbf{x} &= b^4\int_{[b^{\tau_1},1-b^{\tau_1}]^d}\{B_{JLN,d}^{(B)}(\mathbf{x};f)\}^2 d\mathbf{x} + o(n^{-1}b^{-d/2}+b^4)\\ &\quad + O((n^{-1}b^{-d/2})^2\{\log(1/b)\}^d)\\ &= b^4\int_{[0,1]^d}\{B_{JLN,d}^{(B)}(\mathbf{x};f)\}^2 d\mathbf{x} + o(n^{-1}b^{-d/2}+b^4),\\ \int_{[b^{\tau_2},1-b^{\tau_2}]^d} V[\widehat{f}_{b,JLN}^{(B)\Pi}(\mathbf{x})]d\mathbf{x} &= n^{-1}b^{-d/2}\lambda_d\int_{[b^{\tau_2},1-b^{\tau_2}]^d}\frac{f(\mathbf{x})}{\prod_{j=1}^d(2\sqrt{\pi\psi(x_j)})}d\mathbf{x} + o(n^{-1}b^{-d/2}+b^4)\\ &= n^{-1}b^{-d/2}\lambda_d\int_{[0,1]^d}\frac{f(\mathbf{x})}{\prod_{j=1}^d(2\sqrt{\pi\psi(x_j)})}d\mathbf{x} + o(n^{-1}b^{-d/2}+b^4),\end{aligned}$$

and that[15]

$$\begin{aligned}\int_{[0,1]^d\backslash[b^{\tau_1},1-b^{\tau_1}]^d} Bias^2[\widehat{f}_{b,JLN}^{(B)\Pi}(\mathbf{x})]d\mathbf{x} &= \sum_{d_L=1}^d O(b^{4+d_L\tau_1} + (n^{-1}b^{-d_L-(d-d_L)/2})^2 b^{d_L\tau_1}\{\log(1/b)\}^{d-d_L})\\ &= o(n^{-1}b^{-d/2}+b^4),\\ \int_{[0,1]^d\backslash[b^{\tau_2},1-b^{\tau_2}]^d} V[\widehat{f}_{b,JLN}^{(B)\Pi}(\mathbf{x})]d\mathbf{x} &= \sum_{d_L=1}^d O(n^{-1}b^{-d_L-(d-d_L)/2+d_L\tau_1+(d-d_L)\tau_1/2} + b^{5+d_L\tau_1})\\ &\quad + \sum_{d_M=1}^d O(n^{-1}b^{-d/2+d_M\tau_1/2+(d-d_M)\tau_2/2} + b^{5+d_M\tau_2})\\ &= o(n^{-1}b^{-d/2}+b^4). \quad \square\end{aligned}$$

6. Simulation studies

We illustrate, through the simulations, the finite sample performance of the JLN-type bias-corrected Amoroso/IG/BS/RIG/LN[−1/2] KDEs, together with their uncorrected estimators (Examples 1–3).

[15]The subset $[0,1]^d\backslash[b^{\tau_2},1-b^{\tau_2}]^d$ consists of the following two patterns:
(I) For $d_L = 1,\ldots,d$, the d_L components belong to $[0,b^{\tau_1}]\bigcup[1-b^{\tau_1},1]$, and the remaining $d-d_L$ components belong to $[b^{\tau_1},1-b^{\tau_1}]$, and
(II) for $d_M = 1,\ldots,d$, the d_M components belong to $[b^{\tau_1},b^{\tau_2}]\bigcup[1-b^{\tau_2},1-b^{\tau_1}]$, and the remaining $d-d_M$ components belong to $[b^{\tau_2},1-b^{\tau_2}]$.
Note that the subset $[0,1]^d\backslash[b^{\tau_1},1-b^{\tau_1}]^d$ consists of the pattern (I) only.

We generated 1000 replicate samples of $n = 100, 300$ from the five densities:

$$\text{A. } f(x) = \frac{1}{2}\Big(\frac{e^{-x/3}}{3} + \frac{xe^{-x/3}}{9}\Big), \quad \text{B. } f(x) = \frac{e^{-x/3}}{3}, \quad \text{C. } f(x) = \frac{1}{2}\Big(\frac{e^{-x/10}}{10} + xe^{-x}\Big),$$

$$\text{D. } f(x) = \frac{1}{2}\Big[\frac{1}{\sqrt{2\pi}0.8x}\exp\Big\{-\frac{(\log x - 1)^2}{2(0.8)^2}\Big\} + \frac{1}{\sqrt{2\pi}0.4x}\exp\Big\{-\frac{(\log x - 2)^2}{2(0.4)^2}\Big\}\Big],$$

$$\text{E. } f(x) = \frac{2}{\sqrt{2\pi}}\exp\Big(-\frac{x^2}{2}\Big),$$

and calculated the integrated squared error (ISE); $ISE_k = \int_0^\infty \{\widetilde{f}^{[k]}(x) - f(x)\}^2 dx$ for the kth sample. Each smoothing parameter b was chosen using the least squared cross-validation method (here, we chose $\epsilon = 0.000001 \times b^2$). Tables 1–5 show that the average ISEs; $(1/1000)\sum_{k=1}^{1000} ISE_k$ decreased, as the sample size n increased. Overall, the JLN-type bias-corrected KDEs $\widehat{f}_{b,JLN}^{(\#)}$, using the linear ρ-function; $\rho(t) = t + 1$, outperformed the uncorrected estimators $\widehat{f}_b^{(\#)}$, except for cases C and D ($n = 100$). As expected, when the shoulder condition $f^{(1)}(0) = 0$ is satisfied (cases A and E), the two-regime version, denoted by $\widetilde{f}_{b,JLN}^{\star(\#)}$, using $r_{1/4}(t) = (5/4)(t/2)^{8/5} + 1$ (for the Amoroso case, $r_{1/(4|\gamma|)}(t) = \{1/(4|\gamma|)+1\}(t/2)^{2/(1/(4|\gamma|)+1)}+1$); see Subsection 3.2, also worked well, whereas, for cases B and C, having $f^{(1)}(0) \neq 0$, $\widetilde{f}_{b,JLN}^{\star(\#)}$ almost behaved worse than $\widehat{f}_{b,JLN}^{(\#)}$. Additionally, we compared the JLN-type with other bias corrections (see Igarashi and Kakizawa (2018b)), reviewed in Introduction. We observe from Table 6; the simulation results ($n = 300$) of the SS_1-type and JF_1-type bias-corrected Amoroso KDEs, that, for case B, the JLN-type outperformed the SS_1-type and JF_1-type. On the other hand, for cases A and E, since the shoulder condition $f^{(1)}(0) = 0$ is satisfied, the two-regime JLN-type worked very well. For cases C and D, it may be difficult to decide whether the JLN-type was superior or inferior to the SS_1-type and JF_1-type, since the numerical results contradicted with the graph of the AMISE asymptotic efficiency; this may be caused by the small sample size n. In summary, our simulation results confirm the bias reduction. It is worthwhile to note that the two-regime version, especially, the choice of $c = 1/4$ ($c = 1/(4|\gamma|)$ for the Amoroso case), had the advantage over the linear ρ-function, i.e., the average ISEs of $\widetilde{f}_b^{\star(\#)}$ were almost smaller than those of $\widehat{f}_b^{(\#)}$. However, the present simulation results indicate that, without the shoulder condition $f^{(1)}(0) = 0$, such a two-regime formulation is incompatible with the bias correction; this was already pointed out by Igarashi and Kakizawa (2015, 2018a) for the SS/TS/JF-types.

Appendix

For simplicity, we write $\zeta_\beta(x) = \widetilde{f}_\beta(x) - f(x) + \epsilon$, and

$$U_{x,\beta,i} = \frac{f(x) + \zeta_\beta(x)}{f(X_i) + \zeta_\beta(X_i)} K(X_i; x, \beta), \quad i = 1, \ldots, n.$$

Then,

$$E[\widetilde{f}_{\beta,JLN}(x)] = \frac{1}{n}\sum_{i=1}^{n} E[U_{x,\beta,i}] = E[U_{x,\beta,1}],$$

$$V[\widetilde{f}_{\beta,JLN}(x)] = \frac{1}{n^2}\sum_{i,j=1}^{n} Cov[U_{x,\beta,i}, U_{x,\beta,j}] = \frac{1}{n}V[U_{x,\beta,1}] + \frac{n-1}{n}Cov[U_{x,\beta,1}, U_{x,\beta,2}],$$

where, by virtue of the law of total variance/covariance,

$$V[U_{x,\beta,1}] = E\big[V[U_{x,\beta,1}|X_1]\big] + V\big[E[U_{x,\beta,1}|X_1]\big],$$
$$Cov[U_{x,\beta,1}, U_{x,\beta,2}] = E\big[Cov[U_{x,\beta,1}, U_{x,\beta,2}|X_1, X_2]\big] + Cov\big[E[U_{x,\beta,1}|X_1, X_2], E[U_{x,\beta,2}|X_1, X_2]\big].$$

Define, for $\ell = 1, 2$,

$$\mathcal{P}_\ell^{[0]}(x) = \frac{f(x)}{f(X_\ell)} \quad \text{and} \quad \mathcal{P}_\ell^{[j]}(x) = -\zeta_\beta(x)\zeta_\beta^{j-1}(X_\ell) + \frac{f(x)\zeta_\beta^j(X_\ell)}{f(X_\ell)}, \quad j = 1, 2, \ldots.$$

We use the stochastic expansion

$$\frac{f(x) + \zeta_\beta(x)}{f(X_\ell) + \zeta_\beta(X_\ell)} = \sum_{j=0}^{m}(-1)^j \frac{\mathcal{P}_\ell^{[j]}(x)}{f^j(X_\ell)} + \mathcal{R}_{[m],\ell}^{(JLN)}(x) \quad \text{for } m = 0, 2, 4.$$

To complete the proofs below, we must deal with the integrals involving the unbounded function of the power of $1/f$, as well as the asymptotic negligibility of the remainder term $\mathcal{R}_{[m],\ell}^{(JLN)}(x)$. Details are in supplemental issue: Supplemental appendix to "Multiplicative bias correction for asymmetric kernel density estimators revisited", Faculty of Economics, Hokkaido University, Discussion Paper Series A: No. 2018–328.

Proof of Proposition 1 Using the stochastic expansion (we set $m = 2$), we have

$$\begin{aligned}
E[U_{x,\beta,1}] - f(x) &= E\left[\frac{K(X_1;x,\beta)}{f(X_1)}E[\zeta_\beta(x)|X_1]\right] - f(x)E\left[\frac{K(X_1;x,\beta)}{f^2(X_1)}E[\zeta_\beta(X_1)|X_1]\right] \\
&\quad -E\left[\frac{K(X_1;x,\beta)}{f^2(X_1)}E[\zeta_\beta(x)\zeta_\beta(X_1)|X_1]\right] + f(x)E\left[\frac{K(X_1;x,\beta)}{f^3(X_1)}E[\zeta_\beta^2(X_1)|X_1]\right] \\
&\quad +E\big[K(X_1;x,\beta)E[\mathcal{R}_{[2],1}^{(JLN)}(x)|X_1]\big] \\
&= I_{1,1}(x) - f(x)I_{1,2}(x) - I_{1,3}(x) + f(x)I_{1,4}(x) + I_{1,5}(x) \quad \text{(say)}.
\end{aligned}$$

It is shown that, given constants $\eta \in [0, 1)$ and $0 < c_L < c_R$, for all sufficiently small $\beta > 0$,

$$I_{1,1}(x) - \{\epsilon + B_\beta^{(K),f}(x)\} = \begin{cases} O(n^{-1}(\beta x)^{-1/2}) & \text{for } x \in [c_L\beta^\eta, c_U], \\ O(n^{-1}\beta^{-1}) & \text{for } x = \beta\kappa, \end{cases}$$

$$I_{1,2}(x) - \int_{\mathcal{S}} \frac{K(s;x,\beta)}{f(s)}\{\epsilon + B_\beta^{(K),f}(s)\}ds = \begin{cases} O(n^{-1}(\beta x)^{-1/2}) & \text{for } x \in [c_L\beta^\eta, c_U], \\ O(n^{-1}\beta^{-1}) & \text{for } x = \beta\kappa, \end{cases}$$

$$I_{1,3}(x) - B_\beta^{(K),f}(x)\int_{\mathcal{S}} \frac{K(s;x,\beta)}{f(s)}B_\beta^{(K),f}(s)ds = \begin{cases} o(\beta^2) + O(n^{-1}(\beta x)^{-1/2}) & \text{for } x \in [c_L\beta^\eta, c_U], \\ o(\beta^2) + O(n^{-1}\beta^{-1}) & \text{for } x = \beta\kappa, \end{cases}$$

$$I_{1,4}(x) - \int_{\mathcal{S}} \frac{K(s;x,\beta)}{f^2(s)}\{B_\beta^{(K),f}(s)\}^2 ds = \begin{cases} o(\beta^2) + O(n^{-1}(\beta x)^{-1/2}) & \text{for } x \in [c_L\beta^\eta, c_U], \\ o(\beta^2) + O(n^{-1}\beta^{-1}) & \text{for } x = \beta\kappa, \end{cases}$$

$$I_{1,5}(x) = \begin{cases} o(\beta^2 + n^{-1}(\beta x)^{-1/2}) & \text{for } x \in [c_L\beta^\eta, c_U], \\ o(\beta^2 + n^{-1}\beta^{-1}) & \text{for } x = \beta\kappa. \end{cases}$$

The result follows from

$$I_{1,1}(x) - f(x)I_{1,2}(x) = \begin{cases} -\beta^2 f(x)\gamma_1\Big(x; \dfrac{\gamma_1(\cdot; f)}{f(\cdot)}\Big) + o(\beta^2) + O(n^{-1}(\beta x)^{-1/2}) & \text{for } x \in [c_L\beta^\eta, c_U], \\ \beta^2\Big[\{\mu_1'(0)\}^2 \dfrac{\{f^{(1)}(0)\}^2}{f(0)} + \Big\{-\dfrac{1}{2}\displaystyle\int_{\mathcal{S}} \mu_2'(y)p_K(y;\kappa)dy + \mu_2'(\kappa) - \dfrac{\kappa^2}{2}\Big\}f^{(2)}(0)\Big] \\ \qquad +o(\beta^2) + O(n^{-1}\beta^{-1}) \quad \text{for } x = \beta\kappa, \end{cases}$$

$$I_{1,3}(x) - f(x)I_{1,4}(x) = \begin{cases} o(\beta^2) + O(n^{-1}(\beta x)^{-1/2}) & \text{for } x \in [c_L\beta^\eta, c_U], \\ o(\beta^2) + O(n^{-1}\beta^{-1}) & \text{for } x = \beta\kappa, \end{cases}$$

noting that (7), (8),

$$\int_{\mathcal{S}} \frac{K(s;x,\beta)}{f(s)} B_\beta^{(K),f}(s)ds = \begin{cases} \beta\dfrac{\gamma_1(x;f)}{f(x)} + \beta^2\Big\{\dfrac{\gamma_2(x;f)}{f(x)} + \gamma_1\Big(x; \dfrac{\gamma_1(\cdot;f)}{f(\cdot)}\Big)\Big\} + o(\beta^2) & \text{for } x \in [c_L\beta^\eta, c_U], \\ \beta\mu_1'(0)\dfrac{f^{(1)}(0)}{f(0)} + \beta^2\Big[-\mu_1'(0)\{\mu_1'(0)+\kappa\}\Big\{\dfrac{f^{(1)}(0)}{f(0)}\Big\}^2 \\ \qquad +\dfrac{1}{2}\Big\{\displaystyle\int_{\mathcal{S}} \mu_2'(y)p_K(y;\kappa)dy - \mu_2'(\kappa)\Big\}\dfrac{f^{(2)}(0)}{f(0)}\Big] + O(\beta^3) & \text{for } x = \beta\kappa, \end{cases}$$

$$\int_{\mathcal{S}} \frac{K(s;x,\beta)}{f^2(s)} \{B_\beta^{(K),f}(s)\}^2 ds = \begin{cases} \beta^2\Big\{\dfrac{\gamma_1(x;f)}{f(x)}\Big\}^2 + o(\beta^2) & \text{for } x \in [c_L\beta^\eta, c_U], \\ \beta^2\{\mu_1'(0)\}^2\Big\{\dfrac{f^{(1)}(0)}{f(0)}\Big\}^2 + O(\beta^3) & \text{for } x = \beta\kappa, \end{cases}$$

$$\sup_{x\in[0,c_U]}\Big|\int_{\mathcal{S}} \frac{K(s;x,\beta)}{f(s)}ds - \frac{1}{f(x)}\Big| = O(\beta). \quad \square$$

Remark A.1 The vanishing of the $O(\beta^j)$-terms in $I_{1,2j-1}(x)$ and $I_{1,2j}(x)$, for $j = 1, 2$ and $x = \beta\kappa$, is ensured by Assumption A1.3. Actually, if Assumption A1.3 is replaced by A1.3$^\sharp$, then, for $x = \beta\kappa$,

$$\begin{aligned} &I_{1,1}(x) - f(x)I_{1,2}(x) \\ &= \beta\Big[\{\mu_1'(\kappa) - \kappa\} - \int_{\mathcal{S}} \{\mu_1'(y) - y\}p_K(y;\kappa)dy\Big]f^{(1)}(0) \\ &\quad +\beta^2\Big[\int_{\mathcal{S}} (y-\kappa)\{\mu_1'(y) - y\}p_K(y;\kappa)dy \frac{\{f^{(1)}(0)\}^2}{f(0)} + \Big\{-\frac{1}{2}\int_{\mathcal{S}} \mu_2'(y)p_K(y;\kappa)dy + \mu_2'(\kappa) - \frac{\kappa^2}{2}\Big\}f^{(2)}(0)\Big] \\ &\quad +o(\beta^2) + O(n^{-1}\beta^{-1}), \\ &I_{1,3}(x) - f(x)I_{1,4}(x) \\ &= \beta^2\Big[\{\mu_1'(\kappa) - \kappa\}\int_{\mathcal{S}} \{\mu_1'(y) - y\}p_K(y;\kappa)dy - \int_{\mathcal{S}} \{\mu_1'(y) - y\}^2 p_K(y;\kappa)dy\Big]\frac{\{f^{(1)}(0)\}^2}{f(0)} \\ &\quad +o(\beta^2) + O(n^{-1}\beta^{-1}). \end{aligned}$$

Proof of Proposition 2 It is easy to see that, given constants $\eta \in [0, 1)$ and $0 < c_L < c_R$, for all sufficiently small $\beta > 0$,

$$\begin{aligned} \frac{1}{n}E\big[V[U_{x,\beta,1}|X_1]\big] &\le \frac{1}{n}\Big\{\sup_{s\in\mathcal{S}} K(s;x,\beta)\Big\}E\big[K(X_1;x,\beta)E[\{\mathcal{R}_{[0],1}^{(JLN)}(x)\}^2|X_1]\big] \\ &= \begin{cases} o(n^{-1}(\beta x)^{-1/2}) & \text{for } x \in [c_L\beta^\eta, c_U], \\ o(n^{-1}\beta^{-1}) & \text{for } x = \beta\kappa. \end{cases} \end{aligned}$$

Using the stochastic expansion ($m = 0, 2, 4$), it can be shown that, for all sufficiently small $\beta > 0$,

$$\frac{1}{n}V\big[E[U_{x,\beta,1}|X_1]\big] - \frac{1}{n}\mathcal{I}_1(x) = \begin{cases} o(n^{-1}(\beta x)^{-1/2}) & \text{for } x \in [c_L\beta^\eta, c_U], \\ o(n^{-1}\beta^{-1}) & \text{for } x = \beta\kappa, \end{cases}$$

$$E\big[Cov[U_{x,\beta,1}, U_{x,\beta,2}|X_1, X_2]\big] - \frac{1}{n}\mathcal{I}_2(x) = \begin{cases} o(n^{-1}(\beta x)^{-1/2}) + O(\beta^5) & \text{for } x \in [c_L\beta^\eta, c_U], \\ o(n^{-1}\beta^{-1}) + O(\beta^5) & \text{for } x = \beta\kappa, \end{cases}$$

$$Cov\big[E[U_{x,\beta,1}|X_1, X_2], E[U_{x,\beta,2}|X_1, X_2] - \frac{1}{n}\mathcal{I}_3(x) = \begin{cases} o(n^{-1}(\beta x)^{-1/2}) + O(\beta^5) & \text{for } x \in [c_L\beta^\eta, c_U], \\ o(n^{-1}\beta^{-1}) + O(\beta^5) & \text{for } x = \beta\kappa, \end{cases}$$

where

$$\begin{aligned} \mathcal{I}_1(x) &= f^2(x)\int_{\mathcal{S}} \frac{K^2(s;x,\beta)}{f(s)}ds, \\ \mathcal{I}_2(x) &= \int_{\mathcal{S}} K^2(s;x,\beta)f(s)ds - 2f(x)\int_{\mathcal{S}}\frac{K(t;x,\beta)}{f(t)}\int_{\mathcal{S}} K(s;x,\beta)K(s;t,\beta)f(s)dsdt \\ &\quad + f^2(x)\int_{\mathcal{S}}\int_{\mathcal{S}}\frac{K(t;x,\beta)}{f(t)}\frac{K(u;x,\beta)}{f(u)}\int_{\mathcal{S}} K(s;t,\beta)K(s;u,\beta)f(s)dsdtdu, \\ \mathcal{I}_3(x) &= 2f(x)\int_{\mathcal{S}} K^2(s;x,\beta)ds - 2f^2(x)\int_{\mathcal{S}}\frac{K(t;x,\beta)}{f(t)}\int_{\mathcal{S}} K(s;x,\beta)K(s;t,\beta)dsdt \end{aligned}$$

(note that $n^{-1}\sum_{j=1}^3 \mathcal{I}_j(x) = O(n^{-1}(\beta x)^{-1/2})$ for $x \in [c_L\beta^\eta, c_U]$, if $\eta \in [0,1)$). The result follows from

$$\begin{aligned} & n^{-1}\sum_{j=1}^{3}\mathcal{I}_j(x) \\ &= \begin{cases} n^{-1}\beta^{-1/2}\Big(4 - 4\frac{\sqrt{2}}{\sqrt{3}} + \frac{1}{\sqrt{2}}\Big)\frac{f(x)}{2\sqrt{\pi x}} + o(n^{-1}(\beta x)^{-1/2}) & \text{for } x \in [c_L\beta^\eta, c_U], \text{ if } \eta \in [0, 1/4), \\ n^{-1}\beta^{-1}\{4J_1(\kappa) - 4J_2(\kappa) + J_3(\kappa)\}f(0) + o(n^{-1}\beta^{-1}) & \text{for } x = \beta\kappa. \end{cases} \qquad \square \end{aligned}$$

References

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.

Chen, S.-M., Hsu, Y.-S. and Liaw, J.-T. (2009) "On kernel estimators of density ratio", *Statistics*, 43, 463–479.

Funke, B. and Kawka, R. (2015) "Nonparametric density estimation for multivariate bounded data using two non-negative multiplicative bias correction methods", *Computational Statistics and Data Analysis*, 92, 148–162.

Hagmann, M. and Scaillet, O. (2007) "Local multiplicative bias correction for asymmetric kernel density estimators", *Journal of Econometrics*, 141, 213–249.

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”, To appear in *Journal of Nonparametric Statistics.*

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.

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.

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.

Koul, H. L. and Song, W. (2013) “Large sample results for varying kernel regression estimates”, *Journal of Nonparametric Statistics*, 25, 829–853.

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.

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.

Silverman, B. W. (1986) *Density Estimation for Statistics and Data Analysis*, London: Chapman & Hall.

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

Wand, M. P. and Jones, M. C. (1995) *Kernel Smoothing*, London: Chapman & Hall.

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.

Zougab, N., Harfouche, L., Ziane, Y. and Adjabi, S. (2018) “Multivariate generalized Birnbaum–Saunders kernel density estimators”, *Communications in Statistics - Theory and Methods*, 47, 4534–4555.

Table 1: Case A. The average ISEs$\times 10^6$.
The number in the parentheses stands for the standard deviation$\times 10^6$ of the ISEs.

n	linear	A_2	$A_{1.5}$	A_1	$A_{0.5}$	$A_{-0.5}$	A_{-1}	$A_{-1.5}$	A_{-2}	IG	BS	RIG	$LN[-1/2]$
100	$\widehat{f}_b^{(\#)}$	3389	3263	3006	2914	3201	3392	3735	4135	3316	3539	3751	3288
		(2946)	(3216)	(3144)	(3206)	(3235)	(2898)	(2674)	(2941)	(3346)	(3411)	(3468)	(3317)
	$\widehat{f}_{b,JLN}^{(\#)}$	2105	2255	2239	2497	2639	2763	3160	3583	2482	2569	2591	2515
		(2403)	(2781)	(2673)	(2738)	(2271)	(2424)	(2746)	(3035)	(2591)	(2936)	(2996)	(2741)
300	$\widehat{f}_b^{(\#)}$	1483	1399	1240	1193	1252	1450	1633	1772	1356	1460	1615	1358
		(1172)	(1269)	(1166)	(1485)	(1142)	(1220)	(1154)	(1088)	(1280)	(1287)	(1351)	(1279)
	$\widehat{f}_{b,JLN}^{(\#)}$	888	895	909	1003	1086	1099	1237	1379	990	1017	1061	995
		(887)	(989)	(1127)	(1329)	(1068)	(897)	(935)	(1030)	(905)	(999)	(985)	(936)
n	two-regime	A_2	$A_{1.5}$	A_1	$A_{0.5}$	$A_{-0.5}$	A_{-1}	$A_{-1.5}$	A_{-2}	IG	BS	RIG	$LN[-1/2]$
100	$\widetilde{f}_b^{\star(\#)}$	2422	2479	2596	2979	3258	2920	2799	2850	2710	2768	2650	2698
		(2825)	(2881)	(2997)	(3058)	(2921)	(2807)	(2537)	(2571)	(2772)	(2896)	(2764)	(2806)
	$\widetilde{f}_{b,JLN}^{\star(\#)}$	2025	2017	2000	2258	2401	2305	2417	2581	2189	2434	2565	2146
		(2795)	(2798)	(2803)	(2616)	(2604)	(2386)	(2412)	(2512)	(2706)	(2759)	(2724)	(2637)
300	$\widetilde{f}_b^{\star(\#)}$	931	961	1032	1193	1332	1154	1095	1100	1104	1084	1034	1082
		(1059)	(1105)	(1120)	(1138)	(1236)	(1032)	(919)	(935)	(1121)	(1174)	(1035)	(1095)
	$\widetilde{f}_{b,JLN}^{\star(\#)}$	728	705	677	842	935	832	867	937	761	873	937	744
		(947)	(1006)	(926)	(975)	(998)	(855)	(842)	(880)	(859)	(959)	(873)	(864)

Table 2: Case B. The average ISEs$\times 10^6$.
The number in the parentheses stands for the standard deviation$\times 10^6$ of the ISEs.

n	linear	A_2	$A_{1.5}$	A_1	$A_{0.5}$	$A_{-0.5}$	A_{-1}	$A_{-1.5}$	A_{-2}	IG	BS	RIG	$LN[-1/2]$
100	$\widehat{f}_b^{(\#)}$	6452	6480	6049	5578	5924	6652	7505	7983	6565	7184	7466	6478
		(5871)	(7440)	(7483)	(7131)	(6366)	(6154)	(6561)	(6147)	(7350)	(8495)	(7694)	(7071)
	$\widehat{f}_{b,JLN}^{(\#)}$	3168	3639	3899	4390	4548	4455	4543	5041	4088	4288	4010	3921
		(4128)	(5740)	(5521)	(6380)	(5223)	(4549)	(3780)	(4184)	(4366)	(5226)	(4941)	(4401)
300	$\widehat{f}_b^{(\#)}$	2821	2634	2374	2113	2217	2628	3043	3309	2474	2756	3014	2474
		(2599)	(2804)	(2787)	(2642)	(2062)	(2245)	(2339)	(1980)	(2352)	(2752)	(2828)	(2368)
	$\widehat{f}_{b,JLN}^{(\#)}$	1256	1296	1407	1519	1637	1696	1868	2052	1538	1562	1568	1493
		(1475)	(1524)	(1752)	(1790)	(1506)	(1388)	(1464)	(1535)	(1502)	(1706)	(1764)	(1476)
n	two-regime	A_2	$A_{1.5}$	A_1	$A_{0.5}$	$A_{-0.5}$	A_{-1}	$A_{-1.5}$	A_{-2}	IG	BS	RIG	$LN[-1/2]$
100	$\widetilde{f}_b^{\star(\#)}$	5346	5368	5242	5814	6140	5592	5688	5789	5551	5848	5783	5603
		(6183)	(7209)	(6333)	(7612)	(6582)	(5225)	(4956)	(4781)	(6274)	(6198)	(5209)	(6648)
	$\widetilde{f}_{b,JLN}^{\star(\#)}$	4684	4592	4437	4429	4749	4869	5270	5505	4718	5658	6110	4845
		(4748)	(5822)	(5583)	(6522)	(5649)	(4664)	(4780)	(4708)	(4868)	(5101)	(5023)	(6106)
300	$\widetilde{f}_b^{\star(\#)}$	2053	2043	1948	2106	2232	2128	2221	2324	2041	2179	2222	2028
		(2081)	(2408)	(2038)	(2515)	(1975)	(1870)	(1966)	(1969)	(1961)	(2056)	(1919)	(1968)
	$\widetilde{f}_{b,JLN}^{\star(\#)}$	1925	1774	1619	1521	1690	1909	2119	2288	1807	2182	2417	1817
		(1899)	(1710)	(1601)	(1685)	(1752)	(1777)	(1742)	(1836)	(1777)	(1825)	(1783)	(1840)

Table 3: Case C. The average ISEs$\times 10^6$.
The number in the parentheses stands for the standard deviation$\times 10^6$ of the ISEs.

n	linear	A_2	$A_{1.5}$	A_1	$A_{0.5}$	$A_{-0.5}$	A_{-1}	$A_{-1.5}$	A_{-2}	IG	BS	RIG	$LN[-1/2]$
100	$\widehat{f}_b^{(\#)}$	7178	6656	6045	5411	5763	6688	7454	8060	6462	6926	7545	6482
		(4902)	(4722)	(4604)	(4325)	(4565)	(4967)	(4944)	(4936)	(4970)	(5060)	(5310)	(4912)
	$\widehat{f}_{b,JLN}^{(\#)}$	8002	7052	6077	5182	5840	8709	11444	13437	7415	7734	9594	7480
		(4845)	(4818)	(4563)	(4328)	(4469)	(5297)	(6533)	(7794)	(4959)	(5287)	(5300)	(5011)
300	$\widehat{f}_b^{(\#)}$	3264	2969	2594	2234	2323	2787	3213	3577	2699	2985	3332	2703
		(2000)	(1911)	(1785)	(1697)	(1726)	(1870)	(2049)	(2178)	(1844)	(1897)	(2072)	(1852)
	$\widehat{f}_{b,JLN}^{(\#)}$	3155	2757	2343	1941	2166	3920	7310	9415	2624	2877	3394	2622
		(2172)	(1963)	(1818)	(1542)	(1782)	(3653)	(5676)	(6497)	(2149)	(2122)	(2474)	(2124)
n	two-regime	A_2	$A_{1.5}$	A_1	$A_{0.5}$	$A_{-0.5}$	A_{-1}	$A_{-1.5}$	A_{-2}	IG	BS	RIG	$LN[-1/2]$
100	$\widetilde{f}_b^{\star(\#)}$	6131	6032	5826	5539	5853	6587	6914	7138	6243	6268	6386	6276
		(3885)	(4154)	(4294)	(4308)	(4244)	(4016)	(3607)	(3426)	(4094)	(3874)	(3400)	(4027)
	$\widetilde{f}_{b,JLN}^{\star(\#)}$	6132	5920	5735	5460	6054	6881	7212	7293	6359	6356	6389	6433
		(3493)	(3657)	(3951)	(4217)	(4218)	(3473)	(3284)	(2948)	(3712)	(3621)	(3351)	(3706)
300	$\widetilde{f}_b^{\star(\#)}$	2743	2576	2379	2257	2351	2614	3219	3745	2485	2655	3008	2514
		(1729)	(1716)	(1716)	(1739)	(1765)	(1853)	(2056)	(1939)	(1758)	(1762)	(1801)	(1793)
	$\widetilde{f}_{b,JLN}^{\star(\#)}$	3028	2756	2440	2050	2325	3162	4071	4619	2906	3173	3755	2952
		(1712)	(1688)	(1676)	(1486)	(1697)	(1963)	(1891)	(1625)	(1920)	(1593)	(1276)	(1990)

Table 4: Case D. The average ISEs$\times 10^6$.
The number in the parentheses stands for the standard deviation$\times 10^6$ of the ISEs.

n	linear	A_2	$A_{1.5}$	A_1	$A_{0.5}$	$A_{-0.5}$	A_{-1}	$A_{-1.5}$	A_{-2}	IG	BS	RIG	$LN[-1/2]$
100	$\widehat{f}_b^{(\#)}$	4665	4230	3669	3168	3155	3800	4600	5371	3708	4170	4705	3721
		(2851)	(2744)	(2491)	(2295)	(2316)	(2659)	(3011)	(3326)	(2599)	(2736)	(2907)	(2586)
	$\widehat{f}_{b,JLN}^{(\#)}$	4914	4269	3584	3227	3253	4441	5817	6725	3708	4177	5083	3838
		(3030)	(2843)	(2587)	(2344)	(2493)	(3470)	(3855)	(3956)	(2794)	(2955)	(3523)	(2954)
300	$\widehat{f}_b^{(\#)}$	2124	1908	1677	1490	1476	1690	1949	2187	1662	1886	2098	1665
		(1188)	(1108)	(1022)	(965)	(961)	(1051)	(1138)	(1244)	(1020)	(1113)	(1183)	(1022)
	$\widehat{f}_{b,JLN}^{(\#)}$	2106	1828	1554	1467	1421	1648	2929	4104	1543	1772	2079	1562
		(1270)	(1144)	(1044)	(1023)	(984)	(1271)	(2836)	(3330)	(1050)	(1132)	(1272)	(1066)
n	two-regime	A_2	$A_{1.5}$	A_1	$A_{0.5}$	$A_{-0.5}$	A_{-1}	$A_{-1.5}$	A_{-2}	IG	BS	RIG	$LN[-1/2]$
100	$\widetilde{f}_b^{\star(\#)}$	4373	3882	3463	3208	3261	3865	5052	5618	3549	3976	5193	3662
		(2454)	(2448)	(2383)	(2332)	(2354)	(2719)	(2693)	(2400)	(2524)	(2466)	(2321)	(2622)
	$\widetilde{f}_{b,JLN}^{\star(\#)}$	5004	4699	4237	3875	3814	4597	5170	5507	4285	5594	5864	4471
		(1951)	(2123)	(2414)	(2432)	(2375)	(2595)	(2397)	(2364)	(2710)	(1689)	(1560)	(2741)
300	$\widetilde{f}_b^{\star(\#)}$	1660	1586	1516	1484	1499	1538	1874	2897	1511	1590	1838	1509
		(1033)	(1026)	(1010)	(987)	(1002)	(1013)	(1543)	(2281)	(1005)	(1076)	(1417)	(1010)
	$\widetilde{f}_{b,JLN}^{\star(\#)}$	1896	1724	1623	1580	1628	1741	2328	2966	1635	3160	4715	1637
		(1186)	(1089)	(1085)	(1086)	(1073)	(1158)	(1659)	(1850)	(1114)	(2090)	(1375)	(1153)

Table 5: Case E. The average ISEs$\times 10^6$.
The number in the parentheses stands for the standard deviation$\times 10^6$ of the ISEs.

n	linear	A_2	$A_{1.5}$	A_1	$A_{0.5}$	$A_{-0.5}$	A_{-1}	$A_{-1.5}$	A_{-2}	IG	BS	RIG	$LN[-1/2]$
100	$\widehat{f}_b^{(\#)}$	18180	17507	16301	16554	17644	18206	20264	21747	17373	18362	19808	17458
		(19993)	(20660)	(20420)	(20261)	(20074)	(18769)	(19102)	(17432)	(20823)	(21422)	(22564)	(21431)
	$\widehat{f}_{b,JLN}^{(\#)}$	15038	15448	15117	16095	17172	17181	20203	28981	15994	16648	17890	15809
		(15160)	(19084)	(18934)	(17961)	(16453)	(13954)	(15101)	(26322)	(14895)	(18077)	(18604)	(14917)
300	$\widehat{f}_b^{(\#)}$	7431	7027	6492	6565	7097	7403	8287	8966	6984	7361	7982	6961
		(5803)	(5940)	(5727)	(5919)	(5937)	(5527)	(5801)	(5579)	(5824)	(5826)	(6148)	(5830)
	$\widehat{f}_{b,JLN}^{(\#)}$	6287	6271	6085	6485	7235	6946	7813	11065	6472	6778	7246	6451
		(5597)	(6471)	(6578)	(6224)	(5925)	(5207)	(6337)	(11816)	(5591)	(6354)	(6789)	(5720)
n	two-regime	A_2	$A_{1.5}$	A_1	$A_{0.5}$	$A_{-0.5}$	A_{-1}	$A_{-1.5}$	A_{-2}	IG	BS	RIG	$LN[-1/2]$
100	$\widetilde{f}_b^{(\#)}$	12778	13282	14334	16933	18391	15339	14276	14089	14853	14236	13404	15064
		(18000)	(18217)	(18190)	(18570)	(18115)	(12745)	(12241)	(12211)	(15453)	(15446)	(16331)	(17311)
	$\widetilde{f}_{b,JLN}^{(\#)}$	8689	9066	9955	13492	13781	11304	11146	11441	10955	11325	10642	10437
		(13676)	(13839)	(11844)	(12669)	(12878)	(10943)	(10453)	(10544)	(11499)	(12542)	(10778)	(11068)
300	$\widetilde{f}_b^{(\#)}$	4993	5263	5840	6954	7584	6560	6069	5822	6297	5849	5511	6312
		(5666)	(5640)	(5844)	(6635)	(6110)	(5136)	(4730)	(4584)	(5846)	(5747)	(5769)	(6078)
	$\widetilde{f}_{b,JLN}^{(\#)}$	3229	3594	4303	6235	6623	4897	4519	4502	4627	4416	4211	4413
		(3626)	(4450)	(4638)	(5506)	(5144)	(4023)	(3841)	(3843)	(4372)	(4371)	(3936)	(4060)

Table 6: The average ISEs×10^6.
The number in the parentheses stands for the standard deviation×10^6 of the ISEs.

		linear	A_2	$A_{1.5}$	A_1	$A_{0.5}$	$A_{-0.5}$	A_{-1}	$A_{-1.5}$	A_{-2}
$n = 100$	Case A	$\widehat{f}^{(\#)}_{b,SS_1}$	2289	2289	2211	2393	2504	2544	2963	3471
			(2684)	(3121)	(2861)	(2847)	(2255)	(2305)	(2452)	(2733)
		$\widehat{f}^{(\#),+}_{b,SS_1}$	2288	2288	2221	2375	2504	2544	2963	3480
			(2684)	(3121)	(2935)	(2807)	(2260)	(2302)	(2451)	(2760)
		$\widehat{f}^{(\#)}_{b,JF_1}$	2120	2022	2123	2353	2630	2414	2973	3674
			(2487)	(2488)	(2945)	(2879)	(2393)	(2463)	(2682)	(3054)
	Case B	$\widehat{f}^{(\#)}_{b,SS_1}$	4088	4297	4368	4659	4673	4693	5334	5969
			(4789)	(5453)	(5791)	(6653)	(5257)	(4184)	(4556)	(4762)
		$\widehat{f}^{(\#),+}_{b,SS_1}$	4087	4299	4371	4657	4672	4692	5334	6000
			(4789)	(5460)	(5801)	(6643)	(5260)	(4184)	(4557)	(4863)
		$\widehat{f}^{(\#)}_{b,JF_1}$	3458	3656	4023	4447	4700	4159	4815	5667
			(4123)	(4990)	(5803)	(6333)	(5270)	(4240)	(4835)	(5169)
	Case C	$\widehat{f}^{(\#)}_{b,SS_1}$	7796	7204	6295	5308	5847	7164	8343	9218
			(4918)	(4930)	(4578)	(4407)	(4453)	(4503)	(4943)	(5446)
		$\widehat{f}^{(\#),+}_{b,SS_1}$	7807	7223	6298	5318	5848	7163	8350	9216
			(4900)	(4929)	(4571)	(4415)	(4454)	(4499)	(4940)	(5441)
		$\widehat{f}^{(\#)}_{b,JF_1}$	9043	8059	6805	5372	5720	6704	7870	8864
			(4567)	(5084)	(4782)	(4385)	(4379)	(4576)	(5235)	(6167)
	Case D	$\widehat{f}^{(\#)}_{b,SS_1}$	5195	4601	3841	3133	3047	3825	4617	5182
			(2818)	(2964)	(2719)	(2333)	(2384)	(2911)	(3037)	(3100)
		$\widehat{f}^{(\#),+}_{b,SS_1}$	5192	4603	3831	3120	3046	3822	4608	5153
			(2825)	(2982)	(2711)	(2301)	(2388)	(2911)	(3037)	(3097)
		$\widehat{f}^{(\#)}_{b,JF_1}$	5734	5099	4175	3275	3174	3933	4665	5334
			(2956)	(3097)	(2793)	(2355)	(2426)	(2883)	(2985)	(3206)
	Case E	$\widehat{f}^{(\#)}_{b,SS_1}$	11855	12568	13174	14023	14789	13752	14890	17189
			(12877)	(18628)	(20412)	(18530)	(15457)	(12444)	(10726)	(11400)
		$\widehat{f}^{(\#),+}_{b,SS_1}$	11866	12563	13174	13995	14779	13763	14890	17189
			(13002)	(18599)	(20359)	(18571)	(15464)	(12603)	(10726)	(11400)
		$\widehat{f}^{(\#)}_{b,JF_1}$	12321	12416	12837	14622	14962	12936	14003	18814
			(12239)	(15400)	(18734)	(18552)	(14210)	(8751)	(12315)	(18493)

Table 6: Continued.

		linear	A_2	$A_{1.5}$	A_1	$A_{0.5}$	$A_{-0.5}$	A_{-1}	$A_{-1.5}$	A_{-2}
$n=300$	Case A	$\widehat{f}^{(\#)}_{b,SS_1}$	925	879	888	949	1040	1007	1180	1357
			(979)	(1060)	(1179)	(1498)	(1080)	(816)	(883)	(883)
		$\widehat{f}^{(\#),+}_{b,SS_1}$	925	879	888	952	1040	1007	1180	1357
			(979)	(1060)	(1179)	(1503)	(1079)	(816)	(891)	(882)
		$\widehat{f}^{(\#)}_{b,JF_1}$	858	827	850	960	1044	946	1057	1269
			(916)	(926)	(1169)	(1461)	(1031)	(829)	(871)	(989)
	Case B	$\widehat{f}^{(\#)}_{b,SS_1}$	1719	1720	1622	1620	1717	1895	2183	2567
			(1719)	(2122)	(1948)	(1899)	(1752)	(1695)	(1593)	(1815)
		$\widehat{f}^{(\#),+}_{b,SS_1}$	1719	1719	1621	1620	1717	1894	2183	2567
			(1719)	(2114)	(1947)	(1899)	(1752)	(1694)	(1593)	(1815)
		$\widehat{f}^{(\#)}_{b,JF_1}$	1446	1457	1468	1604	1685	1642	1878	2174
			(1642)	(1851)	(1849)	(1985)	(1622)	(1461)	(1497)	(1572)
	Case C	$\widehat{f}^{(\#)}_{b,SS_1}$	3341	2950	2504	2049	2197	3089	4523	5527
			(2266)	(2084)	(1882)	(1679)	(1765)	(2626)	(3566)	(4027)
		$\widehat{f}^{(\#),+}_{b,SS_1}$	3346	2961	2505	2049	2197	3091	4525	5526
			(2267)	(2111)	(1884)	(1680)	(1767)	(2626)	(3565)	(4028)
		$\widehat{f}^{(\#)}_{b,JF_1}$	3540	3079	2583	2034	2200	3248	4506	5384
			(2367)	(2137)	(1917)	(1608)	(1780)	(2757)	(3680)	(4166)
	Case D	$\widehat{f}^{(\#)}_{b,SS_1}$	2252	1955	1655	1438	1408	1524	2095	2810
			(1386)	(1273)	(1113)	(979)	(988)	(1088)	(1723)	(2211)
		$\widehat{f}^{(\#),+}_{b,SS_1}$	2231	1934	1652	1438	1408	1521	2085	2804
			(1374)	(1218)	(1111)	(981)	(988)	(1092)	(1727)	(2228)
		$\widehat{f}^{(\#)}_{b,JF_1}$	2421	2105	1756	1488	1456	1657	2208	2830
			(1387)	(1277)	(1109)	(1046)	(1021)	(1102)	(1639)	(2002)
	Case E	$\widehat{f}^{(\#)}_{b,SS_1}$	5178	5137	5118	5638	6215	5765	6371	7274
			(4871)	(5745)	(6155)	(6237)	(5479)	(4447)	(4541)	(4872)
		$\widehat{f}^{(\#),+}_{b,SS_1}$	5177	5135	5116	5640	6210	5765	6371	7274
			(4871)	(5743)	(6164)	(6246)	(5480)	(4447)	(4541)	(4872)
		$\widehat{f}^{(\#)}_{b,JF_1}$	5793	5611	5389	5948	6428	6345	7158	9156
			(5171)	(5872)	(6277)	(6336)	(5521)	(4002)	(5913)	(11047)