

Department of Policy and Planning Sciences

Discussion Paper Series

No.1340

**Multivariate density estimation using a multivariate
weighted log-normal kernel**

by

Gaku IGARASHI

Feb 2016

UNIVERSITY OF TSUKUBA

Tsukuba, Ibaraki 305-8573
JAPAN

Multivariate density estimation using a multivariate weighted log-normal kernel

Gaku Igarashi

Division of Policy and Planning Sciences, Faculty of Engineering, Information and Systems,
University of Tsukuba, 1-1-1 Tennodai, Tsukuba, Ibaraki 305-8573, Japan

Abstract

This paper suggests a multivariate asymmetric kernel density estimator using a multivariate weighted log-normal (LN) kernel for non-negative multivariate data. Asymptotic properties of the multivariate weighted LN kernel estimator are studied. Simulation studies are also conducted to investigate the finite sample performance of the multivariate weighted LN kernel estimator in the bivariate situation.

Keywords: nonparametric density estimation; boundary problem; asymmetric kernel; multivariate log-normal density

MSC: 62G07; 62G20

1. Introduction

The kernel density estimation (Rosenblatt (1956)) is a well-known method to estimate a density nonparametrically. For multivariate data, the product kernel and the spherically symmetric kernel (see, e.g., Wand and Jones (1995)) are used. However, if the density has a compact or semi-infinite support, the kernel density estimator is inconsistent due to the so-called boundary bias. The boundary bias is caused by the kernel that creates a mass outside the support of the density. Some remedies for such a boundary bias problem in the univariate case were discussed by means of renormalization, reflection, generalized jackknifing, and transformation (see, e.g., Jones (1993) and Marron and Ruppert (1994)).

The asymmetric kernel density estimation is another boundary bias-free density estimation in which an asymmetric kernel, whose support matches the support of the density to be estimated, is used, unlike the standard kernel density estimation. Several univariate asymmetric kernel estimators have been introduced in the literature. Silverman (1986, p.28) mentioned the possibility of using a gamma or log-normal (LN) kernel. Chen (1999, 2000) studied beta and gamma kernel estimators. Jin and Kawczak (2003) discussed Birnbaum–Saunders (BS) and LN kernel estimators. Scaillet (2004) suggested inverse

Email: g-igarashi@sk.tsukuba.ac.jp (G. Igarashi).

Gaussian (IG) and reciprocal inverse Gaussian (RIG) kernel estimators. Mnatsakanov and Sarkisian (2012), Koul and Song (2013), Marchant et al. (2013), and Saulo et al. (2013) studied other asymmetric kernel estimators. Igarashi and Kakizawa (2014) indicated the boundary problem of the BS, IG, and RIG kernel estimators, and re-formulated these estimators to avoid the problem. Also, Igarashi (2016) pointed out the boundary problem of the LN kernel estimator and suggested a weighted LN kernel estimator that does not have the boundary problem. The multivariate asymmetric kernel density estimation was also studied. Bouezmarni and Rombouts (2010) discussed multivariate beta and gamma kernel estimators as well as a multivariate local linear kernel estimator using the product kernels.

In this paper, we suggest a multivariate asymmetric kernel estimator using a multivariate weighted LN kernel that is generally not a product kernel, and then study the asymptotic properties of the new estimator in Section 2. Simulation studies are conducted in Section 3.

2. Multivariate weighted LN kernel estimator

Throughout this paper, we make the following assumptions:

- A1. $\mathbf{X}_i = (X_{i1}, \dots, X_{id}), i = 1, \dots, n$, is a random sample from an unknown d -variate density f with non-negative support.
- A2. $b_j = b_j(n) > 0, j = 1, \dots, d$, are smoothing parameters satisfying $b_j = O(b), j = 1, \dots, d$, where $b \rightarrow 0$ and $nb^d \rightarrow \infty (n \rightarrow \infty)$.
- A3. f is of class C^2 . $f, f_j = \frac{\partial f}{\partial x_j}, f_{jk} = \frac{\partial^2 f}{\partial x_k \partial x_j}, j, k = 1, \dots, d$, are bounded, i.e., $C = \sup_{\mathbf{x} \in [0, \infty)^d} f(\mathbf{x}), C_j = \sup_{\mathbf{x} \in [0, \infty)^d} |f_j(\mathbf{x})|, C_{jk} = \sup_{\mathbf{x} \in [0, \infty)^d} |f_{jk}(\mathbf{x})|$ are finite, where $\mathbf{x} = (x_1, \dots, x_d)'$. Also, $f_{jk}, j, k = 1, \dots, d$, are Hölder continuous, i.e., there exist $L > 0$ and $\eta \in (0, 1]$ such that $|f_{jk}(\mathbf{s}) - f_{jk}(\mathbf{t})| \leq L\{(\mathbf{s} - \mathbf{t})'(\mathbf{s} - \mathbf{t})\}^{\eta/2}$ for any $\mathbf{s}, \mathbf{t} \in [0, \infty)^d$.
- A4. $\int_{[0, \infty)^d} f_j^2(\mathbf{x}) d\mathbf{x}, \int_{[0, \infty)^d} x_j x_k f_{jk}^2(\mathbf{x}) d\mathbf{x} (j, k = 1, \dots, d)$, and $\int_{[0, \infty)^d} (\prod_{j=1}^d x_j^{-1/2}) f(\mathbf{x}) d\mathbf{x}$ are finite.
- A5. $\int_{[0, \infty)^{d_0}} (\prod_{j=1}^{d_0} x_{0,j}^{q+1}) f(\mathbf{x}) d\mathbf{x}_0$ and $\int_{[0, \infty)^{d_0}} f_j^2(\mathbf{x}) d\mathbf{x}_0 (j = 1, \dots, d)$ are uniformly bounded for any remaining variables of $\mathbf{x} = (x_1, \dots, x_d)'$, where $\mathbf{x}_0 = (x_{0,1}, \dots, x_{0,d_0})' (d_0 \leq d)$ consists of arbitrary elements of \mathbf{x} , and $q > \max\{5d - 2, (d - 1)(d + 2), (1 + 2/\eta)d + 4/\eta, d^2/(2\eta) + (3/2 +$

$1/\eta)d - 2\}$, where η is given in the assumption A3.

The assumptions A2 and A3 are standard in the density estimation. The assumptions A4 and A5 ensure the convergence of the integrated squared bias and the integrated variance of the d -variate weighted LN kernel estimator (these assumptions are analogs of the assumptions in Igarashi and Kakizawa (2014) and Igarashi (2016)).

As usual, the d -variate LN density is defined as

$$K_{\boldsymbol{\mu}, \boldsymbol{\Sigma}}^{(LN)}(\mathbf{s}) = \frac{\prod_{j=1}^d s_j^{-1}}{(2\pi)^{d/2} |\boldsymbol{\Sigma}|^{1/2}} \exp\left\{-\frac{1}{2}(\log \mathbf{s} - \boldsymbol{\mu})' \boldsymbol{\Sigma}^{-1}(\log \mathbf{s} - \boldsymbol{\mu})\right\}, \quad \mathbf{s} = \begin{pmatrix} s_1 \\ \vdots \\ s_d \end{pmatrix} \in (0, \infty)^d,$$

where

$$\log \mathbf{s} = \begin{pmatrix} \log s_1 \\ \vdots \\ \log s_d \end{pmatrix}, \quad \boldsymbol{\mu} = \begin{pmatrix} \mu_1 \\ \vdots \\ \mu_d \end{pmatrix}, \quad \boldsymbol{\Sigma} = \begin{pmatrix} \sigma_{11} & \cdots & \sigma_{1d} \\ \vdots & \ddots & \vdots \\ \sigma_{d1} & \cdots & \sigma_{dd} \end{pmatrix}.$$

Also, in the same way as in the univariate case (see Igarashi (2016)), using the concept of weighted distribution (see Rao (1965) and Patil and Rao (1978)), we consider the d -variate weighted LN density

$$K_{\boldsymbol{\mu}, \boldsymbol{\Sigma}, \boldsymbol{\nu}}(\mathbf{s}) = K_{\boldsymbol{\mu}, \boldsymbol{\Sigma}}^{(LN)}(\mathbf{s}) \exp\left(-\boldsymbol{\nu}' \boldsymbol{\mu} - \frac{1}{2} \boldsymbol{\nu}' \boldsymbol{\Sigma} \boldsymbol{\nu}\right) \prod_{j=1}^d s_j^{\nu_j} = K_{\boldsymbol{\mu} + \boldsymbol{\Sigma} \boldsymbol{\nu}, \boldsymbol{\Sigma}}^{(LN)}(\mathbf{s}) \quad (\text{say}),$$

where $\boldsymbol{\nu} = (\nu_1, \dots, \nu_d)' \in \mathbb{R}^d$. Then, in this paper, we consider the d -variate weighted LN kernel estimator

$$\hat{f}_{\mathbf{b}, \rho, \boldsymbol{\nu}}(\mathbf{x}) = \frac{1}{n} \sum_{i=1}^n K_{\boldsymbol{\mu}_x, \boldsymbol{\Sigma}_x, \boldsymbol{\nu}}(\mathbf{X}_i), \quad \mathbf{x} \in [0, \infty)^d, \quad (1)$$

where

$$\mathbf{b} = \begin{pmatrix} b_1 \\ \vdots \\ b_d \end{pmatrix}, \quad \boldsymbol{\mu}_x = \begin{pmatrix} \mu_1(\mathbf{x}) \\ \vdots \\ \mu_d(\mathbf{x}) \end{pmatrix}, \quad \boldsymbol{\Sigma}_x = \begin{pmatrix} \sigma_{11}(x_1) & \cdots & \sigma_{1d}(x_1, x_d) \\ \vdots & \ddots & \vdots \\ \sigma_{d1}(x_d, x_1) & \cdots & \sigma_{dd}(x_d) \end{pmatrix},$$

$$\mu_j(\mathbf{x}) = \log(x_j + b_j) - \sum_{k \neq j}^* \nu_k \sigma_{jk}(x_j, x_k), \quad \sigma_{jj}(x_j) = \log\left(1 + \frac{b_j}{x_j + b_j}\right),$$

$$\sigma_{jk}(x_j, x_k) = \rho \{\sigma_{jj}(x_j) \sigma_{kk}(x_k)\}^{1/2}, \quad j \neq k, \quad \rho \in (-1/(d-1), 1).$$

Here, $\boldsymbol{\Sigma}_x$ has the intraclass correlation structure, i.e.,

$$\boldsymbol{\Sigma}_x = \begin{pmatrix} \sigma_{11}^{1/2}(x_1) & 0 & \cdots & 0 \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & \sigma_{dd}^{1/2}(x_d) \end{pmatrix} \begin{pmatrix} 1 & \rho & \cdots & \rho \\ \rho & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \rho \\ \rho & \cdots & \rho & 1 \end{pmatrix} \begin{pmatrix} \sigma_{11}^{1/2}(x_1) & 0 & \cdots & 0 \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & \sigma_{dd}^{1/2}(x_d) \end{pmatrix}.$$

Therefore, we can see that

$$|\Sigma_{\mathbf{x}}| = h_d(\rho) \prod_{j=1}^d \sigma_{jj}(x_j),$$

where

$$h_d(\rho) = \begin{vmatrix} 1 & \rho & \cdots & \rho \\ \rho & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \rho \\ \rho & \cdots & \rho & 1 \end{vmatrix} = (1 - \rho)^{d-1} \{1 + (d-1)\rho\}.$$

Note that by construction, the j th element in $\boldsymbol{\mu}_{\mathbf{x}} + \Sigma_{\mathbf{x}}\boldsymbol{\nu}$ is independent of x_k , ν_k , and b_k , $k \neq j$, i.e.,

$$\boldsymbol{\mu}_{\mathbf{x}} + \Sigma_{\mathbf{x}}\boldsymbol{\nu} = \begin{pmatrix} \log(x_1 + b_1) + \nu_1\sigma_{11}(x_1) \\ \vdots \\ \log(x_d + b_d) + \nu_d\sigma_{dd}(x_d) \end{pmatrix} = \begin{pmatrix} \tilde{\mu}_1(x_1) \\ \vdots \\ \tilde{\mu}_d(x_d) \end{pmatrix} = \tilde{\boldsymbol{\mu}}_{\mathbf{x}} \quad (\text{say}),$$

and that the kernel $K_{\boldsymbol{\mu}_{\mathbf{x}}, \Sigma_{\mathbf{x}}, \boldsymbol{\nu}}$ is not the product kernel unless $\rho = 0$. Figure 1 shows the shape of $K_{\boldsymbol{\mu}_{\mathbf{x}}, \Sigma_{\mathbf{x}}, \boldsymbol{\nu}}$, when $\mathbf{x} = (2, 2)$ and $\mathbf{b} = (0.5, 0.5)$. We observe that the shape varies depending on ρ and $\boldsymbol{\nu}$.

Now, we establish the asymptotic properties of the d -variate weighted LN kernel estimator. First, we derive the bias and variance of the estimator (1).

Theorem 1 *Under assumptions A1–A3, the bias of the estimator (1) is given by*

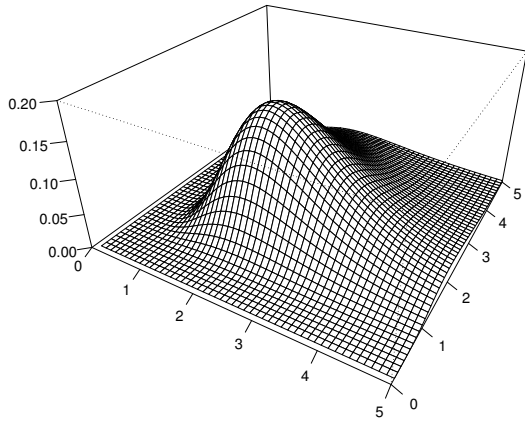
$$\begin{aligned} \text{Bias}[\hat{f}_{\mathbf{b}, \rho, \boldsymbol{\nu}}(\mathbf{x})] &= \sum_{j=1}^d \left\{ b_j \gamma_{1,j}(\mathbf{x}) + \sum_{k \neq j}^* (b_j b_k)^{1/2} \gamma_{2,j,k}(\mathbf{x}) \right\} \\ &\quad + \sum_{j=1}^d \left\{ B_{1,j}(x_j) + \sum_{k \neq j}^* B_{2,j,k}(x_j, x_k) + B_{3,j}(x_j) \right\} \end{aligned}$$

($\sum_{k \neq j}^*$ is the summation over $k = 1, \dots, d$ such that $k \neq j$), where

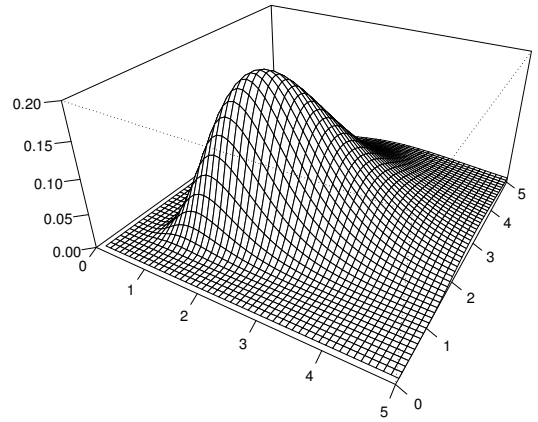
$$\begin{aligned} \gamma_{1,j}(\mathbf{x}) &= \begin{cases} \left(\nu_j + \frac{3}{2} \right) f_j(\mathbf{x}) + \frac{1}{2} x_j f_{jj}(\mathbf{x}), & \frac{x_j}{b_j} \rightarrow \infty, \\ \left\{ (\kappa_j + 1) \left(1 + \frac{1}{\kappa_j + 1} \right)^{\nu_j + 1/2} - \kappa_j \right\} f_j(\mathbf{x}), & \frac{x_j}{b_j} \rightarrow \kappa_j, \end{cases} \\ \gamma_{2,j,k}(\mathbf{x}) &= \begin{cases} \frac{\rho}{2} (x_j x_k)^{1/2} f_{jk}(\mathbf{x}), & \frac{x_j}{b_j} \rightarrow \infty, \quad \frac{x_k}{b_k} \rightarrow \infty, \\ 0, & \frac{x_j}{b_j} \rightarrow \kappa_j \text{ or } \frac{x_k}{b_k} \rightarrow \kappa_k, \end{cases} \\ B_{1,j}(x_j) &= \begin{cases} O(b_j^2 (x_j + b_j)^{-1} + b_j^2), & \frac{x_j}{b_j} \rightarrow \infty, \\ o(b_j), & \frac{x_j}{b_j} \rightarrow \kappa_j, \end{cases} \end{aligned}$$

Figure 1: Shape of $K_{\mu_x, \Sigma_x, \nu}$ when $\mathbf{x} = (2, 2)$ and $\mathbf{b} = (0.5, 0.5)$

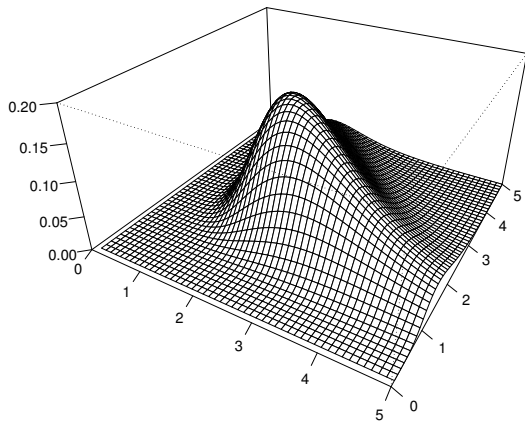
(i) $\rho = 0, \nu = (0, 0)$



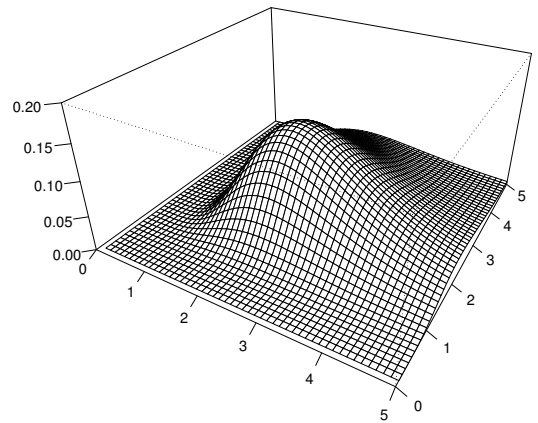
(ii) $\rho = 0.5, \nu = (0, 0)$



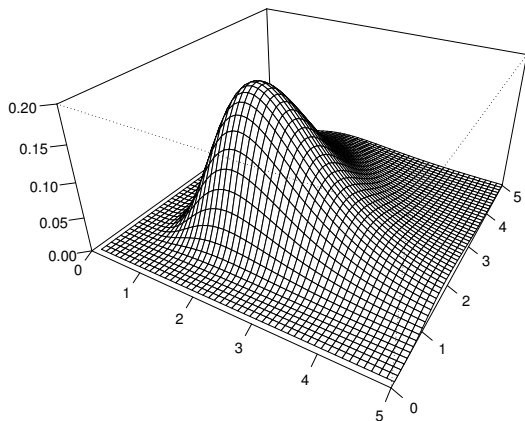
(iii) $\rho = -0.5, \nu = (0, 0)$



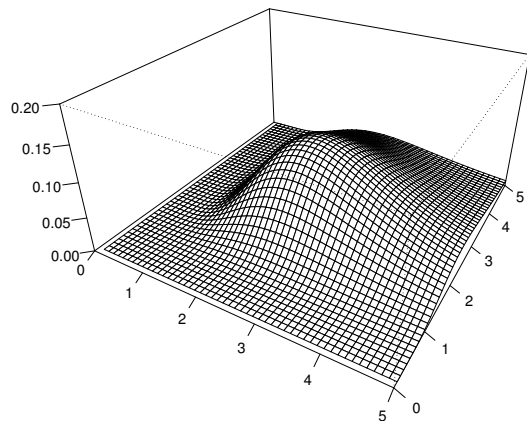
(iv) $\rho = 0, \nu = (0.5, 0.5)$



(v) $\rho = 0, \nu = (-0.5, -0.5)$



(vi) $\rho = 0, \nu = (1, 1)$



$$B_{2,j,k}(x_j, x_k) = \begin{cases} O(b_j^2 x_j^{-1} x_k + b_k^2 x_j x_k^{-1}), & \frac{x_j}{b_j} \rightarrow \infty, \quad \frac{x_k}{b_k} \rightarrow \infty, \\ O(b_j^{1/2} b_k (x_j + b_j)^{1/2}), & \frac{x_j}{b_j} \rightarrow \infty, \quad \frac{x_k}{b_k} \rightarrow \kappa_k, \\ O(b_j b_k^{1/2} (x_k + b_k)^{1/2}), & \frac{x_j}{b_j} \rightarrow \kappa_j, \quad \frac{x_k}{b_k} \rightarrow \infty, \\ O(b_j b_k), & \frac{x_j}{b_j} \rightarrow \kappa_j, \quad \frac{x_k}{b_k} \rightarrow \kappa_k, \end{cases}$$

$$B_{3,j}(x_j) = \begin{cases} O(\{b_j(x_j + b_j)\}^{\eta/2+1}), & \frac{x_j}{b_j} \rightarrow \infty, \\ O(b_j^{\eta+2}), & \frac{x_j}{b_j} \rightarrow \kappa_j, \end{cases}$$

$\kappa_j, \kappa_k \geq 0, j, k = 1, \dots, d$ are constants. Note that $O(b_j^2(x_j + b_j)^{-1})$ in $B_{1,j}(x_j)$ disappears if $\nu_j = \pm 1/2$.

Further, the variance of the estimator (1) is given by

$$V[\hat{f}_{\mathbf{b},\rho,\nu}(\mathbf{x})] = \frac{n^{-1} \prod_{j=1}^d \sigma_{j,\mathbf{b}}^2(\mathbf{x})}{2^d \pi^{d/2} h_d^{1/2}(\rho)} \left\{ f(\mathbf{x}) + \sum_{j=1}^d B_{4,j}(x_j) \right\} \left\{ 1 + \sum_{j=1}^d \sum_{k=1}^d B_{5,j,k}(x_j, x_k) \right\} + O(n^{-1}),$$

where

$$\sigma_{j,\mathbf{b}}^2(\mathbf{x}) = \begin{cases} \frac{b_j^{-1/2}}{x_j^{1/2}} \prod_{k \neq j}^* U_{j,k}, & \frac{x_j}{b_j} \rightarrow \infty, \\ \frac{b_j^{-1} (1 + \frac{1}{\kappa_j+1})^{-\nu_j+1/4}}{\{\log(1 + \frac{1}{\kappa_j+1})\}^{1/2} (\kappa_j + 1)} \prod_{k \neq j}^* U_{j,k}, & \frac{x_j}{b_j} \rightarrow \kappa_j, \end{cases}$$

$$U_{j,k} = \begin{cases} 1, & \frac{x_j}{b_j} \rightarrow \infty \text{ or } \frac{x_k}{b_k} \rightarrow \infty, \\ \exp\left[\frac{\rho}{4} \left\{ \log\left(1 + \frac{1}{\kappa_j+1}\right) \log\left(1 + \frac{1}{\kappa_k+1}\right) \right\}^{1/2}\right], & \frac{x_j}{b_j} \rightarrow \kappa_j, \quad \frac{x_k}{b_k} \rightarrow \kappa_k, \end{cases}$$

$$B_{4,j}(x_j) = \begin{cases} O(\{b_j(x_j + b_j)\}^{1/2}), & \frac{x_j}{b_j} \rightarrow \infty, \\ O(b_j), & \frac{x_j}{b_j} \rightarrow \kappa_j, \end{cases}$$

$$B_{5,j,k}(x_j, x_k) = \begin{cases} O(b_j^{1/2} b_k^{1/2} (x_j + b_j)^{-1/2} (x_k + b_k)^{-1/2}), & \frac{x_j}{b_j} \rightarrow \infty, \quad \frac{x_k}{b_k} \rightarrow \infty, \\ O(b_j^{1/2} (x_j + b_j)^{-1/2}), & \frac{x_j}{b_j} \rightarrow \infty, \quad \frac{x_k}{b_k} \rightarrow \kappa_k, \\ O(b_k^{1/2} (x_k + b_k)^{-1/2}), & \frac{x_j}{b_j} \rightarrow \kappa_j, \quad \frac{x_k}{b_k} \rightarrow \infty, \\ o(1), & \frac{x_j}{b_j} \rightarrow \kappa_j, \quad \frac{x_k}{b_k} \rightarrow \kappa_k \end{cases}$$

($\prod_{k \neq j}^*$ is the production over $k = 1, \dots, d$ such that $k \neq j$).

Next, we study the mean integrated squared error (MISE) of the estimator (1);

$$MISE[\hat{f}_{\mathbf{b},\rho,\nu}] = \int_{[0,\infty)^d} E\{[\hat{f}_{\mathbf{b},\rho,\nu}(\mathbf{x}) - f(\mathbf{x})]^2\} d\mathbf{x} = \int_{[0,\infty)^d} [\{Bias[\hat{f}_{\mathbf{b},\rho,\nu}(\mathbf{x})]\}^2 + V[\hat{f}_{\mathbf{b},\rho,\nu}(\mathbf{x})]] d\mathbf{x}.$$

Theorem 2 In addition to assumptions A1–A5, we assume that $b_j = c_j b$, $j = 1, \dots, d$, where c_j 's are positive constants. Let

$$\gamma(\mathbf{x}) = \sum_{j=1}^d \left[c_j \left\{ \left(\nu_j + \frac{3}{2} \right) f_j(\mathbf{x}) + \frac{1}{2} x_j f_{jj}(\mathbf{x}) \right\} + \frac{\rho}{2} \sum_{k \neq j}^* (c_j c_k)^{1/2} (x_j x_k)^{1/2} f_{jk}(\mathbf{x}) \right],$$

$$\sigma^2(\mathbf{x}) = f(\mathbf{x}) \frac{\prod_{j=1}^d c_j^{-1/2} x_j^{-1/2}}{2^d \pi^{d/2} h_d^{1/2}(\rho)}.$$

Then, the MISE of the estimator (1) is given by

$$MISE[\hat{f}_{\mathbf{b}, \rho, \nu}] = AMISE[\hat{f}_{\mathbf{b}, \rho, \nu}] + o(b^2 + n^{-1} b^{-d/2}),$$

where

$$AMISE[\hat{f}_{\mathbf{b}, \rho, \nu}] = b^2 \int_{[0, \infty)^d} \gamma^2(\mathbf{x}) d\mathbf{x} + n^{-1} b^{-d/2} \int_{[0, \infty)^d} \sigma^2(\mathbf{x}) d\mathbf{x}. \quad (2)$$

Minimizing (2) with respect to b , we have

$$AMISE_{opt}[\hat{f}_{\mathbf{b}, \rho, \nu}] = \frac{d+4}{d^{d/(d+4)} 4^{4/(d+4)}} \left\{ \int_{[0, \infty)^d} \gamma^2(\mathbf{x}) d\mathbf{x} \right\}^{d/(d+4)} \left\{ \int_{[0, \infty)^d} \sigma^2(\mathbf{x}) d\mathbf{x} \right\}^{4/(d+4)} n^{-4/(d+4)},$$

provided that $\gamma(\mathbf{x}) \not\equiv 0$. The optimal b is given by

$$b^* = \left\{ \frac{d \int_{[0, \infty)^d} \sigma^2(\mathbf{x}) d\mathbf{x}}{4 \int_{[0, \infty)^d} \gamma^2(\mathbf{x}) d\mathbf{x}} \right\}^{2/(d+4)} n^{-2/(d+4)}. \quad (3)$$

The best implemented AMISE of the d -variate weighted LN kernel estimator is $O(n^{-4/(d+4)})$ which is the same order as the best implemented AMISEs of the multivariate gamma and local linear kernel estimators (see Bouezmarni and Rombouts (2010)). Note that if $\nu = (-1/2, \dots, -1/2)'$ and $\rho = 0$, then the AMISE of the estimator (1) is equivalent to that of the multivariate gamma kernel estimator of Bouezmarni and Rombouts (2010).

3. Simulation results

In this section, we study, by simulation, the finite sample performance of the bivariate weighted LN kernel estimator $\hat{f}_{\mathbf{b}, \rho, \nu}$. We generated a sample of size $n = 100, 200$ from the following densities:

(A). bivariate LN density; $f(x_1, x_2) = \frac{(x_1 x_2)^{-1}}{3^{1/2} \pi} \exp \left[-\frac{2}{3} \{ (\log x_1)^2 - (\log x_1)(\log x_2 - 2) + (\log x_2 - 2)^2 \} \right],$

(B). bivariate (independent) exponential density; $f(x_1, x_2) = e^{-(x_1 + x_2)},$

(C). bivariate (independent) gamma and exponential density; $f(x_1, x_2) = \frac{x_1^3 e^{-(x_1/2 + x_2)}}{2^4 \Gamma(4)}.$

Table 1: Average ISEs $\times 10^6$ of $\hat{f}_{\mathbf{b},\rho,\boldsymbol{\nu}}$ in case (A).

The number in the parentheses indicates the standard deviation $\times 10^6$ of the ISE.

The bold-faced number indicates the smallest average ISE in each row.

n	$\boldsymbol{\nu}$	$\rho = -0.2$	$\rho = 0$	$\rho = 0.2$	$\rho = 0.4$	$\rho = 0.6$
100	(0,0)	650 (213)	631 (212)	630 (214)	649 (221)	702 (233)
	(-1/2,-1/2)	583 (168)	564 (172)	562 (177)	579 (186)	629 (200)
	(-5/4,-5/4)	527 (194)	507 (193)	504 (196)	521 (203)	572 (219)
	(-3/2,-3/2)	538 (198)	520 (197)	521 (201)	543 (209)	602 (227)
	(0,0)	428 (121)	414 (121)	412 (122)	422 (125)	454 (133)
200	(-1/2,-1/2)	378 (112)	363 (111)	359 (112)	367 (115)	394 (122)
	(-5/4,-5/4)	334 (107)	320 (105)	316 (106)	324 (110)	351 (118)
	(-3/2,-3/2)	339 (109)	326 (108)	324 (109)	335 (113)	367 (122)

We used the parameters $\rho = -0.2, 0, 0.2, 0.4, 0.6$ (in case (A)), $\rho = -0.2, -0.1, 0, 0.1, 0.2$ (in cases (B) and (C)), and $\boldsymbol{\nu} = (0, 0)', (-1/2, -1/2)', (-5/4, -5/4)', (-3/2, -3/2)'$. $\boldsymbol{\nu} = (-3/2, -3/2)'$ implies that the asymptotic integrated squared bias of the estimator (1) is independent of the first-order derivatives f_j , $j = 1, \dots, d$ (see Theorem 2). Also, in the univariate case, the AMISE of Igarashi's univariate LN $[-3/2]$ kernel estimator is smaller than that of the LN $[-1/2]$ kernel estimator. Further, the AMISE of the LN $[\nu]$ kernel estimator is theoretically minimized at $\nu = -5/4$, if $x\{f'(x)\}^2 \rightarrow 0$ ($x \rightarrow \infty$) (see Igarashi (2016)).

In the simulation (1000 replications), optimal smoothing parameter (3) with $c_1, c_2 = 1$, i.e., $(b_1, b_2) = (b^*, b^*)$, was used. We computed the average integrated squared error (ISE); $\frac{1}{1000} \sum_{k=1}^{1000} \int_{[0, \infty)^d} \{\hat{f}_{\mathbf{b},\rho,\boldsymbol{\nu}}^{[k]}(\mathbf{x}) - f(\mathbf{x})\}^2 d\mathbf{x}$, with $\hat{f}_{\mathbf{b},\rho,\boldsymbol{\nu}}^{[k]}$ being computed from the k th sample. The results of Tables 1–3 are summarized as follows:

First, as the sample size n increased, the average ISEs decreased in all cases.

Secondly, we focus on the effect of ρ . In Table 1, the average ISEs were minimum at $\rho = 0.2$ except for the sample size $n = 100$ with $\boldsymbol{\nu} = (-3/2, -3/2)'$. We guess that the small sample size caused this exception. On the other hand, in Table 2, the average ISEs were minimum at $\rho = 0, 0.1, -0.2$ when $\boldsymbol{\nu} = (-1/2, -1/2)'$ (and $(0, 0)'), (-5/4, -5/4)', (-3/2, -3/2)'$, respectively. Also, in Table 3, the

Table 2: Average ISEs $\times 10^5$ of $\hat{f}_{\mathbf{b},\rho,\nu}$ in case (B).

The number in the parentheses indicates the standard deviation $\times 10^5$ of the ISE.

The bold-faced number indicates the smallest average ISE in each row.

n	ν	$\rho = -0.2$	$\rho = -0.1$	$\rho = 0$	$\rho = 0.1$	$\rho = 0.2$
100	(0,0)	2462 (823)	2438 (822)	2430 (824)	2438 (828)	2462 (834)
	(-1/2,-1/2)	2144 (810)	2124 (811)	2119 (813)	2129 (819)	2154 (827)
	(-5/4,-5/4)	1806 (962)	1750 (953)	1720 (945)	1716 (940)	1738 (937)
	(-3/2,-3/2)	1755 (846)	1822 (853)	1910 (867)	2019 (888)	2152 (918)
200	(0,0)	1657 (510)	1639 (511)	1632 (513)	1635 (516)	1649 (520)
	(-1/2,-1/2)	1416 (497)	1401 (499)	1395 (502)	1399 (506)	1414 (512)
	(-5/4,-5/4)	1136 (583)	1100 (580)	1081 (578)	1079 (578)	1093 (579)
	(-3/2,-3/2)	1096 (505)	1140 (509)	1196 (515)	1265 (525)	1347 (537)

Table 3: Average ISEs $\times 10^6$ of $\hat{f}_{\mathbf{b},\rho,\nu}$ in case (C).

The number in the parentheses indicates the standard deviation $\times 10^6$ of the ISE.

The bold-faced number indicates the smallest average ISE in each row.

n	ν	$\rho = -0.2$	$\rho = -0.1$	$\rho = 0$	$\rho = 0.1$	$\rho = 0.2$
100	(0,0)	3585 (1154)	3539 (1148)	3519 (1146)	3523 (1147)	3552 (1152)
	(-1/2,-1/2)	3096 (1121)	3055 (1116)	3038 (1115)	3043 (1118)	3073 (1124)
	(-5/4,-5/4)	2497 (1154)	2473 (1156)	2473 (1162)	2495 (1170)	2541 (1182)
	(-3/2,-3/2)	2429 (1161)	2430 (1164)	2456 (1170)	2507 (1179)	2584 (1192)
200	(0,0)	2453 (718)	2419 (713)	2402 (710)	2402 (709)	2419 (710)
	(-1/2,-1/2)	2096 (698)	2065 (694)	2050 (692)	2050 (691)	2067 (694)
	(-5/4,-5/4)	1667 (738)	1648 (737)	1645 (737)	1658 (739)	1686 (744)
	(-3/2,-3/2)	1600 (735)	1598 (733)	1611 (734)	1641 (738)	1689 (743)

optimal ρ was 0 when $\boldsymbol{\nu} = (-1/2, -1/2)', (-5/4, -5/4)'$. When $\boldsymbol{\nu} = (-3/2, -3/2)'$ ($\boldsymbol{\nu} = (0, 0)'$), the average ISEs were minimum and second minimum at $\rho = -0.2, -0.1$ ($\rho = 0, 0.1$); however, there was little difference between them. The optimal ρ varied depending on $\boldsymbol{\nu}$ and f .

Thirdly, the effect of $\boldsymbol{\nu}$ is of interest. In Tables 1 and 2, we observed that the average ISEs were minimum at $\boldsymbol{\nu} = (-5/4, -5/4)'$, and second minimum at $\boldsymbol{\nu} = (-3/2, -3/2)'$, in cases (A) and (B). These results correspond to the theory in the univariate case (see Igarashi (2016)). However, in Table 3, when $\rho = -0.2, -0.1, 0$, the average ISEs at $\boldsymbol{\nu} = (-3/2, -3/2)'$ were smaller than those at $\boldsymbol{\nu} = (-5/4, -5/4)'$. Also, the average ISEs at $\boldsymbol{\nu} = (0, 0)'$ were the largest in the all cases. The optimal $\boldsymbol{\nu}$ varied depending on f . Furthermore, in this simulation, the bivariate weighted LN kernel estimator with $\boldsymbol{\nu} = (-5/4, -5/4)'$ or $(-3/2, -3/2)'$ performed better.

4. Conclusion

We have discussed the asymptotic properties of the multivariate asymmetric kernel estimator using the multivariate weighted LN kernel that is generally not the product kernel. The MISE of the d -variate weighted LN kernel estimator is shown to be $O(n^{-4/(d+4)})$ if the underlying density f is twice continuously differentiable. Also, we have illustrated the finite sample performance of the bivariate weighted LN kernel estimator by a simulation. From the simulation results, we have observed that the optimal parameter ρ varied depending on parameters $\boldsymbol{\nu}$ and f , and the optimal parameter $\boldsymbol{\nu}$ was $(-5/4, -5/4)'$ or $(-3/2, -3/2)'$ depending on f . In this paper, the problem of the selection of the smoothing parameter is not pursued. In practice, we must choose the smoothing parameter using the smoothing parameter selection method, for example, cross-validation or plug-in (see, e.g., Wand and Jones (1995, chapter 4)). The multivariate LN kernel estimator (with $\boldsymbol{\nu} = (0, \dots, 0)'$) was reported at the 2015 Japanese Joint Statistical Meeting (September, 2015).

Appendix. Proofs of Theorems 1 and 2

We denote by $\mathbf{Y} = (Y_1, \dots, Y_d)'$ and $\mathbf{W} = (W_1, \dots, W_d)'$ random vectors that are distributed according to the densities $K_{\boldsymbol{\mu}_x, \Sigma_x, \boldsymbol{\nu}}$ and $K_{\boldsymbol{\mu}_x, \frac{1}{2}\Sigma_x, 2\boldsymbol{\nu}-\boldsymbol{\iota}}$, respectively, where $\boldsymbol{\iota} = (1, \dots, 1)'$. Also, in this appendix, we use the univariate LN density

$$K_{\mu, \sigma^2}^{(LN)}(s) = \frac{s^{-1}}{\sqrt{2\pi\sigma^2}} \exp\left\{-\frac{1}{2\sigma^2}(\log s - \mu)^2\right\}.$$

To prove Theorems 1 and 2, we prepare the following lemmas.

Lemma A.1 For any $x_j, x_k \in [0, \infty)$, $\nu_j \in \mathbb{R}$, and $j, k = 1, \dots, d$ ($j \neq k$), we have

$$\begin{aligned}
E[Y_j - x_j] &= \begin{cases} b_j \left(\nu_j + \frac{3}{2} \right) + O(b_j^2(x_j + b_j)^{-1}), & \frac{x_j}{b_j} \rightarrow \infty, \\ b_j \left\{ (\kappa_j + 1) \left(1 + \frac{1}{\kappa_j + 1} \right)^{\nu_j + 1/2} - \kappa_j \right\} + o(b_j), & \frac{x_j}{b_j} \rightarrow \kappa_j, \end{cases} \\
E[(Y_j - x_j)^2] &= \begin{cases} b_j x_j + O(b_j^2), & \frac{x_j}{b_j} \rightarrow \infty, \\ O(b_j^2), & \frac{x_j}{b_j} \rightarrow \kappa_j, \end{cases} \\
E[(Y_j - x_j)^4] &= \begin{cases} O(\{b_j(x_j + b_j)\}^2), & \frac{x_j}{b_j} \rightarrow \infty, \\ O(b_j^4), & \frac{x_j}{b_j} \rightarrow \kappa_j, \end{cases} \\
E[(Y_j - x_j)(Y_k - x_k)] &= \begin{cases} \rho(b_j b_k x_j x_k)^{1/2} + O((b_j^3 b_k x_j^{-1} x_k)^{1/2} + (b_j b_k^3 x_j x_k^{-1})^{1/2}), & \frac{x_j}{b_j} \rightarrow \infty, \quad \frac{x_k}{b_k} \rightarrow \infty, \\ O(b_j^{1/2} b_k x_j^{1/2}), & \frac{x_j}{b_j} \rightarrow \infty, \quad \frac{x_k}{b_k} \rightarrow \kappa_k, \\ O(b_j b_k^{1/2} x_k^{1/2}), & \frac{x_j}{b_j} \rightarrow \kappa_j, \quad \frac{x_k}{b_k} \rightarrow \infty, \\ O(b_j b_k), & \frac{x_j}{b_j} \rightarrow \kappa_j, \quad \frac{x_k}{b_k} \rightarrow \kappa_k, \end{cases} \\
E[(W_j - x_j)^2] &= \begin{cases} O(b_j(x_j + b_j)), & \frac{x_j}{b_j} \rightarrow \infty, \\ O(b_j^2), & \frac{x_j}{b_j} \rightarrow \kappa_j, \end{cases}
\end{aligned}$$

where $\kappa_j, \kappa_k \geq 0$ are constants. Especially, for $\nu_j = \pm 1/2$,

$$E[Y_j - x_j] = b_j \left(\nu_j + \frac{3}{2} \right).$$

Proof Use

$$\begin{aligned}
E[Y_j^q] &= \exp \left\{ q \log(x_j + b_j) + \left(\nu_j + \frac{q}{2} \right) q \log \left(1 + \frac{b_j}{x_j + b_j} \right) \right\}, \quad q \in \mathbb{R}, \\
E[Y_j Y_k] &= E[Y_j] E[Y_k] \exp \left[\rho \left\{ \log \left(1 + \frac{b_j}{x_j + b_j} \right) \log \left(1 + \frac{b_k}{x_k + b_k} \right) \right\}^{1/2} \right], \\
E[W_j^q] &= \exp \left[q \log(x_j + b_j) + \left(\nu_j - \frac{1}{2} + \frac{q}{4} \right) q \log \left(1 + \frac{b_j}{x_j + b_j} \right) \right. \\
&\quad \left. - \frac{q}{2} \sum_{k \neq j}^* \rho \left\{ \log \left(1 + \frac{b_j}{x_j + b_j} \right) \log \left(1 + \frac{b_k}{x_k + b_k} \right) \right\}^{1/2} \right], \quad q \in \mathbb{R}. \quad \square
\end{aligned}$$

Lemma A.2 For any $\rho \in (-1/(d-1), 1)$ and $\boldsymbol{\nu} = (\nu_1, \dots, \nu_d)' \in \mathbb{R}^d$, we have

$$K_{\boldsymbol{\mu}_x, \Sigma_x, \boldsymbol{\nu}}(\mathbf{s}) \leq \frac{d^{d/2}}{h_d^{1/2}(\rho)} \prod_{j=1}^d K_{\tilde{\mu}_j(x_j), d\sigma_{jj}(x_j)}^{(LN)}(s_j).$$

Proof Let

$$R = \begin{pmatrix} 1 & \rho & \cdots & \rho \\ \rho & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \rho \\ \rho & \cdots & \rho & 1 \end{pmatrix}, \quad \mathbf{t} = \begin{pmatrix} \frac{\log s_1 - \tilde{\mu}_1(x_1)}{\sigma_{11}^{1/2}} \\ \vdots \\ \frac{\log s_d - \tilde{\mu}_d(x_d)}{\sigma_{dd}^{1/2}} \end{pmatrix}.$$

Noting that the eigenvalues of R are $1 - \rho$ (multiplicity $d - 1$) and $1 + (d - 1)\rho$, we have

$$(\log \mathbf{s} - \tilde{\boldsymbol{\mu}}_{\mathbf{x}})' \Sigma_{\mathbf{x}}^{-1} (\log \mathbf{s} - \tilde{\boldsymbol{\mu}}_{\mathbf{x}}) = \mathbf{t}' R^{-1} \mathbf{t} \geq \frac{\mathbf{t}' \mathbf{t}}{v_{d,\rho}} \geq \frac{\mathbf{t}' \mathbf{t}}{d},$$

where

$$v_{d,\rho} = \begin{cases} 1 - \rho, & -1/(d-1) < \rho < 0, \\ 1 + (d-1)\rho, & 0 \leq \rho < 1. \end{cases}$$

It follows that

$$\begin{aligned} K_{\boldsymbol{\mu}_{\mathbf{x}}, \Sigma_{\mathbf{x}}, \boldsymbol{\nu}}(\mathbf{s}) &\leq \frac{1}{h_d^{1/2}(\rho)} \prod_{j=1}^d \frac{s_j^{-1}}{\{2\pi\sigma_{jj}(x_j)\}^{1/2}} \exp\left[-\frac{\{\log s_j - \tilde{\mu}_j(x_j)\}^2}{2d\sigma_{jj}(x_j)}\right] \\ &= \frac{d^{d/2}}{h_d^{1/2}(\rho)} \prod_{j=1}^d K_{\tilde{\mu}_j(x_j), d\sigma_{jj}(x_j)}^{(LN)}(s_j). \quad \square \end{aligned}$$

Lemma A.3 For any $\boldsymbol{\nu} \in \mathbb{R}^d$, there exists a constant $L_{d,\rho,\boldsymbol{\nu}} > 0$, independent of \mathbf{b} , such that

$$\sup_{\mathbf{x} \in [0, \infty)^d} \sup_{\mathbf{s} \in [0, \infty)^d} K_{\boldsymbol{\mu}_{\mathbf{x}}, \Sigma_{\mathbf{x}}, \boldsymbol{\nu}}(\mathbf{s}) \leq L_{d,\rho,\boldsymbol{\nu}} \prod_{j=1}^d b_j^{-1}.$$

Proof Using Lemma A.2, it suffices to bound $K_{\tilde{\mu}_j(x_j), d\sigma_{jj}(x_j)}^{(LN)}(s_j)$, as in Lemma 4 of Igarashi (2016).

The detail is omitted here. \square

Lemma A.4 For any $\tau \in (0, 1)$, $q > 0$, and $j = 1, \dots, d$, we have

$$\int_{b^{-\tau}}^{\infty} K_{\tilde{\mu}_j(x_j), d\sigma_{jj}(x_j)}^{(LN)}(s_j) dx_j \leq (b^\tau s_j)^{q+1} (1 + b^\tau b_j)^{\{\nu_j - d(q+2)\}^2 / (2d)}, \quad s_j \geq 0.$$

Proof In the spirit of Igarashi (2016), it is easy to see that for $x_j \geq b^{-\tau}$,

$$\begin{aligned} K_{\tilde{\mu}_j(x_j), d\sigma_{jj}(x_j)}^{(LN)}(s_j) &= \frac{s_j^{-1}}{\{2d\pi\sigma_{jj}(x_j)\}^{1/2}} \exp\left[-\frac{1}{2d\sigma_{jj}(x_j)} \{\log s_j - \log(x_j + b_j) - \nu_j\sigma_{jj}(x_j)\}^2\right] \\ &= \frac{s_j(x_j + b_j)^{-1}}{(2d\pi)^{1/2}} \exp\left[-\frac{1}{2d\sigma_{jj}(x_j)} \{\log s_j - \log(x_j + b_j)\}^2\right. \\ &\quad \left. + \left(\frac{\nu_j}{d} - 2\right) \{\log s_j - \log(x_j + b_j)\}\right. \\ &\quad \left. - \log b_j + G(x_j/b_j + 1) - \frac{\nu_j^2}{2d}\sigma_{jj}(x_j)\right] \\ &\leq \frac{b^\tau s_j (1 + b^\tau b_j)^{\{\nu_j - 2d\}^2 / (2d)} (x_j + b_j)^{-1}}{\{2d\pi \log(1 + b^\tau b_j)\}^{1/2}} \end{aligned}$$

$$\begin{aligned}
& \times \exp \left[- \frac{\{\log s_j - \log(x_j + b_j) - (\nu_j - 2d) \log(1 + b^\tau b_j)\}^2}{2d \log(1 + b^\tau b_j)} \right] \\
& = b^\tau s_j (1 + b^\tau b_j)^{(\nu_j - 2d)^2 / (2d)} K_{\log s_j - (\nu_j - 2d) \log(1 + b^\tau b_j), d \log(1 + b^\tau b_j)}^{(LN)}(x_j + b_j),
\end{aligned}$$

where

$$G(u) = -\log u - \frac{1}{2} \log\{\log(1 + 1/u)\}$$

is a decreasing function. It follows that

$$\begin{aligned}
& \int_{b^{-\tau}}^{\infty} K_{\tilde{\mu}_j(x_j), d\sigma_{jj}(x_j)}^{(LN)}(s_j) dx_j \\
& \leq b^\tau s_j (1 + b^\tau b_j)^{(\nu_j - 2d)^2 / (2d)} \int_{b^{-\tau}}^{\infty} K_{\log s_j - (\nu_j - 2d) \log(1 + b^\tau b_j), d \log(1 + b^\tau b_j)}^{(LN)}(y) dy \\
& \leq b^{\tau(q+1)} s_j (1 + b^\tau b_j)^{(\nu_j - 2d)^2 / (2d)} \int_0^{\infty} y^q K_{\log s_j - (\nu_j - 2d) \log(1 + b^\tau b_j), d \log(1 + b^\tau b_j)}^{(LN)}(y) dy \\
& = (b^\tau s_j)^{q+1} (1 + b^\tau b_j)^{\{\nu_j - d(q+2)\}^2 / (2d)}. \quad \square
\end{aligned}$$

Proof of Theorem 1 From Lemma A.1, we have

$$\begin{aligned}
& E[\hat{f}_{\mathbf{b}, \rho, \nu}(\mathbf{x})] \\
& = f(\mathbf{x}) + \sum_{j=1}^d f_j(\mathbf{x}) E[Y_j - x_j] + \frac{1}{2} \sum_{j=1}^d \sum_{k=1}^d f_{jk}(\mathbf{x}) E[(Y_j - x_j)(Y_k - x_k)] \\
& \quad + \sum_{j=1}^d \sum_{k=1}^d \int_{[0, \infty)^d} (s_j - x_j)(s_k - x_k) K_{\mu_{\mathbf{x}}, \Sigma_{\mathbf{x}}, \nu}(\mathbf{s}) \int_0^1 \{f_{jk}(\mathbf{x} + \theta(\mathbf{s} - \mathbf{x})) - f_{jk}(\mathbf{x})\} (1 - \theta) d\theta ds \\
& = f(\mathbf{x}) + \sum_{j=1}^d \left\{ b_j \gamma_{1,j}(\mathbf{x}) + \sum_{k \neq j}^* (b_j b_k)^{1/2} \gamma_{2,j,k}(\mathbf{x}) \right\} + \sum_{j=1}^d \left\{ B_{1,j}(x_j) + \sum_{k \neq j}^* B_{2,j,k}(x_j, x_k) + B_{3,j}(x_j) \right\},
\end{aligned}$$

since

$$\begin{aligned}
& \left| \sum_{j=1}^d \sum_{k=1}^d \int_{[0, \infty)^d} (s_j - x_j)(s_k - x_k) K_{\mu_{\mathbf{x}}, \Sigma_{\mathbf{x}}, \nu}(\mathbf{s}) \int_0^1 \{f_{jk}(\mathbf{x} + \theta(\mathbf{s} - \mathbf{x})) - f_{jk}(\mathbf{x})\} (1 - \theta) d\theta ds \right| \\
& \leq \frac{L}{2} E \left[\{(\mathbf{Y} - \mathbf{x})'(\mathbf{Y} - \mathbf{x})\}^{\eta/2} \sum_{j=1}^d \sum_{k=1}^d |Y_j - x_j| |Y_k - x_k| \right] \\
& \leq \frac{L}{4} E \left[\{(\mathbf{Y} - \mathbf{x})'(\mathbf{Y} - \mathbf{x})\}^{\eta/2} \sum_{j=1}^d \sum_{k=1}^d \{(Y_j - x_j)^2 + (Y_k - x_k)^2\} \right] \\
& \leq \frac{Ld^{\eta/2+1}}{2} \sum_{j=1}^d E[(Y_j - x_j)^{\eta+2}] \\
& \leq \frac{Ld^{\eta/2+1}}{2} \sum_{j=1}^d \{E[(Y_j - x_j)^4]\}^{(\eta+2)/4} \\
& = \sum_{j=1}^d B_{3,j}(x_j).
\end{aligned}$$

On the other hand, we can see that

$$\begin{aligned}
& V[\hat{f}_{\mathbf{b},\rho,\nu}(\mathbf{x})] \\
&= n^{-1} \int_{[0,\infty)^d} \{K_{\boldsymbol{\mu}_{\mathbf{x}},\Sigma_{\mathbf{x}},\nu}(\mathbf{s})\}^2 f(\mathbf{s}) d\mathbf{s} + O(n^{-1}) \\
&= n^{-1} \frac{\exp\{-\boldsymbol{\nu}'(\boldsymbol{\mu}_{\mathbf{x}} + \Sigma_{\mathbf{x}}\boldsymbol{\nu}) + \frac{1}{4}\boldsymbol{\nu}'\Sigma_{\mathbf{x}}\boldsymbol{\nu}\}}{2^d \pi^{d/2} |\Sigma_{\mathbf{x}}|^{1/2}} \int_{[0,\infty)^d} K_{\boldsymbol{\mu}_{\mathbf{x}},\frac{1}{2}\Sigma_{\mathbf{x}},2\nu-\boldsymbol{\nu}}(\mathbf{s}) f(\mathbf{s}) d\mathbf{s} + O(n^{-1}) \\
&= \frac{n^{-1}}{2^d \pi^{d/2} h_d^{1/2}(\rho)} \int_{[0,\infty)^d} K_{\boldsymbol{\mu}_{\mathbf{x}},\frac{1}{2}\Sigma_{\mathbf{x}},2\nu-\boldsymbol{\nu}}(\mathbf{s}) f(\mathbf{s}) d\mathbf{s} \prod_{j=1}^d \frac{(1 + \frac{b_j}{x_j + b_j})^{-\nu_j + 1/4} \exp\{\frac{1}{4} \sum_{k \neq j}^* \sigma_{jk}(x_j, x_k)\}}{\sigma_{jj}^{1/2}(x_j)(x_j + b_j)} \\
&\quad + O(n^{-1}).
\end{aligned}$$

The variance of $\hat{f}_{\mathbf{b},\rho,\nu}(\mathbf{x})$ follows from

$$\begin{aligned}
& \left| \int_{[0,\infty)^d} K_{\boldsymbol{\mu}_{\mathbf{x}},\frac{1}{2}\Sigma_{\mathbf{x}},2\nu-\boldsymbol{\nu}}(\mathbf{s}) \sum_{j=1}^d (s_j - x_j) \int_0^1 f_j(\mathbf{x} + \theta(\mathbf{s} - \mathbf{x})) d\theta d\mathbf{s} \right| \\
&\leq \sum_{j=1}^d C_j E[|W_j - x_j|] \\
&\leq \sum_{j=1}^d C_j \{E[(W_j - x_j)^2]\}^{1/2} \\
&= \sum_{j=1}^d B_{4,j}(x_j)
\end{aligned}$$

(we used Lemma A.1), where

$$\begin{aligned}
& \frac{(1 + \frac{b_j}{x_j + b_j})^{-\nu_j + 1/4} \exp\{\frac{1}{4} \sum_{k \neq j}^* \sigma_{jk}(x_j, x_k)\}}{\sigma_{jj}^{1/2}(x_j)(x_j + b_j)} \\
&= \frac{(1 + \frac{b_j}{x_j + b_j})^{-\nu_j + 1/4}}{\sigma_{jj}^{1/2}(x_j)(x_j + b_j)} \exp\left[\frac{1}{4} \sum_{k \neq j}^* \rho \left\{ \log\left(1 + \frac{b_j}{x_j + b_j}\right) \log\left(1 + \frac{b_k}{x_k + b_k}\right) \right\}^{1/2}\right] \\
&= \begin{cases} \frac{b_j^{-1/2}}{(x_j + b_j)^{1/2}} \{1 + O(b_j(x_j + b_j)^{-1})\} \prod_{k \neq j}^* U_{j,k} \{1 + B_{5,j,k}(x_j, x_k)\}, & \frac{x_j}{b_j} \rightarrow \infty, \\ \frac{b_j^{-1} (1 + \frac{1}{\kappa_j + 1})^{-\nu_j + 1/4}}{\{\log(1 + \frac{1}{\kappa_j + 1})\}^{1/2} (\kappa_j + 1)} \{1 + o(1)\} \prod_{k \neq j}^* U_{j,k} \{1 + B_{5,j,k}(x_j, x_k)\}, & \frac{x_j}{b_j} \rightarrow \kappa_j, \end{cases}
\end{aligned}$$

noting that for $x_j/b_j \rightarrow \infty$,

$$\begin{aligned}
\sigma_{jj}(x_j) &= \log\left(1 + \frac{b_j}{x_j + b_j}\right) = \frac{b_j}{x_j + b_j} \{1 + O(b_j(x_j + b_j)^{-1})\}, \\
\sigma_{jk}(x_j, x_k) &= \rho \left\{ \log\left(1 + \frac{b_k}{x_k + b_k}\right) \right\}^{1/2} \left(\frac{b_j}{x_j + b_j}\right)^{1/2} \{1 + O(b_j(x_j + b_j)^{-1})\}, \quad j \neq k. \quad \square
\end{aligned}$$

Proof of Theorem 2 Let $S_b = [b^{\tau_1}, b^{-\tau_2}]^d$ for $\tau_1 \in (2/3, 1)$ and $\tau_2 \in (\max\{2/(q+2-d), d/\{2(q+2-d)\}\}, \min\{1/(2d), \eta/(\eta+d+2)\})$, where η and q are given in the assumptions A3 and A5, respectively.

Then,

$$MISE[\hat{f}_{\mathbf{b}, \rho, \nu}] = \left(\int_{S_b} + \int_{[0, \infty)^d \setminus S_b} \right) [\{Bias[\hat{f}_{\mathbf{b}, \rho, \nu}(\mathbf{x})]\}^2 + V[\hat{f}_{\mathbf{b}, \rho, \nu}(\mathbf{x})]] d\mathbf{x}.$$

In view of Theorem 1, it is shown that

$$\left| \int_{S_b} V[\hat{f}_{\mathbf{b}, \rho, \nu}(\mathbf{x})] d\mathbf{x} - n^{-1} b^{-d/2} \int_{[0, \infty)^d} \sigma^2(\mathbf{x}) d\mathbf{x} \right| \leq o(n^{-1} b^{-d/2}) + n^{-1} b^{-d/2} \int_{[0, \infty)^d \setminus S_b} \sigma^2(\mathbf{x}) d\mathbf{x} \\ = o(n^{-1} b^{-d/2}),$$

and that

$$\int_{S_b} \mathcal{B}^2(\mathbf{x}) d\mathbf{x} = \int_{S_b} \left[\sum_{j=1}^d \left\{ B_{1,j}(x_j) + \sum_{k \neq j}^* B_{2,j,k}(x_j, x_k) + B_{3,j}(x_j) \right\} \right]^2 d\mathbf{x} = o(b^2),$$

where

$$\mathcal{B}(\mathbf{x}) = Bias[\hat{f}_{\mathbf{b}, \rho, \nu}(\mathbf{x})] - b\gamma(\mathbf{x}) \quad \text{for } \mathbf{x} \in S_b.$$

Hence, we can see that

$$\left| \int_{S_b} \{Bias[\hat{f}_{\mathbf{b}, \rho, \nu}(\mathbf{x})]\}^2 d\mathbf{x} - b^2 \int_{[0, \infty)^d} \gamma^2(\mathbf{x}) d\mathbf{x} \right| \\ = \left| \int_{S_b} \{Bias[\hat{f}_{\mathbf{b}, \rho, \nu}(\mathbf{x})] - b\gamma(\mathbf{x})\} \{Bias[\hat{f}_{\mathbf{b}, \rho, \nu}(\mathbf{x})] + b\gamma(\mathbf{x})\} d\mathbf{x} - b^2 \int_{[0, \infty)^d \setminus S_b} \gamma^2(\mathbf{x}) d\mathbf{x} \right| \\ = \left| \int_{S_b} \mathcal{B}(\mathbf{x}) \{2b\gamma(\mathbf{x}) + \mathcal{B}(\mathbf{x})\} d\mathbf{x} - b^2 \int_{[0, \infty)^d \setminus S_b} \gamma^2(\mathbf{x}) d\mathbf{x} \right| \\ \leq 2b \left\{ \int_{S_b} \gamma^2(\mathbf{x}) d\mathbf{x} \int_{S_b} \mathcal{B}^2(\mathbf{x}) d\mathbf{x} \right\}^{1/2} + \int_{S_b} \mathcal{B}^2(\mathbf{x}) d\mathbf{x} + b^2 \int_{[0, \infty)^d \setminus S_b} \gamma^2(\mathbf{x}) d\mathbf{x} \\ = o(b^2).$$

It remains to evaluate $\int_{[0, \infty)^d \setminus S_b} [\{Bias[\hat{f}_{\mathbf{b}, \rho, \nu}(\mathbf{x})]\}^2 + V[\hat{f}_{\mathbf{b}, \rho, \nu}(\mathbf{x})]] d\mathbf{x}$. Let $\mathcal{X}_l = [0, b^{\tau_1}]^{d_l}$, $\mathcal{X}_m = [b^{\tau_1}, b^{-\tau_2}]^{d_m}$, and $\mathcal{X}_u = [b^{-\tau_2}, \infty)^{d_u}$, where $d_l + d_m + d_u = d$. In what follows, for simplicity, we consider the case $\mathbf{x}_{(l)} = (x_1, \dots, x_{d_l})'$, $\mathbf{x}_{(m)} = (x_{d_l+1}, \dots, x_{d_l+d_m})'$, $\mathbf{x}_{(u)} = (x_{d_l+d_m+1}, \dots, x_d)'$ only, since we can deal with other patterns consisting of any permutation of the d indices. If $d_l \geq 1$ and $d_u = 0$, then we have

$$\int_{\mathcal{X}_l} \int_{\mathcal{X}_m} \{Bias[\hat{f}_{\mathbf{b}, \rho, \nu}(\mathbf{x})]\}^2 d\mathbf{x}_{(m)} d\mathbf{x}_{(l)}$$

$$\begin{aligned}
&= \int_{\mathcal{X}_l} \int_{\mathcal{X}_m} \left[\sum_{j=1}^{d_l} f_j(\mathbf{x}) E[Y_j - x_j] \right. \\
&\quad + \sum_{j=1}^{d_l} \sum_{k=1}^{d_l} \int_{[0, \infty)^d} (s_j - x_j)(s_k - x_k) K_{\boldsymbol{\mu}_{\mathbf{x}}, \Sigma_{\mathbf{x}}, \boldsymbol{\nu}}(\mathbf{s}) \int_0^1 f_{jk}(\mathbf{x} + \theta(\mathbf{s} - \mathbf{x}))(1 - \theta) d\theta d\mathbf{s} \\
&\quad + 2 \sum_{j=1}^{d_l} \sum_{k=d_l+1}^d \int_{[0, \infty)^d} (s_j - x_j)(s_k - x_k) K_{\boldsymbol{\mu}_{\mathbf{x}}, \Sigma_{\mathbf{x}}, \boldsymbol{\nu}}(\mathbf{s}) \int_0^1 f_{jk}(\mathbf{x} + \theta(\mathbf{s} - \mathbf{x}))(1 - \theta) d\theta d\mathbf{s} \\
&\quad + \sum_{j=d_l+1}^d \left\{ b_j \gamma_{1,j}(\mathbf{x}) + \sum_{k \neq j}^{**} (b_j b_k)^{1/2} \gamma_{2,j,k}(\mathbf{x}) \right\} \\
&\quad \left. + \sum_{j=d_l+1}^d \left\{ B_{1,j}(x_j) + \sum_{k \neq j}^{**} B_{2,j,k}(x_j, x_k) + B_{3,j}(x_j) \right\} \right]^2 d\mathbf{x}_{(m)} d\mathbf{x}_{(l)} \\
&= O(b^{(2+d_l)\tau_1} + b^{(4+d_l)\tau_1 - d_m\tau_2} + b^{(2+d_l)\tau_1 + 1 - (d_m+1)\tau_2}) + o(b^2) \\
&= o(b^2)
\end{aligned}$$

($\sum_{k \neq j}^{**}$ is the summation over $k = d_l + 1, \dots, d$ such that $k \neq j$), and

$$\begin{aligned}
&\int_{\mathcal{X}_l} \int_{\mathcal{X}_m} V[\hat{f}_{\mathbf{b}, \rho, \boldsymbol{\nu}}(\mathbf{x})] d\mathbf{x}_{(m)} d\mathbf{x}_{(l)} \\
&\leq n^{-1} \int_{\mathcal{X}_l} \int_{\mathcal{X}_m} \int_{[0, \infty)^d} \{K_{\boldsymbol{\mu}_{\mathbf{x}}, \Sigma_{\mathbf{x}}, \boldsymbol{\nu}}(\mathbf{s})\}^2 \left\{ f(\mathbf{x}) + \sum_{j=1}^d (s_j - x_j) \int_0^1 f_j(\mathbf{x} + \theta(\mathbf{s} - \mathbf{x})) d\theta \right\} d\mathbf{s} d\mathbf{x}_{(m)} d\mathbf{x}_{(l)} \\
&\leq n^{-1} \int_{\mathcal{X}_l} \int_{\mathcal{X}_m} \frac{\exp\{\frac{1}{4} \sum_{j=1}^d \sum_{k \neq j}^* \sigma_{jk}(x_j, x_k)\} \prod_{j=1}^d (1 + \frac{b_j}{x_j + b_j})^{-\nu_j + 1/4}}{2^d \pi^{d/2} \{h_d(\rho) \prod_{j=1}^d \sigma_{jj}(x_j)\}^{1/2} \prod_{j=1}^d (x_j + b_j)} \\
&\quad \times \left\{ f(\mathbf{x}) + \sum_{j=1}^d C_j E[|W_j - x_j|] \right\} d\mathbf{x}_{(m)} d\mathbf{x}_{(l)} \\
&\leq \frac{n^{-1} \exp\{\frac{1}{4} d(d-1) \rho \log 2\} \prod_{j=1}^d c_{\nu_j}}{2^{d/2} \pi^{d/2} h_d^{1/2}(\rho) (\prod_{j=1}^d b_j^{1/2})} \\
&\quad \times \int_{\mathcal{X}_l} \int_{\mathcal{X}_m} \left(\prod_{j=1}^d x_j^{-1/2} \right) \left\{ f(\mathbf{x}) + \sum_{k=1}^{d_l} O(b_k^{\tau_1}) + \sum_{k=d_l+1}^d O(\{b_k(x_k + b_k)\}^{1/2}) \right\} d\mathbf{x}_{(m)} d\mathbf{x}_{(l)} \\
&= o(n^{-1} b^{-d/2}) + O(n^{-1} b^{-d/2 + d_l \tau_1 / 2 - d_m \tau_2 / 2} (b^{\tau_1} + b^{(1-\tau_2)/2})) \\
&= o(n^{-1} b^{-d/2}),
\end{aligned}$$

where

$$c_{\nu} = \begin{cases} 2^{-\nu + 1/4}, & \nu < \frac{1}{4}, \\ 1, & \nu \geq \frac{1}{4}, \end{cases}$$

since, in addition to Lemma A.1, we have

$$E[Y_j - x_j] = (x_j + b_j) \left(1 + \frac{b_j}{x_j + b_j}\right)^{\nu_j + 1/2} - x_j$$

$$\begin{aligned}
&= O(b^{\tau_1}) \quad \text{for } x_j \leq b_j^{\tau_1}, \\
E[(Y_j - x_j)^2] &= (x_j + b_j)^2 \left(1 + \frac{b_j}{x_j + b_j}\right)^{2(\nu_j+1)} - 2x_j(x_j + b_j) \left(1 + \frac{b_j}{x_j + b_j}\right)^{\nu_j+1/2} + x_j^2 \\
&= O(b^{2\tau_1}) \quad \text{for } x_j \leq b_j^{\tau_1}, \\
E[|Y_j - x_j||Y_k - x_k|] &\leq \{E[(Y_j - x_j)^2]E[(Y_k - x_k)^2]\}^{1/2} \\
&= \begin{cases} O(b^{2\tau_1}) & \text{for } x_j, x_k \leq b_j^{\tau_1}, \\ O(b^{\tau_1+(1-\tau_2)/2}) & \text{for } x_j \leq b_j^{\tau_1} \text{ and } x_k \in [b_j^{\tau_1}, b_j^{-\tau_2}], \end{cases} \\
E[(W_j - x_j)^2] &= O(b^{2\tau_1}) \quad \text{for } x_j \leq b_j^{\tau_1},
\end{aligned}$$

and $t/2 \leq \log(1+t) \leq \log 2$ for $t \in [0, 1]$. Also, if $d_u \geq 1$, then we can see that

$$\begin{aligned}
&\int_{\mathcal{X}_l} \int_{\mathcal{X}_m} \int_{\mathcal{X}_u} \{Bias[\hat{f}_{\mathbf{b}, \rho, \nu}(\mathbf{x})]\}^2 d\mathbf{x}_{(u)} d\mathbf{x}_{(m)} d\mathbf{x}_{(l)} \\
&= \int_{\mathcal{X}_l} \int_{\mathcal{X}_m} \int_{\mathcal{X}_u} \left[\int_{[0, \infty)^d} K_{\boldsymbol{\mu}_{\mathbf{x}}, \Sigma_{\mathbf{x}}, \nu}(\mathbf{s}) \{f(\mathbf{s}) - f(\mathbf{x})\} d\mathbf{s} \right]^2 d\mathbf{x}_{(u)} d\mathbf{x}_{(m)} d\mathbf{x}_{(l)} \\
&\leq \int_{\mathcal{X}_l} \int_{\mathcal{X}_m} \int_{\mathcal{X}_u} \int_{[0, \infty)^d} K_{\boldsymbol{\mu}_{\mathbf{x}}, \Sigma_{\mathbf{x}}, \nu}(\mathbf{s}) \{f(\mathbf{s}) - f(\mathbf{x})\}^2 d\mathbf{s} d\mathbf{x}_{(u)} d\mathbf{x}_{(m)} d\mathbf{x}_{(l)} \\
&\leq \frac{2C d^{d/2}}{h_d^{1/2}(\rho)} \int_{\mathcal{X}_l} \int_{\mathcal{X}_m} \int_{\mathcal{X}_u} \int_{[0, \infty)^d} \left\{ \prod_{j=1}^d K_{\tilde{\mu}_j(x_j), d\sigma_{jj}(x_j)}^{(LN)}(s_j) \right\} f(\mathbf{s}) d\mathbf{s} d\mathbf{x}_{(u)} d\mathbf{x}_{(m)} d\mathbf{x}_{(l)} \\
&\quad + 2C \int_{\mathcal{X}_l} \int_{\mathcal{X}_m} \int_{\mathcal{X}_u} f(\mathbf{x}) d\mathbf{x}_{(u)} d\mathbf{x}_{(m)} d\mathbf{x}_{(l)} \\
&\leq \frac{2C d^{d/2}}{h_d^{1/2}(\rho)} \int_{\mathcal{X}_l} \int_{\mathcal{X}_m} \int_{[0, \infty)^d} \left\{ \prod_{j=1}^{d_l+d_m} K_{\tilde{\mu}_j(x_j), d\sigma_{jj}(x_j)}^{(LN)}(s_j) \right\} \\
&\quad \times \left\{ b^{d_u \tau_2 (q+1)} \prod_{j=d_l+d_m+1}^d s_j^{q+1} (1 + o(1)) \right\} f(\mathbf{s}) d\mathbf{s} d\mathbf{x}_{(m)} d\mathbf{x}_{(l)} \\
&\quad + 2C b^{d_u \tau_2 (q+1)} \int_{\mathcal{X}_l} \int_{\mathcal{X}_m} \int_{\mathcal{X}_u} \left(\prod_{j=d_l+d_m+1}^d x_j^{q+1} \right) f(\mathbf{x}) d\mathbf{x}_{(u)} d\mathbf{x}_{(m)} d\mathbf{x}_{(l)} \\
&= O(b^{d_l \tau_1 - d_m \tau_2 + d_u \tau_2 (q+1)}) \\
&= o(b^2),
\end{aligned}$$

and

$$\begin{aligned}
&\int_{\mathcal{X}_l} \int_{\mathcal{X}_m} \int_{\mathcal{X}_u} V[\hat{f}_{\mathbf{b}, \rho, \nu}(\mathbf{x})] d\mathbf{x}_{(u)} d\mathbf{x}_{(m)} d\mathbf{x}_{(l)} \\
&\leq n^{-1} \int_{\mathcal{X}_l} \int_{\mathcal{X}_m} \int_{\mathcal{X}_u} \int_{[0, \infty)^d} \{K_{\boldsymbol{\mu}_{\mathbf{x}}, \Sigma_{\mathbf{x}}, \nu}(\mathbf{s})\}^2 f(\mathbf{s}) d\mathbf{s} d\mathbf{x}_{(u)} d\mathbf{x}_{(m)} d\mathbf{x}_{(l)} \\
&\leq n^{-1} L_{d, \rho, \nu} \left(\prod_{j=1}^d b_j^{-1} \right) \int_{\mathcal{X}_l} \int_{\mathcal{X}_m} \int_{\mathcal{X}_u} \int_{[0, \infty)^d} \left\{ \prod_{j=1}^d K_{\tilde{\mu}_j(x_j), d\sigma_{jj}(x_j)}^{(LN)}(s_j) \right\} f(\mathbf{s}) d\mathbf{s} d\mathbf{x}_{(u)} d\mathbf{x}_{(m)} d\mathbf{x}_{(l)}
\end{aligned}$$

$$\begin{aligned}
&\leq n^{-1} L_{d,\rho,\nu} \left(\prod_{j=1}^d b_j^{-1} \right) \int_{\mathcal{X}_l} \int_{\mathcal{X}_m} \int_{[0,\infty)^d} \left\{ \prod_{j=1}^{d_l+d_m} K_{\tilde{\mu}_j(x_j), d\sigma_{jj}(x_j)}^{(LN)}(s_j) \right\} \\
&\quad \times \left\{ b^{d_u \tau_2(q+1)} \prod_{j=d_l+d_m+1}^d s_j^{q+1} (1+o(1)) \right\} f(\mathbf{s}) d\mathbf{s} d\mathbf{x}_{(m)} d\mathbf{x}_{(l)} \\
&= O(n^{-1} b^{-d+d_l \tau_1 - d_m \tau_2 + d_u \tau_2(q+1)}) \\
&= o(n^{-1} b^{-d/2}),
\end{aligned}$$

using Lemmas A.2–A.4. It follows that

$$\int_{[0,\infty)^d \setminus S_b} [\{Bias[\hat{f}_{\mathbf{b},\rho,\nu}(\mathbf{x})]\}^2 + V[\hat{f}_{\mathbf{b},\rho,\nu}(\mathbf{x})]] d\mathbf{x} = o(b^2 + n^{-1} b^{-d/2}). \quad \square$$

Acknowledgments

The author thanks Professor Yoshihide Kakizawa for his advice. This work was partially supported by the Japan Society for the Promotion of Science (JSPS); Grant-in-Aid for Research Activity Start-up [Grant Number 15H06068].

References

- Bouezmarni, T. and Rombouts, J. V. K. (2010) “Nonparametric density estimation for multivariate bounded data”, *Journal of Statistical Planning and Inference*, 140, 139–152.
- 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.
- Igarashi, G. (2016) “Weighted log-normal kernel density estimation”, *To appear in Communications in Statistics - Theory and Methods*, DOI:10.1080/03610926.2014.963623.
- Igarashi, G. and Kakizawa, Y. (2014) “Re-formulation of inverse Gaussian, reciprocal inverse Gaussian, and Birnbaum–Saunders kernel estimators”, *Statistics and Probability Letters*, 84, 235–246.
- 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.
- Koul, H. L. and Song, W. (2013) “Large sample results for varying kernel regression estimates”, *Journal of Nonparametric Statistics*, 25, 829–853.
- 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.
- Mnatsakanov, R. and Sarkisian, K. (2012) “Varying kernel density estimation on R_+ ”, *Statistics and Probability Letters*, 82, 1337–1345.
- Patil, G. P. and Rao, C. R. (1978) “Weighted distributions and size-biased sampling with applications to wildlife populations and human families”, *Biometrics*, 34, 179–184.
- Rao, C. R. (1965) “On discrete distributions arising out of methods of ascertainment”, *Sankhyā, Series A*, 27, 311–324.
- 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.
- Silverman, B. W. (1986) *Density Estimation for Statistics and Data Analysis*, London: Chapman & Hall.
- Wand, M. P. and Jones, M. C. (1995) *Kernel Smoothing*, London: Chapman & Hall.