Uniform Convergence Rates for Density Derivative Estimators Using Asymmetric Kernels

Benedikt Funke* Masayuki Hirukawa†
TH Köln Ryukoku University

May 30, 2024

Abstract

This paper is concerned with the problem of estimating partial derivatives of an unknown joint density with support either on the upper half-space or on the unit hypercube. Weak and strong uniform convergence rates of nonparametric first-order partial derivative estimators smoothed by the asymmetric gamma and beta kernels are derived. The results are useful for asymptotic analysis of indirect average derivative estimation of nonparametric regression curves, for instance.

JEL Classification Codes: C13; C14.

MSC 2000 Codes: 62G7; 62G8; 62G20.

Keywords: Beta kernel; density derivative estimation; gamma kernel; rates of convergence; uniform consistency.

* *Corresponding Author*; Institute for Insurance Studies, TH Köln - University of Applied Sciences, Gustav-Heinemann-Ufer 54, 50968 Köln, Germany; e-mail: benedikt.funke@th-koeln.de.

† Faculty of Economics, Ryukoku University, 67 Tsukamoto-cho, Fukakusa, Fushimi-ku, Kyoto 612-8577, Japan; e-mail: hirukawa@econ.ryukoku.ac.jp.

1 Introduction

This paper demonstrates weak and strong uniform consistency with rates of kernel-smoothed estimators for partial derivatives of an unknown joint density. From theoretical and practical standpoints, we focus on first-order density derivatives among all orders. In fact, first-order density derivative estimators are applied in many directions. A prominent example is that the estimators can be used to estimate density scores, which are a key building block for indirect average derivative estimation of nonparametric regression curves (e.g., Härdle and Stoker, 1989; Powell et al., 1989; Rilstone, 1991); see Funke and Hirukawa (2024a; abbreviated as "FH24a" hereinafter) for other applications.

Throughout it is assumed that the support of the joint density has at least one boundary in each dimension. Let each dimension be either $\mathbb{R}_+$ or $[0,1]$, to be more precise. There are indeed many economic and financial variables values of which are either nonnegative or limited within a certain interval by definition. Examples of the former are wages, incomes, consumption expenditures, and insurance claims (or financial losses), whereas those of the latter are typically expressed in the forms of shares or proportions, including expenditure and budget shares, unemployment rates, and default and recovery rates. It has been known that asymmetric kernels are viable devices that can capture density and regression curves of such economic and financial variables well. Appealing properties of asymmetric kernels continue to hold for density derivative estimation. For univariate cases, FH24a propose first-order density derivative estimators smoothed by asymmetric kernels and report their superior finite-sample performances to those using standard symmetric kernels.

In this paper, we extend the first-order derivative estimators by FH24a to multivariate cases. Before proceeding, joint density estimators must be defined. Let

$K_{\mathcal{J}(x,b)}(u)$ be a univariate asymmetric kernel indexed by $\mathcal{J}$ that depends on the data point u, the design point x and the smoothing parameter $b\,(>0)$. Among all asymmetric kernels, we specialize in the gamma kernel ($\mathcal{J}=G$)

$$K_{G(x,b)}(u)=\frac{u^{x/b}\exp(-u/b)}{b^{x/b+1}\Gamma(x/b+1)}\mathbf{1}\{u\geq 0\}$$

for $x\in\mathbb{R}_+$ by Chen (2000), and the beta kernel ($\mathcal{J}=B$)

$$K_{B(x,b)}(u)=\frac{u^{x/b}(1-u)^{(1-x)/b}}{B\{x/b+1,(1-x)/b+1\}}\mathbf{1}\{u\in[0,1]\}$$

for $x\in[0,1]$ by Chen (1999), where $\Gamma(z)=\int_0^\infty t^{z-1}\exp(-t)\,dt$ for $z>0$ is the gamma function, $B(\alpha,\beta)=\int_0^1 y^{\alpha-1}(1-y)^{\beta-1}dy$ for $\alpha,\beta>0$ is the beta function, and $\mathbf{1}\{\cdot\}$ denotes an indicator function. Our preference on these kernels is based on three reasons. First, as reported by Hirukawa (2018), the gamma and beta kernels are frequently applied to empirical models in economics and finance due to their favorable evidence. Second, these kernels have been discussed as examples of boundary kernels in textbooks (e.g., Racine, 2019). Third, when convergence properties of asymmetric kernel estimators are explored, kernel-specific and thus diversified approaches must be taken. Inevitably, analytical tractability of the estimators is a key issue. The gamma function is an essential component for the gamma and beta kernels, and there is rich literature on approximation techniques to the gamma and related functions.

To cope with multivariate problems, we consider tensor product kernels

$$\begin{aligned}\mathbb{K}_{\mathcal{J}(\mathbf{x},\mathbf{b})}(\mathbf{u})&=\prod_{j=1}^d K_{\mathcal{J}(x_j,b_j)}(u_j)\\&=\begin{cases}\prod_{j=1}^d\frac{u^{x_j/b_j}\exp(-u_j/b_j)}{b_j^{x_j/b_j+1}\Gamma(x_j/b_j+1)}\mathbf{1}\{u_j\geq 0\} & \text{for } \mathcal{J}=G\\ \prod_{j=1}^d\frac{u_j^{x_j/b_j}(1-u_j)^{(1-x_j)/b_j}}{B\{x_j/b_j+1,(1-x_j)/b_j+1\}}\mathbf{1}\{u_j\in[0,1]\} & \text{for } \mathcal{J}=B\end{cases},\end{aligned}$$

where $\mathbf{u}:=(u_1,\ldots,u_d)^\top$, $\mathbf{x}:=(x_1,\ldots,x_d)^\top$ and $\mathbf{b}:=(b_1,\ldots,b_d)^\top$ are d-dimensional vectors of data points, design points and smoothing parameters, respectively. Now let $f(\mathbf{x})$ be the joint density of $\mathbf{X}$ with support on $\mathbb{R}_+^d$ or $[0,1]^d$. Then, the estimator of $f(\mathbf{x})$ using the product kernel $\mathcal{J}$ is defined as

$$\hat{f}_{\mathcal{J}}(\mathbf{x}) := \frac{1}{n}\sum_{i=1}^{n} \mathbb{K}_{\mathcal{J}(\mathbf{x},\mathbf{b})}(\mathbf{X}_i).$$

Also let $\mathsf{D}_{\mathbf{x}} = \partial/\partial\mathbf{x} = (\partial/\partial x_1, \ldots, \partial/\partial x_d)^{\top}$ denote the d-dimensional first-order partial derivative (or gradient) operator. The vector of d partial derivatives of $f(\mathbf{x})$ can be expressed as $\mathsf{D}_{\mathbf{x}} f(\mathbf{x}) = (\partial f(\mathbf{x})/\partial x_1, \ldots, \partial f(\mathbf{x})/\partial x_d)^{\top} =: \left(f_1^{(1)}(\mathbf{x}), \ldots, f_d^{(1)}(\mathbf{x})\right)^{\top}$. A natural estimator of $f_p^{(1)}(\mathbf{x})$, the pth element of $\mathsf{D}_{\mathbf{x}} f(\mathbf{x})$ for $p \in \{1, \ldots, d\}$, smoothed by the product kernel $\mathcal{J}$ can be defined as

$$\hat{f}_{\mathcal{J},p}^{(1)}(\mathbf{x}) := \frac{\partial \hat{f}_{\mathcal{J}}(\mathbf{x})}{\partial x_p} = \frac{1}{n}\sum_{i=1}^{n} \frac{\partial \mathbb{K}_{\mathcal{J}(\mathbf{x},\mathbf{b})}(\mathbf{X}_i)}{\partial x_p},$$

where $\partial \mathbb{K}_{\mathcal{J}(\mathbf{x},\mathbf{b})}(\mathbf{u})/\partial x_p = (1/b_p)\,\mathcal{L}_{\mathcal{J}(x_p,b_p)}(u_p)\,\mathbb{K}_{\mathcal{J}(\mathbf{x},\mathbf{b})}(\mathbf{u})$,

$$\mathcal{L}_{\mathcal{J}(x,b)}(u) := \begin{cases} \ln u - \ln b - \Psi\left(\frac{x}{b}+1\right) & \text{for } \mathcal{J} = G \\ \ln\left(\frac{u}{1-u}\right) - \Psi\left(\frac{x}{b}+1\right) + \Psi\left(\frac{1-x}{b}+1\right) & \text{for } \mathcal{J} = B \end{cases},$$

and $\Psi(z) = d\ln\Gamma(z)/dx = \Gamma^{(1)}(z)/\Gamma(z)$ is the digamma function. When $d = 1$, $\hat{f}_{\mathcal{J},p}^{(1)}(\mathbf{x})$ collapses to the univariate density derivative estimator studied by FH24a.

Weak and strong uniform convergences with rates of $\hat{f}_{\mathcal{J},p}^{(1)}(\mathbf{x})$ are established on a d-hyperrectangle

$$\mathbb{S}_{\mathbf{X}}^{\mathcal{J}} = \mathbb{S}_{\mathbf{X}}^{\mathcal{J}}(\eta) := \begin{cases} \prod_{j=1}^{d}\left[\eta_j, \eta_j^{-1}\right] \subseteq \mathbb{R}_{+}^{d} & \text{for } \mathcal{J} = G \\ \prod_{j=1}^{d}\left[\eta_j, 1-\eta_j\right] \subseteq [0,1]^{d} & \text{for } \mathcal{J} = B \end{cases},$$

where $\eta := (\eta_1, \ldots, \eta_d)^{\top}$ and the boundary parameter $\eta_j\ (>0)$ for $j \in \{1, \ldots, d\}$ either is fixed or shrinks to zero at a suitable rate. Observe that in the latter scenario, $\mathbb{S}_{\mathbf{X}}^{\mathcal{J}}$ expands to the d-dimensional upper half-space ($\mathcal{J} = G$) or unit hypercube ($\mathcal{J} = B$). This framework enables us to employ Stirling's approximation to the gamma function as a workhorse in technical proofs; see Section 3 for more details.

This paper can be positioned as a complement to Hirukawa et al. (2022; abbreviated as "HMP22" hereinafter) and Funke and Hirukawa (2024b; abbreviated as "FH24b" hereinafter). HMP22 and FH24b provide sets of uniform convergence results with rates for nonparametric density and regression estimators using the beta and gamma kernels, respectively. Furthermore, there is growing literature on joint density estimation using asymmetric kernels. Examples include Bouezmarni and Rombouts (2010), Funke and Kawka (2015), Ouimet (2021, 2022), Ouimet and Tolosana-Delgado (2022), and Bertin et al. (2023), as well as HMP22 and FH24b. This paper can be viewed as yet another contribution to the literature in this class. It is worth emphasizing that this paper is the first work on uniform consistency of partial derivative estimators of a joint density using asymmetric kernels, to the best of our knowledge.

The remainder of this paper is organized as follows. Sections 2 delivers weak and strong uniform consistency with rates of $\hat{f}^{(1)}_{G,p}(\mathbf{x})$ and $\hat{f}^{(1)}_{B,p}(\mathbf{x})$. All proofs are given in Section 3.

This paper adopts the following notational conventions: '$a_n = o(b_n)$' signifies that a_n/b_n converges to 0; '$a_n = O(b_n)$' means that a_n/b_n is bounded; we say that '$a_n \asymp b_n$' if there exist constants $0 < c_1 < c_2 < \infty$ so that $c_1 a_n \leq b_n \leq c_2 a_n$; $f^{(2)}_{pq}(\mathbf{x}) = \frac{\partial^2 h(\mathbf{x})}{\partial x_p \partial x_q}$ and $f^{(3)}_{pqr}(\mathbf{x}) = \frac{\partial^3 f(\mathbf{x})}{\partial x_p \partial x_q \partial x_r}$ denote the second- and third-order partial derivatives of the joint density $f(\mathbf{x})$, respectively; $\Psi_1(z) = \Psi^{(1)}(z) = d\Psi(z)/dz$ is the trigamma function; '*a.s.*' means "almost surely"; $\|\mathbf{A}\|$ is the Frobenius norm of matrix $\mathbf{A}$, i.e., $\|\mathbf{A}\| = \left\{\operatorname{tr}\left(\mathbf{A}^\top \mathbf{A}\right)\right\}^{1/2}$; and the expression '$X \overset{d}{=} Y$' reads "A random variable X obeys the distribution Y."

2 Main Results

2.1 Weak and Strong Uniform Consistency of $\hat{f}^{(1)}_{G,p}(\mathbf{x})$

We start from demonstrating weak uniform consistency of $\hat{f}^{(1)}_{G,p}(\mathbf{x})$ on $\mathbb{S}^G_{\mathbf{X}}$. This result is built on the following regularity conditions.

Assumption G1. $\{\mathbf{X}_i\}_{i=1}^n \in \mathbb{R}^d_+$ are *i.i.d.* random vectors.

Assumption G2. There are constants $C_1, C_2, C_3, \delta > 0$ that satisfy the followings.

(i) $\left| f^{(3)}_{jk\ell}(\mathbf{x}) \right| \le C_1 \left[C_2^{-2(1+\delta)} \mathbf{1}\left\{ \max(x_k, x_\ell) < C_2 \right\} + (x_k x_\ell)^{-(1+\delta)} \mathbf{1}\left\{ \max(x_k, x_\ell) \ge C_2 \right\} \right]$ for all $\mathbf{x} \in \mathbb{R}^d_+$ and for all $j, k, \ell \in \{1, \ldots, d\}$.

(ii) $\left| f^{(3)}_{jk\ell}(\mathbf{x}) - f^{(3)}_{jk\ell}(\mathbf{x}') \right| \le C_3 \|\mathbf{x} - \mathbf{x}'\|$ for all $\mathbf{x}, \mathbf{x}' \in \mathbb{R}^d_+$ and for all $j, k, \ell \in \{1, \ldots, d\}$.

Assumption G3. Sequences $b_j (= b_j(n)), \eta_j (= \eta_j(n)) > 0$ satisfy the followings as $n \to \infty$.

(i) $b_j, \eta_j \to 0$ for all $j \in \{1, \ldots, d\}$.

(ii) There is a sequence $\rho (= \rho(n)) > 0$ that satisfies $b_j / \eta_j \asymp \rho$ for all $j \in \{1, \ldots, d\}$, $\rho \to 0$ and $\rho = o\left(\min_{1 \le j \le d} \eta_j^2 \right)$.

(iii) $\ln n / \left(n b_p \eta_p \sqrt{\prod_{j=1}^d b_j \eta_j} \right) \to 0$ for all $p \in \{1, \ldots, d\}$.

Assumption G2 is a suitably strengthened version of Assumption 2 in FH24b. This assumption implies that third-order partial derivatives of $f(\mathbf{x})$ are Lipschitz continuous and uniformly bounded on $\mathbb{R}^d_+$. It also follows from this assumption and integrability of $f(\mathbf{x})$ that there is a constant $C_0 > 0$ that satisfies

$$\sup_{\mathbf{x} \in \mathbb{R}^d_+} f(\mathbf{x}) \le C_0. \tag{1}$$

The condition (i) in Assumption G2 also regulates tail decay rates of third-order partial derivatives of $f(\mathbf{x})$. As given in Lemma 3 in Section 3, second- and higher-order moments of the univariate gamma kernel around the design point x depend on x, and the value of x is unbounded from above. Assumption G2(i) helps control the order of magnitude in the leading bias term of $\hat{f}^{(1)}_{G,p}(\mathbf{x})$, because it ensures uniform boundedness of $\left| f^{(1)}_j(\mathbf{x}) \right|$, $\left| f^{(2)}_{jk}(\mathbf{x}) \right|$, $\left| f^{(3)}_{jjj}(\mathbf{x}) x_j \right|$, and $\left| f^{(3)}_{jkk}(\mathbf{x}) x_k \right|$ on $\mathbb{R}^d_+$.

Three conditions on the boundary parameter η_j in Assumption G3 are intended for the case in which $\mathbb{S}^G_{\mathbf{X}}$ is expanding. The conditions (i) and (ii) are the same as those in Assumption 4W of FH24b. These jointly mean that η_j shrinks to zero more slowly than b_j does. As will be revealed in Section 3, this is crucial for Stirling's approximation to the gamma function. In addition, $b_j/\eta_j \asymp \rho$ in Assumption G3(ii) merely indicates that the shrinkage rate of the ratio of b_j to η_j is identical across j. This does not automatically guarantee that the ratio b_j/η_j itself, the numerator b_j, or the denominator η_j are identical across j. Moreover, $\rho = o\left(\min_{1\le j\le d} \eta_j^2\right)$ is a technical requirement for controlling orders of magnitude in the remainder terms of the bias. An obvious sufficient condition for Assumption G3 is that sequences $b\left(= b\left(n\right)\right), \eta\left(= \eta\left(n\right)\right) > 0$ satisfy $b_1, \ldots, b_d \asymp b$, $\eta_1, \ldots, \eta_d \asymp \eta$, and $b + \eta + b/\eta^3 + \ln n / \left\{n\left(b\eta\right)^{d/2}\right\} \to 0$ as $n \to \infty$.

The next theorem documents weak uniform consistency of $\hat{f}_{G,p}^{(1)}\left(\mathbf{x}\right)$.

Theorem 1. *If Assumptions G1-G3 hold, then, as $n \to \infty$,*

$$\sup_{\mathbf{x}\in\mathbb{S}_{\mathbf{X}}^{G}} \left|\hat{f}_{G,p}^{(1)}\left(\mathbf{x}\right) - f_{p}^{(1)}\left(\mathbf{x}\right)\right| = O_p\left(\sum_{j=1}^{d} b_j + \sqrt{\frac{\ln n}{n b_p \eta_p \sqrt{\prod_{j=1}^{d} b_j \eta_j}}}\right)$$

for all $p \in \{1, \ldots, d\}$.

Now we turn to strong uniform consistency of $\hat{f}_{G,p}^{(1)}\left(\mathbf{x}\right)$. This result can be obtained by suitably strengthening the condition (iii) of Assumption G3 while all others are left unchanged.

Theorem 2. *Let the condition (iii) in Assumption G3 be replaced by the following stronger one: there is a constant $\kappa \in [0, 1)$ that satisfies*

$$\left(\frac{\ln n}{n b_p \eta_p \sqrt{\prod_{j=1}^{d} b_j \eta_j}}\right)\left(\sum_{j=1}^{d} \frac{1}{b_j \eta_j}\right)^{1-\kappa} = O\left(1\right)$$

for all $p \in \{1, \ldots, d\}$. If Assumptions G1-G3 hold, then, as $n \to \infty$, the statement in Theorem 1 can be strengthened to almost sure convergence.

2.2 Weak and Strong Uniform Consistency of $\hat{f}_{B,p}^{(1)}(\mathbf{x})$

We move on to weak uniform consistency of $\hat{f}_{B,p}^{(1)}(\mathbf{x})$ on $\mathbb{S}_{\mathbf{X}}^{B}$. Before proceeding, a set of regularity conditions are provided.

Assumption B1. $\{\mathbf{X}_i\}_{i=1}^{n} \in [0,1]^d$ are *i.i.d.* random variables.

Assumption B2. Third-order derivatives of $f(\mathbf{x})$ are Lipschitz continuous and uniformly bounded on $\mathbf{x} \in [0,1]^d$.

Assumption B3. Sequences $b_j (= b_j(n)), \eta_j (= \eta_j(n)) > 0$ satisfy the followings as $n \to \infty$.

(i) $b_j, \eta_j \to 0$ for all $j \in \{1, \dots, d\}$.

(ii) $\max_{1 \le j \le d} b_j / \min_{1 \le j \le d} \sqrt{b_j \eta_j} \to 0$.

(iii) $\ln n / \left(n b_p \eta_p \sqrt{\prod_{j=1}^{d} b_j \eta_j} \right) \to 0$ for all $p \in \{1, \dots, d\}$.

These assumptions are similar to Assumptions G1-G3 above. Assumption B2 and compactness of $[0,1]^d$ again ensure that there is a constant $L > 0$ that satisfies

$$\sup_{\mathbf{x} \in [0,1]^d} f(\mathbf{x}) \le L. \tag{2}$$

Observe that Assumption B2 does not contain an equivalent of the condition (ii) in Assumption G2. While moments of the univariate beta kernel around the design point x also depend on x (see, e.g., the proof of Lemma 1 in HMP22), the value of x is confined within the unit interval and thus this type of condition is unnecessary.

Furthermore, the condition (ii) in Assumption B3 helps control orders of magnitude in the remainder terms of $E\left\{\hat{f}_{B,p}^{(1)}(\mathbf{x})\right\}$. As with the condition (ii) in Assumption G3, this condition leads to $b_j/\eta_j \to 0$ for all $j \in \{1, \dots, d\}$. However, it does not automatically guarantee that either $b_1, \dots, b_d$ or $\eta_1, \dots, \eta_d$ are identical. Once again, a sufficient condition for Assumption B3 is that sequences $b\,(=b\,(n))\,,\eta\,(=\eta\,(n))>0$ satisfy $b_1,\ldots,b_d \asymp b$, $\eta_1,\ldots,\eta_d \asymp \eta$, and $b+\eta+b/\eta+\ln n/\left\{n\,(b\eta)^{d/2+1}\right\} \to 0$ as $n\to\infty$.

The next two theorems formally demonstrate weak and strong uniform consistency with rates of $\hat{f}^{(1)}_{B,p}\,(\mathbf{x})$. In particular, strong uniform consistency can be obtained with the condition (iii) of Assumption B3 replaced by the same one as given in Theorem 2. Also observe that weak and strong uniform convergence rates of $\hat{f}^{(1)}_{B,p}\,(\mathbf{x})$ concur with those of $\hat{f}^{(1)}_{G,p}\,(\mathbf{x})$.

Theorem 3. *If Assumptions B1-B3 hold, then, as $n\to\infty$,*

$$\sup_{\mathbf{x}\in\mathbb{S}^B_{\mathbf{X}}}\left|\hat{f}^{(1)}_{B,p}\,(\mathbf{x})-f^{(1)}_{p}\,(\mathbf{x})\right| = O_p\left(\sum_{j=1}^{d} b_j + \sqrt{\frac{\ln n}{nb_p\eta_p\sqrt{\prod_{j=1}^{d} b_j\eta_j}}}\right)$$

for all $p\in\{1,\ldots,d\}$.

Theorem 4. *Let the condition (iii) in Assumption B3 be replaced by the following stronger one: there is a constant $\kappa\in[0,1)$ that satisfies*

$$\left(\frac{\ln n}{nb_p\eta_p\sqrt{\prod_{j=1}^{d} b_j\eta_j}}\right)\left(\sum_{j=1}^{d}\frac{1}{b_j\eta_j}\right)^{1-\kappa} = O\,(1)$$

for all $p\in\{1,\ldots,d\}$. If Assumptions B1-B3 hold, then, as $n\to\infty$, the statement in Theorem 3 can be strengthened to almost sure convergence.

2.3 Optimal Uniform Convergence Rates

It is possible to derive the optimal uniform convergence rates of $\hat{f}^{(1)}_{\mathcal{J},p}\,(\mathbf{x})$ for $\mathcal{J}\in\{G,B\}$ when $\mathbb{S}^{\mathcal{J}}_{\mathbf{X}}$ is fixed and a single smoothing parameter b is employed for each dimension. In this scenario, both weak and strong uniform convergence rates of $\hat{f}^{(1)}_{\mathcal{J},p}\,(\mathbf{x})$ reduce to $b+\sqrt{\ln n/\left(nb^{d/2+1}\right)}$. It can be found that $b^*\asymp(\ln n/n)^{2/(6+d)}$ balances two terms. The optimal weak and strong uniform convergence rates of the estimator under such b^* become $(\ln n/n)^{2/(6+d)}$. This rate coincides with Stone's (1983) optimal global rate for nonparametric first-order density derivative estimation.

3 Proofs

3.1 Proof of Theorem 1

This proof requires the following lemmata. Lemma 1 refers to a uniform version of Stirling's approximation to the gamma function, and uniform approximations to the digamma and trigamma functions. Lemma 2 documents uniform bounds of the univariate gamma kernel and its first- and second-order derivatives with respect to the design point x. Lemma 3 presents moments of the univariate gamma kernel around the design point x. Some odd moments are omitted because these are not used in the proofs. Part (ii) in this lemma helps control orders of magnitude in remainder terms of $E\left\{\hat{f}_{G,p}^{(1)}(\mathbf{x})\right\}$. Lemma 4 states Bernstein's inequality.

Lemma 1. *Suppose that sequences $b\,(=b\,(n))\,,\eta\,(=\eta\,(n))>0$ satisfy $b,\eta\to 0$ and $b/\eta\to 0$ as $n\to\infty$. Then, the followings hold true as $n\to\infty$.*

$$\sup_{x\in[\eta,\eta^{-1}]}\left|\frac{\Gamma(x/b+1)}{\sqrt{2\pi}\,(x/b)^{x/b+1/2}\exp(-x/b)}-1\right|=O\left(\frac{b}{\eta}\right).$$

$$\sup_{x\in[\eta,\eta^{-1}]}\left|\frac{\Psi(x/b+1)-\ln(x/b)}{b/(2x)}-1\right|=O\left(\frac{b}{\eta}\right).$$

$$\sup_{x\in[\eta,\eta^{-1}]}\left|\frac{\Psi_1(x/b+1)}{b/x}-1\right|=O\left(\frac{b}{\eta}\right).$$

Lemma 2. *Under the same condition as in Lemma 1, the followings hold true as $n\to\infty$.*

$$\sup_{(x,u)\in[\eta,\eta^{-1}]\times\mathbb{R}_+}K_{G(x,b)}(u)\le\sqrt{\frac{2}{\pi}}b^{-\frac{1}{2}}\eta^{-\frac{1}{2}}. \tag{3}$$

$$\sup_{(x,u)\in[\eta,\eta^{-1}]\times\mathbb{R}_+}\left|\frac{\partial K_{G(x,b)}(u)}{\partial x}\right|\le 4\sqrt{\frac{2}{\pi}}b^{-\frac{3}{2}}\eta^{-\frac{3}{2}}. \tag{4}$$

$$\sup_{(x,u)\in[\eta,\eta^{-1}]\times\mathbb{R}_+}\left|\frac{\partial^2 K_{G(x,b)}(u)}{\partial x^2}\right|=O\left(b^{-\frac{5}{2}}\eta^{-\frac{5}{2}}\right). \tag{5}$$

Lemma 3. *Let* $\xi_x \stackrel{d}{=} G(x/b+1, b)$.

(i) *The followings hold true.*

$$
\begin{aligned}
&(a) \quad E(\xi_x - x) = b,\ E(\xi_x - x)^2 = xb + 2b^2, \\
&\qquad E(\xi_x - x)^4 = 3x^2b^2 + 26xb^3 + 24b^4, \ \text{and} \\
&\qquad E(\xi_x - x)^6 = 15x^3b^3 + 340x^2b^4 + 1044xb^5 + 720b^6. \\
&(b) \quad E\left\{\mathcal{L}_{G(x,b)}(\xi_x)\right\} = 0, E\left\{\mathcal{L}_{G(x,b)}(\xi_x)(\xi_x - x)\right\} = b, \\
&\qquad E\left\{\mathcal{L}_{G(x,b)}(\xi_x)(\xi_x - x)^2\right\} = 3b^2, \ \text{and} \\
&\qquad E\left\{\mathcal{L}_{G(x,b)}(\xi_x)(\xi_x - x)^3\right\} = 3xb^2 + 11b^3.
\end{aligned}
$$

(ii) *Under the same condition as in Lemma 1, the followings also hold true as* $n \to \infty$*, uniformly on* $x \in [\eta, \eta^{-1}]$.

$$
\begin{aligned}
&E\left\{\mathcal{L}^2_{G(x,b)}(\xi_x)\right\} = O\left(\tfrac{b}{\eta}\right), \ E\left\{\mathcal{L}^2_{G(x,b)}(\xi_x)(\xi_x - x)^2\right\} = O(b^2), \\
&E\left\{\mathcal{L}^2_{G(x,b)}(\xi_x)(\xi_x - x)^4\right\} = O\left(\tfrac{b^3}{\eta}\right), \ E\left\{\mathcal{L}^2_{G(x,b)}(\xi_x)(\xi_x - x)^6\right\} = O\left(\tfrac{b^4}{\eta^2}\right), \ \text{and} \\
&E\left\{\mathcal{L}^2_{G(x,b)}(\xi_x)(\xi_x - x)^8\right\} = O\left(\tfrac{b^5}{\eta^3}\right).
\end{aligned}
$$

Lemma 4. (Van der Vaart and Wellner, 1996, Lemma 2.2.9) *Let* $X_1, \dots, X_n$ *be independent random variables with bounded ranges* $[-M, M]$ *and zero means. Then,*

$$
\Pr\left(\left|\sum_{i=1}^n X_i\right| > x\right) \le 2\exp\left\{-\frac{x^2}{2(v + Mx/3)}\right\}
$$

for all x *and* $v \ge Var\left(\sum_{i=1}^n X_i\right)$.

3.1.1 Proof of Lemma 1

First two statements are the same as Lemma 2 of FH24b. The third statement is established by the double inequality for the trigamma function in Theorem 4 of Gordon (1994), equation (A2) of FH24a and $x \in [\eta, \eta^{-1}]$. ■

3.1.2 Proof of Lemma 2

Because first two statements are the same as Lemma 3 of FH24b, we only show (5). Observe that

$$
\left|\frac{\partial^2 K_{G(x,b)}(u)}{\partial x^2}\right| = \frac{1}{b^2}\left|\mathcal{L}^2_{G(x,b)}(u) - \Psi_1\left(\frac{x}{b} + 1\right)\right| K_{G(x,b)}(u),
$$

where

$$\left|\mathcal{L}^2_{G(x,b)}(u) - \Psi_1\left(\frac{x}{b}+1\right)\right|$$
$$\leq |\ln u|^2 + 2\left|\ln b + \Psi\left(\frac{x}{b}+1\right)\right| |\ln u| + \left\{\ln b + \Psi\left(\frac{x}{b}+1\right)\right\}^2 + \left|\Psi_1\left(\frac{x}{b}+1\right)\right|.$$

Using the fact that $|\ln z|^2 \leq \max\{z, z^{-1}\}$ for $z > 0$ and Lemma 1, jointly with the arguments in the proof of Lemma 3(ii) in FH24b, we can reach the following conclusions as $n \to \infty$, uniformly on $(x,u) \in [\eta, \eta^{-1}] \times \mathbb{R}_+$.

$$|\ln u|^2 K_{G(x,b)}(u) = O\left(b^{-\frac{1}{2}}\eta^{-\frac{3}{2}}\right).$$
$$\left|\ln b + \Psi\left(\frac{x}{b}+1\right)\right| |\ln u| K_{G(x,b)}(u) = O\left(b^{-\frac{1}{2}}\eta^{-\frac{5}{2}}\right).$$
$$\left\{\ln b + \Psi\left(\frac{x}{b}+1\right)\right\}^2 K_{G(x,b)}(u) = O\left(b^{-\frac{1}{2}}\eta^{-\frac{3}{2}}\right).$$
$$\left|\Psi_1\left(\frac{x}{b}+1\right)\right| K_{G(x,b)}(u) = O\left(b^{\frac{1}{2}}\eta^{-\frac{3}{2}}\right).$$

Then, the result immediately follows. ■

3.1.3 Proof of Lemma 3

Part(i)-(a) is the same as Lemma 1 of FH24b, and Part(i)-(b) is given in the proof of Theorem 2.1(i) by FH24a. For Part (ii), it follows from Lemma A.1 of FH24a and a property of the gamma function that

$$E\left\{\mathcal{L}^2_{G(x,b)}(\xi_x)\,\xi_x^m\right\}$$
$$= b^m \prod_{k=1}^{m}\left(\frac{x}{b}+k\right)\left[\left\{\Psi\left(\frac{x}{b}+m+1\right) - \Psi\left(\frac{x}{b}+1\right)\right\}^2 + \Psi_1\left(\frac{x}{b}+m+1\right)\right]$$

for $m \in \mathbb{N}$. Equations (A1) and (A2) of FH24a can further simplify this quantity as

$$E\left\{\mathcal{L}^2_{G(x,b)}(\xi_x)\,\xi_x^m\right\}$$
$$= b^m \prod_{k=1}^{m}\left(\frac{x}{b}+k\right)\left\{\Psi_1\left(\frac{x}{b}+1\right) + \left(\sum_{k=1}^{m}\frac{1}{x/b+k}\right)^2 - \sum_{k=1}^{m}\frac{1}{(x/b+k)^2}\right\}.$$

The stated results can be established by straightforward but tedious calculations, Lemma 1 and $x \in [\eta, \eta^{-1}]$. ■

3.1.4 Proof of Theorem 1

The notations

$$a_{np} = \sqrt{\frac{\ln n}{n b_p \eta_p \sqrt{\prod_{j=1}^{d} b_j \eta_j}}},\ N_{np} = a_{np}^{-1} \frac{1}{b_p \eta_p} \left(\prod_{j=1}^{d} b_j \eta_j \right)^{-\frac{1}{2}} \left(\sum_{j=1}^{d} \frac{1}{b_j \eta_j} \right),\ \text{and}$$

$$\varsigma_{inp}^{\mathcal{J}} (\mathbf{x}) = \frac{1}{n b_p} \left[\mathcal{L}_{\mathcal{J}(x_p, b_p)} (X_{pi}) \mathbb{K}_{\mathcal{J}(\mathbf{x},\mathbf{b})} (\mathbf{X}_i) - E \left\{ \mathcal{L}_{\mathcal{J}(x_p, b_p)} (X_{pi}) \mathbb{K}_{\mathcal{J}(\mathbf{x},\mathbf{b})} (\mathbf{X}_i) \right\} \right]$$

for $\mathcal{J} \in \{G, B\}$ are adopted, where X_{pi} is the pth element of the ith observation $\mathbf{X}_i$. Below we demonstrate the next two statements.

$$\sup_{\mathbf{x} \in \mathbb{S}_{\mathbf{X}}^{G}} \left| E \left\{ \hat{f}_{G,p}^{(1)} (\mathbf{x}) \right\} - f_p^{(1)} (\mathbf{x}) \right| = O \left(\sum_{j=1}^{d} b_j \right). \tag{6}$$

$$\sup_{\mathbf{x} \in \mathbb{S}_{\mathbf{X}}^{G}} \left| \hat{f}_{G,p}^{(1)} (\mathbf{x}) - E \left\{ \hat{f}_{G,p}^{(1)} (\mathbf{x}) \right\} \right| = O_p (a_{np}). \tag{7}$$

Proof of (6). We may write

$$\begin{aligned} E \left\{ \hat{f}_{G,p}^{(1)} (\mathbf{x}) \right\} &= \frac{1}{b_p} \int_{\mathbb{R}_+^d} \mathcal{L}_{G(x_p, b_p)} (u_p) \mathbb{K}_{G(\mathbf{x},\mathbf{b})} (\mathbf{u}) f (\mathbf{u}) \mathbf{du} \\ &= \frac{1}{b_p} E \left\{ \mathcal{L}_{G(x_p, b_p)} \left(\xi_{x_p} \right) f (\xi_{\mathbf{x}}) \right\}, \end{aligned}$$

where $\xi_{\mathbf{x}} := (\xi_{x_1}, \ldots, \xi_{x_d})^{\top}$, $\xi_{x_j} \overset{d}{=} G (x_j / b_j + 1, b_j)$ and $\xi_{x_j} \perp\!\!\!\perp \xi_{x_k}$ for all $j \neq k$. Then, by a third-order Taylor expansion of $f (\xi_{\mathbf{x}})$ around $\xi_{\mathbf{x}} = \mathbf{x}$,

$$\begin{aligned} & E \left\{ \hat{f}_{G,p}^{(1)} (\mathbf{x}) \right\} \\ &= \frac{1}{b_p} f (\mathbf{x}) E \left\{ \mathcal{L}_{G(x_p, b_p)} \left(\xi_{x_p} \right) \right\} \\ &+ \frac{1}{b_p} \sum_{j=1}^{d} f_j^{(1)} (\mathbf{x}) E \left\{ \mathcal{L}_{G(x_p, b_p)} \left(\xi_{x_p} \right) \left(\xi_{x_j} - x_j \right) \right\} \\ &+ \frac{1}{2 b_p} \sum_{j=1}^{d} \sum_{k=1}^{d} f_{jk}^{(2)} (\mathbf{x}) E \left\{ \mathcal{L}_{G(x_p, b_p)} \left(\xi_{x_p} \right) \left(\xi_{x_j} - x_j \right) (\xi_{x_k} - x_k) \right\} \\ &+ \frac{1}{6 b_p} \sum_{j=1}^{d} \sum_{k=1}^{d} \sum_{\ell=1}^{d} f_{jk\ell}^{(3)} (\mathbf{x}) E \left\{ \mathcal{L}_{G(x_p, b_p)} \left(\xi_{x_p} \right) \left(\xi_{x_j} - x_j \right) (\xi_{x_k} - x_k) (\xi_{x_\ell} - x_\ell) \right\} \\ &+ \frac{1}{6 b_p} \sum_{j=1}^{d} \sum_{k=1}^{d} \sum_{\ell=1}^{d} E \left[\left\{ f_{jk\ell}^{(3)} (\bar{\mathbf{x}}) - f_{jk\ell}^{(3)} (\mathbf{x}) \right\} \mathcal{L}_{G(x_p, b_p)} \left(\xi_{x_p} \right) \left(\xi_{x_j} - x_j \right) (\xi_{x_k} - x_k) (\xi_{x_\ell} - x_\ell) \right] \\ &= D_1 + D_2 + D_3 + D_4 + D_5 \text{ (say)} \end{aligned}$$

for some $\bar{\mathbf{x}}$ joining $\xi_{\mathbf{x}}$ and $\mathbf{x}$.

By Lemma 3, $D_1 = 0$ and $D_2 = f_p^{(1)}(\mathbf{x})$. Since $\left|f_{jk}^{(2)}(\mathbf{x})\right|$, $\left|f_{jjj}^{(3)}(\mathbf{x})\,x_j\right|$ and $\left|f_{jkk}^{(3)}(\mathbf{x})\,x_k\right|$ are all uniformly bounded, Lemma 3 also establishes that $|D_3|, |D_4| = O\left(\sum_{j=1}^{d} b_j\right)$ uniformly on $\mathbf{x} \in \mathbb{S}_{\mathbf{X}}^{G}$. Finally, by combining Assumption G2(ii), the Cauchy-Schwarz inequality, Lemma 3, and Assumption G3(ii), it can be shown that $|D_5| = O\left(\rho^{3/2}\right) = o\left(\sum_{j=1}^{d} b_j\right)$ uniformly on $\mathbf{x} \in \mathbb{S}_{\mathbf{X}}^{G}$. Then, (6) is demonstrated.

Proof of (7). As in the proof of Theorem 1 in FH24b, our proof takes the following two steps. The truncation step in the proof of Theorem 1 in FH24b is unnecessary.

1. Split each edge of the d-hyperrectangle $\mathbb{S}_{\mathbf{X}}^{G}$ into N_{np} equally-spaced grids to create N_{np}^{d} sub-hyperrectangles, and replace the supremum with a maximization over the finite N_{np}^{d} sub-hyperrectangles.

2. Employ Lemma 4 (Bernstein's inequality) to bound the remainder term.

Step 1. Let $\mathbf{A}_h$ be the hth sub-hyperrectangle for $h \in \left\{1, \ldots, N_{np}^{d}\right\}$. Also let $\mathbf{x}_h$ be the most distant point from the origin in $\mathbf{A}_h$, i.e., $\mathbf{x}_h := \arg\max_{\mathbf{x} \in \mathbf{A}_h} \|\mathbf{x}\|$. Suppose that the design point $\mathbf{x}$ falls into $\mathbf{A}_h$. Then, the order of magnitude in $\sup_{\mathbf{x} \in \mathbf{A}_h} \left|\sum_{i=1}^{n} \varsigma_{inp}^{G}(\mathbf{x}) - \sum_{i=1}^{n} \varsigma_{inp}^{G}(\mathbf{x}_h)\right|$ is determined by

$$\frac{1}{b_p}\left|\mathcal{L}_{G(x_p,b_p)}(X_{pi})\,\mathbb{K}_{G(\mathbf{x},\mathbf{b})}(\mathbf{X}_i) - \mathcal{L}_{G\left(x_{ph},b_p\right)}(X_{pi})\,\mathbb{K}_{G(\mathbf{x}_h,\mathbf{b})}(\mathbf{X}_i)\right|,$$

where x_{ph} is the pth element of $\mathbf{x}_h$. Now, by the mean-value theorem,

$$\begin{aligned}
&\frac{1}{b_p}\left|\mathcal{L}_{G(x_p,b_p)}(u_p)\,\mathbb{K}_{G(\mathbf{x},\mathbf{b})}(\mathbf{u}) - \mathcal{L}_{G\left(x_{ph},b_p\right)}(u_p)\,\mathbb{K}_{G(\mathbf{x}_h,\mathbf{b})}(\mathbf{u})\right| \\
&= \frac{1}{b_p}\left|\left\{\mathsf{D}_{\mathbf{x}}\mathcal{L}_{G(\tilde{x}_p,b_p)}(u_p)\,\mathbb{K}_{G(\tilde{\mathbf{x}},\mathbf{b})}(\mathbf{u})\right\}^{\top}(\mathbf{x}-\mathbf{x}_h)\right| \\
&\le \frac{1}{b_p}\sup_{(\mathbf{x},\mathbf{u}) \in \mathbf{A}_h \times \mathbb{R}_{+}^{d}}\left\|\mathsf{D}_{\mathbf{x}}\mathcal{L}_{G(x_p,b_p)}(u_p)\,\mathbb{K}_{G(\mathbf{x},\mathbf{b})}(\mathbf{u})\right\|\sup_{\mathbf{x} \in \mathbf{A}_h}\|\mathbf{x}-\mathbf{x}_h\|
\end{aligned}$$

for some $\tilde{\mathbf{x}}$ joining $\mathbf{x}$ and $\mathbf{x}_h$.

By (3) and (4), the jth element of $(1/b_p)\,\mathsf{D}_{\mathbf{x}}\mathcal{L}_{G(x_p,b_p)}(u_p)\,\mathbb{K}_{G(\mathbf{x},\mathbf{b})}(\mathbf{u})$ for $j \neq p$ satisfies

$$\sup_{(\mathbf{x},\mathbf{u})\in\mathbf{A}_h\times\mathbb{R}_+^d}\left|\frac{1}{b_pb_j}\mathcal{L}_{G(x_p,b_p)}(u_p)\,\mathcal{L}_{G(x_j,b_j)}(u_j)\,\mathbb{K}_{G(\mathbf{x},\mathbf{b})}(\mathbf{u})\right| = O\left\{\frac{1}{b_p\eta_pb_j\eta_j}\left(\prod_{j=1}^d b_j\eta_j\right)^{-\frac{1}{2}}\right\}.$$

Combining this with (5) yields

$$\begin{aligned}
&\frac{1}{b_p}\sup_{(\mathbf{x},\mathbf{u})\in\mathbf{A}_h\times\mathbb{R}_+^d}\left\|\mathsf{D}_{\mathbf{x}}\mathcal{L}_{G(x_p,b_p)}(u_p)\,\mathbb{K}_{G(\mathbf{x},\mathbf{b})}(\mathbf{u})\right\| \\
&= O\left\{\frac{1}{b_p\eta_p}\left(\prod_{j=1}^d b_j\eta_j\right)^{-\frac{1}{2}}\left(\sum_{j=1}^d \frac{1}{b_j^2\eta_j^2}\right)^{\frac{1}{2}}\right\} \\
&\leq O\left\{\frac{1}{b_p\eta_p}\left(\prod_{j=1}^d b_j\eta_j\right)^{-\frac{1}{2}}\left(\sum_{j=1}^d \frac{1}{b_j\eta_j}\right)\right\}.
\end{aligned}$$

It follows from $\sup_{\mathbf{x}\in\mathbf{A}_h}\|\mathbf{x}-\mathbf{x}_h\| = O\left(N_{np}^{-1}\right)$ and the definition of N_{np} that uniformly on $(\mathbf{x},\mathbf{u})\in\mathbf{A}_h\times\mathbb{R}_+^d$,

$$\begin{aligned}
&\frac{1}{b_p}\left|\mathcal{L}_{G(x_p,b_p)}(u_p)\,\mathbb{K}_{G(\mathbf{x},\mathbf{b})}(\mathbf{u}) - \mathcal{L}_{G\left(x_{ph},b_p\right)}(u_p)\,\mathbb{K}_{G(\mathbf{x}_h,\mathbf{b})}(\mathbf{u})\right| \\
&\leq O\left\{N_{np}^{-1}\frac{1}{b_p\eta_p}\left(\prod_{j=1}^d b_j\eta_j\right)^{-\frac{1}{2}}\left(\sum_{j=1}^d \frac{1}{b_j\eta_j}\right)\right\} = O\left(a_{np}\right).
\end{aligned}$$

Therefore,

$$\max_{1\leq h\leq N_{np}^d}\sup_{\mathbf{x}\in\mathbf{A}_h}\left|\sum_{i=1}^n \varsigma_{inp}^G(\mathbf{x}) - \sum_{i=1}^n \varsigma_{inp}^G(\mathbf{x}_h)\right| = O\left(a_{np}\right).$$

Step 2. Two bounds M and v for Lemma 4 can be specified as follows. First, (3), (4) and $\int_{\mathbb{R}_+^d} f(\mathbf{u})\,\mathbf{du} = 1$ lead to

$$\frac{1}{b_p}\left|\mathcal{L}_{G(x_p,b_p)}(X_{pi})\,\mathbb{K}_{G(\mathbf{x},\mathbf{b})}(\mathbf{X}_i)\right| \leq 4\left(\frac{2}{\pi}\right)^{\frac{d}{2}}\frac{1}{b_p\eta_p}\left(\prod_{j=1}^d b_j\eta_j\right)^{-\frac{1}{2}}$$

and

$$\left|E\left[\frac{1}{b_p}\left\{\mathcal{L}_{G(x_p,b_p)}(X_{pi})\,\mathbb{K}_{G(\mathbf{x},\mathbf{b})}(\mathbf{X}_i)\right\}\right]\right| \leq 4\left(\frac{2}{\pi}\right)^{\frac{d}{2}}\frac{1}{b_p\eta_p}\left(\prod_{j=1}^d b_j\eta_j\right)^{-\frac{1}{2}}.$$

Then, by the definition of a_{np},

$$\left|\varsigma_{inp}^{G}(\mathbf{x})\right| \leq \frac{8\,(2/\pi)^{d/2}}{nb_p\eta_p\sqrt{\prod_{j=1}^{d} b_j\eta_j}} = 8\left(\frac{2}{\pi}\right)^{\frac{d}{2}}\frac{a_{np}^2}{\ln n} =: M.$$

Second, by (1),

$$\begin{aligned} Var\left\{\sum_{i=1}^{n}\varsigma_{inp}^{G}(\mathbf{x})\right\} &= \sum_{i=1}^{n} Var\left\{\varsigma_{inp}^{G}(\mathbf{x})\right\} \\ &\leq \frac{1}{n^2 b_p^2}\sum_{i=1}^{n} E\left\{\mathcal{L}_{G(x_p,b_p)}^{2}(X_{pi})\,\mathbb{K}_{G(\mathbf{x},\mathbf{b})}^{2}(\mathbf{X}_i)\right\} \\ &\leq \frac{C_0}{nb_p^2}\int_{\mathbb{R}_+^d}\mathcal{L}_{G(x_p,b_p)}^{2}(u_p)\,\mathbb{K}_{G(\mathbf{x},\mathbf{b})}^{2}(\mathbf{u})\,\mathbf{du}. \end{aligned} \tag{8}$$

Following the argument in the proof of Theorem 2.1(ii) in FH24a, we may write $\int_{\mathbb{R}_+}\mathcal{L}_{G(x,b)}^{2}(u)\,K_{G(x,b)}^{2}(u)\,du =: A_{G(b,x)}\Lambda_{G(b,x)}$, where

$$A_{G(b,x)} := \frac{b^{-1}\Gamma(2x/b+1)}{2^{2x/b+1}\Gamma^2(x/b+1)} \leq \frac{b^{-1/2}x^{-1/2}}{2\sqrt{\pi}}$$

by equation (3.4) of Chen (2000), and

$$\Lambda_{G(b,x)} = \left[\left\{\Psi\left(\frac{x}{b}+1\right)-\Psi\left(\frac{2x}{b}+1\right)+\ln 2\right\}^2+\Psi_1\left(\frac{2x}{b}+1\right)\right] \leq \frac{b}{2x}\left\{1+o(1)\right\}$$

uniformly on $x \in [\eta, \eta^{-1}]$ by Lemma 1. Then, by $o(1) \leq 1$ for a sufficiently large n and $x \in [\eta, \eta^{-1}]$,

$$\int_{\mathbb{R}_+}\mathcal{L}_{G(x,b)}^{2}(u)\,K_{G(x,b)}^{2}(u)\,du \leq \frac{b^{-1/2}x^{-1/2}}{2\sqrt{\pi}}\left[\frac{b}{2x}\left\{1+o(1)\right\}\right] \leq \frac{b^{1/2}\eta^{-3/2}}{2\sqrt{\pi}}. \tag{9}$$

It follows from (8), (9) and the definition of a_{np} that

$$Var\left\{\sum_{i=1}^{n}\varsigma_{inp}^{G}(\mathbf{x})\right\} \leq \frac{C_0/\left(2\sqrt{\pi}\right)^d}{nb_p\eta_p\sqrt{\prod_{j=1}^{d} b_j\eta_j}} = \frac{C_0}{\left(2\sqrt{\pi}\right)^d}\frac{a_{np}^2}{\ln n} =: v.$$

Lemma 4 establishes that for such M and v and an arbitrarily chosen $K > 0$,

$$\Pr\left\{\left|\sum_{i=1}^{n}\varsigma_{inp}^{G}(\mathbf{x})\right| > K\sqrt{\frac{C_0}{\left(2\sqrt{\pi}\right)^d}}a_{np}\right\} \leq 2\exp\left[-\frac{K^2\ln n}{2\left\{1+\left(\frac{8}{3}\right)\left(\frac{4}{\sqrt{\pi}}\right)^{d/2}\frac{Ka_{np}}{\sqrt{C_0}}\right\}}\right].$$

Then, it follows from $a_{np} = o(1)$ that $(8/3)\left(4/\sqrt{\pi}\right)^{d/2} K a_{np}/\sqrt{C_0} \leq 1$ for a sufficiently large n, and as a result,

$$\Pr\left\{\left|\sum_{i=1}^{n} \varsigma_{inp}^{G}(\mathbf{x})\right| > K\sqrt{\frac{C_0}{\left(2\sqrt{\pi}\right)^d}} a_{np}\right\} \leq 2n^{-\frac{K^2}{4}}$$

is the case. It holds that

$$\begin{aligned}
&\Pr\left\{\max_{1\leq h\leq N_{np}^d}\left|\sum_{i=1}^{n} \varsigma_{inp}^{G}(\mathbf{x}_h)\right| > K\sqrt{\frac{C_0}{\left(2\sqrt{\pi}\right)^d}} a_{np}\right\} \\
&\leq \sum_{h=1}^{N_{np}^d} \Pr\left\{\left|\sum_{i=1}^{n} \varsigma_{inp}^{G}(\mathbf{x}_h)\right| > K\sqrt{\frac{C_0}{\left(2\sqrt{\pi}\right)^d}} a_{np}\right\} \\
&\leq N_{np}^d \times \max_{1\leq h\leq N_{np}^d} \Pr\left\{\left|\sum_{i=1}^{n} \varsigma_{inp}^{G}(\mathbf{x}_h)\right| > K\sqrt{\frac{C_0}{\left(2\sqrt{\pi}\right)^d}} a_{np}\right\} \\
&= O\left(N_n^d n^{-\frac{K^2}{4}}\right). \qquad (10)
\end{aligned}$$

Putting $K = 2\sqrt{5d}$ and using the definitions of N_{np} and a_{np}, we have

$$N_{np}^d n^{-\frac{K^2}{4}} = \left[a_{np}^9 (\ln n)^{-5} (b_p\eta_p)^4 \left(\prod_{j=1}^{d} b_j\eta_j\right)\left\{\sum_{j=1}^{d}\left(\prod_{k=1,k\neq j}^{d} b_k\eta_k\right)\right\}\right]^d \to 0$$

as $n \to \infty$. Hence,

$$\max_{1\leq h\leq N_{np}^d}\left|\sum_{i=1}^{n} \varsigma_{inp}^{G}(\mathbf{x}_h)\right| = O_p(a_{np}).$$

The results from Steps 1 and 2 establishes (7). This completes the proof. ■

3.2 Proof of Theorem 2

While the definition of N_{np} is changed to

$$N_{np} = n^{1+\epsilon}\frac{1}{b_p\eta_p}\left(\prod_{j=1}^{d} b_j\eta_j\right)^{-\frac{1}{2}}\left(\sum_{j=1}^{d}\frac{1}{b_j\eta_j}\right)$$

for an arbitrarily small $\epsilon > 0$, all other notations used in the proof of Theorem 1 remain unchanged. Then, it suffices to demonstrate that

$$\sup_{\mathbf{x}\in\mathbb{S}_{\mathbf{X}}^{G}}\left|\hat{f}_{G,p}^{(1)}(\mathbf{x}) - E\left\{\hat{f}_{G,p}^{(1)}(\mathbf{x})\right\}\right| = O(a_{np}) \ a.s. \qquad (11)$$

By the definition of a_{np}, it holds that $n^{-(1+\epsilon)} \leq O\left(n^{-1/2}\right) \leq o\left(a_{np}\right) \leq O\left(a_{np}\right)$. Hence,

$$\begin{aligned}
&\max_{1\leq h\leq N_{np}^d}\sup_{\mathbf{x}\in\mathbf{A}_h}\left|\sum_{i=1}^n \varsigma_{inp}^G\left(\mathbf{x}\right)-\sum_{i=1}^n \varsigma_{inp}^G\left(\mathbf{x}_h\right)\right| \\
&= O\left\{N_{np}^{-1}\frac{1}{b_p\eta_p}\left(\prod_{j=1}^d b_j\eta_j\right)^{-\frac{1}{2}}\left(\sum_{j=1}^d\frac{1}{b_j\eta_j}\right)\right\} \\
&= O\left(a_{np}\right). \qquad (12)
\end{aligned}$$

The strengthened condition (iii) in Assumption G3 also leads to

$$\frac{1}{b_p\eta_p}\left(\prod_{j=1}^d b_j\eta_j\right)^{-\frac{1}{2}}\left(\sum_{j=1}^d\frac{1}{b_j\eta_j}\right)=O\left\{n^{\frac{1}{1-\kappa}}\left(\frac{\left(b_p\eta_p\right)^\kappa\left(\prod_{j=1}^d b_j\eta_j\right)^{\frac{\kappa}{2}}}{\ln n}\right)^{\frac{1}{1-\kappa}}\right\}\leq O\left(n^{\frac{1}{1-\kappa}}\right),$$

because $\left(b_p\eta_p\right)^\kappa\left(\prod_{j=1}^d b_j\eta_j\right)^{\kappa/2}/\ln n = o(1)$. Putting $K = 2\sqrt{(d+1)(1+\epsilon)+d/(1-\kappa)}$ gives $N_{np}^d n^{-K^2/4} = O\left\{n^{-(1+\epsilon)}\right\}$. Finally, by (10),

$$\sum_{n=1}^\infty \Pr\left\{\max_{1\leq h\leq N_{np}^d}\left|\sum_{i=1}^n\varsigma_{inp}^G\left(\mathbf{x}_h\right)\right| > K\sqrt{\frac{C_0}{\left(2\sqrt{\pi}\right)^d}}a_{np}\right\}\leq\sum_{n=1}^\infty O\left\{n^{-(1+\epsilon)}\right\}<\infty,$$

and thus

$$\max_{1\leq h\leq N_{np}^d}\left|\sum_{i=1}^n\varsigma_{inp}^G\left(\mathbf{x}_h\right)\right| = O\left(a_{np}\right)\ \ a.s. \qquad (13)$$

by the Borel-Cantelli lemma. The stated result can be established by recognizing that (12) and (13) suffice for (11). ■

3.3 Proof of Theorem 3

This proof requires the following lemmata. Lemma 5 corresponds to Lemma 1, except that the range of the design point x is modified. Lemma 6 documents uniform bounds of the univariate beta kernel and its first- and second-order derivatives with respect to the design point x. In particular, the $O\left(b^{-3/2}\eta^{-3/2}\right)$ bound in (15) is sharper than the $O\left(b^{-5/2}\eta^{-1/2}\right)$ bound in Lemma 3 of HMP22, because it follows from $b/\eta = o(1)$ that $b^{-3/2}\eta^{-3/2} = o\left(b^{-5/2}\eta^{-1/2}\right)$ holds. Lemma 7 presents moments of the univariate beta kernel around the design point x. As before, some odd moments are not used in the proofs and thus omitted.

Lemma 5. *Under the same condition as in Lemma 1, the followings hold true as $n \to \infty$.*

$$\sup_{x\in[\eta,1-\eta]}\left|\frac{\Gamma\left(x/b+1\right)}{\sqrt{2\pi}\left(x/b\right)^{x/b+1/2}\exp\left(-x/b\right)}-1\right|=O\left(\frac{b}{\eta}\right).$$

$$\sup_{x\in[\eta,1-\eta]}\left|\frac{\Psi\left(x/b+1\right)-\ln\left(x/b\right)}{b/\left(2x\right)}-1\right|=O\left(\frac{b}{\eta}\right).$$

$$\sup_{x\in[\eta,1-\eta]}\left|\frac{\Psi_1\left(x/b+1\right)}{b/x}-1\right|=O\left(\frac{b}{\eta}\right).$$

Lemma 6. *Under the same condition as in Lemma 1, the followings hold true as $n \to \infty$.*

$$\sup_{(x,u)\in[\eta,1-\eta]\times[0,1]}K_{B(x,b)}\left(u\right)\leq\frac{9}{4\sqrt{\pi}}b^{-\frac{1}{2}}\eta^{-\frac{1}{2}}. \tag{14}$$

$$\sup_{(x,u)\in[\eta,1-\eta]\times[0,1]}\left|\frac{\partial K_{B(x,b)}\left(u\right)}{\partial x}\right|\leq\frac{13}{\sqrt{\pi}}b^{-\frac{3}{2}}\eta^{-\frac{3}{2}}. \tag{15}$$

$$\sup_{(x,u)\in[\eta,1-\eta]\times[0,1]}\left|\frac{\partial^2 K_{B(x,b)}\left(u\right)}{\partial x^2}\right|=O\left(b^{-\frac{5}{2}}\eta^{-\frac{5}{2}}\right). \tag{16}$$

Lemma 7. *Let $\theta_x \overset{d}{=} Beta\left(x/b+1,b\right)$.*

(i) *The followings hold true as $n \to \infty$, uniformly on $x \in [0,1]$.*

$$\begin{aligned}(a)\quad & \left|E\left(\theta_x-x\right)\right|=O\left(b\right),\ E\left(\theta_x-x\right)^2=O\left(b\right),\ E\left(\theta_x-x\right)^4=O\left(b^2\right),\\ & E\left(\theta_x-x\right)^6=O\left(b^3\right),\ \text{and}\ E\left(\theta_x-x\right)^8=O\left(b^4\right).\\ (b)\quad & E\left\{\mathcal{L}_{B(x,b)}\left(\theta_x\right)\right\}=0,\ \left|b^{-1}E\left\{\mathcal{L}_{B(x,b)}\left(\theta_x\right)\left(\theta_x-x\right)\right\}-1\right|=O\left(b\right),\\ & \left|E\left\{\mathcal{L}_{B(x,b)}\left(\theta_x\right)\left(\theta_x-x\right)^2\right\}\right|=O\left(b^2\right),\ \text{and}\ \left|E\left\{\mathcal{L}_{B(x,b)}\left(\theta_x\right)\left(\theta_x-x\right)^3\right\}\right|=O\left(b^2\right).\end{aligned}$$

(ii) *Under the same condition as in Lemma 1, the followings also hold true as $n \to \infty$, uniformly on $x \in [\eta, 1-\eta]$.*

$$\begin{aligned}& E\left\{\mathcal{L}^2_{B(x,b)}\left(\theta_x\right)\right\}=O\left(b/\eta\right),\ E\left\{\mathcal{L}^2_{B(x,b)}\left(\theta_x\right)\left(\theta_x-x\right)^2\right\}=O\left(b^2\right),\\ & E\left\{\mathcal{L}^2_{B(x,b)}\left(\theta_x\right)\left(\theta_x-x\right)^4\right\}=O\left(b^3\right),\ E\left\{\mathcal{L}^2_{B(x,b)}\left(\theta_x\right)\left(\theta_x-x\right)^6\right\}=O\left(b^4\right),\ \text{and}\\ & E\left\{\mathcal{L}^2_{B(x,b)}\left(\theta_x\right)\left(\theta_x-x\right)^8\right\}=O\left(b^5\right).\end{aligned}$$

3.3.1 Proof of Lemma 5

This lemma can be established as minor modifications of Lemma 1. ■

3.3.2 Proof of Lemma 6

Because (14) is the same as Lemma 2 of HMP22, we only need to show (15) and (16).

Proof of (15). We consider the cases with $u = 0, 1$ and $u \in (0, 1)$ separately. If $u = 0, 1$, then $x/b, (1-x)/b > 0$ holds for $x \in [\eta, 1-\eta]$, and thus $K_{B(x,b)}(0) = K_{B(x,b)}(1) = 0$. Consequently, $\partial K_{B(x,b)}(0)/\partial x = \partial K_{B(x,b)}(1)/\partial x = 0$, and the result trivially holds.

For $u \in (0, 1)$, observe that

$$\begin{aligned} b\left|\frac{\partial K_{B(x,b)}(u)}{\partial x}\right| &\leq |\ln u|\, K_{B(x,b)}(u) + \left|\ln b + \Psi\left(\frac{x}{b}+1\right)\right| K_{B(x,b)}(u) \\ &\quad + |\ln(1-u)|\, K_{B(x,b)}(u) + \left|\ln b + \Psi\left(\frac{1-x}{b}+1\right)\right| K_{B(x,b)}(u) \\ &= F_1 + F_2 + F_3 + F_4 \text{ (say)}. \end{aligned}$$

For F_1, using $|\ln z| \leq z^{-1}$ for $z \in (0, 1)$ yields $|\ln u|\, K_{B(x,b)}(u) \leq u^{-1} K_{B(x,b)}(u)$. Now, $u^{x/b-1}(1-u)^{(1-x)/b}$ is maximized at $u = (x-b)/(1-b)$. Then, by equation (A2) of HMP22, Lemma 5, $(x-b)^{x/b-1} = x^{x/b-1}e^{-1}\{1+o(1)\}$, $(1-b)^{1/b} = e^{-1}\{1+o(1)\}$, and $x \in [\eta, 1-\eta]$,

$$\begin{aligned} u^{-1}K_{B(x,b)}(u) &\leq \frac{\{(x-b)/(1-b)\}^{x/b-1}\{1-(x-b)/(1-b)\}^{(1-x)/b}}{B\{x/b+1, (1-x)/b+1\}} \\ &= \frac{b^{-1/2}(1-b^2)\{1+o(1)\}}{\sqrt{2\pi}x^{3/2}(1-x)^{1/2}} \leq \frac{b^{-1/2}\{1+o(1)\}}{\sqrt{2\pi}\eta^{3/2}(1-\eta)^{1/2}}. \end{aligned}$$

Finally, taking $o(1) \leq 1$ and $\eta \leq 1/2$ for a sufficiently large n leads to $F_1 \leq (2/\sqrt{\pi})\, b^{-1/2}\eta^{-3/2}$.

For F_2, it follows from Lemma 5, $x \in [\eta, 1-\eta]$ and $|\ln z| \leq z^{-1}$ for $z \in (0, 1)$ that $|\ln b + \Psi(x/b+1)| \leq |\ln x| + o(1) \leq x^{-1} + o(1) \leq \eta^{-1}\{1+o(\eta)\}$. Then, putting $o(\eta) \leq 1$ for a sufficiently large n and using (14), we have $F_2 \leq 2\eta^{-1} \times \{9/(4\sqrt{\pi})\}\, b^{-1/2}\eta^{-1/2} = \{9/(2\sqrt{\pi})\}\, b^{-1/2}\eta^{-3/2}$.

We can also show that $F_3 \leq (2/\sqrt{\pi})\, b^{-1/2}\eta^{-3/2}$ and $F_4 \leq \{9/(2\sqrt{\pi})\}\, b^{-1/2}\eta^{-3/2}$ by realizing that $|\ln(1-z)| \leq (1-z)^{-1}$ for $z \in (0,1)$. Then, the result immediately holds.

Proof of (16). Notice that

$$\frac{\partial^2 K_{B(x,b)}(u)}{\partial x^2} = \frac{1}{b^2}\left[\mathcal{L}^2_{B(x,b)}(u) - \left\{\Psi_1\left(\frac{x}{b}+1\right) + \Psi_1\left(\frac{1-x}{b}+1\right)\right\}\right] K_{B(x,b)}(u).$$

The result can be established by taking the same steps as in the proof of (15), recognizing that $|\ln z|, \ln^2 z \leq z^{-1}$ and $|\ln(1-z)|, \ln^2(1-z) \leq (1-z)^{-1}$ for $z \in (0,1)$, and utilizing Lemma 5 and (14) repeatedly. ■

3.3.3 Proof of Lemma 7

Part (i)-(a) comes from non-central moments of the beta random variable, and Part(i)-(b) is implied by the argument in the proof of Theorem 2.1(i) by FH24a. Part (ii) can be obtained by

$$\begin{aligned} E\left\{\mathcal{L}^2_{B(x,b)}(\theta_x)\,\theta_x^m\right\} = \frac{\prod_{k=1}^m (x/b+k)}{\prod_{k=1}^m (1/b+1+k)} &\left\{\Psi_1\left(\frac{x}{b}+1\right) + \Psi_1\left(\frac{1-x}{b}+1\right)\right. \\ &\left. + \left(\sum_{k=1}^m \frac{1}{x/b+k}\right)^2 - \sum_{k=1}^m \frac{1}{(x/b+k)^2}\right\} \end{aligned}$$

for $m \in \mathbb{N}$, straightforward but tedious calculations, Lemma 5, and $x \in [\eta, 1-\eta]$. ■

3.3.4 Proof of Theorem 3

The notations used in the proof of Theorem 1 are maintained. In what follows, we demonstrate the next two statements.

$$\sup_{\mathbf{x}\in\mathbb{S}^B_{\mathbf{X}}} \left|E\left\{\hat{f}^{(1)}_{B,p}(\mathbf{x})\right\} - f^{(1)}_p(\mathbf{x})\right| = O\left(\sum_{j=1}^d b_j\right). \quad (17)$$

$$\sup_{\mathbf{x}\in\mathbb{S}^B_{\mathbf{X}}} \left|\hat{f}^{(1)}_{B,p}(\mathbf{x}) - E\left\{\hat{f}^{(1)}_{B,p}(\mathbf{x})\right\}\right| = O_p(a_{np}). \quad (18)$$

Proof of (17). Observe that

$$E\left\{\hat{f}^{(1)}_{B,p}(\mathbf{x})\right\} = \frac{1}{b_p}\int_{[0,1]^d} \mathcal{L}_{B(x_p,b_p)}(u_p)\,\mathbb{K}_{B(\mathbf{x},\mathbf{b})}(\mathbf{u})\, f(\mathbf{u})\,\mathbf{du} = \frac{1}{b_p} E\left\{\mathcal{L}_{B(x_p,b_p)}\left(\theta_{x_p}\right) f(\theta_{\mathbf{x}})\right\},$$

where $\theta_{\mathbf{x}} := (\theta_{x_1}, \ldots, \theta_{x_d})^\top$, $\theta_{x_j} \stackrel{d}{=} Beta(x_j/b_j + 1, b_j)$ and $\theta_{x_j} \perp\!\!\!\perp \theta_{x_k}$ for all $j \neq k$. Then, by a third-order Taylor expansion of $f(\theta_{\mathbf{x}})$ around $\theta_{\mathbf{x}} = \mathbf{x}$,

$$\begin{aligned}
&E\left\{\hat{f}^{(1)}_{B,p}(\mathbf{x})\right\} \\
&= \frac{1}{b_p} f(\mathbf{x})\, E\left\{\mathcal{L}_{B(x_p,b_p)}\left(\theta_{x_p}\right)\right\} + \frac{1}{b_p}\sum_{j=1}^{d} f^{(1)}_j(\mathbf{x})\, E\left\{\mathcal{L}_{B(x_p,b_p)}\left(\theta_{x_p}\right)\left(\theta_{x_j} - x_j\right)\right\} \\
&+ \frac{1}{2b_p}\sum_{j=1}^{d}\sum_{k=1}^{d} f^{(2)}_{jk}(\mathbf{x})\, E\left\{\mathcal{L}_{B(x_p,b_p)}\left(\theta_{x_p}\right)\left(\theta_{x_j} - x_j\right)\left(\theta_{x_k} - x_k\right)\right\} \\
&+ \frac{1}{6b_p}\sum_{j=1}^{d}\sum_{k=1}^{d}\sum_{\ell=1}^{d} f^{(3)}_{jk\ell}(\mathbf{x})\, E\left\{\mathcal{L}_{B(x_p,b_p)}\left(\theta_{x_p}\right)\left(\theta_{x_j} - x_j\right)\left(\theta_{x_k} - x_k\right)\left(\theta_{x_\ell} - x_\ell\right)\right\} \\
&+ \frac{1}{6b_p}\sum_{j=1}^{d}\sum_{k=1}^{d}\sum_{\ell=1}^{d} E\left[\left\{f^{(3)}_{jk\ell}(\breve{\mathbf{x}}) - f^{(3)}_{jk\ell}(\mathbf{x})\right\}\mathcal{L}_{B(x_p,b_p)}\left(\theta_{x_p}\right)\left(\theta_{x_j} - x_j\right)\left(\theta_{x_k} - x_k\right)\left(\theta_{x_\ell} - x_\ell\right)\right] \\
&= H_1 + H_2 + H_3 + H_4 + H_5 \text{ (say)}
\end{aligned}$$

for some $\breve{\mathbf{x}}$ joining $\theta_{\mathbf{x}}$ and $\mathbf{x}$.

By Lemma 7, $H_1 = 0$ and $H_2 = f^{(1)}_p(\mathbf{x}) + O(b_p)$ uniformly on $\mathbf{x} \in \mathbb{S}^B_{\mathbf{X}}$. This lemma also establishes that $|H_3|, |H_4| = O\left(\sum_{j=1}^d b_j\right)$ uniformly on $\mathbf{x} \in \mathbb{S}^B_{\mathbf{X}}$. Finally, it follows from Assumption B2, the Cauchy-Schwarz inequality and Lemma 7 that $|H_5| \leq O\left\{(b_p\eta_p)^{-1/2}\left(\sum_{j=1}^d b_j\right)^2\right\}$. This bound is shown to be $o\left(\sum_{j=1}^d b_j\right)$ uniformly on $\mathbf{x} \in \mathbb{S}^B_{\mathbf{X}}$ by Assumption B3(ii). Therefore, (17) is shown.

Proof of (18). The proof takes the same steps as in the proof of (7). Following the argument in Step 1 and using Lemma 6, we can show that

$$\begin{aligned}
&\frac{1}{b_p}\left|\mathcal{L}_{B(x_p,b_p)}(X_{pi})\,\mathbb{K}_{B(\mathbf{x},\mathbf{b})}(\mathbf{X}_i) - \mathcal{L}_{B\left(x_{ph},b_p\right)}(X_{pi})\,\mathbb{K}_{B(\mathbf{x}_h,\mathbf{b})}(\mathbf{X}_i)\right| \\
&\leq O\left\{N_{np}^{-1}\frac{1}{b_p\eta_p}\left(\prod_{j=1}^d b_j\eta_j\right)^{-\frac{1}{2}}\left(\sum_{j=1}^d \frac{1}{b_j\eta_j}\right)\right\} = O(a_{np})
\end{aligned}$$

uniformly on $(\mathbf{x},\mathbf{u}) \in \mathbf{A}_h \times [0,1]^d$. Therefore,

$$\max_{1\le h\le N_{np}^d} \sup_{\mathbf{x}\in\mathbf{A}_h} \left| \sum_{i=1}^n \varsigma_{inp}^G(\mathbf{x}) - \sum_{i=1}^n \varsigma_{inp}^G(\mathbf{x}_h) \right| = O(a_{np})$$

is established.

In Step 2, two bounds M and v for Lemma 4 can be found as follows. First, (14), (15) and $\int_{[0,1]^d} f(\mathbf{u})\,\mathbf{du} = 1$ yield

$$\frac{1}{b_p} \left| \mathcal{L}_{B(x_p,b_p)}(X_{pi}) \mathbb{K}_{B(\mathbf{x},\mathbf{b})}(\mathbf{X}_i) \right| \le \frac{13}{\sqrt{\pi}} \left(\frac{9}{4\sqrt{\pi}} \right)^{d-1} \frac{1}{b_p \eta_p} \left(\prod_{j=1}^d b_j \eta_j \right)^{-\frac{1}{2}}$$

and

$$\left| E\left[\frac{1}{b_p} \left\{ \mathcal{L}_{B(x_p,b_p)}(X_{pi}) \mathbb{K}_{B(\mathbf{x},\mathbf{b})}(\mathbf{X}_i) \right\} \right] \right| \le \frac{13}{\sqrt{\pi}} \left(\frac{9}{4\sqrt{\pi}} \right)^{d-1} \frac{1}{b_p \eta_p} \left(\prod_{j=1}^d b_j \eta_j \right)^{-\frac{1}{2}}.$$

Then,

$$\left| \varsigma_{inp}^B(\mathbf{x}) \right| \le \frac{26}{\sqrt{\pi}} \left(\frac{9}{4\sqrt{\pi}} \right)^{d-1} \frac{1}{n b_p \eta_p \sqrt{\prod_{j=1}^d b_j \eta_j}} = \frac{26}{\sqrt{\pi}} \left(\frac{9}{4\sqrt{\pi}} \right)^{d-1} \frac{a_{np}^2}{\ln n} =: M.$$

Moreover, by (2),

$$Var\left\{ \sum_{i=1}^n \varsigma_{inp}^B(\mathbf{x}) \right\} \le \frac{L}{n b_p^2} \int_{[0,1]^d} \mathcal{L}^2_{B(x_p,b_p)}(u_p) \mathbb{K}^2_{B(\mathbf{x},\mathbf{b})}(\mathbf{u})\,\mathbf{du}.$$

Now, by the proof of Theorem 2.1(ii) in FH24a, $\int_0^1 \mathcal{L}^2_{B(x,b)}(u) K^2_{B(x,b)}(u)\,du =: A_{B(b,x)} \Lambda_{B(b,x)}$, where

$$A_{B(b,x)} = \frac{B\{2x/b+1, 2(1-x)/b+1\}}{B^2\{x/b+1, (1-x)/b+1\}},$$

and

$$\Lambda_{B(b,x)} = \left[\Psi\left(\frac{x}{b}+1\right) - \Psi\left(\frac{1-x}{b}+1\right) - \Psi\left(\frac{2x}{b}+1\right) + \Psi\left\{\frac{2(1-x)}{b}+1\right\} \right]^2 + \Psi_1\left(\frac{2x}{b}+1\right) + \Psi_1\left\{\frac{2(1-x)}{b}+1\right\}.$$

Using Lemma of Chen (1999) and $\eta \le x$ and picking $b \le 1$ and $\eta \le 1/2$ for a sufficiently large n, we have

$$A_{B(b,x)} \le \frac{b^{-1/2}(1+b)^{3/2}}{2\sqrt{\pi}\sqrt{x(1-x)}} \le \frac{2}{\sqrt{\pi}} b^{-\frac{1}{2}} \eta^{-\frac{1}{2}}.$$

Applying Lemma 5 and a similar argument also yields $\Lambda_{B(b,x)} \leq b/\eta$, and thus $A_{B(b,x)}\Lambda_{B(b,x)} \leq (2/\sqrt{\pi})\, b^{1/2}\eta^{-3/2}$ uniformly on $x \in [\eta, 1-\eta]$. Recognizing that $\int_0^1 K_{B(x,b)}^2 (u)\, du = A_{B(b,x)}$ and using the definition of a_{np}, we finally have

$$Var\left\{\sum_{i=1}^{n} \varsigma_{inp}^{B}(\mathbf{x})\right\} \leq \frac{(2/\sqrt{\pi})^d L}{nb_p\eta_p\sqrt{\prod_{j=1}^{d} b_j\eta_j}} = \left(\frac{2}{\sqrt{\pi}}\right)^d L\frac{a_{np}^2}{\ln n} =: v.$$

Lemma 4 implies that for M and v given above and an arbitrarily chosen $K > 0$,

$$\Pr\left\{\left|\sum_{i=1}^{n} \varsigma_{inp}^{B}(\mathbf{x})\right| > K\sqrt{\left(\frac{2}{\sqrt{\pi}}\right)^d La_{np}}\right\} \leq 2\exp\left[-\frac{K^2 \ln n}{2\left\{1 + \left(\frac{26}{3}\right)\frac{(9/4)^{d-1}Ka_{np}}{\left(2\sqrt{\pi}\right)^{d/2}\sqrt{L}}\right\}}\right].$$

Then, it follows from $a_{np} = o(1)$ that $(26/3)\,(9/4)^{d-1}\, Ka_{np}/\left\{(2\sqrt{\pi})^{d/2}\sqrt{L}\right\} \leq 1$ holds for a sufficiently large n, and as a result,

$$\Pr\left\{\left|\sum_{i=1}^{n} \varsigma_{inp}^{B}(\mathbf{x})\right| > K\sqrt{\left(\frac{2}{\sqrt{\pi}}\right)^d La_{np}}\right\} \leq 2n^{-\frac{K^2}{4}}.$$

so that

$$\Pr\left\{\max_{1\leq h\leq N_{np}^d}\left|\sum_{i=1}^{n} \varsigma_{inp}^{B}(\mathbf{x})\right| > K\sqrt{\left(\frac{2}{\sqrt{\pi}}\right)^d La_{np}}\right\} = O\left(N_{np}^d n^{-\frac{K^2}{4}}\right).$$

As before, $K = 2\sqrt{5d}$ and the definitions of N_{np} and a_{np} lead to $N_{np}^d n^{-K^2/4} = o(1)$. Then,

$$\max_{1\leq h\leq N_{np}^d}\left|\sum_{i=1}^{n} \varsigma_{inp}^{B}(\mathbf{x}_h)\right| = O_p(a_{np})$$

is also demonstrated. This completes the proof. ■

3.4 Proof of Theorem 4

The proof is similar to that of Theorem 2, and thus details are omitted. ■

Competing Interests

The authors report that there are no competing interests to declare.

References

[1] Bertin, K., C. Genest, N. Klutchnikoff, and F. Ouimet (2023): "Minimax Properties of Dirichlet Kernel Density Estimators," *Journal of Multivariate Analysis*, 195, Article No.105158.

[2] Bouezmarni, T., and J. V. K. Rombouts (2010): "Nonparametric Density Estimation for Multivariate Bounded Data," *Journal of Statistical Planning and Inference*, 140, 139 - 152.

[3] Chen, S. X. (1999): "Beta Kernel Estimators for Density Functions," *Computational Statistics & Data Analysis*, 31, 131 - 145.

[4] Chen, S. X. (2000): "Probability Density Function Estimation Using Gamma Kernels," *Annals of the Institute of Statistical Mathematics*, 52, 471 - 480.

[5] Funke, B., and M. Hirukawa (2024a): "Density Derivative Estimation Using Asymmetric Kernels," *Journal of Nonparametric Statistics*, forthcoming.

[6] Funke, B., and M. Hirukawa (2024b): "On Uniform Consistency of Nonparametric Estimators Smoothed by the Gamma Kernel," Working Paper.

[7] Funke, B., and R. Kawka (2015): "Nonparametric Density Estimation for Multivariate Bounded Data Using Two Non-Negative Multiplicative Bias Correction Methods," *Computational Statistics & Data Analysis*, 92, 148 - 162.

[8] Gordon, L. (1994): "A Stochastic Approach to the Gamma Function," *American Mathematical Monthly*, 101, 858 - 865.

[9] Härdle, W., and T. M. Stoker (1989): "Investigating Smooth Multiple Regression by the Method of Average Derivatives," *Journal of the American Statistical Association*, 84, 986 - 995.

[10] Hirukawa, M. (2018): *Asymmetric Kernel Smoothing: Theory and Applications in Economics and Finance*. Singapore: Springer.

[11] Hirukawa, M., I. Murtazashvili, and A. Prokhorov (2022): "Uniform Convergence Rates for Nonparametric Estimators Smoothed by the Beta Kernel," *Scandinavian Journal of Statistics*, 49, 1353 - 1382.

[12] Ouimet, F. (2021): "Asymptotic Properties of Bernstein Estimators on the Simplex," *Journal of Multivariate Analysis*, 185, Article No.104874.

[13] Ouimet, F. (2022): "A Symmetric Matrix-Variate Normal Local Approximation for the Wishart Distribution and Some Applications," *Journal of Multivariate Analysis*, 189, Article No.104923.

[14] Ouimet, F., and R. Tolosana-Delgado (2022): "Asymptotic Properties of Dirichlet Kernel Density Estimators," *Journal of Multivariate Analysis*, 187, Article No.104832.

[15] Powell, J. L., J. H. Stock, and T. M. Stoker (1989): "Semiparametric Estimation of Index Coefficients," *Econometrica*, 57, 1403 - 1430.

[16] Racine, J. S. (2019): *An Introduction to the Advanced Theory of Nonparametric Econometrics: A Replicable Approach Using R.* Cambridge, UK: Cambridge University Press.

[17] Rilstone, P. (1991): "Nonparametric Hypothesis Testing with Parametric Rates of Convergence," *International Economic Review*, 32, 209 - 227.

[18] Stone, C. J. (1983): "Optimal Uniform Rate of Convergence for Nonparametric Estimators of a Density Function or Its Derivatives," in M. H. Rizvi, J. S. Rustagi, and D. Siegmund (eds.), *Recent Advances in Statistics: Papers in Honor of Herman Chernoff on His Sixtieth Birthday.* New York: Academic Press, 393 - 406.

[19] Van der Vaart, A. W., and J. A. Wellner (1996): *Weak Convergence and Empirical Processes: With Applications to Statistics.* New York: Springer-Verlag.