



HAL
open science

Kernel density estimation for stationary random fields with values in a finite-dimensional Riemannian manifold

Wiem Nefzi, Anne-Françoise Yao, Salah KHARDANI

► **To cite this version:**

Wiem Nefzi, Anne-Françoise Yao, Salah KHARDANI. Kernel density estimation for stationary random fields with values in a finite-dimensional Riemannian manifold. 2025. hal-04896382

HAL Id: hal-04896382

<https://hal.science/hal-04896382v1>

Preprint submitted on 19 Jan 2025

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Kernel density estimation for stationary random fields with values in a finite-dimensional Riemannian manifold

Nefzi Wiem¹, Khardani Salah¹ and Yao Anne-Françoise^{2,3}

¹ Faculté des Sciences Tunis, Université El-Manar, Laboratoire de Modélisation mathématique, Statistique et Analyse stochastique M2ΦSAS, Tunisia.

khardani_salah@yahoo.fr, nefziwiem24@gmail.com

² Laboratoire de Mathématiques Blaise Pascal, CNRS UMR 6620, Campus des Cézeaux, 63171 Aubière Cedex, France.

anne.yao@uca.fr

³ Centre de Mathématiques Appliquées, Ecole polytechnique Route de Saclay 91128 Palaiseau Cedex.

anne.yao@uca.fr

Abstract

This paper investigates some asymptotic properties of the kernel spatial density estimation for stationary α -mixing process on a finite-dimensional Riemannian manifold without boundary. The results extend beyond the classical independently and identically distributed (i.i.d.) data, focusing on the case where the manifold is known and extending the classical theory to random fields.

Keywords: Kernel density estimator, Riemannian manifolds, Spatial data, Mixing condition, Random fields, Mean Squared Error, Convergence in probability.

1. Introduction

Kernel density estimation (KDE) is a nonparametric method commonly used to estimate the probability density function. While Euclidean spaces have been extensively studied with the assumption of independent and identically distributed (i.i.d.) data, real-world applications frequently involve spatially dependent data that don't fit flat Euclidean structures. This calls for extensions of KDE to more complex settings, such as random fields on Riemannian manifolds.

Non-Euclidean data structures naturally arise in fields like biology (e.g., protein data in [Mardia et al. \(2008\)](#)), geology, medical imaging ([Pennec \(2006\)](#)), and computer vision, where the data often lies on curved spaces. Applying standard Euclidean methods in such cases can lead to significant errors due to the manifold’s curvature, motivating the need for specialized techniques that account for the underlying geometry.

KDE on Riemannian manifolds poses unique challenges because distances and volume elements must respect the geometry of the space. [Pelletier \(2005\)](#) laid the groundwork for KDE on known Riemannian manifolds, focusing on i.i.d. data. Subsequent research has explored various extensions, including applications to manifolds, as in [Kim and Park \(2013\)](#), [Berry and Sauer \(2017\)](#), [Cleanthous et al. \(2020\)](#), [Berenfeld and Hoffmann \(2021\)](#), [Khardani and Yao \(2022\)](#). Recently, [Abdillahi et al. \(2024\)](#) studied some asymptotic properties of the Kernel density estimation for a stochastic process with values in a Riemannian manifold. [Bouzebda and Taachouche \(2024\)](#) examined the strong uniform consistency of generic kernels, including kernel density estimators, on Riemannian manifolds. The study focused on Riemann integrable isotropic kernels distinct from those belonging to the Vapnik-Chervonenkis class. Building on this, [Bouzebda and Taachouche \(2023\)](#) extended the framework by establishing strong uniform consistency results for general kernels on Riemannian manifolds, specifically addressing conditional U -processes with Riemann integrable kernels. In addition, KDE for spatial data in Euclidean settings, as studied by [Carbon et al. \(1997\)](#), [Hallin et al. \(2004\)](#), [Chen \(2008\)](#), [Dabo-Niang et al. \(2014\)](#), [El Machkouri and Reding \(2021\)](#). However, to our knowledge, no prior work has addressed the case of spatially dependent random fields on manifolds, particularly under the α -mixing condition, which is the focus of this paper.

This paper investigates the asymptotic properties of KDE for stationary random fields on finite-dimensional Riemannian manifolds. Our work builds upon the classical kernel density estimation framework, extending it to handle dependencies in the data through α -mixing condition.

The paper is structured as follows. Section 2 presents the technical background, including key definitions and notations related to Riemannian geometry and kernel density estimation. Section 3 provides the main results, and we give the proofs in Section 4.

2. Technical background

This study pertains to any measurable stationary spatial process $(X_{\mathbf{i}}, \mathbf{i} \in \mathbb{N}^N)$, defined on some probability space $(\Omega, \mathcal{F}, \mathbb{P})$ with values on a Riemannian submanifold $\mathcal{M} (\subset \mathbb{R}^d)$, where $d \geq 2$.

2.1 Notations related to the structure of Riemannian submanifold

In the following, for any $x \in \mathcal{M}$, $T_x\mathcal{M}$ will denote the tangent space to \mathcal{M} at x . We assume that (\mathcal{M}, g) is endowed with a measure (ν_g) and is geodesically complete and compact

without boundary, ensuring, by the Hopf Rinow Theorem, that (\mathcal{M}, d_g) is a complete metric space, where d_g is the metric induced by g . For more details, see [Gallot et al. \(2004\)](#). This allows us to define the exponential map at x , $\exp_x : T_x\mathcal{M} \rightarrow \mathcal{M}$, such that for any $v \in T_x\mathcal{M}$, $\exp_x(v) = \gamma_v(1)$, where γ_v is starting at x with velocity v . Specifically, $\gamma_v(t) = \exp_x(tv)$ for $t \in \mathbb{R}$, with $\gamma_v(0) = x$ and $\dot{\gamma}_v(0) = v$. Its inverse, denoted by \exp_x^{-1} , is a map from the image $(\text{Im}(\exp_x))$ to $T_x\mathcal{M}$. Let $\text{Inj}(\mathcal{M}) > 0$ be the injectivity radius of \mathcal{M} . For more information on these notions, we refer the reader to [Gallot et al. \(2004\)](#).

We will denote by 0_x and μ_x the null vector and Lebesgue measure, respectively, in $T_x\mathcal{M}$. For simplicity, we will write $\int_{T_x\mathcal{M}} J(v)dv = \int_{T_x\mathcal{M}} J(v)d\mu_x(v)$ for any integrable function J defined on $T_x\mathcal{M}$. Let $\langle \cdot, \cdot \rangle$ be the inner product related to g in $T_x\mathcal{M}$. Thus, for any $u, v \in T_x\mathcal{M}$, we have $\langle u, v \rangle = g(u, v)$. The associated norm is denoted by $\|\cdot\|$. $B(x, h) = \{y, d(x, y) \leq h\}$ and $B(0_x, h) = \{u, \|u\| \leq h\}$ are the balls of radius h centered at x and 0_x , respectively. We define $h_* = \min\{\text{Inj}(\mathcal{M}), \frac{\pi}{2\sqrt{\kappa}}\}$, where κ is the supremum of sectional curvatures of \mathcal{M} . We set $\frac{\pi}{2\sqrt{\kappa}} = +\infty$ when $\kappa \leq 0$. Throughout, we only consider regular balls in \mathcal{M} . We recall that a ball $B(x, h)$ is said to be regular (or convex) if $h < h_*$. Under these conditions, one also has $B(x, h) = \exp_x B(0_x, h)$. Below, we use $B(h)$ to denote $B(0_x, h)$.

From the previous properties, we can naturally deduce (the following property which will be helpful for the proofs) that for any continuous function $\psi : \mathcal{M} \rightarrow \mathbb{R}$ with support on $B(x, h)$ with $h < h_*$, we have

$$\nu_g(\psi) = \int_{\mathcal{M}} \psi(y) d\nu_g(y) = \int_{\exp_x(B(h))} \psi(y) d\nu_g(y) = \int_{B(h)} \psi(\exp_x(u)) |g_x(u)|^{1/2} du,$$

where $|g_x(u)|^{1/2}$ denotes the determinant of $g_x(u)$, with g_x being the local expression of g in the coordinate system (x^1, \dots, x^1) , given by $g_x(u) = (g_{ij}(u))$ where $g_{ij}(u) = g(\partial x_i, \partial x_j)$ (or $g_x = \sum_{i,j} g_{ij} dx^i dx^j$). We recall that $|g_x(u)|^{1/2} = \frac{d\mu_g}{d\mu_x}(\exp_x(u)) = \frac{d\mu_{\exp_x^*g}}{d\mu_x}(u)$ is the density of $\mu_{\exp_x^*g}$ with respect to μ_x on $T_x(\mathcal{M})$ (see, for example, [Gallot et al. \(2004\)](#) for more details). We repeatedly use the expansion of $|g_x(u)|^{1/2}$ at any, $u \in T_x(\mathcal{M})$ which can be found in, for example, [Karcher \(1977\)](#), as follows:

$$|g_x(u)|^{1/2} = 1 - \frac{\text{Ric}_x(u, u)}{6} + O(\|u\|^3),$$

where Ric_x is the Ricci tensor at x . We assume that X_i 's are distributed as a random variable X with an unknown stationary density f with respect to ν_g . Then, the transported density of $\exp_x^{-1}(X)$ with respect to the Lebesgue measure in $T_x\mathcal{M}$ is given by $f_{T_x}(v) = f(\exp_x(v)) |g_x(v)|^{1/2}$, $\forall v \in T_x\mathcal{M}$.

Note that $f_{T_x}(0_x) = f(x) |g_x(0_x)|^{1/2} = f(x)$ since $|g_x(0_x)|^{1/2} = 1$.

In what follows, grad and Hess denote, respectively, the gradient and Hessian operators

2.2 Kernel density estimator for stationary random fields

Let $\mathcal{I}_{\mathbf{n}}$ be a rectangular region defined as

$$\mathcal{I}_{\mathbf{n}} = \left\{ \mathbf{i} = (i_1, i_2, \dots, i_N) \in (\mathbb{N}^*)^N, 1 \leq i_k \leq n_k, k = 1, \dots, N \right\},$$

with $\mathbf{n} = (n_1, \dots, n_N)$. As in the previous papers on spatial setting (for example, [Carbon et al. \(1997\)](#)), we will write $\mathbf{n} \rightarrow \infty$ if $\min \{n_k\} \rightarrow +\infty$ and $\left| \frac{n_l}{n_k} \right| < C$ for some $0 < C < \infty$, $1 \leq l, k \leq N$, and set $\widehat{\mathbf{n}} = n_1 \times \dots \times n_N$. Throughout the paper, we will denote by $\|\mathbf{i}\| = \sqrt{i_1^2 + \dots + i_N^2}$, and $0 < C < \infty$ will be some constants.

Now, the spatial kernel density estimator, based on the observations $(X_{\mathbf{i}}, \mathbf{i} \in \mathcal{I}_{\mathbf{n}})$, is defined by

$$\widehat{f}_{\mathbf{n}}(x) = \frac{1}{\widehat{\mathbf{n}}} \sum_{\mathbf{i} \in \mathcal{I}_{\mathbf{n}}} \frac{1}{h_{\mathbf{n}}^d} \theta_x(X_{\mathbf{i}}) K\left(\frac{d_g(x, X_{\mathbf{i}})}{h_{\mathbf{n}}}\right), \quad \forall x \in \mathcal{M},$$

where $K : \mathbb{R}^+ \rightarrow \mathbb{R}^+$ is a bounded integrable function, and the bandwidth $h_{\mathbf{n}}$, with $h_{\mathbf{n}} < h_*$, tends to zero as $\mathbf{n} \rightarrow \infty$.

2.3 Assumptions

To study asymptotic properties of $\widehat{f}_{\mathbf{n}}$, we assume that $(X_{\mathbf{i}}, \mathbf{i} \in \mathbb{N}^N)$ satisfies the following mixing condition: there exists a function $\varphi(t) \rightarrow 0$ as $t \rightarrow \infty$ such that, for finite sets of sites $E, E' \subset \mathbb{N}^N$, we have

$$\begin{aligned} \alpha(\mathcal{B}(E), \mathcal{B}(E')) &= \sup_{A \in \mathcal{B}(E), B \in \mathcal{B}(E')} \left\{ |\mathbb{P}(AB) - \mathbb{P}(A)\mathbb{P}(B)|, A \in \mathcal{B}(E), B \in \mathcal{B}(E') \right\} \\ &\leq s\left(\text{Card}(E), \text{Card}(E')\right) \varphi\left(\text{dist}(E, E')\right), \end{aligned}$$

where $\text{Card}(E)$ (resp. $\text{Card}(E')$) denotes the cardinality of E (resp. E'), $\mathcal{B}(E)$ (resp. $\mathcal{B}(E')$) is the σ -fields generated by $(X_{\mathbf{i}}, \mathbf{i} \in E)$ (resp. $(X_{\mathbf{i}}, \mathbf{i} \in E')$) for all \mathbf{i} belonging to E (resp. E'), and $\text{dist}(E, E')$ denotes the Euclidean distance between E and E' . The function s is a symmetric positive function that is non-decreasing in each variable which satisfies

$$s(n, m) \leq C \min(n, m), \quad \forall n, m \in \mathbb{N}. \quad (2.1)$$

If $s(\cdot, \cdot) \equiv 1$, then $(X_{\mathbf{i}})$ is called strongly mixing. We set $\Psi_{\mathbf{n}} = \sqrt{\frac{\log \widehat{\mathbf{n}}}{\widehat{\mathbf{n}} h_{\mathbf{n}}^d}}$.

Assumption H1 (Kernel Properties):

1. $\int K(\|x\|) dx = 1$ and K satisfies a Lipschitz condition;

2. $\text{supp } K = [0; 1]$;
3. $\int xK(\|x\|)dx = 0$
4. $\int \|x\|^2 K(\|x\|)dx < \infty$.

Assumption H2 (Mixing Conditions):

1. $\varphi(i) \leq Ci^{-\theta}$ for some $\theta > 0$;
2. $\sum_{i=1}^{+\infty} i^{N-1}(\varphi(i))^a < \infty$ for some $0 < a < \frac{1}{2}$.

Assumption H3 (Bandwidth):

$h_n \rightarrow 0$, $\widehat{\mathbf{n}}h_n^d \rightarrow \infty$ and $h_n < \frac{h_*}{2}$ as $\mathbf{n} \rightarrow \infty$.

Assumption H4 (Compactness of the Manifold):

\mathcal{M} is compact, i.e., $\mathcal{M} = \bigcup_{k=1}^{\mu_n} B(x_k, a_n)$, with $a_n = h_*^{-\frac{d}{2}} h_n^{(d+1)} \Psi_n$ and $\mu_n \leq C(h_n^{(d+1)} \Psi_n)^{-d}$.

Assumption H5 (Regularity of the Target Density):

1. The stationary density f is bounded and twice continuously differentiable at any $x \in \mathcal{M}$.
2. $\sup \|\text{Hess}f(x)\|_{HS} = \widetilde{C}$, where $\|\cdot\|_{HS}$ is the Hilbert Schmidt norm and $0 < \widetilde{C} < \infty$ is a positive constant.

Assumption H6: $\forall \mathbf{i}, \mathbf{j}$, the joint stationary density $f_{\mathbf{i}, \mathbf{j}}$ of $(X_{\mathbf{i}}, X_{\mathbf{j}})$ exists and satisfies

$$\sup_{\mathbf{i}, \mathbf{j}} \sup_{u, v \in \mathcal{M} \times \mathcal{M}} |f_{\mathbf{i}, \mathbf{j}}(u, v) - f(u)f(v)| < M.$$

Some comments on these assumptions

We give here some remarks to highlight the assumptions above. These assumptions are essential for deriving the asymptotic properties of the kernel density estimator and ensuring the consistency of the results. Some assumptions are generalizations imposed for spatial data in Euclidean space and kernel density estimation for i.i.d data in a Riemannian manifold. More precisely,

1. **Assumption H1** and **Assumption H5** are classical regularity conditions on K and f needed to derive asymptotic bias of \widehat{f}_n . That is, for all $h > 0$ and $\|v\| \leq 1$, we have

$$f(\exp_x(hv)) = f(x) + h \langle \text{grad}f(x), v \rangle + \frac{h^2}{2} \text{Hess}f(x)(\tau v, \tau v). \quad (2.2)$$

with $\tau \in]0, 1[$, and $\text{Hess}f(x)(\tau v, \tau v) := \langle \tau v, \text{Hess}f(x)\tau v \rangle$.

2. **Assumption H2** and **Assumption H6** are necessary to control the dependence structure in spatial data.
3. **Assumption H3** is the Riemannian manifold counterpart of the classical assumptions on the bandwidth in spatial settings when the process takes values in Euclidean space. In fact, here as mentioned in Section 2.1, we need to impose $h_{\mathbf{n}} < h_*$ to ensure the convexity of any ball $B(x, h_{\mathbf{n}})$, $x \in \mathcal{M}$. Nevertheless, having $h_{\mathbf{n}} < \frac{h_*}{2}$ ensures that x is locally a central point for $f_{x,K}$ (see, for example, [Karcher \(1977\)](#) or [Pelletier \(2005\)](#) for more details).
4. **Assumption H4** will be used to study the uniform convergence of a kernel density estimator. It is the Riemannian manifold counterpart of the condition imposed in the Euclidean case, as stated, for example, in [Carbon et al. \(1997\)](#). But, we need to ensure the convexity of each ball of the covering set of \mathcal{M} . This is satisfied, since, for all \mathbf{n} ,

$$a_{\mathbf{n}} = h_*^{-\frac{d}{2}} h_{\mathbf{n}}^{\left(\frac{d}{2}+1\right)} \left(\frac{\log \widehat{\mathbf{n}}}{\widehat{\mathbf{n}}} \right)^{\frac{1}{2}} < h_*.$$

3. Results

In this section, we present the asymptotic properties of the estimator $\widehat{f}_{\mathbf{n}}(x)$ under the assumptions stated earlier. Specifically, we study its consistency in terms of MSE (Mean Squared Error), MISE (Mean Integrated Squared Error), and in probability meaning.

Proposition 3.1. *Under Assumption H1 and Assumption H5(.1), we have*

$$b(x) := \mathbb{E} \widehat{f}_{\mathbf{n}}(x) - f(x) = \frac{h_{\mathbf{n}}^2}{2} \int_{B(1)} K(\|u\|) \text{Hess} f(x)(\tau u, \tau u) du. \quad (3.1)$$

Additionally, under Assumption H5(.2), if \mathcal{M} is compact (Assumption H4), then,

$$\sup_{x \in \mathcal{M}} |b(x)| \leq \widetilde{C} h_{\mathbf{n}}^2 \int_{B(1)} \|u\|^2 K(\|u\|) du, \quad (3.2)$$

and

$$\int_{\mathcal{M}} b^2(x) dv_g(x) \leq \widetilde{C}^2 h_{\mathbf{n}}^4 \left(\int_{B(1)} \|u\|^2 K(\|u\|) du \right)^2 \text{Vol}(\mathcal{M}), \quad (3.3)$$

where $\text{Vol}(\mathcal{M})$ denotes the volume of \mathcal{M} , defined as $\text{Vol}(\mathcal{M}) = \int_{\mathcal{M}} dv_g(x)$.

Proposition 3.2. *Under Assumptions **H1** to **H5***

$$\text{Var}(\widehat{f}_n(x)) = \frac{1}{\widehat{\mathbf{n}}h_n^d} f(x) \int_{B(1)} K^2(\|u\|) du + o\left(\frac{1}{\widehat{\mathbf{n}}h_n^d}\right). \quad (3.4)$$

Consequently,

$$\lim_{n \rightarrow \infty} \widehat{\mathbf{n}}h_n^d \text{Var}(\widehat{f}_n(x)) = f(x) \int_{B(1)} K^2(\|u\|) du \quad (3.5)$$

Additionally, if \mathcal{M} is compact, then,

$$\sup_{x \in \mathcal{M}} \left| \text{Var}(\widehat{f}_n(x)) \right| \leq 2 \times \max \left\{ \|f\|_\infty \|K\|_\infty, \sup_{x \in \mathcal{M}} C_x \right\} \frac{1}{\widehat{\mathbf{n}}h_n^d} + o\left(\frac{1}{\widehat{\mathbf{n}}h_n^d}\right), \quad (3.6)$$

and

$$\int_{\mathcal{M}} \text{Var}(\widehat{f}_n(x)) dv_g(x) \leq 2 \times \max \left\{ \|K\|_\infty, \text{Vol}(\mathcal{M}) \sup_{x \in \mathcal{M}} C_x \right\} \frac{1}{\widehat{\mathbf{n}}h_n^d} + o\left(\frac{1}{\widehat{\mathbf{n}}h_n^d}\right). \quad (3.7)$$

Using the usual decomposition of MSE in terms of the squared bias and variance and the fact that, $MISE = \int_{\mathcal{M}} MSE(x) dv_g(x)$, we derive the following results.

Theorem 3.1. *Under Assumptions **H1** to **H5**, for each $x \in \mathcal{M}$, the MSE satisfies*

$$MSE(x) := \mathbb{E} \left((\widehat{f}_n(x) - f(x))^2 \right) \leq 3 \times \max \left\{ \widetilde{C}^2, \|f\|_\infty \|K\|_\infty, C_x \right\} \left(h_n^4 + \frac{1}{\widehat{\mathbf{n}}h_n^d} \right) + o\left(\frac{1}{\widehat{\mathbf{n}}h_n^d}\right),$$

and the uniform bound

$$\sup_{x \in \mathcal{M}} MSE(x) \leq 3 \times \max \left\{ \widetilde{C}^2, \|f\|_\infty \|K\|_\infty, \sup_{x \in \mathcal{M}} C_x \right\} \left(h_n^4 + \frac{1}{\widehat{\mathbf{n}}h_n^d} \right) + o\left(\frac{1}{\widehat{\mathbf{n}}h_n^d}\right),$$

Additionally, the $MISE$ satisfies

$$MISE := \int_{\mathcal{M}} MSE(x) dv_g(x) \leq 2 \times \max \left\{ \|K\|_\infty, \text{Vol}(\mathcal{M}) \sup_{x \in \mathcal{M}} C_x \right\} \left(h_n^4 + \frac{1}{\widehat{\mathbf{n}}h_n^d} \right) + o\left(\frac{1}{\widehat{\mathbf{n}}h_n^d}\right).$$

Corollary 3.1. *(Optimal rate) The bandwidth which minimizes the MSE , at each point $x \in \mathcal{M}$, under Assumptions **H1** to **H6**, is given by*

$$h_{n,opt} = C \widehat{\mathbf{n}}^{-\frac{1}{d+4}},$$

and the corresponding MSE is

$$MSE(x) = C_{x,f} \widehat{\mathbf{n}}^{-\frac{4}{d+4}} + o\left(\widehat{\mathbf{n}}^{-\frac{4}{d+4}}\right),$$

where $C_{x,f} = C^4 C_1 + C^{-d} C_2$ with $C = \left(\frac{-dC_2}{4C_1}\right)^{\frac{1}{4+d}}$, $C_1 = \left(\int_{B(1)} \text{Hess}f(x)(\tau v, \tau v) K(\|v\|) dv\right)^2$, with $\tau \in]0, 1[$ and $C_2 = f(x) \int_{B(1)} K^2(\|v\|) dv$.

Next, we give pointwise and uniform rates of convergence in probability under additional conditions.

Theorem 3.2. *Under the assumptions **H1**, **H3** and **H5**, if $\varphi(i) \leq Ci^{-\theta}$ with $\theta > 2N$ and if $\widehat{n}(\log \widehat{n})^{-1} h_n^{\frac{d\theta}{2N-\theta}} \rightarrow \infty$, then for a given $x \in \mathcal{M}$*

$$\left| \widehat{f}_n(x) - f(x) \right| = O\left(h_n^2\right) + o_p\left(\Psi_n\right). \quad (3.8)$$

Corollary 3.2. *Under Assumption **H4**, if $\theta > (d+2)N$ and $\widehat{n}(\log \widehat{n})^{-1} h_n^{\frac{d(\theta+dN+2N)}{\theta-(d+2)N}} \rightarrow \infty$, then*

$$\sup_{x \in \mathcal{M}} \left| \widehat{f}_n(x) - f(x) \right| = O\left(h_n^2\right) + o_p\left(\mu_n \Psi_n\right). \quad (3.9)$$

4. Proofs

In the following, we will denote by $C_g = \sup_{x \in \mathcal{M}} C_g(x)$ with $C_g(x) = \sup_{y \in B(x, h_n)} \theta_x^{-1}(y)$ (we refer the readers [Pelletier \(2005\)](#)(p.303)).

Proof of Proposition 3.1: For all $x \in \mathcal{M}$,

$$\begin{aligned} \mathbb{E} \widehat{f}_n(x) - f(x) &= \frac{1}{\widehat{n}} \sum_{i \in \mathcal{I}_n} \frac{1}{h_n^d} \mathbb{E} \left(\frac{1}{\theta_x(X_i)} K \left(\frac{d_g(x, X_i)}{h_n} \right) \right) - f(x) \\ &= \frac{1}{h_n^d} \int_{\mathcal{M}} \frac{1}{\theta_x(y)} K \left(\frac{d_g(x, y)}{h_n} \right) f(y) d\nu_g(y) - f(x) \\ &= \frac{1}{h_n^d} \int_{\mathcal{M}} \frac{1}{\theta_x(y)} K \left(\frac{d_g(x, y)}{h_n} \right) (f(y) - f(x)) d\nu_g(y). \end{aligned}$$

Taking the integral over $B_{\mathcal{M}}(x, h_n)$, we obtain

$$\begin{aligned} \mathbb{E} \widehat{f}_n(x) - f(x) &= \frac{1}{h_n^d} \int_{B_{\mathcal{M}}(x, h_n)} \frac{1}{\theta_x(y)} K \left(\frac{d_g(x, y)}{h_n} \right) (f(y) - f(x)) d\nu_g(y) \\ &= \frac{1}{h_n^d} \int_{B_{\mathcal{M}}(h_n)} \frac{1}{\theta_x(\exp_x(u))} K \left(\frac{\|u\|}{h_n} \right) (f(\exp_x(u)) - f(x)) |g_x(u)|^{1/2} du \\ &= \int_{B(1)} K(\|u\|) (f(\exp_x(h_n u)) - f(x)) du. \end{aligned}$$

By applying (2.2), under the assumptions **H1** and **H5**(.1), we obtain (3.1).

On the other hand, by Assumption **H5**(.2) and due to the compactness of, \mathcal{M} , we get (3.2).

Furthermore, for the same reasons provided by [Pelletier \(2005\)](#)(p.302), we get (3.3).

Proof of Proposition 3.2 This result is an extension of Lemma 2.2 in [Tran \(1990\)](#) to the case of Riemannian manifold. For all $x \in \mathcal{M}$, we have

$$\text{Var}(\widehat{f}_{\mathbf{n}}(x)) = \text{Var}(\widehat{f}_{\mathbf{n}}(x) - \mathbb{E}\widehat{f}_{\mathbf{n}}(x)) = \text{Var}\left(\frac{1}{\widehat{\mathbf{n}}}\sum_{\mathbf{i} \in \mathcal{I}_{\mathbf{n}}} Z_{\mathbf{i}}(x)\right),$$

where for each $\mathbf{i} \in \mathcal{I}_{\mathbf{n}}$, $Z_{\mathbf{i}}(x) = \frac{1}{h_{\mathbf{n}}^d} \frac{1}{\theta_x(X_{\mathbf{i}})} K\left(\frac{d_g(x, X_{\mathbf{i}})}{h_{\mathbf{n}}}\right) - \frac{1}{h_{\mathbf{n}}^d} \mathbb{E}\left(\frac{1}{\theta_x(X_{\mathbf{i}})} K\left(\frac{d_g(x, X_{\mathbf{i}})}{h_{\mathbf{n}}}\right)\right)$. Then,

$$\begin{aligned} \text{Var}(\widehat{f}_{\mathbf{n}}(x)) &= \frac{1}{\widehat{\mathbf{n}}^2} \sum_{\mathbf{i} \in \mathcal{I}_{\mathbf{n}}} \text{Var}(Z_{\mathbf{i}}(x)) + \frac{1}{\widehat{\mathbf{n}}^2} \sum_{\substack{\mathbf{i} \neq \mathbf{j} \\ \mathbf{i}, \mathbf{j} \in \mathcal{I}_{\mathbf{n}}}} \text{Cov}(Z_{\mathbf{i}}(x), Z_{\mathbf{j}}(x)) \\ &= \frac{1}{\widehat{\mathbf{n}}^2} \sum_{\mathbf{i} \in \mathcal{I}_{\mathbf{n}}} \mathbb{E}(Z_{\mathbf{i}}^2(x)) + \frac{1}{\widehat{\mathbf{n}}^2} \sum_{\substack{\mathbf{i} \neq \mathbf{j} \\ \mathbf{i}, \mathbf{j} \in \mathcal{I}_{\mathbf{n}}}} \mathbb{E}(Z_{\mathbf{i}}(x)Z_{\mathbf{j}}(x)) \\ &:= A(x) + B(x), \end{aligned}$$

where

$$\begin{aligned} A(x) &= \frac{1}{\widehat{\mathbf{n}}h_{\mathbf{n}}^{2d}} \int_{\mathcal{M}} \frac{1}{\theta_x^2(y)} K^2\left(\frac{d_g(x, y)}{h_{\mathbf{n}}}\right) f(y) d\nu_g(y) - \frac{1}{\widehat{\mathbf{n}}h_{\mathbf{n}}^{2d}} \left(\int_{\mathcal{M}} \frac{1}{\theta_x(y)} K\left(\frac{d_g(x, y)}{h_{\mathbf{n}}}\right) f(y) d\nu_g(y) \right)^2 \\ &= I_{1, \mathbf{n}}(x) - I_{2, \mathbf{n}}(x). \end{aligned}$$

This expression has been previously addressed in the literature, notably in [Pelletier \(2005\)](#), our objective here is to give an explicit expression thereof. To start by establishing $I_{1, \mathbf{n}}(x)$, recall that for any $v \in T_x \mathcal{M}$, $\theta_x(\exp_x(v)) = |g_x(v)|^{\frac{1}{2}}$, then

$$\begin{aligned} I_{1, \mathbf{n}}(x) &= \frac{1}{\widehat{\mathbf{n}}h_{\mathbf{n}}^{2d}} \int_{B(x, h_{\mathbf{n}})} \frac{1}{\theta_x^2(y)} K^2\left(\frac{d_g(x, y)}{h_{\mathbf{n}}}\right) f(y) d\nu_g(y) \\ &= \frac{1}{\widehat{\mathbf{n}}h_{\mathbf{n}}^d} \int_{B(1)} \frac{1}{|g_x(h_{\mathbf{n}}u)|^{\frac{1}{2}}} K^2(\|u\|) f(\exp_x(h_{\mathbf{n}}u)) du. \\ &= \frac{1}{\widehat{\mathbf{n}}h_{\mathbf{n}}^d} \int_{B(1)} \frac{1}{|g_x(h_{\mathbf{n}}u)|^{\frac{1}{2}}} K^2(\|u\|) (f(\exp_x(h_{\mathbf{n}}u)) - f(x)) du \\ &\quad + \frac{f(x)}{\widehat{\mathbf{n}}h_{\mathbf{n}}^d} \int_{B(1)} \frac{1}{|g_x(h_{\mathbf{n}}u)|^{\frac{1}{2}}} K^2(\|u\|) du. \\ &= \frac{f(x)}{\widehat{\mathbf{n}}h_{\mathbf{n}}^d} \int_{B(1)} K^2(\|u\|) du + \frac{f(x)}{\widehat{\mathbf{n}}h_{\mathbf{n}}^d} \int_{B(1)} \left(\frac{1}{|g_x(h_{\mathbf{n}}u)|^{\frac{1}{2}}} - 1 \right) K^2(\|u\|) du \quad (4.1) \end{aligned}$$

$$+ \frac{1}{\widehat{\mathbf{n}}h_{\mathbf{n}}^d} \int_{B(1)} \frac{1}{|g_x(h_{\mathbf{n}}u)|^{\frac{1}{2}}} K^2(\|u\|) (f(\exp_x(h_{\mathbf{n}}u)) - f(x)) du \quad (4.2)$$

Since $\left| |g_x(h_n u)|^{\frac{1}{2}} - 1 \right| \leq C_{1,x} h_n^2$ and by (2.2), for $\|u\| \leq 1$, we have

$$\left| I_{1,n}(x) - \frac{1}{\widehat{\mathbf{n}} h_n^d} f(x) \int_{B(1)} K^2(\|u\|) du \right| \leq C_x \frac{h_n}{\widehat{\mathbf{n}} h_n^d}, \quad (4.3)$$

where $C_x = (C_{1,x} f(x) + \|\text{grad} f(x)\|) \|K\|_\infty$. Additionally, under **H1** and **H4**, we have

$$\sup_{x \in \mathcal{M}} \widehat{\mathbf{n}} h_n^d I_{1,n}(x) \leq \|f\|_\infty \|K\|_\infty + \sup_{x \in \mathcal{M}} C_x h_n, \quad (4.4)$$

Similarly, it is easy to see that

$$\begin{aligned} I_{2,n}(x) &= \frac{1}{\widehat{\mathbf{n}} h_n^{2d}} \left(\int_{\mathcal{M}} \frac{1}{\theta_x(y)} K \left(\frac{d_g(x,y)}{h_n} \right) f(y) d\nu_g(y) \right)^2 \\ &= \frac{1}{\widehat{\mathbf{n}} h_n^{2d}} \left(\int_{B(h_n)} \frac{1}{\theta_x(\exp_x(u))} K \left(\frac{\|u\|}{h_n} \right) f(\exp_x(u)) |g_x(u)|^{1/2} du \right)^2 \\ &= \frac{1}{\widehat{\mathbf{n}}} \left(\int_{B(1)} K(\|u\|) f(\exp_x(h_n u)) du \right)^2 \\ &= \frac{1}{\widehat{\mathbf{n}}} \left(\int_{B(1)} K(\|u\|) (f(\exp_x(h_n u)) - f(x)) du + f(x) \int_{B(1)} K(\|u\|) du \right)^2. \end{aligned}$$

Hence, under Assumptions **H1**(.1) and **H5**, and using (2.2), we obtain

$$I_{2,n}(x) \leq 2 \frac{f^2(x) + (C_2 h_n^2)^2}{\widehat{\mathbf{n}}} = \frac{f^2(x)}{\widehat{\mathbf{n}}} + o(1) = O(\widehat{\mathbf{n}}^{-1}).$$

Additionally, under **H4**, we have

$$\sup_{x \in \mathcal{M}} \widehat{\mathbf{n}} I_{2,n}(x) \leq \|f\|_\infty^2. \quad (4.5)$$

Whereas, the expression

$$\begin{aligned} B &= \frac{1}{\widehat{\mathbf{n}}^2} \sum_{(i,j) \in E_1} \mathbb{E}(Z_i(x) Z_j(x)) + \frac{1}{\widehat{\mathbf{n}}^2} \sum_{(i,j) \in E_2} \mathbb{E}(Z_i(x) Z_j(x)) \\ &= J_{1,n}(x) + J_{2,n}(x) \end{aligned}$$

has been studied by [Tran \(1990\)](#) within the context of Euclidean spaces. However, our contribution specifically pertains to the study of $J_{2,n}(x)$.

Let $\beta_{\mathbf{n}} = h_{\mathbf{n}}^{-\frac{d(1-\gamma)}{\theta}}$, with $\theta = -N - \varepsilon + (1 - \gamma)Na^{-1}$, and γ and ε are small positive constants. We require that there exists a real number a , such that $0 < a < \frac{1}{2}$, satisfying the condition $a^{-1} - (N + \varepsilon)(N(1 - \gamma))^{-1} > 1$ and $\theta > N(1 - \gamma)$. Now, by setting

$$E_1 := \{(\mathbf{i}, \mathbf{j}) \in \mathcal{I}_{\mathbf{n}} \times \mathcal{I}_{\mathbf{n}} \mid 0 < d(\mathbf{i}, \mathbf{j}) \leq \beta_{\mathbf{n}}\} \quad \text{and} \quad E_2 := \{(\mathbf{i}, \mathbf{j}) \in \mathcal{I}_{\mathbf{n}} \times \mathcal{I}_{\mathbf{n}} \mid d(\mathbf{i}, \mathbf{j}) > \beta_{\mathbf{n}}\},$$

we have

$$\begin{aligned} \left| \mathbb{E} \left(Z_{\mathbf{i}}(x) Z_{\mathbf{j}}(x) \right) \right| &\leq \int_{B(x, h_{\mathbf{n}})} \int_{B(x, h_{\mathbf{n}})} \frac{1}{h_{\mathbf{n}}^{2d}} \frac{1}{\theta_x(y) \theta_x(z)} K \left(\frac{d_g(x, y)}{h_{\mathbf{n}}} \right) K \left(\frac{d_g(x, z)}{h_{\mathbf{n}}} \right) \\ &\quad \times |f_{\mathbf{i}, \mathbf{j}}(y, z) - f(y) f(z)| dv_g(y) dv_g(z) \\ &\leq M \int_{B(1)} \int_{B(1)} K(\|u\|) K(\|v\|) dudv \quad (\text{by Assumption H6}) \\ &\leq M \quad (\text{by Assumption H1(1)}). \end{aligned}$$

Consequently, $J_{1, \mathbf{n}}(x) \leq M \widehat{\mathbf{n}}^{-1} \beta_{\mathbf{n}}^N$. Since $\theta > N(1 - \gamma)$, we have

$$J_{1, \mathbf{n}}(x) \leq M \widehat{\mathbf{n}}^{-1} h_{\mathbf{n}}^{-\frac{dN(1-\gamma)}{\theta}} = o(\widehat{\mathbf{n}}^{-1} h_{\mathbf{n}}^{-d}).$$

To bound the term $J_{2, \mathbf{n}}(x)$, we set $\delta = \frac{1-\gamma}{\gamma}$, which implies $\gamma = \frac{2}{2+\delta}$ and $\frac{\delta}{2+\delta} = 1 - \gamma$. By applying Lemma 2.1 in [Tran \(1990\)](#) with $r = s = 2 + \delta$, $h = \frac{2+\delta}{\delta}$, we obtain

$$\begin{aligned} \left| \mathbb{E} \left(Z_{\mathbf{i}}(x) Z_{\mathbf{j}}(x) \right) \right| &\leq C \mathbb{E} \left[\left(\frac{1}{h_{\mathbf{n}}^d} \frac{1}{\theta_x(X_{\mathbf{i}})} K \left(\frac{d_g(x, X_{\mathbf{i}})}{h_{\mathbf{n}}} \right) \right)^{2+\delta} \right]^{\frac{\gamma}{2}} \mathbb{E} \left[\left(\frac{1}{h_{\mathbf{n}}^d} \frac{1}{\theta_x(X_{\mathbf{j}})} K \left(\frac{d_g(x, X_{\mathbf{j}})}{h_{\mathbf{n}}} \right) \right)^{2+\delta} \right]^{\frac{\gamma}{2}} \\ &\quad \times (s(1, 1) \varphi(d(\{\mathbf{i}\}, \{\mathbf{j}\})))^{1-\gamma} \\ &\leq C \left(\int_{\mathcal{M}} \left(\frac{1}{h_{\mathbf{n}}^d} \frac{1}{\theta_x(y)} K \left(\frac{d_g(x, y)}{h_{\mathbf{n}}} \right) \right)^{2+\delta} f(y) dv_g(y) \right)^{\gamma} \times [\varphi(\|\mathbf{i} - \mathbf{j}\|)]^{1-\gamma}. \end{aligned}$$

Consequently,

$$\begin{aligned} J_{2, \mathbf{n}}(x) &\leq \frac{C}{\widehat{\mathbf{n}}^2} \sum_{\mathbf{i}, \mathbf{j} \in E_2} \left(\int_{\mathcal{M}} \left(\frac{1}{h_{\mathbf{n}}^d} \frac{1}{\theta_x(y)} K \left(\frac{d_g(x, y)}{h_{\mathbf{n}}} \right) \right)^{2+\delta} f(y) dv_g(y) \right)^{\gamma} \times [\varphi(\|\mathbf{i} - \mathbf{j}\|)]^{1-\gamma} \\ &\leq \frac{C}{\widehat{\mathbf{n}}^2} h_{\mathbf{n}}^{-\gamma d(1+\delta)} \left(\int_{\mathcal{M}} \frac{1}{h_{\mathbf{n}}^d} \left(\frac{1}{\theta_x(y)} K \left(\frac{d_g(x, y)}{h_{\mathbf{n}}} \right) \right)^{2+\delta} f(y) dv_g(y) \right)^{\gamma} \times \sum_{\mathbf{i}, \mathbf{j} \in E_2} (\varphi(\|\mathbf{i} - \mathbf{j}\|))^{1-\gamma}. \end{aligned}$$

Clearly, $\sum_{\mathbf{i}, \mathbf{j} \in E_2} (\varphi(\|\mathbf{i} - \mathbf{j}\|))^{1-\gamma} \leq \widehat{\mathbf{n}} \sum_{\|\mathbf{j}\| > \beta_{\mathbf{n}}} (\varphi(\|\mathbf{j}\|))^{1-\gamma}$ and then

$$\begin{aligned} \widehat{\mathbf{n}} h_{\mathbf{n}}^d J_{2,\mathbf{n}}(x) &\leq h_{\mathbf{n}}^{-d(1-\gamma)} \left(\int_{B_{\mathcal{M}(x, h_{\mathbf{n}})}} \frac{1}{h_{\mathbf{n}}^d} \left(\frac{1}{\theta_x(y)} K \left(\frac{d_g(x, y)}{h_{\mathbf{n}}} \right) \right)^{2+\delta} f(y) dv_g(y) \right)^\gamma \times \sum_{\|\mathbf{j}\| > \beta_{\mathbf{n}}} (\varphi(\|\mathbf{j}\|))^{1-\gamma} \\ &\leq h_{\mathbf{n}}^{-d(1-\gamma)} \left(C_g^{1+\delta} \int_{B(1)} f(\exp_x(h_{\mathbf{n}}v)) K(\|v\|)^{2+\delta} dv \right)^\gamma \times \sum (\varphi(\|\mathbf{j}\|))^{1-\gamma}. \end{aligned} \quad (4.6)$$

Now, **H2** implies $i^{N-1}(\varphi(i))^a = o\left(\frac{1}{i}\right)$, which leads to $\varphi(i) = o\left(i^{-\frac{N}{a}}\right)$ as $i \rightarrow \infty$ (we refer the readers to [Tran \(1990\)](#)). Since φ is a nonincreasing function, we have $\|\mathbf{i}\|^\theta (\varphi(\|\mathbf{i}\|))^{1-\gamma} = \|\mathbf{i}\|^\theta o\left(\|\mathbf{i}\|^{-\frac{N(1-\gamma)}{a}}\right) = o\left(\|\mathbf{i}\|^{-N-\varepsilon}\right)$, for $\theta = -N - \varepsilon + (1-\gamma)Na^{-1}$. Thus,

$$\sum_{\substack{k=1, \dots, N \\ i_k=1}}^{+\infty} \|\mathbf{i}\|^\theta (\varphi(\|\mathbf{i}\|))^{1-\gamma} < +\infty. \quad (4.7)$$

Combining (4.6) and (4.7) and noting that $h_{\mathbf{n}}^{-d(1-\gamma)} \beta_{\mathbf{n}}^{-\theta} = 1$, we have

$$\begin{aligned} \limsup \widehat{\mathbf{n}} h_{\mathbf{n}}^d J_{2,\mathbf{n}}(x) &\leq C \limsup h_{\mathbf{n}}^{-d(1-\gamma)} \sum_{\|\mathbf{i}\| > \beta_{\mathbf{n}}} (\varphi(\|\mathbf{i}\|))^{1-\gamma} \\ &\leq C \limsup h_{\mathbf{n}}^{-d(1-\gamma)} \beta_{\mathbf{n}}^{-\theta} \sum_{\|\mathbf{i}\| > \beta_{\mathbf{n}}} \|\mathbf{i}\|^\theta (\varphi(\|\mathbf{i}\|))^{1-\gamma} \\ &\leq C \limsup \sum_{\|\mathbf{i}\| > \beta_{\mathbf{n}}} \|\mathbf{i}\|^\theta (\varphi(\|\mathbf{i}\|))^{1-\gamma}, \end{aligned}$$

where $C = \left(C_g^{1+\delta} \|K\|_\infty^{1+\delta} \|f\|_\infty \right)^\gamma$. Then, $\limsup \widehat{\mathbf{n}} h_{\mathbf{n}}^d J_{2,\mathbf{n}}$ tends to zero, as $\beta_{\mathbf{n}} \rightarrow \infty$ which completes the proof of (3.4), since

$$J_{1,\mathbf{n}}(x) + J_{2,\mathbf{n}}(x) = o\left(\widehat{\mathbf{n}}^{-1} h_{\mathbf{n}}^{-d}\right). \quad (4.8)$$

Also, we deduce (3.6) and (3.7) by using (4.4), (4.5) and (4.8).

Proof of Theorem 3.2: A key tool for this proof is Lemma 3.2 from [Dabo-Niang and Yao \(2007\)](#). By (3.1), it suffices to show that

$$\left| \widehat{f}_{\mathbf{n}}(x) - \mathbb{E} \left(\widehat{f}_{\mathbf{n}}(x) \right) \right| = o_p(\Psi_{\mathbf{n}}), \quad (4.9)$$

where $\Psi_{\mathbf{n}} = \sqrt{\frac{\log \widehat{\mathbf{n}}}{\widehat{\mathbf{n}} h_{\mathbf{n}}^d}}$.

Let $S_{\mathbf{n}} = \sum_{\mathbf{i} \in \mathcal{I}_{\mathbf{n}}} Z_{\mathbf{i}}(x)$, with $\mathbb{E}(Z_{\mathbf{i}}(x)) = 0$ and $\sup_{\mathbf{i}} |Z_{\mathbf{i}}(x)| \leq bh_{\mathbf{n}}^{-d}$, where $b = C_g(x)\|K\|_{\infty}$ by Assumption **H1**. Let $\zeta_{\mathbf{i}}(x) = Z_{\mathbf{i}}(x)$, $\mathbf{i} \in \mathbb{N}^N$ and let p be an integer. Thus, applying Lemma 3.2 in [Dabo-Niang and Yao \(2007\)](#), for each $\varepsilon_{\mathbf{n}} > 0$, we have

$$\mathbb{P}\left(\left|\widehat{f}_{\mathbf{n}}(x) - \mathbb{E}\left(\widehat{f}_{\mathbf{n}}(x)\right)\right| > \varepsilon_{\mathbf{n}}\right) \leq 2^{N+1} \exp\left(\frac{-\varepsilon_{\mathbf{n}}^2}{4v^2(\mathbf{q})}\widehat{\mathbf{q}}\right) + \frac{2^{N+2}b\alpha(p)}{\varepsilon_{\mathbf{n}}},$$

where $\widehat{\mathbf{n}} = 2^N p^N \widehat{\mathbf{q}}$ and $v^2(\mathbf{q}) = \frac{4}{p^{2N}}\sigma^2(\mathbf{q}) + b\varepsilon_{\mathbf{n}}$, with $\sigma^2(\mathbf{q}) = \text{Var}\left(\sum_{1 \leq i_k \leq p, k=1, \dots, N} Z_{\mathbf{i}}(x)\right)$. Now, from Proposition 3.2, we know that $\sigma^2(\mathbf{q}) \leq \frac{a}{p^N h_{\mathbf{n}}^d}$, where $a = a(K, \|f\|_{\infty}, d)$ does not depend on x . Thus, we obtain

$$v^2(\mathbf{q}) \leq \frac{b\varepsilon_{\mathbf{n}}}{h_{\mathbf{n}}^d} + \frac{4a}{p^N h_{\mathbf{n}}^d}.$$

If we choose $p = \lceil \varepsilon_{\mathbf{n}}^{-1} \rceil$, then we obtain

$$\mathbb{P}\left(\left|\widehat{f}_{\mathbf{n}}(x) - \mathbb{E}\left(\widehat{f}_{\mathbf{n}}(x)\right)\right| > \varepsilon_{\mathbf{n}}\right) \leq 2^{N+1} \exp\left(\frac{-\varepsilon_{\mathbf{n}} \widehat{\mathbf{q}} h_{\mathbf{n}}^d}{4A_0}\right) + \frac{2^{N+2}b\alpha(p)}{\varepsilon_{\mathbf{n}} h_{\mathbf{n}}^d} \quad (4.10)$$

where A_0 is a positive constant.

Now, let $\varepsilon_{\mathbf{n}} = \eta \Psi_{\mathbf{n}}$ with $\eta > 0$, we have $\frac{\alpha(p)}{\varepsilon_{\mathbf{n}} h_{\mathbf{n}}^d} = \frac{\alpha(p)}{\eta h_{\mathbf{n}}^d \Psi_{\mathbf{n}}}$ and setting $p = \left\lceil \Psi_{\mathbf{n}}^{-\frac{1}{N}} \right\rceil \leq \Psi_{\mathbf{n}}^{-\frac{1}{N}}$, we have $\widehat{\mathbf{q}} \geq \frac{\Psi_{\mathbf{n}} \widehat{\mathbf{n}}}{2^N}$ and since $\Psi_{\mathbf{n}} = \sqrt{\frac{\log \widehat{\mathbf{n}}}{\widehat{\mathbf{n}} h_{\mathbf{n}}^d}}$, we have $\frac{\varepsilon_{\mathbf{n}} \widehat{\mathbf{q}} h_{\mathbf{n}}^d}{A_0} \geq \frac{\eta \Psi_{\mathbf{n}}^2 \widehat{\mathbf{n}} h_{\mathbf{n}}^d}{2^N A_0} = \frac{\eta}{2^N A_0} \log \widehat{\mathbf{n}}$, then for $c = \frac{\eta}{2^{N+2} A_0}$, we obtain

$$\mathbb{P}\left(\left|\widehat{f}_{\mathbf{n}}(x) - \mathbb{E}\left(\widehat{f}_{\mathbf{n}}(x)\right)\right| > \varepsilon_{\mathbf{n}}\right) \leq C\left(c_{\widehat{\mathbf{n}}} + \widehat{\mathbf{n}}^{-c}\right),$$

with $c_{\widehat{\mathbf{n}}} = h_{\mathbf{n}}^{-d} s(\widehat{\mathbf{n}}, p^N) \varphi(p) \Psi_{\mathbf{n}}^{-1}$. By Assumption **H2** and the inequality (2.1), we get

$$c_{\widehat{\mathbf{n}}} \leq Ch_{\mathbf{n}}^{-d} p^{N-\theta} \Psi_{\mathbf{n}}^{-1} \sim Ch_{\mathbf{n}}^{-d} \Psi_{\mathbf{n}}^{\frac{\theta-2N}{N}} = C \widehat{\mathbf{n}}^{\frac{2N-\theta}{2N}} (\log \widehat{\mathbf{n}})^{\frac{\theta-2N}{2N}} h_{\mathbf{n}}^{\frac{-d\theta}{2N}}.$$

Consequently,

$$c_{\widehat{\mathbf{n}}} \leq C \left(\widehat{\mathbf{n}} (\log \widehat{\mathbf{n}})^{-1} h_{\mathbf{n}}^{\frac{-d\theta}{2N-\theta}} \right)^{\frac{2N-\theta}{2N}} = o(1), \quad (4.11)$$

and thus we conclude that

$$\mathbb{P}\left(\left|\widehat{f}_{\mathbf{n}}(x) - \mathbb{E}\left(\widehat{f}_{\mathbf{n}}(x)\right)\right| > \varepsilon_{\mathbf{n}}\right) \leq C \left(\widehat{\mathbf{n}}^{-c} + \left(\widehat{\mathbf{n}} (\log \widehat{\mathbf{n}})^{-1} h_{\mathbf{n}}^{\frac{-d\theta}{2N-\theta}} \right)^{\frac{2N-\theta}{2N}} \right). \quad (4.12)$$

Since $\theta > 2N$ and $\widehat{\mathbf{n}}(\log \widehat{\mathbf{n}})^{-1} h_{\widehat{\mathbf{n}}}^{-\frac{d\theta}{2N-\theta}} \rightarrow \infty$ by assumption, we have

$$\mathbb{P}\left(\left|\widehat{f}_{\widehat{\mathbf{n}}}(x) - \mathbb{E}\left(\widehat{f}_{\widehat{\mathbf{n}}}(x)\right)\right| > \varepsilon_{\widehat{\mathbf{n}}}\right) = o(1).$$

Proof of Corollary 3.2: By (3.2), it suffices to show that

$$\sup_{x \in \mathcal{M}} \left| \widehat{f}_{\widehat{\mathbf{n}}}(x) - \mathbb{E}\left(\widehat{f}_{\widehat{\mathbf{n}}}(x)\right) \right| = O_p(\Psi_{\widehat{\mathbf{n}}}).$$

Under Assumption **H4**, we have

$$\begin{aligned} \widehat{f}_{\widehat{\mathbf{n}}}(x) - \mathbb{E}\widehat{f}_{\widehat{\mathbf{n}}}(x) &= \left(\widehat{f}_{\widehat{\mathbf{n}}}(x) - \widehat{f}_{\widehat{\mathbf{n}}}(x_k)\right) + \left(\mathbb{E}\widehat{f}_{\widehat{\mathbf{n}}}(x_k) - \mathbb{E}\widehat{f}_{\widehat{\mathbf{n}}}(x)\right) + \left(\widehat{f}_{\widehat{\mathbf{n}}}(x_k) - \mathbb{E}\widehat{f}_{\widehat{\mathbf{n}}}(x_k)\right) \\ &= S_{1\widehat{\mathbf{n}}} + S_{2\widehat{\mathbf{n}}} + S_{3\widehat{\mathbf{n}}}. \end{aligned}$$

Therefore,

$$\sup_{x \in \mathcal{M}} |\widehat{f}_{\widehat{\mathbf{n}}}(x) - \mathbb{E}\widehat{f}_{\widehat{\mathbf{n}}}(x)| \leq \max_{1 \leq k \leq \mu_{\widehat{\mathbf{n}}}} \sup_{x \in B(x_k, a_{\widehat{\mathbf{n}}})} |S_{1\widehat{\mathbf{n}}} + S_{2\widehat{\mathbf{n}}}| + \max_{1 \leq k \leq \mu_{\widehat{\mathbf{n}}}} |S_{3\widehat{\mathbf{n}}}|. \quad (4.13)$$

For the first term on the right-hand side, we apply the same steps as in [Abdillahi et al. \(2024\)](#). As a result, we obtain

$$\max_{1 \leq k \leq \mu_{\widehat{\mathbf{n}}}} \sup_{x \in B(x_k, a_{\widehat{\mathbf{n}}})} |S_{1\widehat{\mathbf{n}}} + S_{2\widehat{\mathbf{n}}}| = O_p(\Psi_{\widehat{\mathbf{n}}}) \quad (4.14)$$

Now, using (4.12), we obtain

$$\mathbb{P}\left(\left|\widehat{f}_{\widehat{\mathbf{n}}}(x) - \mathbb{E}\left(\widehat{f}_{\widehat{\mathbf{n}}}(x)\right)\right| > \varepsilon_{\widehat{\mathbf{n}}}\right) \leq C\mu_{\widehat{\mathbf{n}}}\left(\widehat{\mathbf{n}}^{-c} + \left(\widehat{\mathbf{n}}(\log \widehat{\mathbf{n}})^{-1} h_{\widehat{\mathbf{n}}}^{-\frac{d\theta}{2N-\theta}}\right)^{\frac{2N-\theta}{2N}}\right). \quad (4.15)$$

We have $\mu_{\widehat{\mathbf{n}}} \leq C\widehat{\mathbf{n}}^{\frac{d}{2}} h_{\widehat{\mathbf{n}}}^{-d(\frac{d}{2}+1)} (\log \widehat{\mathbf{n}})^{-\frac{d}{2}}$. Since $\widehat{\mathbf{n}} h_{\widehat{\mathbf{n}}}^d \rightarrow \infty$ by Assumption **H3**, for $\widehat{\mathbf{n}}$ large enough, there exists $C > 0$ such that $\widehat{\mathbf{n}} h_{\widehat{\mathbf{n}}}^d > C$. It follows that $h_{\widehat{\mathbf{n}}}^{-d(\frac{d}{2}+1)} \leq C\widehat{\mathbf{n}}^{\frac{d}{2}+1}$. Consequently,

$$\mu_{\widehat{\mathbf{n}}}\widehat{\mathbf{n}}^{-c} \leq C\widehat{\mathbf{n}}^{d+1-c} (\log \widehat{\mathbf{n}})^{-\frac{d}{2}}, \quad (4.16)$$

which goes to 0 since $c > d + 1$. Next, using (4.11), we get

$$\mu_{\widehat{\mathbf{n}}} C_{\widehat{\mathbf{n}}} \leq C \left(\widehat{\mathbf{n}}(\log \widehat{\mathbf{n}})^{-1} h_{\widehat{\mathbf{n}}}^{\frac{d(\theta+dN+2N)}{\theta-(d+2)N}} \right)^{\frac{-\theta+(d+2)N}{2N}}. \quad (4.17)$$

The proof is then achieved since $\theta > (d+2)N$ and $\widehat{\mathbf{n}}(\log \widehat{\mathbf{n}})^{-1} h_{\widehat{\mathbf{n}}}^{\frac{d(\theta+dN+2N)}{\theta-(d+2)N}} \rightarrow \infty$.

Acknowledgments: *This work was supported by the University of Tunis El-Manar, Tunis (Tunisia), through the partnership with the University of Clermont Auvergne (France).*

References

- Abdillahi Isman, M., Nefzi, W., Mbaye, P., Khardani, S. and Yao, A.- F. (2024), ‘Kernel density estimation for a stochastic process with values in a Riemannian manifold’, *Journal of Nonparametric Statistics*, 1-20.
- Bouzebda, S. and Taachouche, N. (2023). Rates of the Strong Uniform Consistency for the Kernel-Type Regression Function Estimators with General Kernels on Manifolds. *Math. Methods Statist.*, **32**(1), 27–80.
- Bouzebda, S. and Taachouche, N. (2024). Rates of the strong uniform consistency with rates for conditional U -statistics estimators with general kernels on manifolds. *Math. Methods Statist.*, **33**(1), 95 –152.
- Berenfeld, C. and Hoffmann, M. (2021), ‘Density estimation on an unknown submanifold’, *Electronic Journal of Statistics* **15**(1), 2179 – 2223.
- Berry, T. and Sauer, T. (2017), ‘Density estimation on manifolds with boundary’, *Computational Statistics & Data Analysis* **107**, 1–17.
- Carbon, M., Tran, L. T. and Wu, B. (1997), ‘Kernel density estimation for random fields (density estimation for random fields)’, *Statistics & Probability Letters* **36**(2), 115–125.
- Chen, J. (2008), ‘Asymptotics of kernel density estimators on weakly associated random fields’, *Statistics & Probability Letters* **78**(18), 3230–3237.
- Cleanthous, G., Georgiadis, A. G., Kerkyacharian, G., Petrushev, P. and Picard, D. (2020), ‘Kernel and wavelet density estimators on manifolds and more general metric spaces’, *Bernoulli* **26**(3), 1832 – 1862.
- Dabo-Niang, S., Hamdad, L., Ternynck, C. and Yao, A.-F. (2014), ‘A kernel spatial density estimation allowing for the analysis of spatial clustering. application to monsoon Asia drought atlas data’, *Stochastic Environmental Research and Risk Assessment* **28**(8), 2075–2099.
- Dabo-Niang, S. and Yao, A.-F. (2007), ‘Kernel regression estimation for continuous spatial processes’, *Mathematical Methods of Statistics* **16**(4), 298–317.
- El Machkouri, M. and Reding, L. (2021), ‘On a class of recursive estimators for spatially dependent observations’, *Electronic Journal of Statistics* **15**, 4580 – 4624.
- Gallot, S., Hulin, D. and Lafontaine, J. (2004), ‘*Riemannian Geometry*’, Universitext Springer-Verlag, Berlin.

- Hallin, M., Lu, Z. and Tran, L. T. (2004), ‘Kernel density estimation for spatial processes: the l_1 theory’, *Journal of Multivariate Analysis* **88**(1), 61–75.
- Karcher, H. (1977), ‘Riemannian center of mass and mollifier smoothing’, *Communications on Pure and Applied Mathematics* **30**(5), 509–541.
- Khardani, S. and Yao, A. F. (2022), ‘Nonparametric recursive regression estimation on Riemannian manifolds’, *Statistics & Probability Letters* **182**, 109274.
- Kim, Y. and Park, H. (2013), ‘Geometric structures arising from kernel density estimation on Riemannian manifolds’, *Journal of Multivariate Analysis* **114**, 112–126.
- Mardia, K. V., Hughes, G., Taylor, C. C. and Singh, H. (2008), ‘A multivariate von Mises distribution with applications to bioinformatics’, *Canadian Journal of Statistics* **36**(1), 99–109.
- Pelletier, B. (2005), ‘Kernel density estimation on Riemannian manifolds’, *Statistics & Probability Letters* **73**, 297–304.
- Penneç, X. (2006), ‘Intrinsic statistics on Riemannian manifolds: Basic tools for geometric measurements’, *Journal of Mathematical Imaging and Vision* **25**(1), 127–154.
- Tran, L. T. (1990), ‘Kernel density estimation on random fields’, *Journal of Multivariate Analysis* **34**(1), 37–53.