Journal of Multivariate Analysis 114 (2013) 112–126

Contents lists available at SciVerse ScienceDirect

Journal of Multivariate Analysis

journal homepage: www.elsevier.com/locate/jmva

Geometric structures arising from kernel density estimation on Riemannian manifolds☆

Yoon Tae Kim, Hyun Suk Park *

Department of Finance and Information Statistics, Hallym University, Chuncheon 200-702, Republic of Korea

ARTICLE INFO

Article history:
Received 9 July 2011
Available online 16 July 2012

AMS subject classifications:
primary 62G07
secondary 53B20
58J60

Keywords:
Exponential map
Kernel density estimator
Riemannian manifolds
Minimax convergence rate
Relative efficiency

ABSTRACT

Estimating the kernel density function of a random vector taking values on Riemannian manifolds is considered. We make use of the concept of *exponential map* in order to define the kernel density estimator. We study the asymptotic behavior of the kernel estimator which contains geometric quantities (i.e. the curvature tensor and its covariant derivatives). Under a Hölder class of functions defined on a Riemannian manifold with some global losses, the L_2-minimax rate and its relative efficiency are obtained.

1. Introduction

We consider the problem of estimating the kernel density of an independent sample of points observed on the non-Euclidean space. The statistical aspect of the problem on the m-dimensional sphere $\mathbb{S}^m$ or the directional data was foreseen in an early paper of Fisher [4]. A survey of statistical methodologies dealing with the kind of non-Euclidean data may be found in [13,17,3,7,1,11,12,8,10].

The case where the sample space is a compact Riemannian manifold without boundary has been studied by Hendriks [9] and Pelletier [16]. When considering the statistical problem of estimating a density function defined on the non-Euclidean, all procedures used in the Euclidean space cannot be properly applied. For solving this problem, the exponential map will be introduced, which was actually one of the main topics in the dissertation of Park [14]. The kernel density estimator defined by the exponential map on $\mathbb{S}^m$ is considered in [15], where the estimator may be compared with one of [7] or [1]. This paper is to generalize the results of [15] to a complete Riemannian manifold.

Among the works regarding this research area, we already referred to the estimators of Hendriks [9] and Pelletier [16]. Their works lead to consider two points as follows.

- Hendriks' approach used Fourier analysis on a compact Riemannian manifold is analogous to techniques used in the case of the Euclidean space. In comparison with Hendriks' one, we point out that his result does not provide information on geometric structures of a Riemannian manifold M while being only interested in finding the $L^2(M)$-convergence rate.

☆ This research was supported by Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Education, Science and Technology (No. 2012-0002046).

* Corresponding author.
E-mail address: hspark@hallym.ac.kr (H.S. Park).

0047-259X/$ – see front matter © 2012 Elsevier Inc. All rights reserved.
doi:10.1016/j.jmva.2012.07.006

- Pelletier's estimator is based on kernels that are functions of a Riemannian geodesic distance on a manifold. We point out that the kernel function is defined by a map such that $K : \mathbb{R}_+ \to \mathbb{R}$, while the kernel estimator is defined by a map $f_{n,K} : p \in M \to f_{n,K}(p) \in \mathbb{R}$.

In this paper we place the focus on that the kernel function is defined on the *tangent space* $T_p(M)$ of Riemannian manifold M, and the *exponential map* for connecting the manifold with the tangent space is used. The purpose of these works is to establish that the formula of the asymptotic behavior of the bias and the mean square error contains geometric quantities expressed by terms of polynomials in the component of the curvature tensor and its covariant derivatives on a complete Riemannian manifold. Also this formula provides that the behavior depends on whether the manifold is curved positively or negatively. Besides, we state that the estimator achieves an optimal minimax rate over a Hölder class of functions defined on a complete Riemannian manifold under some global losses. For these purposes, we first use the exponential map in order to connect the manifold valued random sample X_i with the argument in the kernel function defined on the tangent space such that $\exp_p : T_p(M) \to M$, for $p \in M$. The exponential map is defined by $\exp_p \xi = \gamma_\xi(1)$, where $\gamma_\xi(t) = \exp_p t\xi$ is the geodesic through $p \in M$ at $t = 0$ with $\frac{d\gamma_\xi}{dt}|_{t=0} = \xi \in T_p(M)$. Using the exponential map, we can find the quantity analogous to $(x - X_i)/h$, which appears in the kernel function defined on the Euclidean space. In the case of $\mathbb{R}^m$, the geodesic γ passing through p tangent to ξ is the straight line of equation $\gamma_\xi(t) = p + t\xi$ for $t \in \mathbb{R}$. Hence $\exp_p \xi = \gamma_\xi(1) = p + \xi$. Therefore the quantity $p - X_i$ is equal to $\exp_p^{-1} X_i$ in the case of $\mathbb{R}^m$. It is natural to replace $p - X_i$ by $\exp_p^{-1} X_i$ if p and X_i are on M. This enables us to define the kernel density estimator by using the exponential map.

Let (M, g) be an m-dimensional complete Riemannian manifold with metric g. Suppose that we have a collection of i.i.d. random variables $X_1, \ldots, X_n$ taking values in M and having a probability density function f with respect to dV_g, where dV_g denote the m-dimensional volume element of M associated with the metric g. We then define the kernel density estimator of f using by the exponential map as follows.

Definition 1. The kernel density estimator with the kernel function defined on m-dimensional tangent space $T_p(M)$, for each $p \in M$ and smoothing parameter $h > 0$ is

$$\widehat{f}_n(p) = \frac{1}{nh^m C_h} \sum_{i=1}^{n} K\left(\frac{1}{h} \exp_p^{-1} X_i\right), \tag{1}$$

where C_h is the positive constant and h is the smoothing parameter such that

$$h^m C_h = \int_M K\left(\frac{1}{h} \exp_p^{-1} x\right) dV_g(p). \tag{2}$$

We will prove that the integral computation of (2) is independent of x, for any $x \in M$ and $C_h \to 1$ as $h \to 0$ in Lemma 2.

The paper is organized as follows. In Section 2, we discuss whether the proposed estimator given by (1) is well-defined. A property of the kernel function is considered and the asymptotic behavior of the estimator is formulated. The lower and upper bounds on the $L^2(M)$ convergence rate are achieved under a certain Hölder condition, and its relative efficiency is discussed in Section 3. Basic definitions and notations for the differential geometry are described in Appendix A, and also the proofs of the main theorems are given in Appendix B.

2. Asymptotic behavior

In this section, throughout, we follow the notations and basic concepts of Riemannian geometry in Appendix A. We suggest a positive kernel function $K(\cdot)$ as a function defined on the m-dimensional tangent space $T_p(M)$ such that

$$\int_{T_p(M)} K(v)dv = \int_{\mathcal{E}_p} \int_0^\infty K(t\xi)t^{m-1} dt d\mu_p(\xi) = 1, \tag{3}$$

and for all $k = 1, 2, \ldots$ and $l = 1, 2,$

$$\int_{\mathcal{E}_p} \int_0^\infty K^l(t\xi)t^{(m-1)+k} dt d\mu_p(\xi) < \infty, \tag{4}$$

where dv is a Lebesgue measure on $T_P(M)$ and $d\mu_p$ denotes the $(m-1)$-dimensional volume element on $\mathcal{E}_p = \{\xi \in T_p(M) : \|\xi\| = 1\}$. Usually K is taken as a radially symmetric unimodal probability function in the case of the m-dimensional Euclidean space. Hence we may choose $K(v) = T(\langle v, v\rangle_p^{1/2})$ for $v \in T_p(M)$, where $\langle \cdot, \cdot \rangle_p$ is the inner product with respect to the Riemannian metric g_p.

The kernel density estimator of f as the exponential map which associates the value $\hat{f}_n(p)$ defined on the manifold is formulated in Introduction. The random vector fields $\exp_p^{-1} X_i,\ i = 1, \ldots, n$, in (1) are well defined in the following sense: for each $\xi \in \mathcal{E}_p$ we define $c(\xi) = \sup\{t > 0 : d(p, \gamma_\xi(t)) = t\}$, where d is a Riemannian distance. Since *cut* *locus* $\partial\mathbf{D}_p$ has zero m-dimensional Riemannian measure in M and $M = \exp_p(\mathcal{D}_p) \cup \partial\mathbf{D}_p$ where the union is disjoint, and $\mathcal{D}_p = \{t\xi \in T_p(M) : 0 \le t < c(\xi),\ \xi \in \mathcal{E}_p\}$, it follows that

$$P(X \in M \setminus \mathbf{D}_p) = \int_{M\setminus\mathbf{D}_p} f(x)dV_g(x) = \int_{\partial\mathbf{D}_p} f(x)dV_g(x) = 0,$$

where $\mathbf{D}_p = \exp_p(\mathcal{D}_p)$. Hence we may consider only random variables X that take values, almost surely, on $\mathbf{D}_p$ of any point in a complete Riemannian manifold M. We shall suppose that $\exp_p^{-1} X_i$, $p \in M$, is defined in this sense.

Remark 1. If $M = \mathbb{R}^m$ the identity map $(x^1, \dots, x^n)$ of $\mathbb{R}^m$, by itself, is an atlas. Considering the usual Euclidean affine connection, we find that

$$\frac{1}{h}\exp_p^{-1} X_i = \frac{p - X_i}{h}, \quad \text{for } p,\ X_i \in \mathbb{R}^m. \quad \square$$

We now describe the asymptotic behavior of bias and variance of $\hat{f}_n(p)$ given in (1). First we need Lemma 1 for the expansions of f. For this we introduce some notations and new symbols as follows (see Appendix A for details of the Riemannian curvature):

$$\rho_{ij}(p) = \mathrm{Ric}_p(e_i, e_j) = \sum_{k=1}^m R_{ikjk}(p), \qquad \langle R(e_i, e_j)e_k, e_l\rangle_p = R_{ijkl}(p),$$

$$S(p) = \sum_{i=1}^m \rho_{ii}(p), \qquad \|R\|^2 = \sum_{i,j,k,l=1}^m R_{ijkl}^2(p), \qquad \|\rho\|^2 = \sum_{ij}^m \rho_{ij}^2(p),$$

$$\Delta S(p) = \sum_{i=1}^m \nabla_{ii}^2 S(p), \qquad \langle \nabla S, \nabla f\rangle = \sum_{j=1}^m \nabla_j S(p)\frac{\partial f}{\partial x^j}(p),$$

$$\langle \rho, \nabla^2 f\rangle = \sum_{i,j=1}^m \rho_{ij}(p)\mathrm{Hess}\, f(p)(e_i, e_j).$$

Let $\gamma_\xi(t) = \exp_p t\xi$. Then under the assumptions $f \in C^\infty(M)$, the following lemma gives the expansion of $f(\gamma_\xi(t))$ with respect to t.

Lemma 1. *We have the expansion of $f(\gamma_\xi(t))$.*

$$f(\gamma_\xi(t)) = \sum_{k=0}^\infty \frac{\gamma_k}{k!} t^k, \tag{5}$$

where $\gamma_k = \nabla_\xi \cdots \nabla_\xi f$. In particular, $\gamma_0 = f(p)$, $\gamma_1 = \nabla_\xi f$ and $\gamma_2 = \nabla_\xi \nabla_\xi f = \mathrm{Hess} f(p)(\xi, \xi)$.

Proof. Let $\mathbf{x} = (x^1, \dots, x^n)$ be a coordinate system in M at p. In local coordinates, γ_ξ is given by $\mathbf{x} \circ \gamma_\xi(t) = (x^1(t), \dots, x^m(t))$. If we write $\xi^i = \xi(x^i)$, we have $\xi = \sum_{i=1}^m \xi^i \frac{\partial}{\partial x^i}$ with $\xi^i = \dot{x}^i(0)$. Therefore, restricting f to γ_ξ, and then continuously differentiating $f(\gamma_\xi(t))$ with respect to t, we can obtain γ_k. $\square$

Now we prove that the integral computation given in (2) is independent of $x \in M$ and $C_h \to 1$ as $h \to 0$ under some conditions.

Lemma 2. *Let*

$$C_h = \frac{1}{h^m}\int_M K\left(\frac{1}{h}\exp_p^{-1} x\right) dV_g(p). \tag{6}$$

Then the integral computation of (6) is independent of $x \in M$. We assume that for some constant κ, Ricci curvature satisfies $\mathrm{Ric}(\xi, \xi) \ge \kappa(m-1)|\xi|^2$ for all $\xi \in T_p(M)$ and for sufficiently small $\delta > 0$,

$$\int_0^\infty T(s)e^{(m-1)\delta s} s^{m-1} ds < \infty. \tag{7}$$

Then $C_h \to 1$ as $h \to 0$.

Proof. For computation of the integral of C_h we may choose any point $x \in M$ by *geodesic spherical coordinates*. Then we have

$$\begin{aligned} C_h &= \frac{1}{h^m}\int_M T\left(\frac{1}{h}\langle \exp_p^{-1} x, \exp_p^{-1} x\rangle_p^{1/2}\right) dV_g(p) \\ &= \frac{1}{h^{m-1}}\int_{\mathcal{E}_x}\int_0^{c(\xi)/h} T(s)\sqrt{\mathbf{g}(hs;\xi)}ds d\mu_x(\xi), \end{aligned} \tag{8}$$

where $c(\xi)$ is the (possibly infinite) distance to the *cut locus* in the direction ξ. The last integral in (8) can be written as $\int_M F(p)dV_g(p)$, where $F: M \to \mathbb{R}$ is a function defined by $F(\exp_x(t\xi/h)) = T(t/h)$ for all $(t\xi/h) \in T_x(M)$. Thus the volume element $dV_g(p)$ in (6) can be replaced by $dV_g(\exp_x(t\xi/h))$ for any $x \in M$ and from (12) in Appendix A the integral of C_h can be formed. Hence the integral in (6) is independent of x. By the comparison theorem for Ricci curvature, we have

$$C_h \le \frac{1}{h^{m-1}}\int_{\mathcal{E}_x}\int_0^{c(\xi)/h} T(s)\mathbf{S}_\kappa^{m-1}(hs)ds d\mu_x(\xi),$$

where

$$\mathbf{S}_\kappa(t) = \begin{cases} (1/\sqrt{\kappa})\sin\sqrt{\kappa}t, & \kappa > 0, \\ t, & \kappa = 0, \\ (1/\sqrt{-\kappa})\sin\sqrt{-\kappa}t, & \kappa < 0. \end{cases}$$

Also note that

$$\lim_{h\to 0}\frac{\sqrt{\mathbf{g}(hs;\xi)}}{(hs)^{m-1}} = 1.$$

Hence from (7) and the dominated convergence theorem,

$$\lim_{h\to 0} C_h = \omega_{m-1}\int_0^\infty T(s)s^{m-1}ds = 1,$$

where ω_{m-1} is the area of the surface of the unit sphere centered at the origin in R^m. Using Lemmas 1 and 3 of Appendix A, we obtain an asymptotic formula of bias and variance of $\hat{f}_n(p)$ in the following theorem.

Hereafter we assume that $C_h \to 1$ as $h \to 0$. □

Theorem 1. *Suppose that the unknown density f is bounded and continuously four times differentiable. We assume that for some constant $c > 0$ and $l = 1, 2$, as $h \to 0$,*

$$\int_{c/h}^\infty T^l(s)s^{(m-1)+k}ds = o(e^{-1/h}) \quad \text{for } k = 1, 2, \ldots. \tag{9}$$

Then we have the bias and variance of the kernel estimator as follows: for fixed $p \in M$,

$$\begin{aligned} \mathrm{Bias}(\widehat{f_n}(p)) &= E\left[\widehat{f_n}(p)\right] - f(p) \\ &= \frac{\omega_{m-1}}{2m}\Delta f(p)\int_0^\infty T(s)s^{m+1}ds\frac{h^2}{C_h} + \left[\frac{\omega_{m-1}}{24m(m+2)}\left\{-4\langle\nabla S, \nabla f\rangle - 2S(p)\Delta f(p)\right.\right. \\ &\quad \left.\left. - 4\langle\rho, \nabla^2 f\rangle + 3\sum_{i,j=1}^m \nabla^4_{ijij} f(p)\right\} \times \int_0^\infty T(s)s^{m+3}ds\right]\frac{h^4}{C_h} + o\left(\frac{h^4}{C_h}\right), \quad \text{as } h \to 0, \end{aligned}$$

and

$$\begin{aligned} \mathrm{Var}(\widehat{f_n}(p)) &= E\left[\widehat{f_n}(p) - E\widehat{f_n}(p)\right]^2 \\ &= \frac{\omega_{m-1}}{nh^m C_h^2}\left[f(p)\int_0^\infty T^2(s)s^{m-1}ds + \frac{1}{2m}\left(\Delta f(p) - \frac{1}{3}f(p)S(p)\right)\int_0^\infty T^2(s)s^{m+1}dsh^2\right. \\ &\quad + \frac{1}{24m(m+2)}\left\{\frac{1}{15}\Big(-3\|R\|^2 + 8\|\rho\|^2 + 5S^2(p) - 18\Delta S(p)\Big)f(p)\right. \\ &\quad \left. - 4\langle\nabla S, \nabla f\rangle - 2(S(p)\Delta f(p) - 4\langle\rho, \nabla^2 f\rangle) + 3\sum_{i,j=1}^m \nabla^4_{ijij} f(p)\right\} \end{aligned}$$

$$\times \int_0^\infty T^2(s)s^{m+3}dsh^4\Bigg] - \frac{1}{n}\left\{f^2(p) + \frac{\omega_{m-1}}{m}f(p)\Delta f(p)\int_0^\infty T^2(s)s^{m+1}ds\frac{h^2}{C_h}\right\}$$
$$+ o\left((nh^{m-4})^{-1} + (nh^{-2})^{-1}\right), \quad \textit{as } nh^m \to \infty \textit{ and } h \to 0.$$

Remark 2. While Pelletier's kernel estimator does not contain geometric quantities in spite of the estimator on the Riemannian manifold, the results of Theorem 1 contain the curvature tensor and its covariant derivatives. These geometric information imply whether the manifold is curved negatively or positively. For example, the Euclidean space $\mathbb{R}^m$, the sphere $\mathbb{S}^{m-1}(\rho) = \{x \in \mathbb{R}^m : \|x\| = \rho\}$ and the hyperbolic space $\mathbb{B}^m(\rho)$ are the simply connected spaces with constant sectional curvatures $\kappa = 0, \kappa = 1/\rho$ and $\kappa = -1/\rho$, respectively. Therefore if $M = \mathbb{R}^m$, the *scalar curvature* $S(p) = 0$ for all $p \in \mathbb{R}^m$. □

Remark 3. From the variance part, $\mathrm{Var}(\widehat{f_n}(p))$ has the term

$$\frac{1}{2m}\left(\Delta f(p) - \frac{1}{3}f(p)S(p)\right).$$

If *scalar curvature* $S(p) > 0$ for $p \in M$, then we have, up to order $\frac{h^2}{nh^m}$, $\mathrm{Var}(\tilde{f}_n(p)) \geq \mathrm{Var}(\widehat{f_n}(p))$, where $\tilde{f}_n$ is a kernel density estimator on $\mathbb{R}^m$, and vice versa. □

Remark 4. Recently the density estimator defined on a manifold (or Lie group) has been used for image processing (e.g. astrophysics, nonparametric clustering mean shift techniques, 3D multiple rigid motion algorithm and so on). The result of Theorem 1 is motivated by the applications of those. To have the geometric quantities in Theorem 1 seems to us a warrant to consider how the manifold is curved in practice. □

3. The rate of $L_2(M)$-convergence

In this section we give the optimal minimax rate of convergence of the proposed estimator. As already mentioned in Introduction, Hendriks [9] is to be noted because he considers a compact Riemannian manifold without boundary to the case of the generalization of estimation with Fourier series, where the theory builds up the eigenfunctions of the Laplace–Beltrami operator on the manifold. His approach used Fourier analysis on $L_2(M)$ is analogy to techniques used in the case of the Euclidean space. Therefore he obtained the rate of convergence of the proposed estimator f_n^* in L_2 and L_∞ senses:

$$E\left(\|f - f_n^*\|_{L_2}^2\right) \leq O(n^{m/(2s+m)}n^{-1}), \qquad E\left(\|f - f_n^*\|_{L_\infty}^2\right) \leq O(n^{2m/(2s+m)}n^{-1}),$$

where suppose that the unknown density f has bounded and continuously $s(s > m/2)$ times differentiable. The $L_2(M)$ convergence rate of Pelletier [16, Theorem 5], is equal to those of Hendriks.

Now we state the lower and the upper bound on the minimax L_2 convergence rate in the case where f is in a certain Hölder class. Suppose f is β times differentiable with β-th covariant derivative. For a real function f belonging to $C^\beta(M)$ where $\beta \geq 0$ is an integer, we define

$$|\nabla^\beta f|^2 = \sum_{j_1,\ldots,j_\beta=1}^{m} \nabla_{j_1}\nabla_{j_2}\cdots\nabla_{j_\beta}f\,\nabla_{j_1}\nabla_{j_2}\cdots\nabla_{j_\beta}f,$$

where $\nabla^\beta f$ means any βth covariant derivative of f. In particular, $|\nabla^0 f| = |f|, |\nabla^1 f|^2 = |\nabla f|^2 = \sum_{i=1}^m \nabla_i f \nabla_i f$. For $0 < \gamma < 1$, we define $C^{\beta+\gamma}(M)$ to be the subspace $C^\beta(M)$ consisting of those functions f for which $\nabla^\beta f$ satisfies a Hölder condition of exponent γ, i.e.,

$$|\nabla^\beta f(x) - \nabla^\beta f(y)| \leq Kd(x, y)^\gamma, \quad \text{for } x, y \in M, \tag{10}$$

and take

$$\Sigma_m(M) = C^{\beta+\gamma}(M) \cap \left\{f : f \geq 0, \int_M f dV_g = 1\right\} \cap L_2.$$

We consider the L_2-norm as the global loss function:

$$L(f, g) = \left(\int_M |f(x) - g(x)|^2 dV_g(x)\right)^{1/2}.$$

We assume that

$$\int_{\mathcal{E}_p}\int_0^{c(\xi)/h} \xi^\alpha K(s\xi)\sqrt{\mathbf{g}(hs;\xi)}dsd\mu_p(\xi) = 0 \quad \text{for } 1 \le |\alpha| \le \beta, \tag{11}$$

where $\xi^\alpha = (\xi^1)^{\alpha_1}\cdots(\xi^m)^{\alpha_m}$, $\alpha_1 + \cdots + \alpha_m = |\alpha|$. To further simplify matters, we shall assume that functions in $\Sigma_m(M)$ may be regarded as taking values on a cube $Q \subseteq M$.

Theorem 2. *Suppose the probability function density f on M is in $\Sigma_m(M)$. Under the condition (11), an estimator of the form (1) with $h = n^{-1/(2\beta+2\gamma+m)}$ has, as $n \to \infty$,*

$$E\left(\|f(p) - \widehat{f}_n(p)\|^2_{L_2}\right) \le O\left(n^{-2(\beta+\gamma)/(2\beta+2\gamma+m)}\right),$$

where m is the dimension of M.

Theorem 3. *Suppose the probability function density f on M is in $\Sigma_m(M)$. The lower rate of convergence for the L_2-norm is given by*

$$\liminf_{n\to\infty}\inf_{\widehat{f}_n}\sup_{f\in\Sigma_m(M)} E\left(\|f(p) - \widehat{f}_n(p)\|^2_{L_2} n^{2(\beta+\gamma)/(2\beta+2\gamma+m)}\right) > 0,$$

where m is the dimension of M.

Remark 5. The rates in Theorems 2 and 3 are improved under the Hölder condition with component γ, which could be compared with those of Hendriks [9] and Pelletier [16], respectively.

The relative performance of the proposed estimator to that of Pelletier's estimator is of good interest. Therefore we shall consider statistical efficiency of the proposed estimator. Pelletier considers the following kernel density estimator on a compact Riemannian manifold M: for fixed $p \in M$

$$f_{n,K}(p) = \frac{1}{n}\sum_{i=1}^n \frac{1}{r^m\theta_{X_i}(p)}K\left(\frac{d_g(p, X_i)}{r}\right),$$

where $\theta_p(q)$ denotes the volume density function. When f is 2-times differentiable, we compute the asymptotic relative efficiency (ARE) of $f_{n,K}$ with respect to $\widehat{f}_n$ defined by

$$\text{ARE}(f_{n,K}, \widehat{f}_n) = \frac{\text{AMISE}(\widehat{f}_n)}{\text{AMISE}(f_{n,K})},$$

where AMISE is the asymptotic mean integrated square error plugging in the optimal bandwidth of each kernel density. By Lemmas 3.2 and 3.3 in [16], the asymptotic relative efficiency is evaluated as follows:

$$\text{ARE}(f_{n,K}, \widehat{f}_n) = \frac{C_1\left(\frac{mC_1}{4C_2}\right)^{-m/(m+4)} + C_2\left(\frac{mC_1}{4C_2}\right)^{4/(m+4)}}{D_1\left(\frac{mD_1}{4D_2}\right)^{-m/(m+4)} + D_2\left(\frac{mD_1}{4D_2}\right)^{4/(m+4)}},$$

where

$$C_1 = \omega_{m-1}\int_0^\infty T^2(s)s^{m-1}ds,$$

$$C_2 = \left(\frac{\omega_{m-1}}{2m}\int_0^\infty T(s)s^{m+1}ds\right)^2\int_M (\Delta f(p))^2 dV_g(p),$$

$$D_1 = \sup_M \sup_{B(q,r_0)} \theta_q^{-1}(p)\text{Vol}(B(0;1))K^2(0),$$

$$D_2 = \left(\int_{\mathcal{B}(0;1)} \|y\|^2 K(\|y\|)dy\right)^2 \int_M (\Delta f(p))^2 dV_g(p).$$

Acknowledgments

The authors would like to thank the reviewers for their careful reading of the manuscript and their constructive comments.

Appendix A. Riemannian differential geometry

We review some basic definitions and notations in Riemannian geometry. Given any tangent vector $\xi \in T_p(M)$ there is a maximal open interval I_ξ in $\mathbb{R}$ about the origin and an unique geodesic γ_ξ in M such that $\gamma_\xi(0) = p,\ \gamma_\xi'(0) = \xi$. We assume that M is *geodesically complete*, that is, $I_\xi = \mathbb{R}$. We define the *exponential map*, viz. $\exp_p : T_p(M) \to M$ by $\exp_p \xi = \gamma_\xi(1)$, thus $\gamma_\xi(t) = \exp_p t\xi$. Now we briefly mention the notions involving cut points. Let $\mathcal{E}_p = \{\xi \in T_p(M) \parallel \xi| = 1\}$. For each $\xi \in \mathcal{E}_p$ we define $c(\xi) = \sup\{t > 0 : d(p, \gamma_\xi(t)) = t\}$, where d is the Riemannian distance. If we set $\mathcal{D}_p = \{t\xi \in T_p(M) | 0 \le t < c(\xi), \xi \in \mathcal{E}_p\}$ and $\mathbf{D}_p = \exp \mathcal{D}_p$, then $\exp_p$ maps $\mathcal{D}_p$ diffeomorphically onto $\mathbf{D}_p$, and $\bar{\mathcal{D}}_p$ is mapped onto all of M. The *cut locus* $\partial\mathbf{D}_p$ is the image by $\exp_p$ of $\partial\mathcal{D}_p$, and for any $p \in M$, the *cut locus* $\partial\mathbf{D}_p$ has zero measure in M. For each $\xi \in \mathcal{E}_p$, let $\xi^\perp$ be the orthogonal complement of $\{\mathbb{R}\xi\}$ in $T_p(M)$, and let $\tau_t : T_p(M) \to T_{\exp_p t\xi}$ denote parallel translation along γ_ξ. Then we define the path of linear transformations $\mathcal{A}(t;\xi) : \xi^\perp \to \xi^\perp$ by $\mathcal{A}(t;\xi)\eta = (\tau_t)^{-1}Y(t)$, where $Y(t)$ is the Jacobi field along γ_ξ such that $Y(0) = 0$ and $(\nabla_t Y)(0) = \eta$, where ∇ is the Riemannian connection of M. The proposition of the linear transformations $\mathcal{A}(t;\xi)$ is given in [2, see p. 66]. We set $\sqrt{\mathbf{g}(t;\xi)} = \det \mathcal{A}(t;\xi)$. Using polar coordinates in $T_p(M)$, the m-dimensional volume element of M, dV_g, is given by

$$dV_g(\exp_p(t\xi)) = \sqrt{\mathbf{g}(t;\xi)}\,dt\,d\mu_p(\xi), \tag{12}$$

where $d\mu_p$ denotes the $(m-1)$-dimensional volume element on the unit sphere $\mathcal{E}_p$ in $(T_p(M), g_p)$.

We turn to the Riemannian curvature. For vector fields X, Y, Z on M, define $R(X,Y)Z = \nabla_Y\nabla_X Z - \nabla_X\nabla_Y Z + \nabla_{[X,Y]}Z$, where $[X,Y] = XY - YX$. R is called the *Riemannian curvature tensor* of ∇. For $p \in M$, the *Ricci tensor* Ric: $T_p(M) \times T_p(M) \to \mathbb{R}$ is defined by $\mathrm{Ric}(\xi,\eta) = \mathrm{trace}(\zeta \mapsto R(\xi,\zeta)\eta)$, and the *scalar curvature* $S(p)$ is defined to be the trace of Ric with respect to the Riemannian metric. Thus if $\{e_1,\ldots,e_m\}$ is an orthonormal basis of $T_p(M)$, we have $\mathrm{Ric}_p(\xi,\eta) = \sum_{i=1}^m \langle R(\xi,e_i)\eta, e_j\rangle$, and $S(p) = \sum_{i,j=1}^m \langle R(e_i,e_j)e_i, e_j\rangle$.

The power series expansions for $\sqrt{\mathbf{g}(t;\xi)}$ can be derived from [5] as follows: if we let $\xi = \sum_{i=1}^m \xi^i e_i$, where $\{e_1,\ldots,e_m\}$ is an orthonormal basis of a tangent space $T_p(M)$, then we have the following.

Lemma 3. *We have the asymptotic expansion in small t:*

$$\sqrt{\mathbf{g}(t;\xi)} = t^{m-1}\sum_{k=0}^{\infty}\frac{\gamma_k(\xi)}{k!}t^k, \tag{13}$$

where

$$\gamma_0 = 1, \qquad \gamma_1 = 0, \qquad \gamma_2 = -\frac{1}{3}\sum_{i,j=1}^m \rho_{ij}\xi^i\xi^j, \qquad \gamma_3 = -\frac{1}{2}\sum_{i,j,k=1}^m \nabla_i\rho_{jk}\xi^i\xi^j\xi^k$$

$$\gamma_4 = \sum_{i,j,k,l=1}^m \left\{-\frac{3}{5}\nabla_{ij}^2\rho_{kl} + \frac{1}{3}\rho_{ij}\rho_{kl} - \frac{2}{15}\sum_{a,b=1}^m R_{iajb}R_{kalb}\right\}\xi^i\xi^j\xi^k\xi^l, \ldots.$$

Appendix B. The proof of theorems

Proof of Theorem 1. Note that for any $p \in M$, $M = \exp_p(\mathcal{D}_p) \cup \partial\mathbf{D}_p$ where the union is disjoint and the *cut locus* $\partial\mathbf{D}_p$ has zero m-dimensional Riemannian measure. Therefore we have

$$\begin{aligned} E(\hat{f}_n(p)) &= \frac{1}{h^m C_h}\int_M K\left(\frac{1}{h}\exp_p^{-1}x\right)f(x)dV_g(x) \\ &= \frac{1}{h^m C_h}\int_{\exp_p(\mathcal{D}_p)} K\left(\frac{1}{h}\exp_p^{-1}x\right)f(x)dV_g(x). \end{aligned} \tag{14}$$

We work with the *geodesic spherical coordinates*, given by $y(t,\xi) = \exp_p t\xi$. If $c(\xi)$ is the (possibly infinite) distance to the *cut locus* in the direction ξ, the last integral in (14) becomes

$$\begin{aligned} &\frac{1}{h^{m-1}C_h}\int_{\mathcal{E}_p}\int_0^{c(\xi)/h} K(s\xi)f(y(hs,\xi))\sqrt{\mathbf{g}(hs;\xi)}\,ds\,d\mu_p(\xi) \\ &= \frac{1}{h^{m-1}C_h}\int_{\mathcal{E}_p}\int_0^{c(\xi)/h} T(s)f(y(hs,\xi))\sqrt{\mathbf{g}(hs;\xi)}\,ds\,d\mu_p(\xi). \end{aligned} \tag{15}$$

The geodesic $s \to \gamma_\xi(hs)$ has initial velocity $h\gamma'_\xi(0) = h\xi$. Hence $\gamma_\xi(hs) = \gamma_{h\xi}(s)$ for all h and s. By (8), we obtain

$$E\hat{f}_n(p) = f(p) + \frac{1}{h^{m-1}C_h}\int_{\mathcal{E}_p}\int_0^{c(\xi)/h} T(s)[f(y(hs,\xi)) - f(p)]\sqrt{\mathbf{g}(hs;\xi)}dsd\mu_p(\xi). \tag{16}$$

Note that for all $\xi \in \mathcal{E}_p$, $c(\xi)$ are bounded below by a strictly positive real number. Using the expansions in Lemmas 3 and 1 and the conditions in (9), we write $E\hat{f}_n(p) - f(p)$ as follows: as $h \to 0$,

$$E\hat{f}_n(p) - f(p) = \frac{1}{C_h}(B_1(h) + B_2(h) + B_3(h) + B_4(h) + o(h^4)),$$

where

$$\begin{aligned}
B_1(h) \;=\; & \int_{\mathcal{E}_p}\int_0^\infty T(s)\langle \mathrm{grad} f(p), \xi\rangle hs^m \Bigg\{1 - \frac{1}{6}\sum_{i,j=1}^m \rho_{ij}\xi^i\xi^j(hs)^2 - \frac{1}{12}\sum_{i,j,k=1}^m \nabla_i\rho_{jk}(hs)^3\xi^i\xi^j\xi^k \\
& + \frac{1}{24}\sum_{i,j,k,l=1}^m\left(-\frac{3}{5}\nabla^2_{ij}\rho_{kl} + \frac{1}{3}\rho_{ij}\rho_{kl} - \frac{2}{15}\sum_{a,b=1}^m R_{iajb}R_{kalb}\right) \times (hs)^4\xi^i\xi^j\xi^k\xi^l\Bigg\}dsd\mu_p(\xi), \\
B_2(h) \;=\; & \frac{1}{2}\int_{\mathcal{E}_p}\int_0^\infty T(s)\mathrm{Hess} f(p)(\xi,\xi)h^2s^{m+1} \times \Bigg\{1 - \frac{1}{6}\sum_{i,j=1}^m \rho_{ij}\xi^i\xi^j(hs)^2 - \frac{1}{12}\sum_{i,j,k=1}^m \nabla_i\rho_{jk}(hs)^3\xi^i\xi^j\xi^k \\
& + \frac{1}{24}\sum_{i,j,k,l=1}^m\left(-\frac{3}{5}\nabla^2_{ij}\rho_{kl} + \frac{1}{3}\rho_{ij}\rho_{kl} - \frac{2}{15}\sum_{a,b=1}^m R_{iajb}R_{kalb}\right)(hs)^4\xi^i\xi^j\xi^k\xi^l\Bigg\}dsd\mu_p(\xi), \\
B_3(h) \;=\; & \frac{1}{6}\int_{\mathcal{E}_p}\int_0^\infty T(s)\nabla^2 df(\xi,\xi,\xi)h^3s^{m+2} \times \Bigg\{1 - \frac{1}{6}\sum_{i,j=1}^m \rho_{ij}\xi^i\xi^j(hs)^2 - \frac{1}{12}\sum_{i,j,k=1}^m \nabla_i\rho_{jk}(hs)^3\xi^i\xi^j\xi^k \\
& + \frac{1}{24}\sum_{i,j,k,l=1}^m\left(-\frac{3}{5}\nabla^2_{ij}\rho_{kl} + \frac{1}{3}\rho_{ij}\rho_{kl} - \frac{2}{15}\sum_{a,b=1}^m R_{iajb}R_{kalb}\right) \times (hs)^4\xi^i\xi^j\xi^k\xi^l\Bigg\}dsd\mu_p(\xi), \\
B_4(h) \;=\; & \frac{1}{24}\int_{\mathcal{E}_p}\int_0^\infty T(s)\nabla^3 df(\xi,\xi,\xi,\xi)h^4s^{m+3} \times \Bigg\{1 - \frac{1}{6}\sum_{i,j=1}^m \rho_{ij}\xi^i\xi^j(hs)^2 - \frac{1}{12}\sum_{i,j,k=1}^m \nabla_i\rho_{jk}(hs)^3\xi^i\xi^j\xi^k \\
& + \frac{1}{24}\sum_{i,j,k,l=1}^m\left(-\frac{3}{5}\nabla^2_{ij}\rho_{kl} + \frac{1}{3}\rho_{ij}\rho_{kl} - \frac{2}{15}\sum_{a,b=1}^m R_{iajb}R_{kalb}\right) \times (hs)^4\xi^i\xi^j\xi^k\xi^l\Bigg\}dsd\mu_p(\xi).
\end{aligned}$$

We write $B_1(h) := B_{1,1}(m)h + B_{1,2}(m)h^3 + B_{1,3}(m)h^4$, where

$$\begin{aligned}
B_{1,1}(m) &= \int_{\mathcal{E}_p}\int_0^\infty T(s)\langle \mathrm{grad} f(p), \xi\rangle s^m dsd\mu_p(\xi), \\
B_{1,2}(m) &= -\frac{1}{6}\sum_{i,j=1}^m \rho_{ij}\int_{\mathcal{E}_p}\int_0^\infty T(s)\langle \mathrm{grad}\, f(p), \xi\rangle s^{m+2}\xi^i\xi^j dsd\mu_p(\xi), \\
B_{1,3}(m) &= -\frac{1}{12}\sum_{i,j,k=1}^m \nabla_i\rho_{jk}(hs)^3\int_{\mathcal{E}_p}\int_0^\infty T(s)\langle \mathrm{grad}\, f(p), \xi\rangle \times s^{m+3}\xi^i\xi^j\xi^k dsd\mu_p(\xi).
\end{aligned}$$

We may choose an orthonormal basis $\{e_1, \ldots, e_m\}$ of $T_p(M)$, and then writing $\xi = \sum_{i=1}^m \xi^i e_i$, we obtain

$$\begin{aligned}
\int_{\mathcal{E}_p}\langle \mathrm{grad}\, f(p), \xi\rangle d\mu_p(\xi) \;&=\; \sum_{i=1}^m \langle \mathrm{grad}\, f(p), e_i\rangle \int_{\mathcal{E}_p}\xi^i d\mu_p(\xi) \\
&= 0.
\end{aligned} \tag{17}$$

Hence $B_{1,1}(m) = 0$. Also $B_{1,2}(m)$ is

$$\begin{aligned}
B_{1,2}(m) \;&=\; -\frac{1}{6}\sum_{i,j,k=1}^m \rho_{ij}\int_{\mathcal{E}_p}\int_0^\infty T(s)\langle \mathrm{grad}\, f(p), e_k\rangle s^{m+2}\xi^i\xi^j\xi^k dsd\mu_p(\xi) \\
&= 0.
\end{aligned}$$

Consider $B_{1,3}(m)$ term. Write

$$\eta_{ijkl} = \nabla_i \rho_{jk} \frac{\partial f}{\partial x^l}(p).$$

First note that $\sum_{i=1}^m \nabla_i \rho_{ij} = 1/2 \nabla_j S$ and $\sum_{j=1}^m \nabla_i \rho_{jj} = \nabla_i S$. Then

$$\begin{aligned}
\sum_{i,j,k,l=1}^m \eta_{ijkl} \int_{\mathcal{E}_p} \xi^i \xi^j \xi^k \xi^l d\mu_p(\xi) &= \frac{3\mathrm{Vol}(\mathcal{E}_p)}{m(m+2)} \sum_{i=1}^m \eta_{iiii} + \frac{\mathrm{Vol}(\mathcal{E}_p)}{m(m+2)} \sum_{i\neq j=1}^m (\eta_{iijj} + \eta_{ijij} + \eta_{ijji}) \\
&= \frac{\mathrm{Vol}(\mathcal{E}_p)}{m(m+2)} \sum_{i,j=1}^m (\eta_{iijj} + \eta_{ijij} + \eta_{ijji}) \\
&= \frac{\mathrm{Vol}(\mathcal{E}_p)}{m(m+2)} \sum_{i,j=1}^m \left\{ 2\nabla_i \rho_{ij} \frac{\partial f}{\partial x^j}(p) + \nabla_i \rho_{jj} \frac{\partial f}{\partial x^i}(p) \right\} \\
&= 2\frac{\mathrm{Vol}(\mathcal{E}_p)}{m(m+2)} \sum_{j=1}^m \nabla_j S \frac{\partial f}{\partial x^j}(p) \\
&= 2\frac{\mathrm{Vol}(\mathcal{E}_p)}{m(m+2)} \langle \nabla S, \nabla f \rangle.
\end{aligned}$$

Hence

$$B_1(h) = -\frac{\mathrm{Vol}(\mathcal{E}_p)}{6m(m+2)} \langle \nabla S, \nabla f \rangle \int_0^\infty T(s) s^{m+3} ds h^4. \tag{18}$$

We write $B_2(h) := B_{2,1}(m)h^2 + B_{2,2}(m)h^4$, where

$$\begin{aligned}
B_{2,1}(m) &= \frac{1}{2} \int_{\mathcal{E}_p} \int_0^\infty T(s) \mathrm{Hess}\, f(p)(\xi, \xi) s^{m+1} ds d\mu_p(\xi), \\
B_{2,2}(m) &= -\frac{1}{12} \sum_{i,j=1}^m \rho_{ij} \int_{\mathcal{E}_p} \int_0^\infty T(s) \mathrm{Hess}\, f(p)(\xi, \xi) s^{m+3} \xi^i \xi^j ds d\mu_p(\xi).
\end{aligned}$$

Diagonalizing the symmetric bilinear form Hess$f$ with respect to an orthonormal basis $\{e_1, \ldots, e_m\}$ of $T_p(M)$, we have

$$\begin{aligned}
\int_{\mathcal{E}_p} \mathrm{Hess}\, f(p)(\xi, \xi) d\mu_p(\xi) &= \sum_{i,j=1}^m \mathrm{Hess}\, f(p)(e_i, e_j) \int_{\mathcal{E}_p} \xi^i \xi^j d\mu_p(\xi) \\
&= \sum_{i=1}^m \mathrm{Hess}\, f(p)(e_i, e_i) \int_{\mathcal{E}_p} (\xi^i)^2 d\mu_p(\xi) + \sum_{i,j=1, i\neq j}^m \mathrm{Hess} f(p)(e_i, e_j) \int_{\mathcal{E}_p} \xi^i \xi^j d\mu_p(\xi) \\
&= \frac{\mathrm{Vol}(\mathcal{E}_p)}{m} \mathrm{tr}\, \mathrm{Hess}\, f(p).
\end{aligned}$$

Thus

$$B_{2,1}(m) = \frac{\mathrm{Vol}(\mathcal{E}_p)}{2m} \Delta f(p) \int_0^\infty T(s) s^{m+1} ds.$$

Write $\kappa_{ijkl} = \rho_{ij} \mathrm{Hess} f(p)(e_k, e_l)$. Then similarly as computations in $B_{1,3}(h)$, we have

$$\begin{aligned}
\sum_{i,j,k,l=1}^m \kappa_{ijkl} \int_{\mathcal{E}_p} \xi^i \xi^j \xi^k \xi^l d\mu_p(\xi) &= \frac{\mathrm{Vol}(\mathcal{E}_p)}{m(m+2)} \sum_{i,j=1}^m (\kappa_{iijj} + \kappa_{ijij} + \kappa_{ijji}) \\
&= \frac{\mathrm{Vol}(\mathcal{E}_p)}{m(m+2)} \sum_{i,j=1}^m (\rho_{ii} \mathrm{Hess}\, f(p)(e_j, e_j) + 2\rho_{ij} \mathrm{Hess}\, f(p)(e_i, e_j)) \\
&= \frac{\mathrm{Vol}(\mathcal{E}_p)}{m(m+2)} (S(p)\Delta f(p) + 2\langle \rho, \nabla^2 f \rangle).
\end{aligned}$$

Hence $B_2(h)$ is given as

$$B_2(h) = \frac{\mathrm{Vol}(\mathcal{E}_p)}{2m}\Delta f(p)\int_0^\infty T(s)s^{m+1}dsh^2 - \frac{\mathrm{Vol}(\mathcal{E}_p)}{12m(m+2)}(S(p)\Delta f(p) + 2\langle\rho, \nabla^2 f\rangle)\int_0^\infty T(s)s^{m+3}dsh^4. \quad (19)$$

Since $\int_{\mathcal{E}_p}\xi^i\xi^j\xi^k\mu_p(\xi) = 0$, we have that

$$B_3(h) = o(h^4) \quad \text{as } h \to 0. \quad (20)$$

Now consider $B_4(h)$. As $h \to 0$, we write $B_4(h) := B_{4,1}(m)h^4 + o(h^4)$, where

$$B_{4,1}(m) = \frac{1}{24}\int_{\mathcal{E}_p}\int_0^\infty T(s)\nabla^3 df(\xi, \xi, \xi, \xi)s^{m+3}dsd\mu_p(\xi) + o(h^4).$$

First note that

$$\begin{aligned}\sum_{i,j,k,l=1}^m \nabla^3 df(e_i, e_j, e_k, e_l)\int_{\mathcal{E}_p}\xi^i\xi^j\xi^k\xi^l d\mu_p(\xi) &= \frac{\mathrm{Vol}(\mathcal{E}_p)}{m(m+2)}\sum_{i,j=1}^m\Big(\nabla^3 df(e_i, e_i, e_j, e_j) \\ &\quad + \nabla^3 df(e_i, e_j, e_i, e_j) + \nabla^3 df(e_i, e_j, e_j, e_i)\Big) \\ &= 3\frac{\mathrm{Vol}(\mathcal{E}_p)}{m(m+2)}\sum_{i,j=1}^m \nabla^4_{ijij}f(p).\end{aligned}$$

Hence we have that

$$B_{4,1}(m) = \frac{\mathrm{Vol}(\mathcal{E}_p)}{8m(m+2)}\sum_{i,j=1}^m \nabla^4_{ijij}f(p)\int_0^\infty T(s)s^{m+3}ds. \quad (21)$$

Now combining the above results (18)–(21), we obtain

$$\begin{aligned}B_1(h) + B_2(h) + B_3(h) + B_4(h) + o(h^4) &= \frac{\mathrm{Vol}(\mathcal{E}_p)}{2m}\Delta f(p)\int_0^\infty T(s)s^{m+1}dsh^2 \\ &\quad + \frac{\mathrm{Vol}(\mathcal{E}_p)}{24m(m+2)}\Big\{-4\langle\nabla S, \nabla f\rangle - 2S(p)\Delta f(p) - 4\langle\rho, \nabla^2 f\rangle \\ &\quad + 3\sum_{i,j=1}^m \nabla^4_{ijij}f(p)\Big\} \times \int_0^\infty T(s)s^{m+3}dsh^4 + o(h^4), \quad \text{as } h \to 0. \quad (22)\end{aligned}$$

From (22), it follows that

$$\begin{aligned}\mathrm{Bias}(\hat{f}_n(p)) &= \frac{\mathrm{Vol}(\mathcal{E}_p)}{2m}\Delta f(p)\int_0^\infty T(s)s^{m+1}ds\frac{h^2}{C_h} + \Bigg[\frac{\mathrm{Vol}(\mathcal{E}_p)}{24m(m+2)}\Big\{-4\langle\nabla S, \nabla f\rangle \\ &\quad - 2(S(p)\Delta f(p) - 4\langle\rho, \nabla^2 f\rangle) + 3\sum_{i,j=1}^m \nabla^4_{ijij}f(p)\Big\} \times \int_0^\infty T(s)s^{m+3}ds\Bigg]\frac{h^4}{C_h} \\ &\quad + o\left(\frac{h^4}{C_h}\right), \quad \text{as } \frac{h}{C_h} \to 0.\end{aligned}$$

Next we compute the asymptotic behavior of the variance. Similarly as before we can prove that as $h \to 0$,

$$\frac{1}{h^{m-1}}\int_{\mathcal{E}_p}\int_0^{c(\xi)/h} T^2(s)f(p)\sqrt{\mathbf{g}(hs; \xi)}dsd\mu_p(\xi) = A_1(m) + A_2(m)h^2 + A_3(m)h^3 + A_4(m)h^4 + o(h^4),$$

where

$$A_1(m) = \int_0^\infty T^2(s)s^{m-1}dsf(p)$$

$$A_2(m) = -\frac{1}{6}\int_{\mathcal{E}_p}\int_0^\infty T^2(s)s^{m+1}\sum_{i,j=1}^m \rho_{ij}\xi^i\xi^j dsd\mu_p(\xi)f(p)$$

$$A_3(m) = -\frac{1}{12}\int_{\mathcal{E}_p}\int_0^\infty T^2(s)s^{m+2}\sum_{i,j,k=1}^m \nabla_i\rho_{jk}\xi^i\xi^j\xi^k dsd\mu_p(\xi)f(p)$$

$$A_4(m) = \frac{1}{24}\int_{\mathcal{E}_p}\int_0^\infty T^2(s)s^{m+3}\sum_{i,j,k,l=1}^m\left(-\frac{3}{5}\nabla^2_{ij}\rho_{kl}+\frac{1}{3}\rho_{ij}\rho_{kl}-\frac{2}{15}\sum_{a,b=1}^m R_{iajb}R_{kalb}\right)\xi^i\xi^j\xi^k\xi^l dsd\mu_p(\xi)f(p).$$

Since $\int_{\mathcal{E}_p}\xi^i\xi^j d\mu_p(\xi) = 0$ and $\sum_{i=1}^m \rho_{ii}(p) = S(p)$, we have

$$\sum_{i,j=1}^m \rho_{ij}\int_{\mathcal{E}_p}\xi^i\xi^j\mu_p(\xi) = S(p)\frac{\mathrm{Vol}(\mathcal{E}_p)}{m}.$$

So

$$A_2(m) = -\frac{1}{6}S(p)\frac{\mathrm{Vol}(\mathcal{E}_p)}{m}\int_0^\infty T^2(s)s^{m+1}dsf(p).$$

Since $\int_{\mathcal{E}_p}\xi^i\xi^j\xi^k\mu_p(\xi) = 0$, we have that $A_3(m) = 0$. Consider $A_4(m)$. Write

$$\zeta_{ijkl} := -\frac{3}{5}\nabla^2_{ij}\rho_{kl}+\frac{1}{3}\rho_{ij}\rho_{kl}-\frac{2}{15}\sum_{a,b=1}^m R_{iajb}R_{kalb}.$$

Using the similar computation as in $B_{1,3}(h)$ and $\sum_{i,j,k,l=1}^m R_{ijkl}R_{ikjjl} = (1/2)\|R\|^2$, we have

$$\sum_{i,j,k,l=1}^m \zeta_{ijkl}\int_{\mathcal{E}_p}\xi^i\xi^j\xi^k\xi^l\mu_p(\xi) = \frac{\mathrm{Vol}(\mathcal{E}_p)}{15m(m+2)}(-3\|R\|^2+8\|\rho\|^2+5S^2(p)-18\Delta S(p)).$$

Hence

$$A_4(m) = \frac{\mathrm{Vol}(\mathcal{E}_p)}{360m(m+2)}\Big(-3\|R\|^2+8\|\rho\|^2+5S^2(p)-18\Delta S(p)\Big)\times\int_0^\infty T^2(s)s^{m+3}dsf(p).$$

The computation of the remaining part is essentially the same as the computation of bias. Thus

$$\begin{aligned}
\mathrm{Var}(\hat{f}_n(p)) &= \frac{1}{nh^{2m}C_h^2}\mathrm{Var}\left(K\left(\frac{1}{h}\exp_p^{-1}X\right)\right)\\
&= \frac{1}{nh^mC_h^2}\left\{\frac{1}{h^m}E\left(K^2\left(\frac{1}{h}\exp_p^{-1}X\right)\right)\right\}-\frac{1}{n}(E(\hat{f}_n(p)))^2\\
&= \frac{\mathrm{Vol}(\mathcal{E}_p)}{nh^mC_h^2}\Bigg[\int_0^\infty T^2(s)s^{m-1}dsf(p)+\quad\frac{1}{2m}\Big(\Delta f(p)-\frac{1}{3}f(p)S(p)\Big)\int_0^\infty T^2(s)s^{m+1}dsh^2\\
&\quad+\frac{1}{24m(m+2)}\Bigg\{\frac{1}{15}(-3\|R\|^2+8\|\rho\|^2+5S^2(p)-18\Delta S(p))f(p)\\
&\quad-4\langle\nabla S,\nabla f\rangle-2(S(p)\Delta f(p)-4\langle\rho,\nabla^2 f\rangle)+3\sum_{i,j=1}^m\nabla^4_{ijij}f(p)\Bigg\}\times\int_0^\infty T^2(s)s^{m+3}dsh^4\Bigg]\\
&\quad-\frac{1}{n}\left\{f^2(p)+\frac{\mathrm{Vol}(\mathcal{E}_p)}{m}f(p)\Delta f(p)\int_0^\infty T^2(s)s^{m+1}ds\frac{h^2}{C_h}\right\}+o\Big((nh^{m-4})^{-1}+(nh^{-2})^{-1}\Big),
\end{aligned}$$

as $nh^m\to\infty$ and $h\to 0$. □

Proof of Theorem 2. From (11) and Lemma 1, we obtain

$$\begin{aligned}
E\hat{f}_n(p)-f(p) &= \frac{1}{h^{m-1}C_h}\int_{\mathcal{E}_p}\int_0^{c(\xi)/h}T(s)[f(y(hs,\xi))-f(p)]\sqrt{\mathbf{g}(hs;\xi)}dsd\mu_p(\xi)\\
&= \frac{h^\beta}{h^{m-1}C_h\beta!}\sum_{j_1,\ldots,j_\beta=1}^m\int_{\mathcal{E}_p}\int_0^{c(\xi)/h}\xi^{j_1}\cdots\xi^{j_\beta}T(s)s^\beta\nabla^{\beta-1}df(y(\theta,\xi))(e_{j_1},\ldots,e_{j_\beta})
\end{aligned}$$

$$\times \sqrt{\mathbf{g}(hs;\xi)}dsd\mu_p(\xi) \tag{23}$$

$$= \frac{h^\beta}{h^{m-1}C_h\beta!}\sum_{j_1,\dots,j_\beta=1}^{m}\int_{\mathcal{E}_p}\int_0^{c(\xi)/h}\xi^{j_1}\cdots\xi^{j_\beta}T(s)s^\beta\Big[\nabla^{\beta-1}df(y(\theta,\xi))(e_{j_1},\dots,e_{j_\beta})$$

$$-\nabla^{\beta-1}df(p)(e_{j_1},\dots,e_{j_\beta})\Big]\sqrt{\mathbf{g}(hs;\xi)}dsd\mu_p(\xi), \tag{24}$$

where θ is a point in $(0, hs)$. Applying the condition (10)–(24), we estimate

$$|E\hat{f}_n(p)-f(p)| \le \frac{Kh^\beta}{h^{m-1}C_h\beta!}\sum_{j_1,\dots,j_\beta=1}^{m}\int_{\mathcal{E}_p}\int_0^{c(\xi)/h}|\xi^{j_1}\cdots\xi^{j_\beta}|T(s)s^\beta d(y(\theta,\xi),p)^\gamma\sqrt{\mathbf{g}(hs;\xi)}dsd\mu_p(\xi)$$

$$\le \frac{Kh^{\beta+\gamma}}{h^{m-1}C_h\beta!}\sum_{j_1,\dots,j_\beta=1}^{m}\int_{\mathcal{E}_p}\int_0^{c(\xi)/h}|\xi^{j_1}\cdots\xi^{j_\beta}|T(s)s^{\beta+\gamma}\sqrt{\mathbf{g}(hs;\xi)}dsd\mu_p(\xi). \tag{25}$$

The right-hand side in (25) can be written as

$$|E\hat{f}_n(p)-f(p)| \le \frac{Kh^{\beta+\gamma}}{h^{m-1}\beta!}\sum_{j_1,\dots,j_\beta=1}^{m}\int_{\mathcal{E}_p}\int_0^{c(\xi)/h}|\xi^{j_1}\cdots\xi^{j_\beta}|T(s)s^{\beta+\gamma}\sqrt{\mathbf{g}(hs;\xi)}dsd\mu_p(\xi)$$

$$\times\left[1+\left(\frac{1}{C_h}-1\right)\right]. \tag{26}$$

The expansion of $\sqrt{\mathbf{g}(hs;\xi)}$ in Lemma 3 and the fact that $C_h \to 1$ as $h \to 0$ yield, from (26), that $|E\hat{f}_n(p)-f(p)|$ can be estimated by

$$|E\hat{f}_n(p)-f(p)| \le \frac{Kh^{\beta+\gamma}}{\beta!}\sum_{j_1,\dots,j_\beta=1}^{m}\int_{\mathcal{E}_p}\int_0^{\infty}|\xi^{j_1}\cdots\xi^{j_\beta}|T(s)s^{\beta+\gamma}dsd\mu_p(\xi)$$

$$\le \frac{Kh^{\beta+\gamma}}{\beta!}\mathrm{Vol}(\mathcal{E}_p)\int_0^\infty T(s)s^{\beta+\gamma+m-1}ds + o(h^{\beta+\gamma}), \tag{27}$$

where the constant A_1 depends only on m and β. On the other hand, the variance part in Theorem 1 can be estimated by

$$\mathrm{Var}(\hat{f}_n(p)) \le \frac{\mathrm{Vol}(\mathcal{E}_p)}{nh^m}f(p)\int_0^\infty T^2(s)s^{m-1}ds + o\Big(\frac{1}{nh^m}\Big). \tag{28}$$

From two inequalities (27) and (28), for sufficiently small $h > 0$, the mean squared error of the kernel density can be estimated by

$$\int_Q E|\hat{f}_n(p)-f(p)|^2dV_g(p) \le C_1h^{2(\beta+\gamma)} + \frac{C_2}{nh^m}, \tag{29}$$

where

$$C_1 = \frac{K^2\omega_{m-1}^2}{(\beta!)^2}\left(\int_0^\infty T(s)s^{\beta+\gamma+m-1}ds\right)^2 V_g(Q),$$

$$C_2 = \frac{\omega_{m-1}}{nh^m}\int_0^\infty T^2(s)s^{m-1}ds.$$

From (29), choosing $h = \left(\frac{mC_2}{2C_1(\beta+\gamma)}\right)^{1/(2\beta+2\gamma+m)} n^{-1/(2\beta+2\gamma+m)}$, then we have

$$\int_Q E|\hat{f}_n(p)-f(p)|^2dV_g(p) \le O(n^{-2(\beta+\gamma)/(2\beta+2\gamma+m)}).$$

So we arrive at the assertion of theorem. □

It is necessary to give new notations for the proof of the following theorems:

$$\mathcal{B}(p;\delta)=\{\xi\in T_p(\mathbb{S}^m):\|\xi\|<\delta\},\qquad B(p;\delta)=\{x\in M:d(p,x)<\delta\},$$
$$S(p;\delta)=\{x\in M:d(p,x)=\delta\},\qquad \mathcal{E}(p;\delta)=\{\xi\in T_p(\mathbb{S}^m):\|\xi\|=\delta\},$$

where d is a Riemannian distance. Note that we have $B(p;\delta)=\exp_p\mathcal{B}(p;\delta)$.

Proof of Theorem 3. We consider a cube $Q\subseteq M$, i.e., Q lies in the domain of an associated, oriented, coordinate neighborhood U,φ and $\varphi(Q)=C=\{x\in\mathbb{R}^m:0\le x^i\le a, i=1,\dots,m\}$ for a constant $a>0$, and a cube of $\mathbb{R}^m$. Thus a cube Q is a compact set. Let K_n be a sequence of positive integers tending to infinity. We begin with K_n^m cubes $C_{j_1,\dots,j_m}=\mathcal{I}_{j_1}\times\cdots\times\mathcal{I}_{j_m}$, having center $x_{j_1,\dots,j_m}, j_1,\dots,j_m=1,\dots,K_n$, where $\mathcal{I}_{j_i}=[a(j_i-1)/K_n, aj_i/K_n), i=1,\dots,m$ and $x_{j_1,\dots,j_m}=(a(2j_1-1)/2K_n,\dots,a(2j_m-1)/2K_n)$. Write C as the disjoint union of $C_{n,\alpha}$ $\alpha=1,\dots,K_n^m$, and the center of $C_{n,\alpha}$ as $x_{n,\alpha}$. Thus we have a partition of Q on M, $Q_{n,\alpha}=\varphi^{-1}(C_{n,\alpha}), \alpha=1,\dots,K_n^m$ and take $p_{n,\alpha}=\varphi^{-1}(x_{n,\alpha})$. Therefore we may choose a ball with the property that $Q_{n,\alpha}\supset\varphi^{-1}(\bar{\mathcal{B}}_{a/2K_n}(x_{n,\alpha}))$, where $\bar{\mathcal{B}}_{a/2K_n}(x_{n,\alpha})$ denotes the closure of the open ball of radius $a/2K_n$ centered at $x_{n,\alpha}$.

Let $x^1,\dots,x^m$ denote the local coordinates and $g_{ij}, i,j=1,\dots,m$, the components of the Riemannian metric tensor g as a function of these coordinates. Since $g_{ij}(x)$ is C^∞ and is positive definite for each x in $\varphi(Q)$, and hence continuous in $x\in\varphi(Q)$, on the compact set $\varphi(Q)$, there exists $\lambda>0$ such that

$$g_{ij}(x)y^iy^j\ge\lambda\|y\|^2$$

for all $y=(y^1,\dots,y^m)\in\mathbb{R}^m$. Hence for all α and n we have

$$g_{ij}(x)y^iy^j\ge\lambda_{\alpha,n}\|y\|^2\ge\lambda\|y\|^2 \tag{30}$$

for all $y=(y^1,\dots,y^m)\in\mathbb{R}^m$ and $x\in\bar{\mathcal{B}}_{a/2K_n}(x_{n,\alpha})$. For each α let $\gamma_\alpha(t), a\le t\le b$, be a piecewise smooth curve lying in $\varphi^{-1}(\bar{\mathcal{B}}_{a/2K_n}(x_{n,\alpha}))\subset Q_{n,\alpha}$ with $\gamma_\alpha(a)=p_{n,\alpha}$ and $\gamma_\alpha(b)=q_{n,\alpha}\in\varphi^{-1}(\{x\in\mathbb{R}^m:\|x-x_{n,\alpha}\|=a/2K_n\})$, where $\|x-x_{n,\alpha}\|$ is the Euclidean distance from x to $x_{n,\alpha}$. Working with the coordinates $(x^1(\gamma_\alpha(t)),\dots,x^m(\gamma(t)))$ we use the following notation

$$\dot{x}^i(t)=\frac{d}{dt}x^i(\gamma_\alpha(t)).$$

Let

$$L(\gamma_\alpha)=\int_a^b\left[\sum_{i,j=1}^m g_{ij}(x(\gamma_\alpha(t)))\dot{x}^i(t)\dot{x}^j(t)\right]^{1/2}dt$$

denote the length of the curve $\gamma_\alpha(t), a\le t\le b$. From (30), we have

$$0<\lambda\|\varphi(q_{n,\alpha})-x_{n,\alpha}\|\le L(\gamma_\alpha).$$

Since the curve $\gamma_\alpha(t)$ was arbitrarily chosen for each α, we have

$$0<\lambda\|\varphi(q_{n,\alpha})-x_{n,\alpha}\|\le d(p_{n,\alpha},q_{n,\alpha})\quad\text{for all }\alpha=1,\dots,K_n^m.$$

So

$$\frac{a\lambda}{2K_n}\le\inf_\alpha d(p_{n,\alpha},q_{n,\alpha}). \tag{31}$$

Let $\bar{\mathcal{B}}(p,\delta)=\{q\in M|d(p,q)\le\delta\}$. Therefore it follows from (31) that for all α, we have $\bar{\mathcal{B}}(p_{n,\alpha},a\lambda/2K_n)\subset Q_{n,\alpha}$. For each n set $a_n=a\lambda/2K_n$.

We choose $f_0\in\Sigma_m$ such that $f_0(x)=C_0>0$ on Q and take a bounded function $H_{n,\alpha}$ defined on $T_{p_{n,\alpha}}(M)$ such that

$$\int_0^1\int_{\mathcal{E}_{p_{n,\alpha}}}H^2(t\xi)t^{m-1}dtd\mu_p(\xi)=c, \tag{32}$$

where c is independent of n and α. Define $g_{n,\alpha}$ on Q by

$$g_{n,\alpha}(x)=C_1a_n^\beta H_{n,\alpha}(a_n^{-1}\exp_{p_{n,\alpha}}^{-1}(x)).$$

We may assume that for all α and n it has its support in $\bar{\mathcal{B}}(p_{n,\alpha},a_n)$ and

$$\int_M g_{n,\alpha}(x)dV_g(x)=0.$$

We consider a sequence $\epsilon_n = \{\epsilon_{n,\alpha}\}$ taking values $\{0, 1\}$ and then set $f_n^{\epsilon_n}(x) = f_0(x) + \sum_{\alpha=1}^{K_n^m} \epsilon_{n,\alpha} g_{n,\alpha}(x)$. By suitable choice of C_1, f_0 and $H_{n,\alpha}$, the function $f_n^{\epsilon_n}$ will be contained in Σ_m. Let $\epsilon(\alpha, 0) = (\epsilon_1, \epsilon_2, \ldots, \epsilon_{\alpha-1}, 0, \epsilon_{\alpha+1}, \ldots, \epsilon_{K_n^m})$ and $\epsilon(\alpha, 1) = (\epsilon_1, \epsilon_2, \ldots, \epsilon_{\alpha-1}, 1, \epsilon_{\alpha+1}, \ldots, \epsilon_{K_n^m})$. Denote by $\Lambda_{n,\epsilon,\alpha}(1, 0)$ the Radon–Nikodym derivative of $P_{f_n^{\epsilon(\alpha,1)}}$ with respect to $P_{f_n^{\epsilon(\alpha,0)}}$. Let

$$c_n = \{4 \sup_{\epsilon,\alpha} E_{f_n^{\epsilon(\alpha,0)}} \Lambda_{n,\epsilon,\alpha}(1, 0)^2\}^{-1}.$$

Suppose $\liminf_{n\to\infty} c_n = c(0 < c < \infty)$, from the standard arguments to get the lower rate of convergence for the nonparametric density in the case of the Euclidean space (see, for example, Section 3 in [6]), we obtain the inequality

$$\liminf_{n\to\infty} \inf_{\hat{f}_n} \sup_{f\in\Sigma_m(M)} P_f\left\{L(\hat{f}_n, f) \geq \left(\frac{c}{c+1}\right)^{1/2} K_n^{m/2}\delta_n\right\} \geq c^2. \tag{33}$$

Now we consider the sequence $\{c_n\}$. Since $f_0(\omega) = C_0 > 0$ on Q and $\mathrm{supp}(g_{n,\alpha}) = \bar{\mathcal{B}}(p_{n,\alpha}, a_n)$, we have

$$\begin{aligned} E_{f_n^{\epsilon(\alpha,0)}} \Lambda_{n,\epsilon,\alpha}(1, 0)^2 &= \left\{\int_M (f_n^{\epsilon(\alpha,1)})^2(x)(f_n^{\epsilon(\alpha,0)})^{-1}(x)dV_g(x)\right\}^n \\ &= \left\{1 + \int_M (f_n^{\epsilon(\alpha,1)}(x) - f_n^{\epsilon(\alpha,0)})^2(x)(f_n^{\epsilon(\alpha,0)})^{-1}(x)dV_g(x)\right\}^n \\ &= \left\{1 + \frac{C_1^2}{C_0} a_n^{2(\beta+\gamma)} \int_{\bar{\mathcal{B}}(p_{n,\alpha},a_n)} H_{n,\alpha}^2\left(\frac{1}{a_n}\exp_{p_{n,\alpha}}^{-1}(x)\right)dV_g(x)\right\}^n. \end{aligned} \tag{34}$$

Introducing the *geodesic spherical coordinates* on $\bar{\mathcal{B}}(p_{n,\alpha}, a_n)$ as used in Section 3, then the last term in (34) can be written as

$$\left\{1 + \left(\frac{C_1^2}{C_0}\right) a_n^{2\beta+2\gamma+1} \int_0^1 \int_{\mathcal{E}_{p_{n,\alpha}}} H_{n,\alpha}^2(t\xi)\sqrt{\mathbf{g}(a_n t; \xi)}d\mu_{p_{n,\alpha}}(\xi)dt\right\}^n. \tag{35}$$

If we let $\xi = \sum_{i=1}^n \xi e_i$, from Lemma 3, it follows that

$$\sqrt{\mathbf{g}(t; \xi)} = t^{m-1} + t^{m-1}\sum_{k=1}^{\infty} \frac{t^k}{k!} \sum_{j_1,\ldots,j_k=1}^{m} \Omega_{j_1,\ldots,j_k}(p_{n,\alpha})\xi^{j_1}\ldots\xi^{j_k}. \tag{36}$$

Since the Riemannian curvature tensor is a C^∞ covariant tensor field and $\Omega_{j_1,\ldots,j_k}(p_{n,\alpha})$, the coefficients of $\xi^{j_1}\cdots\xi^{j_k}$ for k even, can be expressed in terms of the curvature tensor and its covariant derivatives, then on the compact set Q, each term in (36) is

$$\Omega_{j_1,\ldots,j_k}(p)\xi^{j_1}\cdots\xi^{j_k} \leq A$$

for all $p \in Q$ and $\xi \in \mathcal{E}_p$, where A is independent of n. Therefore, from this fact and (32), we see that

$$E_{f_n^{\epsilon(\alpha,0)}} \Lambda_{n,\epsilon,\alpha}(1, 0)^2 = \left\{1 + \left(\frac{C_1^2}{C_0}\right) a_n^{2(\beta+\gamma)+m} \int_0^1 \int_{\mathcal{E}_{p_{n,\alpha}}} H_{n,\alpha}^2(t\xi)t^{m-1}dtd\mu_{p_{n,\alpha}}(\xi) + o(a_n^{2(\beta+\gamma)+m})\right\}^n. \tag{37}$$

If we take $a_n = n^{-(1/2\beta+2\gamma+m)}$, then $\lim_{n\to\infty} c_n = c < \infty$. Now we compute δ_n. As we have just seen,

$$\begin{aligned} L(0, g_{n,\alpha}) &= \left\{\int_{\bar{\mathcal{B}}(p_{n,\alpha},a_n)} |g_{n,\alpha}|^2(x)dV_g(x)\right\}^{1/2} \\ &= C_1 a_n^{\beta+\gamma}\left\{\int_{\bar{\mathcal{B}}(p_{n,\alpha},a_n)} H_{n,\alpha}^2\left(\frac{1}{a_n}\exp_{p_{n,\alpha}}^{-1}(x)\right)dV_g(x)\right\}^{1/2} \\ &= C_1 a_n^{\beta+\gamma+(m/2)}\left\{\int_0^1 \int_{\mathcal{E}_{p_{n,\alpha}}} H_{n,\alpha}^2(t\xi)t^{m-1}dtd\mu_{p_{n,\alpha}}(\xi) + o(a_n)\right\}^{1/2} \\ &= Ca_n^{\beta+\gamma+(m/2)} + o_\alpha(a_n^{\beta+\gamma+(m/2)}), \end{aligned}$$

where $o_\alpha(a_n^{\beta+\gamma+(m/2)})$ may depend on α. But it is obvious that $\inf_\alpha o_\alpha(a_n^{\beta+\gamma+(m/2)}) = o(a_n^{\beta+\gamma+(m/2)})$. From this it follows that

$$K_n^{m/2}\delta_n = Cn^{-(\beta+\gamma)/(2\beta+2\gamma+m)}(1+o(1)),$$

and hence from (33), $n^{-(\beta+\gamma)/(2\beta+2\gamma+m)}$ is a lower rate of convergence. This result immediately gives lower bounds to convergence rates in $L^2(M)$ metrics:

$$\liminf_{n\to\infty} \inf_{\hat{f}_n} \sup_{f\in\Sigma_m(M)} \mathbb{E}\left(\|f(p) - \hat{f}_n(p)\|_{L_2}^2 n^{-2(\beta+\gamma)/(2\beta+2\gamma+m)} \right) > 0. \quad \square$$

References

[1] Z.D. Bai, C.R. Rao, L.C. Zaho, Kernel estimators of density function of directional data, Journal of Multivariate Analysis 27 (1988) 24–39.
[2] I. Chavel, Eigenvalues in Riemannian Geometry, Academic Press, New York, 1984.
[3] L. Devroye, L. Györfi, Nonparametric Density Estimation: The L_1-View, Wiley, New York, 1985.
[4] R.A. Fisher, Dispersion on a sphere, Proceedings of the Royal Society of London, Series A 217 (1953) 295–305.
[5] A. Gray, L. Vanhecke, Riemannian geometry as determined by the volumes of small geodesic balls, Acta Mathematica, Uppsala 142 (1979) 157–198.
[6] P. Hall, On convergence rates in nonparametric problems, International Statistical Review 57 (1) (1989) 45–58.
[7] P. Hall, G. Watson, J. Cabrera, Kernel density estimation with spherical data, Biometrika 74 (1987) 751–762.
[8] D. Healy, H. Hendriks, P. Kim, Spherical deconvolution, Journal of Multivariate Analysis 67 (1998) 1–22.
[9] H. Hendriks, Nonparametric estimation of a probability density on a Riemannian manifold using Fourier expansions, Annals of Statistics 18 (1990) 832–849.
[10] H. Hendriks, Application of fast spherical Fourier transform to density estimation, Journal of Multivariate Analysis 84 (2003) 209–221.
[11] P.E. Jupp, K.V. Mardia, A unified view of the theory of directional statistics; 1975–1988, International Statistical Review 57 (1989) 261–294.
[12] P.E. Jupp, K.V. Mardia, Directional Statistics, J. Wiley, Chichester, 2000.
[13] K.V. Mardia, Statistics of Directional Data, Academic Press, New York, 1972.
[14] H.S. Park, Mean distance for Brownian motion on Riemannian manifold and density estimation on Riemannian manifold, Ph.D. Dissertation, Hallym University, 2004.
[15] H.S. Park, Asymptotic behavior of kernel density estimation from geometry viewpoint, Communications in Statistics—Theory and Methods 41 (18) (2012).
[16] B. Pelletier, Kernel density estimation on Riemannian manifolds, Statistics & Probability Letters 73 (2005) 297–304.
[17] G.S. Watson, Statistics on Spheres, Wiley, New York, 1983.