PRÉ-PUBLICAÇÕES DMUC

DMUC Preprint 26-36, 15 Jul 2026

On a wrapped kernel class of density estimators for circular data

Carlos Tenreiro*

CMUC, Department of Mathematics, University of Coimbra, Coimbra, Portugal

15 July 2026

Abstract

The wrapped kernel density estimator studied in this paper is a type of delta sequence density estimator whose delta function sequence is obtained by wrapping a scaled version of a probability density function defined on the real line. An asymptotic expansion for the mean integrated squared error (MISE) of this wrapped kernel density estimator is derived allowing for a direct comparison with other density estimators from the literature, such as the standard kernel density estimator and the Parzen-Rosenblatt type estimator. The wrapped Cauchy kernel density estimator receives particular attention.

KEYWORDS: Circular data; Density estimation; Mean integrated squared error.

AMS 2020 SUBJECT CLASSIFICATIONS: 62H11, 62G07

*ORCID ID: 0000-0002-5495-6644 E-mail: tenreiro@mat.uc.pt URL: http://www.mat.uc.pt/∼tenreiro/ Postal address: Departamento de Matemática da FCTUC, Largo D. Dinis, 3000-143 Coimbra, Portugal.

[v1] Wed, 15 Jul 2026

1 Introduction

Given an independent and identically distributed sample of angles $X_1, \ldots, X_n \in [0, 2\pi)$ from some absolutely continuous random variable X with unknown probability density function f, a new nonparametric density estimator for circular data, called Parzen-Rosenblatt type estimator, was introduced and studied in Tenreiro (2022, 2024, 2025). This estimator, which is close in spirit to the Parzen-Rosenblatt estimator for linear data (see Rosenblatt, 1956, and Parzen, 1962), has a simple structure being defined, for $\theta \in [0, 2\pi)$, by

$$\hat{f}_{\mathrm{PR}}(\theta; K, h) = \frac{d_h(K)}{n} \sum_{i=1}^{n} K_h(\theta - X_i), \tag{1}$$

where $h = h_n$ is a sequence of strictly positive real numbers converging to zero as n tends to infinity, K_h is a real-valued periodic function on $\mathbb{R}$, with period 2π, such that

$$K_h(\theta) = h^{-1} K(h^{-1}\theta), \text{ for } \theta \in [-\pi, \pi),$$

where K is a kernel on $\mathbb{R}$, that is, a nonnegative, bounded and integrable real-valued function on $\mathbb{R}$, and $d_h(K)$ is a normalizing constant depending on the kernel K and the bandwidth h which is chosen so that $\hat{f}_{\mathrm{PR}}$ integrates to unity.

The previous estimator is closely related to the so-called standard kernel density estimator for circular data, initially studied in Beran (1979), Hall, Watson, and Cabrera (1987), and Bai, Rao, and Zhao (1988), and more recently by García-Portugués (2013) and García-Portugués et al. (2013). It is defined, for $\theta \in [0, 2\pi)$, by

$$\hat{f}_{\mathrm{ST}}(\theta; L, g) = \frac{c_g(L)}{n} \sum_{i=1}^{n} L\big(g^{-2}(1 - \cos(\theta - X_i))\big), \tag{2}$$

where $L : [0, \infty) \to \mathbb{R}$ is a nonnegative and bounded function satisfying some additional conditions, $g = g_n$ is a sequence of strictly positive numbers such that $g \to 0$, as $n \to \infty$, and $c_g(L)$, depending on the kernel L and the bandwidth g, is chosen so that $\hat{f}_{\mathrm{ST}}$ integrates to unity. Under some additional assumptions on the kernel K, the estimator $\hat{f}_{\mathrm{ST}}$, with kernel $L(t) = K(\sqrt{t})$ and bandwidth $g = h/\sqrt{2}$, is asymptotically equivalent to $\hat{f}_{\mathrm{PR}}$ in the sense of the mean integrated squared error, for a large class of density functions f (see Tenreiro, 2022, Theorems 3.3 and 3.4, pp. 385, 387). Namely, when K is the normal or gaussian kernel, $K(u) = \exp(-u^2)$, $u \in \mathbb{R}$, $L(t) = K(\sqrt{t}) = e^{-t}$, $t \geq 0$, is the so-called von Mises kernel, and the density estimator (2), denoted henceforth by $\hat{f}_{\mathrm{vM}}$, is the estimator considered in Taylor (2008) and Oliveira, Crujeiras, and Rodríguez-Casal (2012). When $g = h/\sqrt{2}$ the estimator $\hat{f}_{\mathrm{vM}}$ is a combination of circular or von Mises densities with mean directions X_i and concentration parameters equal to $2h^{-2}$ as it takes the form

$$\hat{f}_{\mathrm{vM}}(\theta; h) = \frac{1}{n} \sum_{i=1}^{n} f_{\mathrm{vM}}(\theta; X_i, 2h^{-2}), \tag{3}$$

where

$$f_{\mathrm{vM}}(\theta;\mu,\kappa)=\frac{1}{2\pi I_0(\kappa)}\exp\big(\kappa\cos(\theta-\mu)\big),$$

is the von Mises density with mean direction $\mu\in[0,2\pi)$ and concentration parameter $\kappa\geq 0$, and $I_r(\nu)$ is, for $\nu\geq 0$ and $r\geq 0$, the modified Bessel function of order r defined by

$$I_r(\nu)=\frac{1}{2\pi}\int_0^{2\pi}\cos(r\theta)\exp(\nu\cos\theta)d\theta.$$

In this work we are especially interested in an alternative estimator of f, called wrapped kernel density estimator, which is defined, for $\theta\in[0,2\pi)$, by

$$\hat{f}_{\mathrm{WK}}(\theta;K,h)=\frac{1}{n}\sum_{i=1}^{n}K_n^w\big(\theta-X_i\big), \tag{4}$$

where K_n^w is the periodic extension to $\mathbb{R}$ of the circular density function over the interval $[-\pi,\pi)$ obtained by wrapping the probability density function on the real line defined by

$$x\mapsto h^{-1}K\big(h^{-1}x\big)\left(\int_{-\infty}^{\infty}K(u)du\right)^{-1},\ x\in\mathbb{R}, \tag{5}$$

where, as defined before, K is a kernel and $h=h_n$ is a sequence of strictly positive real numbers converging to zero as n tends to infinity. Examples of such estimators, frequently discussed in the literature, are those obtained by selecting the normal kernel for K, in which case

$$K_n^w(y)=\frac{1}{2\pi}\left(1+2\sum_{k=1}^{\infty}e^{-h^2k^2/4}\cos(ky)\right),\ y\in\mathbb{R}, \tag{6}$$

is called wrapped normal kernel (see Mardia and Jupp, 2000, p. 50), or by taking the Cauchy kernel $K(u)=(1+u^2)^{-1}$, $u\in\mathbb{R}$, where

$$K_n^w(y)=\frac{1}{2\pi}\left(1+2\sum_{k=1}^{\infty}e^{-hk}\cos(ky)\right),\ y\in\mathbb{R}, \tag{7}$$

is called wrapped Cauchy kernel or *Poisson kernel* (see Mardia and Jupp, 2000, p. 51; Bhatia, 2005, pp. 20–23). This type of density estimators is considered in Di Marzio et al. (2009, 2011), Tsuruta and Sagae (2017), and Tenreiro (2022), and more recently in Ameijeiras-Alonso (2024) where the density estimators based on the wrapped normal and wrapped Cauchy kernels are used to build estimators for the derivatives of f.

When K is equal to zero outside of a finite interval, K_n^w defined by (5) is almost everywhere equal to $d_h(K)K_h$, for n large enough, and the wrapped kernel density estimator based on K_n^w coincides with the Parzen-Rosenblatt type estimator with kernel K. Naturally, this does not apply to every kernel K. Even so, we will establish in this paper that, under some additional assumptions on the kernel K, $\hat{f}_{\mathrm{WK}}$ is asymptotically equivalent in the sense of the mean integrated squared error (MISE) to both estimators $\hat{f}_{\mathrm{PR}}$ and $\hat{f}_{\mathrm{ST}}$, for a large class of probability density functions f.

This is the primary goal of this work. This property, which is valid when K is the normal kernel, enables us to conclude that the estimator $\hat{f}_{\mathrm{WK}}$ with K_n^w given by (6) is asymptotically equivalent in the sense of the MISE to both estimators $\hat{f}_{\mathrm{PR}}$ and $\hat{f}_{\mathrm{vM}}$. To the best of our knowledge, this is the first time that this fact is noticed in the literature.

The remainder of this paper is organised as follows. In Section 2 we derive an asymptotic expansion for the MISE of a delta sequence estimator of f, a general class of density estimators considered in Tenreiro (2022) and closely connected with the class of density estimators studied in Di Marzio, Panzera, and Taylor (2009), that includes, under some conditions on the kernel, the estimators $\hat{f}_{\mathrm{PR}}$, $\hat{f}_{\mathrm{ST}}$ and $\hat{f}_{\mathrm{WK}}$. This general MISE expansion is used in Sections 2 and 3 to obtain asymptotic expansions for the MISE of the previous density estimators, which allows us to conclude that, under some assumptions on the kernel function, the estimators $\hat{f}_{\mathrm{PR}}$, $\hat{f}_{\mathrm{ST}}$ and $\hat{f}_{\mathrm{WK}}$ are asymptotically equivalent in the sense of the MISE. The interesting case of the Cauchy kernel, a kernel not fulfilling the mentioned assumptions, is studied in detail in Section 4. We will see that $\hat{f}_{\mathrm{WK}}$ and $\hat{f}_{\mathrm{ST}}$ are asymptotically equivalent in the sense of the MISE when K is the Cauchy kernel but, somewhat surprisingly, both estimators are overcome by the Parzen-Rosenblatt type estimator $\hat{f}_{\mathrm{PR}}$. Taking into account the asymptotic nature of the previous results, in Section 5 we present a comparative finite-sample analysis of the performance of the estimators under consideration when K is taken to be the normal kernel or the Cauchy kernel. Finally, we present some overall conclusions in Section 6, and all the proofs are gathered in Section 7. The numerical results and plots shown in this article were carried out using the R software (R Core Team, 2024).

2 A MISE expansion for delta sequence estimators

In this section we derive an asymptotic expansion for the mean integrated squared error (MISE) of the delta sequence estimator of f defined by

$$\hat{f}_n(\theta) = \frac{1}{n} \sum_{i=1}^{n} \delta_n(\theta - X_i), \ \theta \in [0, 2\pi), \tag{8}$$

where $\delta_n : \mathbb{R} \to [0, +\infty)$, for $n \in \mathbb{N}$, is a delta function sequence as defined in Tenreiro (2022, p. 377), that is, a sequence of periodic functions with period 2π satisfying the conditions

(Δ.1) $\displaystyle\int_{-\pi}^{\pi} \delta_n(y)dy = 1$, for all n;

(Δ.2) $\displaystyle\sup_{\lambda<|y|\leq\pi} \delta_n(y) \to 0$, as $n \to +\infty$, for all $0 < \lambda < \pi$;

(Δ.3) $\displaystyle\int_{-\pi}^{\pi} \delta_n(y)^2 dy < \infty$, for all n.

The reader is referred to Watson and Leadbetter (1964) for the concept of delta function sequence in the context of data on the real line. As discussed in Tenreiro (2022, pp. 377–378), if the delta function sequence is symmetric, that is, $\delta_n(-y) = \delta_n(y)$, for all $y \in \mathbb{R}$ and $n \in \mathbb{N}$, a condition we are not assuming here, the previous class of density estimators is closely related with the subclass of nonnegative circular estimators (of second sin-order) considered in Di Marzio et al. (2009) (see also Di Marzio, Panzera, and Taylor, 2011, Definitions 1 and 2, pp. 2157–2158). Conditions (ii) and (iii) of Definition 1 in Di Marzio et al. (2009, p. 2067) are trivially fulfilled under assumptions (Δ.1) and (Δ.2), being this last condition stronger than condition (iii). Although the results presented in this paper may be obtained under this weaker condition, we decided to adopt the definition of delta function sequence given in Tenreiro (2022, p. 377), where condition (Δ.2) plays an important role for establishing an uniform asymptotic expansion for the variance of the delta sequence estimator for all density f continuous on $[0, 2\pi]$. However, under the previous set of conditions the delta function sequence does not necessarily admit a pointwise convergent Fourier series representation as assumed in condition (i) of Definition 1 in Di Marzio et al. (2009). Nevertheless, from assumptions (Δ.1) and (Δ.3) we know that the delta function sequence admits the $L^2([-\pi, \pi])$ Fourier series representation

$$\delta_n(y) = \frac{1}{2\pi}\left(1 + 2\sum_{k=1}^{\infty}\left\{a_k(\delta_n)\cos(ky) + b_k(\delta_n)\sin(ky)\right\}\right), \tag{9}$$

where $a_k(\delta_n) = \int_{-\pi}^{\pi}\delta_n(y)\cos(ky)dy$ and $b_k(\delta_n) = \int_{-\pi}^{\pi}\delta_n(y)\sin(ky)dy$, are, up to a constant, the real Fourier coefficients of δ_n (Butzer and Nessel, 1971, Definition 1.2.1 and Proposition 4.2.3, pp. 40, 175).

The derivation of asymptotic expansions for the mean integrated squared error of a delta sequence estimator of f was studied in Di Marzio et al. (2009) and, more recently, in Tenreiro (2022). Although Theorem 3.1 in Tenreiro (2022) is fully adequate for deriving MISE expansions for the classes of density estimators (1) and (2), the same cannot be said for the class of density estimators of the form (4), as can be concluded from the arguments presented in Tenreiro (2022, p. 384). For this reason, the following result presents a generalisation of Theorem 3.1 in Tenreiro (2022) that will be especially useful in Section 3. Moreover, it also improves and corrects Theorem 3.1 in Di Marzio et al. (2009) as discussed below.

Throughout this paper, we denote by f the probability density function of the circular random variable X as well as its periodic extension to the real line. Moreover, we denote by ϕ_{δ_n} and ϕ_f the Fourier transforms of δ_n and f given respectively by $\phi_{\delta_n}(k) = \int_{-\pi}^{\pi} e^{\mathrm{i}ky}\delta_n(y)dy = a_k(\delta_n) + \mathrm{i}b_k(\delta)$, and $\phi_f(k) = \int_0^{2\pi} e^{\mathrm{i}ky}f(y)dy$, for $k \in \mathbb{Z}$.

Theorem 1. *Under assumptions (Δ.1)–(Δ.3), let us assume that for some $\nu > 0$ and $C > 0$, there exist a sequence of strictly positive real numbers $\gamma(\delta_n)$ and $n_0 \in \mathbb{N}$, such that for all $k \in \mathbb{N}$, we have*

$$|\phi_{\delta_n}(k) - 1| = \psi_k k^\nu \gamma(\delta_n)(1 + o(1)), \ \textit{as}\ n \to \infty,$$

and

$$|\phi_{\delta_n}(k) - 1| \leq C k^\nu \gamma(\delta_n), \ \textit{for}\ n \geq n_0,$$

where $\psi_k \in \mathbb{R}$, $k \in \mathbb{N}$, and $\psi_k > 0$, for some $k \in \mathbb{N}$. If $f^{(p-1)}$ is absolutely continuous on $[0, 2\pi]$ with $f^{(p)} \in L^2([0, 2\pi))$, where p is the smallest integer greater than or equal to ν, then

$$\mathrm{MISE}(f; \hat{f}_n, n) = n^{-1}\alpha(\delta_n) + \gamma(\delta_n)^2 \frac{1}{\pi} \sum_{k=1}^{\infty} \psi_k^2 k^{2\nu} |\phi_f(k)|^2 + o\big(n^{-1}\alpha(\delta_n) + \gamma(\delta_n)^2\big),$$

where

$$\alpha(\delta_n) = \int_{-\pi}^{\pi} \delta_n(y)^2 dy = \frac{1}{2\pi}\bigg(1 + 2\sum_{k=1}^{\infty} |\phi_{\delta_n}(k)|^2\bigg).$$

Moreover, if $p = \nu$ and $\psi_k = 1$, for all $k \in \mathbb{N}$, then

$$\mathrm{MISE}(f; \hat{f}_n, n) = n^{-1}\alpha(\delta_n) + \gamma(\delta_n)^2 \theta_p(f) + o\big(n^{-1}\alpha(\delta_n) + \gamma(\delta_n)^2\big), \tag{10}$$

where

$$\theta_p(f) = \int_0^{2\pi} f^{(p)}(\theta)^2 d\theta = \frac{1}{\pi} \sum_{k=1}^{\infty} k^{2p} |\phi_f(k)|^2.$$

Under the previous conditions, it can be proved that the sequence $\gamma(\delta_n)$ converges to zero as n tends to infinity. When δ_n is a delta function sequence with $\int_{-\pi}^{\pi} y\delta_n(y)dy = 0$ and $\int_{-\pi}^{\pi} |y|^3\delta_n(y)dy = o(\beta(\delta_n))$, where $\beta(\delta_n) := \int_{-\pi}^{\pi} y^2\delta_n(y)dy = 4\sum_{k=1}^{\infty}(-1)^k(a_k(\delta_n) - 1)/k^2$ (see assumptions (Δ.4)–(Δ.5) in Tenreiro, 2022, p. 382), then the assumptions of Theorem 1 are fulfilled with $\gamma(\delta_n) = \beta(\delta_n)/2$, $\nu = p = 2$ and $\psi_k = 1$, for all $k \in \mathbb{N}$. In this case, but under weaker conditions on the underlying density f, the expansion (10) for $\mathrm{MISE}(f; \hat{f}_n, n)$ agrees with the one given in Tenreiro (2022, Theorem 3.1, p. 383). Moreover, when the delta function sequence δ_n is symmetric, from the $L^2([-\pi, \pi])$ representation (9) we have $\alpha(\delta_n) = (2\pi)^{-1}\big(1 + 2\sum_{k=1}^{\infty} a_k(\delta_n)^2\big)$, and $\beta(\delta_n) = (1 - a_2(\delta_n)) + o(\beta(\delta_n))$ (see Tenreiro, 2022, p. 404). Therefore, the leading terms of expansion (10) agree with the first two terms of the MISE expansion presented in Di Marzio et al. (2009, Theorem 1, p. 2068). However, from Theorem 1 we can further conclude that the sum of such leading terms is asymptotically equivalent to the MISE of $\hat{f}_n$, a fact not established in Di Marzio et al. (2009). In this case, the previous result can be viewed as an improvement of Theorem 3.1 in Di Marzio et al. (2009). The delta sequence estimator based on the wrapped Cauchy kernel (7) gives us an example where the previous result corrects Theorem 3.1 in Di Marzio et al. (2009). This delta sequence function fulfils the conditions of the Di Marzio et al.'s result but the corresponding MISE expansion stated there is not correct as we can confirm from Theorem 3.2 in Tenreiro (2022, p. 384).

We finish this section with a first application of Theorem 1 to derive asymptotic expansions for the mean integrated squared errors of the estimators $\hat{f}_{\mathrm{PR}}$ and $\hat{f}_{\mathrm{ST}}$ defined by equations (1) and (2), respectively. These expansions, which have already been obtained in Tenreiro (2022, Theorems 3.3 and 3.4, pp. 385, 387), are presented below in a unified form, assuming that $\hat{f}_{\mathrm{ST}}$ is defined with kernel $L(t) = K(\sqrt{t})$ and bandwidth $g = h/\sqrt{2}$. Based on them, we conclude that the estimators $\hat{f}_{\mathrm{PR}}$ and $\hat{f}_{\mathrm{ST}}$ are asymptotically equivalent in the sense of the MISE, that is, $\mathrm{MISE}(f; \hat{f}_{\mathrm{PR}}, n) = \mathrm{MISE}(f; \hat{f}_{\mathrm{ST}}, n)\,(1 + o(1))$, for all density f with f' absolutely continuous on $[0, 2\pi]$ and $f'' \in L^2([0, 2\pi))$.

Theorem 2. *Let K be a nonnegative, bounded and symmetric function such that $\lim_{u\to+\infty} u^3 K(u) = 0$, $\int_{-\infty}^{\infty} K(u)du > 0$, and $\int_{-\infty}^{\infty} u^2 K(u)du < \infty$, and assume that f' is absolutely continuous on $[0, 2\pi]$ with $f'' \in L^2([0, 2\pi))$. Let us denote by $\hat{f}$ one of the estimators $\hat{f}_{\mathrm{PR}}$ or $\hat{f}_{\mathrm{ST}}$, where $\hat{f}_{\mathrm{ST}}$ is defined with $L(t) = K(\sqrt{t})$ and $g = h/\sqrt{2}$. If $h \to 0$, as $n \to +\infty$, we have*

$$\mathrm{MISE}(f; \hat{f}, n) = \frac{1}{nh}\boldsymbol{d}_1(K) + \frac{h^4}{4}\boldsymbol{d}_2(K)\,\theta_2(f) + o\left(\frac{1}{nh} + h^4\right),$$

with

$$\boldsymbol{d}_1(K) = \int_{-\infty}^{\infty} K(u)^2 du \left(\int_{-\infty}^{\infty} K(u)du\right)^{-2}$$

and

$$\boldsymbol{d}_2(K) = \left(\int_{-\infty}^{\infty} u^2 K(u)du\right)^2 \left(\int_{-\infty}^{\infty} K(u)du\right)^{-2}.$$

From the proof of this result given in Section 7, it follows that the technical condition $\lim_{u\to\infty} u^3 K(u) = 0$ can be replaced by the weaker assumption $\lim_{u\to\infty} uK(u) = 0$ in the case of the estimator $\hat{f}_{\mathrm{PR}}$.

3 A MISE expansion for $\hat{f}_{\mathrm{WK}}$

As stated in the introductory section, we denote by K_n^w the periodic extension to $\mathbb{R}$ of the circular probability density function over the interval $[-\pi, \pi)$ obtained by wrapping the probability density function on the real line defined by (5). Thus, K_n^w can be written as

$$K_n^w(y) = \sum_{k\in\mathbb{Z}} h^{-1} K\big(h^{-1}(y + 2k\pi)\big) \left(\int_{-\infty}^{\infty} K(u)du\right)^{-1}, \tag{11}$$

for almost all $y \in [-\pi, \pi)$ (see Mardia and Jupp, 2000, p. 48). This fact is used in the following result to prove that, under some general conditions on K, K_n^w is, for n large enough, a delta sequence of functions.

Theorem 3. *If K is a nonnegative, bounded, and symmetric function such that $\int_{-\infty}^{\infty} K(u)du > 0$ and $\limsup_{u\to+\infty} u^2 K(u) < \infty$, and the sequence $h = h_n$ converges to zero as n tends to infinity, then K_n^w is, for n large enough, a delta sequence of functions.*

Using the Theorem 1 we can now obtain an asymptotic expansion for the mean integrated squared error of the wrapped kernel density estimator $\hat{f}_{\mathrm{WK}}$ defined by (4). Note that, from the definition of K_n^w we know that

$$\phi_{K_n^w}(k) = \phi_K(kh) \left(\int_{-\infty}^{\infty} K(u)du\right)^{-1}, \; k \in \mathbb{Z}, \tag{12}$$

where $\phi_{K_n^w}$ and ϕ_K are the Fourier transforms of K_n^w and K, respectively, given by $\phi_{K_n^w}(k) = \int_{-\pi}^{\pi} e^{\,\mathrm{i}\,ky} K_n^w(y)dy$, for $k \in \mathbb{Z}$, and $\phi_K(t) = \int_{-\infty}^{\infty} e^{\,\mathrm{i}\,ty} K(y)dy$, for $t \in \mathbb{R}$ (see Mardia and Jupp, 2000, p. 48).

Theorem 4. *Let K be a nonnegative, bounded and symmetric function such that $\int_{-\infty}^{\infty} K(u)du > 0$, $\limsup_{u\to+\infty} u^2K(u) < \infty$, $\int_{-\infty}^{\infty} u^2K(u)du < \infty$, and $\limsup_{t\to+\infty} t^{1+\epsilon}|\phi_K(t)|^2 < \infty$, for some $\epsilon > 0$, and assume that f' is absolutely continuous on $[0, 2\pi]$ with $f'' \in L^2([0, 2\pi))$. If h is such that $h \to 0$, as $n \to +\infty$, then*

$$\mathrm{MISE}(f; \hat{f}_{\mathrm{WK}}, n) = \frac{1}{nh}\boldsymbol{d}_1(K) + \frac{h^4}{4}\boldsymbol{d}_2(K)\,\theta_2(f) + o\bigg(\frac{1}{nh} + h^4\bigg),$$

where $\boldsymbol{d}_1(K)$ and $\boldsymbol{d}_2(K)$ are defined in Theorem 2.

Under the previous assumptions on the kernel K, from Theorems 2 and 4 we conclude that $\hat{f}_{\mathrm{WK}}$ is asymptotically equivalent in the sense of the MISE to both density estimators $\hat{f}_{\mathrm{PR}}$ and $\hat{f}_{\mathrm{ST}}$, for all density f with f' absolutely continuous on $[0, 2\pi]$ and $f'' \in L^2([0, 2\pi))$, where $\hat{f}_{\mathrm{ST}}$ is defined with $L(t) = K(\sqrt{t})$ and $g = h/\sqrt{2}$. These assumptions on the kernel K are fulfilled by the normal kernel in which case K_n^w is given by (6) and $\hat{f}_{\mathrm{ST}}$ is the estimator $\hat{f}_{\mathrm{vM}}$ defined by (3).

4 The Cauchy kernel case

The assumptions of Theorem 4 are not met by the Cauchy kernel, as its moment of first order does not exist. In this case, the wrapped kernel estimator $\hat{f}_{\mathrm{WK}}$ is studied in Tenreiro (2022) where it is established that its mean integrated squared error admits the expansion

$$\mathrm{MISE}(f; \hat{f}_{\mathrm{WK}}, n) = \frac{1}{2\pi}\frac{1}{nh} + h^2\,\theta_1(f) + o\bigg(\frac{1}{nh} + h^2\bigg),$$

whenever $h \to 0$, as $n \to +\infty$, and f is absolutely continuous on $[0, 2\pi]$ with $f' \in L^2([0, 2\pi))$ where

$$\theta_1(f) = \frac{1}{\pi}\sum_{k=1}^{\infty} k^2|\phi_f(k)|^2$$

(see Tenreiro, 2022, Theorem 3.2). From this and the corresponding expansions given in Theorems 2 and 4 for estimators $\hat{f}_{\mathrm{PR}}$, $\hat{f}_{\mathrm{ST}}$ and $\hat{f}_{\mathrm{WK}}$ when K is a kernel satisfying the conditions considered there, we conclude that the optimal convergence rate $n^{-2/3}$ we get for the MISE of the estimator $\hat{f}_{\mathrm{WK}}$ with K the Cauchy kernel is overcome by the corresponding optimal convergence rate $n^{-4/5}$ obtained for all the previous estimators when f' is absolutely continuous on $[0, 2\pi]$ with $f'' \in L^2([0, 2\pi))$. From the following result we will see that the same conclusion applies to the estimators $\hat{f}_{\mathrm{PR}}$ and $\hat{f}_{\mathrm{ST}}$ when K is the Cauchy kernel, the latter estimator being defined with $L(t) = K(\sqrt{t})$ and $g = h/\sqrt{2}$. Moreover, we conclude that $\hat{f}_{\mathrm{WK}}$ and $\hat{f}_{\mathrm{ST}}$ are asymptotically equivalent in the sense of the MISE, but, surprisingly, these estimators are overcome by the Parzen-Rosenblatt type estimator $\hat{f}_{\mathrm{PR}}$ when K is the Cauchy kernel.

Theorem 5. *Let K be the Cauchy kernel, and assume that f is absolutely continuous on $[0, 2\pi]$ with $f' \in L^2([0, 2\pi))$. If $h \to 0$, as $n \to +\infty$, then*

$$\mathrm{MISE}(f; \hat{f}_{\mathrm{PR}}, n) = \frac{1}{2\pi}\frac{1}{nh} + h^2\,\theta_{1,\psi}(f) + o\bigg(\frac{1}{nh} + h^2\bigg),$$

where

$$\theta_{1,\psi}(f) = \frac{1}{\pi}\sum_{k=1}^{\infty} \psi_k^2 k^2 |\phi_f(k)|^2,$$

with

$$\psi_k = \frac{2}{\pi}\int_0^{k\pi} (1-\cos y)y^{-2}\,dy,\ k\in\mathbb{N},$$

and

$$\mathrm{MISE}(f;\hat f_{\mathrm{ST}},n) = \frac{1}{2\pi}\frac{1}{nh} + h^2\,\theta_1(f) + o\Big(\frac{1}{nh}+h^2\Big),$$

where $\hat f_{\mathrm{ST}}$ *is defined with* $L(t)=K(\sqrt{t})$ *and* $g = h/\sqrt{2}$.

Assuming that f is not the uniform density in the circle, and that the asymptotic optimal bandwidths (that is, the bandwidths that minimise the leading terms of the MISE expansions) are used in each one of the estimators $\hat f_{\mathrm{PR}}$ and $\hat f$, where $\hat f$ refers to one of the estimators $\hat f_{\mathrm{ST}}$ or $\hat f_{\mathrm{WK}}$, we find

$$\mathrm{MISE}(f;\hat f_{\mathrm{PR}},n) = C_{f,\psi}\,\mathrm{MISE}(f;\hat f,n)\,(1+o(1)),$$

where

$$C_{f,\psi} = \left(\sum_{k=1}^{\infty}\psi_k^2 k^2|\phi_f(k)|^2\right)^{1/3}\left(\sum_{k=1}^{\infty} k^2|\phi_f(k)|^2\right)^{-1/3},$$

for all absolutely continuous density f with $f'\in L^2([0,2\pi))$. As $\psi_k<1$, for all $k\in\mathbb{N}$, we conclude that $\hat f_{\mathrm{PR}}$ is better than both $\hat f_{\mathrm{ST}}$ and $\hat f_{\mathrm{WK}}$ in the sense of the MISE for this class of densities.

A more informative measure for the asymptotic comparison of the previous (oracle) estimators is the efficiency of $\hat f$ with respect to $\hat f_{\mathrm{PR}}$ in the sense of the MISE, which is defined by

$$\mathrm{eff}(\hat f;\hat f_{\mathrm{PR}}) := \lim_{n\to+\infty}\frac{m(n)}{n} = C_{f,\psi}^{3/2}, \tag{13}$$

where the sample size $m(n)$ is such that $\mathrm{MISE}(f;\hat f_{\mathrm{PR}},m(n)) = \mathrm{MISE}(f;\hat f,n)(1+o(1))$. Taking into account that the sequence (ψ_k) is increasing, we conclude that $\mathrm{eff}(\hat f;\hat f_{\mathrm{PR}})\geq\psi_1 = 0.7737$. This lower bound is achieved by the cardioid density, which corresponds to the perturbation of the uniform density by a cosine function (see Mardia and Jupp, 2000, Section 5.5.5). Taking for f the von Mises density $f_{\mathrm{vM}}(\cdot;\mu,\kappa)$, where $|\phi_f(k)| = I_k(\kappa)/I_0(\kappa)$, $k\in\mathbb{N}$, the efficiencies $\mathrm{eff}(\hat f;\hat f_{\mathrm{PR}})$ for different values of the concentration parameter κ are reported in Table 1.

κ	0.1	0.5	1	5	10	20
$\mathrm{eff}(\hat f;\hat f_{\mathrm{PR}})$	0.7740	0.7819	0.8011	0.8984	0.9283	0.9493

Table 1: *Efficiencies of* $\hat f_{\mathrm{ST}}$ *and* $\hat f_{\mathrm{WK}}$ *with respect to* $\hat f_{\mathrm{PR}}$ *for several von Mises densities* $f_{\mathrm{vM}}(\cdot;\mu,\kappa)$*, when* K *is the Cauchy kernel.*

5 A finite-sample analysis

In this section we will consider the density estimators $\hat{f}_{\rm PR}$, $\hat{f}_{\rm ST}$, and $\hat{f}_{\rm WK}$ defined by (1), (2), and (4), respectively, when K is taken to be the normal kernel $K(u) = \exp(-u^2)$, $u \in \mathbb{R}$, or the Cauchy kernel $K(u) = (1+u^2)^{-1}$, $u \in \mathbb{R}$. In both cases we will assume that the standard density estimator $\hat{f}_{\rm ST}$ is defined with $L(t) = K(\sqrt{t})$ and $g = h/\sqrt{2}$. In Sections 2 and 3 we have proved that, under some general conditions on the kernel K, which are fulfilled by the normal kernel, the previous density estimators are asymptotically equivalent in the sense of the MISE, for a large class of density functions f. The asymptotic nature of this result raises the question of whether the previous estimators perform similarly for finite values of n. The same question arises when K is the Cauchy kernel. In this case, based on the asymptotic theory developed in Section 4, we expect to obtain estimators that are less efficient than those obtained for the normal kernel, with the estimators $\hat{f}_{\rm ST}$ and $\hat{f}_{\rm WK}$ being equally performing, but both less efficient than the Parzen-Rosenblatt type estimator $\hat{f}_{\rm PR}$.

In order to address these questions we consider the set of 19 circular distributions M2-M20 in Oliveira et al. (2012), which includes the von Mises (M2), the cardioid (M4), the wrapped normal (M3), the wrapped Cauchy (M5), and the wrapped skew-normal (M6) distributions, and mixtures of them (M7-M20). This set of distributions is very rich, containing densities with a wide variety of distribution features such as multimodality, skewness and/or peakedness. For a careful description of the different models and the plots of the corresponding circular densities see Oliveira et al. (2012, pp. 3901, 3902, 3907). Although we have used self-programmed code written in R and functions from the 'circular' package in R (Agostinelli and Lund, 2025) for generating data from the previous models, this can also be done by using the function rcircmix from the 'NPCirc' package (Oliveira et al., 2014).

For each one of the previous distributions and for different sample sizes n, we are interested in the distribution of the integrated squared error

$$\mathrm{ISE}(f;\hat{f},n) = \int_0^{2\pi} \{\hat{f}(\theta;\hat{h}) - f(\theta)\}^2 d\theta,$$

where $\hat{f}$ is one of the previous density estimators, and we take for $\hat{h}$ the corresponding Fourier series-based plug-in bandwidth selector following the approach in Tenreiro (2022, Sec. 4, pp. 389–393). In order to define such bandwidth selectors, we will denote by $\hat{\theta}_r$, $r = 1,2$, and $\hat{\theta}_{1,\psi}$, the Fourier series-based or projection estimators of $\theta_r(f)$, $r = 1,2$, and $\theta_{1,\psi}(f)$, given by $\hat{\theta}_r = \pi^{-1}\sum_{k=1}^m k^{2r}(\hat{a}_k^2 + \hat{b}_k^2)$, $r = 1,2$, and $\hat{\theta}_{1,\psi} = \pi^{-1}\sum_{k=1}^m \psi_k^2 k^2(\hat{a}_k^2 + \hat{b}_k^2)$, respectively, with $\hat{a}_k = n^{-1}\sum_{i=1}^n \cos(kX_i)$, $\hat{b}_k = n^{-1}\sum_{i=1}^n \sin(kX_i)$, and the number m of Fourier terms is selected according to the data-dependent method proposed in Tenreiro (2022, pp. 391–392), whose implementation depends on a correction parameter $\gamma = 0.5$, and deterministic sequences $L_n = \lfloor C_1 n^\xi \rfloor + 1$ and $U_n = \lfloor C_2 n^\xi \rfloor$, with $C_1 = 0.25$, $C_2 = 25$, and $\xi = 1/(4r+3)$. Taking into account the expressions for the asymptotic optimal bandwidths of the density estimators under consideration we can derive from Theorems 2, 4, and 5, the Fourier series-based plug-in bandwidth selector

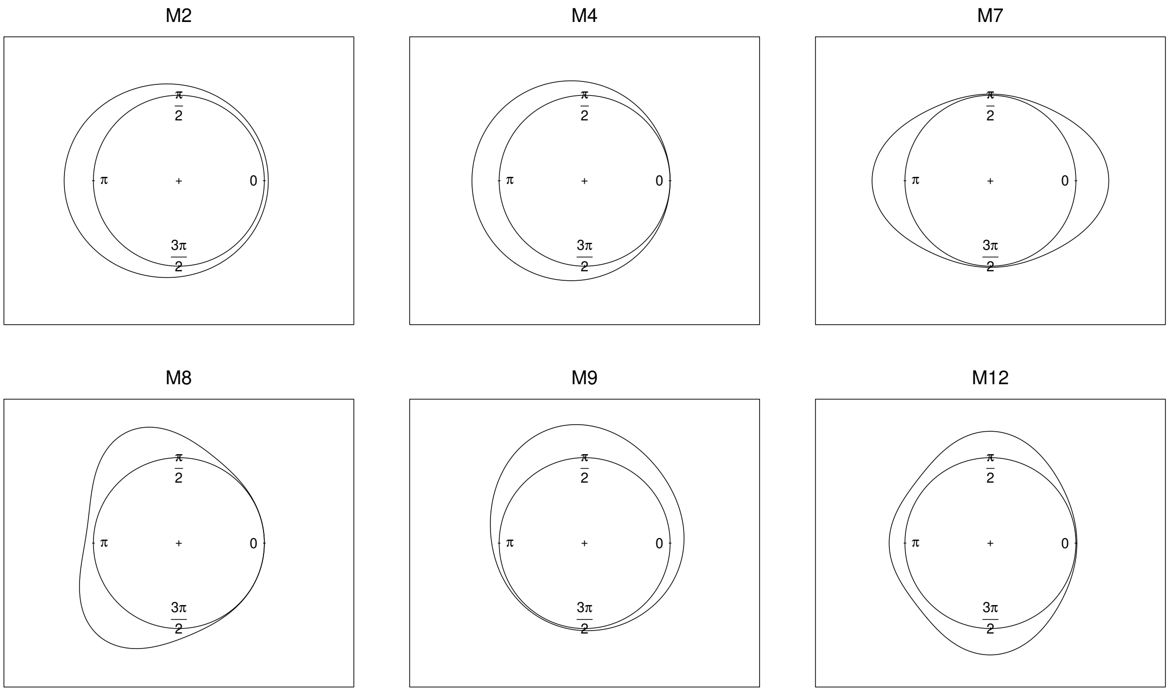

Figure 1: Probability densities of models M2, M4, M7, M8, M9, and M12 in Oliveira et al. (2012).

$\hat{h}$ is given by: 1) $\hat{h} = (8/\pi)^{1/10}\,(\sqrt{\pi/8}\,\hat{\theta}_2\,n)^{-1/5}$, if $\hat{f}$ is one of the density estimators based on the normal kernel; 2) $\hat{h} = (4\pi\,\hat{\theta}_{1,\psi}\,n)^{-1/3}$, if $\hat{f}$ is Parzen-Rosenblatt type estimator based on the Cauchy kernel; 3) $\hat{h} = (4\pi\,\hat{\theta}_1\,n)^{-1/3}$, if $\hat{f}$ is the standard or the wrapped kernel estimator based on the Cauchy kernel.

Although a larger set of densities was considered in our study, here we present the results obtained for the models M2, M4, M7, M8, M9, and M12 whose probability densities are depicted in Figure 1. We used the R-package 'circular' (Agostinelli and Lund, 2025) to produce theses graphics. In Figure 2 we describe the empirical distribution of the integrated squared error $\mathrm{ISE}(f;\hat{f},n) = \int_0^{2\pi}\{\hat{f}(\theta;\hat{h}) - f(\theta)\}^2 d\theta$, based on 500 simulated samples of size $n = 100, 400, 800$ from the density f. A logarithmic scale is used on the vertical axes of the boxplots. The previous integral is evaluated by using a grip of equally spaced 1501 points and the composite Simpson's rule. We include a polygonal line going through the sample mean values of these distributions, thus giving an approximation of $\mathrm{E}(\mathrm{ISE}(f;\hat{f},n))$.

As predicted by the asymptotic theory, we see that the density estimators based on the normal kernel perform similarly and always better than the estimators based on the Cauchy kernel, for all the considered models. In this latter case, the three estimators perform similarly for the models M7, M8, and M12, while for the models M2, M4, and M9, the Parzen-Rosenblatt type estimators performs slightly better than the other two estimators. This can be explained by the greater efficiency of $\hat{f}_{\mathrm{ST}}$ and $\hat{f}_{\mathrm{WK}}$ with respect to $\hat{f}_{\mathrm{PR}}$ for the former models than for the latter. In fact, from (13) and denoting by $\hat{f}$ one of the estimators $\hat{f}_{\mathrm{ST}}$, and $\hat{f}_{\mathrm{WK}}$, we have

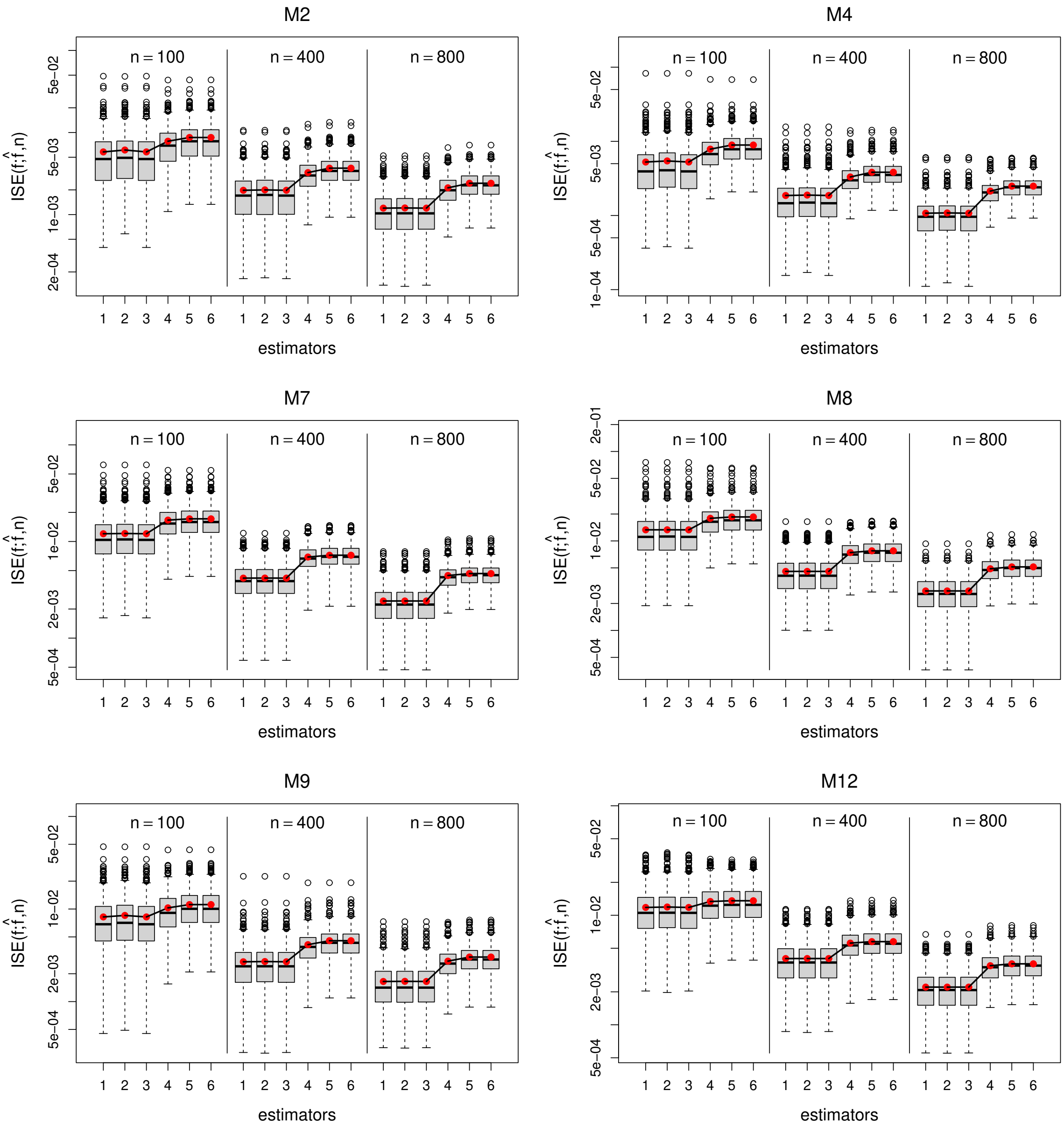

Figure 2: Empirical distribution of $\mathrm{ISE}(f;\hat{f},n)$ for models M2, M4, M7, M8, M9, and M12 in Oliveira et al. (2012). Estimators 1, 2, 3: $\hat{f}_{\mathrm{PR}}$, $\hat{f}_{\mathrm{ST}}$, and $\hat{f}_{\mathrm{WK}}$, with K the normal kernel; Estimators 4, 5, 6: $\hat{f}_{\mathrm{PR}}$, $\hat{f}_{\mathrm{ST}}$, and $\hat{f}_{\mathrm{WK}}$, with K the Cauchy kernel. The number of replications is 500.

$\mathrm{eff}(\hat{f};\hat{f}_{\mathrm{PR}}) = 0.9109, 0.9138$, and 0.9183, for the models M7, M8, and M12, respectively, and $\mathrm{eff}(\hat{f};\hat{f}_{\mathrm{PR}}) = 0.8011, 0.7737$, and 0.8350, for the models M2, M4, and M9, respectively. With regard to the other models considered in our study, the efficiency of $\hat{f}_{\mathrm{ST}}$ and $\hat{f}_{\mathrm{WK}}$ with respect to $\hat{f}_{\mathrm{PR}}$ ranges between 0.8978, obtained for the model M3, and 0.9790, obtained for the model M20, which explains why the three estimators based on the Cauchy kernel exhibit similar performance for these models (not shown here).

6 Conclusions

In this paper we studied a class of nonparametric density estimators for circular data, called wrapped kernel density estimators. These estimators are delta sequence estimators whose delta function sequence is obtained by wrapping a scaled version of a probability density function on the real line that is based on a kernel K and a sequence $h = h_n$ of scale parameters that converges to zero as n tends to infinity. Under some general assumptions on K, we conclude that the estimators of this class are asymptotically equivalent in the sense of the MISE to the standard kernel estimators of Hall et al. (1987) and Bai et al. (1988), and to the Parzen-Rosenblatt type estimators introduced in Tenreiro (2022). For density estimators based on the Cauchy kernel, a kernel that does not fulfill the mentioned assumptions, we conclude that the Parzen-Rosenblatt type estimator may be more efficient in the sense of the MISE than the other two density estimators.

7 Proofs

Recall that f denotes both the probability density function of the circular random variable X and its periodic extension to the real line.

Proof of Theorem 1: Taking into account that the Fourier transforms of $\hat{f}_n$ and $\mathrm{E}\hat{f}_n = \delta_n * f$, where $*$ denotes the convolution operator, are given respectively by $\phi_{\delta_n}\phi_n$ and $\phi_{\delta_n}\phi_f$, where ϕ_n is the sample characteristic function defined by $\phi_n(k) = n^{-1}\sum_{j=1}^n \exp(-\mathrm{i}kX_j)$, for $k \in \mathbb{Z}$, from the Parseval's identity (see Butzer and Nessel, 1971, Proposition 4.2.2, p. 175) and assumptions $(\Delta.1)$ and $(\Delta.3)$ we get

$$\begin{aligned}\mathrm{IV}(f;\hat{f}_n,n) &:= \int_0^{2\pi} \mathrm{Var}\hat{f}_n(\theta)d\theta = \mathrm{E}\int_0^{2\pi}\{\hat{f}_n(\theta) - \mathrm{E}\hat{f}_n(\theta)\}^2 d\theta \\ &= \frac{1}{2\pi}\sum_{k\in\mathbb{Z}} |\phi_{\delta_n}(k)|^2 \mathrm{E}|\phi_n(k) - \phi_f(k)|^2 \\ &= n^{-1}\frac{1}{2\pi}\sum_{k\in\mathbb{Z}} |\phi_{\delta_n}(k)|^2(1 - |\phi_f(k)|^2) \\ &= n^{-1}\alpha(\delta_n) + O\big(n^{-1}\big),\end{aligned}$$

and

$$\mathrm{ISB}(f;\hat{f}_n,n) := \int_0^{2\pi}\{\mathrm{E}\hat{f}_n(\theta) - f(\theta)\}^2 d\theta = \frac{1}{\pi}\sum_{k=1}^{\infty} |\phi_{\delta_n}(k) - 1|^2|\phi_f(k)|^2.$$

Reasoning as in Watson and Leadbetter (1964, Proof of Lemma 3, p. 104), under the assumptions $(\Delta.1)$-$(\Delta.3)$ we have $\alpha(\delta_n) \to +\infty$, from which we deduce that

$$\mathrm{IV}(f;\hat{f}_n,n) = n^{-1}\alpha(\delta_n) + o\big(n^{-1}\alpha(\delta_n)\big). \tag{14}$$

Moreover, as $f^{(p-1)}$ is absolutely continuous on $[0,2\pi]$ with $f^{(p)} \in L^2([0,2\pi))$, from Propositions 4.1.8 and 4.2.2 of Butzer and Nessel (1971, pp. 172, 175) we know that

$$\frac{1}{\pi}\sum_{k=1}^{\infty} k^{2p}|\phi_f(k)|^2 = \int_0^{2\pi} f^{(p)}(\theta)^2 d\theta < \infty. \tag{15}$$

Therefore, from the hypotheses on $|\phi_{\delta_n}(k)-1|$, and the dominated convergence theorem we get

$$\mathrm{ISB}(f;\hat{f}_n,n) = \gamma(\delta_n)^2 \frac{1}{\pi}\sum_{k=1}^{\infty} \psi_k^2 k^{2\nu}|\phi_f(k)|^2\,(1+o(1)). \tag{16}$$

Taking into account that $\mathrm{MISE}(f;\hat{f}_n,n) = \mathrm{IV}(f;\hat{f}_n,n) + \mathrm{ISB}(f;\hat{f}_n,n)$, the stated expansions for the mean integrated squared error of $\hat{f}_n$ follow from (14), (15) and (16). ■

Proof of Theorem 2: We begin by assuming that $\delta_n(y) = d_h(K)K_h(y)$, $y \in \mathbb{R}$. From Tenreiro (2022, p. 381) we know that δ_n is a delta function sequence with

$$\alpha(\delta_n) = h^{-1}\int_{-\infty}^{\infty} K(u)^2 du \left(\int_{-\infty}^{\infty} K(u)du\right)^{-2}(1+o(1)) = h^{-1}\boldsymbol{d}_1(K)\,(1+o(1)), \tag{17}$$

whenever $\lim_{u\to+\infty} uK(u) = 0$ and $0 < \int_{-\infty}^{\infty} K(u)du < \infty$. Moreover, for $k \in \mathbb{N}$ we have

$$\begin{aligned}
\phi_{\delta_n}(k) &= \int_{-\pi}^{\pi} \delta_n(y)\cos(ky)dy \\
&= 1 - k^2\int_{-\pi}^{\pi} y^2\delta_n(y)\int_0^1 (1-t)\cos(tky)dtdy \\
&= 1 - d_h(K)k^2\int_{-\pi}^{\pi} y^2 K_h(y)\int_0^1 (1-t)\cos(tky)dtdy \\
&= 1 - d_h(K)k^2h^2\int_{-\pi/h}^{\pi/h}\int_0^1 u^2K(u)(1-t)\cos(tkuh)dtdu,
\end{aligned}$$

where $d_h(K)^{-1} = \int_{-\infty}^{\infty} K(u)du\,(1+o(1))$. Therefore, using the dominated convergence theorem together with the hypothesis $\int_{-\infty}^{\infty} u^2K(u)du < \infty$, the stated expansion for $\mathrm{MISE}(f;\hat{f}_{\mathrm{PR}},n)$ follows from Theorem 1 with $\nu = 2$, $\psi_k = 1$, for all $k \in \mathbb{N}$, and

$$\gamma(\delta_n)^2 = \frac{h^4}{4}\left(\int_{-\infty}^{\infty} u^2K(u)du\right)^2\left(\int_{-\infty}^{\infty} K(u)du\right)^{-2}(1+o(1)) = \frac{h^4}{4}\boldsymbol{d}_2(K)\,(1+o(1)).$$

Let us now assume that $\delta_n(y) = c_g(L)L(g^{-2}(1-\cos y))$, $y \in \mathbb{R}$, where $L(t) = K(\sqrt{t})$ and $g = h/\sqrt{2}$. From Tenreiro (2022, p. 381) we know that δ_n is a delta function sequence with

$$\alpha(\delta_n) = g^{-1}2^{-1/2}\int_0^{\infty} t^{-1/2}L(t)^2dt\left(\int_0^{\infty} t^{-1/2}L(t)dt\right)^{-2}(1+o(1)) = h^{-1}\boldsymbol{d}_1(K)\,(1+o(1)), \tag{18}$$

as $\lim_{t\to+\infty} t^{1/2}L(t) = 0$ and $0 < \int_0^{\infty} t^{-1/2}L(t)dt < \infty$.

Moreover, for $k \in \mathbb{N}$ we have

$$\begin{aligned}
\phi_{\delta_n}(k) &= 1 - c_g(L)k^2 \int_{-\pi}^{\pi} y^2 L(g^{-2}(1-\cos y)) \int_0^1 (1-s)\cos(sky)dsdy \\
&= 1 - 2c_g(L)k^2 \int_{-1}^{1}\int_0^1 (\arccos x)^2 L(g^{-2}(1-x))(1-s) \\
&\qquad\qquad \cos(sk\arccos x)(1-x^2)^{-1/2}dsdx \\
&= 1 - 4c_g(L)k^2 g^3 \int_0^{2g^{-2}}\int_0^1 \Big(\frac{\arccos(1-tg^2)}{(2tg^2)^{1/2}}\Big)^2 t^{1/2}L(t)(1-s) \\
&\qquad\qquad \cos(sk\arccos(1-tg^2))(2-tg^2)^{-1/2}dsdt,
\end{aligned}$$

where $c_g(L)^{-1} = 2^{1/2}g\int_0^\infty t^{-1/2}L(t)dt\,(1+o(1))$. Therefore, using the hypotheses $\int_0^\infty t^{1/2}L(t)dt < \infty$ and $\lim_{t\to+\infty} t^{3/2}L(t) = 0$, and reasoning as in García-Portugués et al. (2013, proof of Lemma 1, p. 172), we conclude the stated expansion for $\mathrm{MISE}(f;\hat{f}_{\mathrm{PR}},n)$ follows from Theorem 1 with $\nu = 2$, $\psi_k = 1$, for all $k \in \mathbb{N}$, and

$$\gamma(\delta_n)^2 = g^4\Big(\int_0^\infty t^{1/2}L(t)dt\Big)^2\Big(\int_0^\infty t^{-1/2}L(t)dt\Big)^{-2}(1+o(1)) = \frac{h^4}{4}\boldsymbol{d}_2(K)\,(1+o(1)). \qquad \blacksquare$$

Proof of Theorem 3: Taking into account that K_n^w is a circular density function on $[-\pi,\pi)$, K_n^w fulfills assumption $(\Delta.1)$ for all $n \in \mathbb{N}$. In order to establish assumption $(\Delta.2)$, let us take $\lambda \in (0,\pi)$. From the hypothesis $\limsup_{u\to+\infty} u^2K(u) < \infty$ we know that there exist strictly positive constants K_0, M_0 such that

$$0 \le K(u) \le K_0 C(u), \tag{19}$$

for all $|u| > M_0$, where $C(u) = (1+u^2)^{-1}$, $u \in \mathbb{R}$, is the Cauchy kernel. Thus, using the fact that $|h^{-1}(y+2k\pi)| > h^{-1}\lambda$, for all $\lambda < |y| \le \pi$ and $k \in \mathbb{Z}$, from the representation (11) we conclude that for n large enough such that $h < \lambda/M_0$, we have

$$0 \le K_n^w(y) \le k_0^{-1}K_0 \sum_{k\in\mathbb{Z}} h^{-1}C\big(h^{-1}(y+2k\pi)\big) = k_0^{-1}K_0\pi C_n^w(y),$$

for almost all $\lambda < |y| \le \pi$, where $k_0 = \int_{-\infty}^{\infty} K(u)du$, and C_n^w is the wrapped Cauchy density function given by (7). As C_n^w is a delta sequence of functions (see Bhatia, 2005, pp. 20–23), from the previous inequality we deduce that K_n^w satisfies assumption $(\Delta.2)$.

Assuming that $0 < h < \pi/M_0$, which occurs for n large enough, for all $y \in [-\pi,\pi]$ we have $h^{-1}(y+2k\pi) \ge \pi/h > M_0$, for $k > 0$, and $h^{-1}(y+2k\pi) \le -\pi/h < -M_0$, for $k < 0$. Thus, from the representation (11) of K_n^w, and the double inequality (19), we get

$$\begin{aligned}
0 \le K_n^w(y) &\le k_0^{-1}h^{-1}K(h^{-1}y) + k_0^{-1}K_0 \sum_{k\in\mathbb{Z}\setminus\{0\}} h^{-1}C\big(h^{-1}(y+2k\pi)\big) \\
&\le k_0^{-1}h^{-1}K(h^{-1}y) + k_0^{-1}K_0\pi C_n^w(y),
\end{aligned}$$

for almost all $|y| \leq \pi$, from which we deduce that $K_n^w \in L^2([-\pi,\pi))$ for n large enough. ∎

Proof of Theorem 4: Under the assumptions on K and h, from Theorem 3 we know that $\delta_n = K_n^w$ is a delta function sequence for n large enough. In order to use Theorem 1 to establish the stated expansion for $\mathrm{MISE}(f;\hat{f}_{\mathrm{WK}},n)$, we begin by obtaining an asymptotic expansion for

$$\alpha(\delta_n) = \frac{1}{2\pi}\left(1+2\sum_{k=1}^{\infty}|\phi_{\delta_n}(k)|^2\right) = \frac{1}{2\pi}\sum_{k\in\mathbb{Z}}|\phi_K(kh)|^2\left(\int_{-\infty}^{\infty}K(u)du\right)^{-2}.$$

Using the fact that $\limsup_{t\to+\infty} t^{1+\epsilon}|\phi_K(t)|^2 < \infty$, for some $\epsilon > 0$, and reasoning as in Bertrand-Retali (1978, Lemme 2, pp. 378–380), we conclude that there exist two sequences of integrable functions φ_n^- and φ_n^+ such that

$$\varphi_n^-(t) \leq |\phi_K(t)|^2 \leq \varphi_n^+(t),\ t\in\mathbb{R},$$

where φ_n^- and φ_n^+ are constant on each interval $[kh,(k+1)h)$, $k\in\mathbb{Z}$, and

$$\lim_{n\to+\infty}\int_{-\infty}^{\infty}\varphi_n^-(t)dt = \lim_{n\to+\infty}\int_{-\infty}^{\infty}\varphi_n^+(t)dt = \int_{-\infty}^{\infty}|\phi_K(t)|^2dt.$$

Thus, for all $k\in\mathbb{Z}$ we have

$$\varphi_n^-(kh) \leq |\phi_K(kh)|^2 \leq \varphi_n^+(kh),$$

and

$$h^{-1}\int_{-\infty}^{\infty}\varphi_n^-(y)dy = \sum_{k\in\mathbb{Z}}\varphi_n^-(kh) \leq \sum_{k\in\mathbb{Z}}|\phi_K(kh)|^2 \leq \sum_{k\in\mathbb{Z}}\varphi_n^+(kh) = h^{-1}\int_{-\infty}^{\infty}\varphi_n^+(y)dy,$$

from which we conclude that

$$\frac{1}{2\pi}\sum_{k\in\mathbb{Z}}|\phi_K(kh)|^2 = h^{-1}\frac{1}{2\pi}\int_{-\infty}^{\infty}|\phi_K(t)|^2dt\,(1+o(1)) = h^{-1}\int_{-\infty}^{\infty}K(u)^2du\,(1+o(1)),$$

and

$$\alpha(\delta_n) = h^{-1}\int_{-\infty}^{\infty}K(u)^2du\left(\int_{-\infty}^{\infty}K(u)du\right)^{-2}(1+o(1)) = h^{-1}\boldsymbol{d}_1(K)\,(1+o(1)).$$

As the kernel K is such that $\int_{-\infty}^{\infty}u^2K(u)du < \infty$, we know that its Fourier transform is twice continuously differentiable on $\mathbb{R}$ with $|\phi_K''(t)| \leq \int_{-\infty}^{\infty}u^2K(u)du$, for all $t\in\mathbb{R}$, and $\phi_K''(0) = -\int_{-\infty}^{\infty}u^2K(u)du < 0$. Using a Taylor expansion of second order and equality (12), for $k\in\mathbb{N}$ and $h>0$ we get

$$\phi_K(kh) = \phi_K(0) + k^2h^2\int_0^1(1-t)\phi_K''(tkh)dt,$$

and

$$|\phi_{\delta_n}(k)-1| = |\phi_K(kh)\phi_K(0)^{-1}-1| = k^2h^2\left(\int_0^1(1-t)\phi_K''(tkh)dt\right)\phi_K(0)^{-1}.$$

Therefore, the conditions of Theorem 1 are fulfilled with $\nu=2$, $\psi_k=1$, for all $k\in\mathbb{N}$, and

$$\gamma(\delta_n)^2 = \frac{h^4}{4}\phi_K''(0)^2\phi_K(0)^{-2} = \frac{h^4}{4}\boldsymbol{d}_2(K)\,(1+o(1)).$$ ∎

Proof of Theorem 5: In order to obtain the expansion given for $\mathrm{MISE}(f;\hat{f}_{\mathrm{PR}},n)$, let us define $\delta_n(y)=d_h(K)K_h(y)$, $y\in\mathbb{R}$. From (17) we have $\alpha(\delta_n)=(h2\pi)^{-1}(1+o(1))$. Moreover, for $k\in\mathbb{N}$ we have

$$\begin{aligned}
\phi_{\delta_n}(k) &= \int_{-\pi}^{\pi}\delta_n(y)\cos(ky)dy \\
&= d_h(K)\int_{-\pi/h}^{\pi/h}K(x)\cos(kxh)dx \\
&= d_h(K)\int_{-\pi/h}^{\pi/h}K(x)\Big(1-kxh\int_0^1\sin(tkxh)dt\Big)dx \\
&= 1-2d_h(K)kh\int_0^{k\pi}\int_0^1 y\sin(ty)dt\,\frac{1}{k^2h^2+y^2}dy \\
&= 1-2d_h(K)kh\int_0^{k\pi}\frac{1-\cos y}{k^2h^2+y^2}dy,
\end{aligned}$$

and

$$|\phi_{\delta_n}(k)-1| = 2d_h(K)kh\int_0^{k\pi}\frac{1-\cos y}{k^2h^2+y^2}dy,$$

where $d_h(K)=\pi^{-1}\,(1+o(1))$ and

$$0\le\int_0^{k\pi}\frac{1-\cos y}{k^2h^2+y^2}dy\le\int_0^{\infty}(1-\cos y)y^{-2}dy=\frac{\pi}{2}.$$

Therefore, the stated expansion for $\mathrm{MISE}(f;\hat{f}_{\mathrm{PR}},n)$ follows from Theorem 1 with $\nu=1$, $\gamma(\delta_n)=h$, and $\psi_k=2\pi^{-1}\int_0^{k\pi}(1-\cos y)y^{-2}dy$, for all $k\in\mathbb{N}$.

In order to obtain the expansion given for $\mathrm{MISE}(f;\hat{f}_{\mathrm{ST}},n)$, let us define $\delta_n(y)=c_g(L)\,L(g^{-2}(1-\cos y))$, $y\in\mathbb{R}$, where $L(t)=K(\sqrt{t})$ and $g=h/\sqrt{2}$. From (18) we have $\alpha(\delta_n)=(h2\pi)^{-1}(1+o(1))$. Moreover, for $k\in\mathbb{N}$ we have

$$\begin{aligned}
\phi_{\delta_n}(k) &= 1-c_g(L)\int_{-\pi}^{\pi}L(g^{-2}(1-\cos y))(1-\cos(ky))dy \\
&= 1-c_g(L)g^2\int_{-\pi}^{\pi}\frac{1-\cos(ky)}{g^2+1-\cos y}dy,
\end{aligned}$$

and

$$|\phi_{\delta_n}(k)-1| = c_g(L)g^2\int_{-\pi}^{\pi}\frac{1-\cos(ky)}{g^2+1-\cos y}dy,$$

where $c_g(L)=2^{-1/2}g^{-1}\pi^{-1}(1+o(1))$ and

$$0\le\int_{-\pi}^{\pi}\frac{1-\cos(ky)}{g^2+1-\cos y}dy\le\int_{-\pi}^{\pi}\frac{1-\cos(ky)}{1-\cos y}dy=2k\pi$$

(see Bhatia, 2005, p. 31). Therefore, the stated expansion for $\mathrm{MISE}(f;\hat{f}_{\mathrm{ST}},n)$ follows from Theorem 1 with $\nu=1$, $\gamma(\delta_n)=2^{1/2}g=h$, and $\psi_k=1$, for all $k\in\mathbb{N}$. ■

Funding

Research partially supported by the Centre for Mathematics of the University of Coimbra (CMUC, `https://doi.org/10.54499/UID/00324/2025`) under the Portuguese Foundation for Science and Technology (FCT).

References

Agostinelli, C., Lund, U. (2025). R package 'circular': Circular Statistics (version 0.5-2). URL `https://CRAN.R-project.org/package=circular`

Ameijeiras-Alonso, J. (2024). A reliable data-based smoothing parameter selection method for circular kernel estimation. *Stat. Comput.* 34, 73.

Bai, Z.D., Rao, C.R., Zhao, L.C. (1988). Kernel estimators of density function of directional data. *J. Multivariate Anal.* 27, 24–39.

Beran, R. (1979). Exponential models for directional data. *Ann. Statist.* 7, 1162–1178.

Bertrand-Retali, M. (1978). Convergence uniforme d'un estimateur de la densité par la méthode du noyau. *Rev. Roumaine Math. Pures Appl.* 13, 361–385.

Bhatia, R. (2005). *Fourier series.* Washington: Mathematical Association of America.

Butzer, P.L., and Nessel, R.J. (1971). *Fourier Analysis and Approximation. One-Dimensional Theory.* Basel: Birkhauser.

Di Marzio, M, Panzera, A., Taylor, C.C. (2009). Local polynomial regression for circular predictors. *Statis. Probab. Lett.* 79, 2066–2075.

Di Marzio, M., Panzera, A., Taylor, C.C. (2011). Kernel density estimation on the torus. *J. Statist. Plann. Inference* 141, 2156–2173.

García-Portugués, E. (2013). Exact risk improvement of bandwidth selectors for kernel density estimation with directional data. *Electron. J. Stat.* 7, 1655–1685.

García-Portugués, E., Crujeiras, R.M., González-Manteiga, W. (2013). Kernel density estimation for directional-linear data. *J. Multivariate Anal.* 121, 152–175.

Hall, P., Watson, G.S., Cabrera, J. (1987). Kernel density estimation with spherical data. *Biometrika* 74, 751–762.

Mardia, K.V., Jupp, P.E. (2000). *Directional statistics.* Chischester: John Wiley.

Oliveira, M., Crujeiras, R.M., Rodríguez-Casal, A. (2012). A plug-in rule for bandwidth selection in circular density estimation. *Comp. Stat. Data Anal.* 56, 3898–3908.

Oliveira, M., Crujeiras, R.M., Rodríguez-Casal, A. (2014). NPCirc: An R package for nonparametric circular methods. *Journal of Statistical Software* 61 (9), 1–26.

Parzen, E. (1962). On estimation of a probability density function and mode. *Ann. Math. Statist.* 33, 1065–1076.

R Core Team (2024). *R: a language and environment for statistical computing.* R Foundation for Statistical Computing, Vienna, Austria. URL `https://www.R-project.org/`

Rosenblatt, M. (1956). Remarks on some non-parametric estimates of a density function. *Ann. Math. Statist.* 27, 832–837.

Taylor, C.C. (2008). Automatic bandwidth selection for circular density estimation. *Comput. Statist. Data Anal.* 52, 3493–3500.

Tenreiro, C. (2022). Kernel density estimation for circular data: a Fourier series-based plug-in approach for bandwidth selection. *J. Nonparametr. Stat.* 34, 377–406.

Tenreiro, C. (2024). A Parzen-Rosenblatt type density estimator for circular data: asymptotic and exact optimal bandwidths. *Comm. Statist. Theory Methods* 53, 7436–7452.

Tenreiro, C. (2025). A note on a Parzen–Rosenblatt type density estimator for circular data. In *New Frontiers in Statistics and Data Science (SPE2023, Guimarães, Portugal, October 11-14)* Henriques-Rodrigues, L. et al. (eds), Springer Proceedings in Mathematics & Statistics 469, 1–11.

Tsuruta, Y., Sagae, M. (2017). Asymptotic property of wrapped Cauchy kernel density estimation on the circle. *Bulletin of Informatics and Cybernetics* 49, 1–10.

Watson, G.S., and Leadbetter, M.R. (1964). Hazard analysis II. *Sankhyā: The Indian Journal of Statistics*, Ser. A, 26, 101–116.