On a family of two–piece circular distributions

Ameijeiras-Alonso, Jose; Gijbels, Irene & VERHASSELT, Anneleen (2022) On a family of two–piece circular distributions. In: COMPUTATIONAL STATISTICS & DATA ANALYSIS, 168 (Art N° 107403).

DOI: 10.1016/j.csda.2021.107403
Handle: http://hdl.handle.net/1942/36809

On a family of two–piece circular distributions

Jose Ameijeiras-Alonso[a], Irène Gijbels[b], Anneleen Verhasselt[c]

[a] *Department of Statistics, Mathematical Analysis and Optimization, Universidade de Santiago de Compostela, Santiago de Compostela (A Coruña), Spain*
[b] *Department of Mathematics and Leuven Statistics Research Center (LStat), KU Leuven, Leuven, Belgium*
[c] *Center for Statistics, Data Science Institute, Hasselt University, Hasselt, Belgium*

Abstract

A new way of constructing flexible and unimodal circular models, focusing on the modal direction, is proposed. Starting from a base symmetric density and a weight function, a two–piece four parameters density is introduced. The proposed density provides an extension of the base density to allow for sharply peaked and flat–topped unimodal distributions as well as a wide range of skewness. In particular, it generalizes some well–known peakedness–free models such as the Batschelet and Papakonstantinou densities. The four parameters of the model have a clear interpretation: modal direction, concentration, peakedness at the left and at the right of the modal direction. Symmetric submodels are obtained when the peakedness parameters are equal. The main properties related to the shape of the new density are presented and asymptotic results for maximum likelihood estimators are derived. An illustrative application concerning the flight orientation of migrating raptors is investigated.

Keywords: Circular Statistics, Flexible Modeling, Peakedness, Skewness, Unimodality.

1. Introduction

Circular statistics became an area of particular relevance in many applied fields as, in many examples, data can be represented on a circumference taking

*Corresponding author
Email address: jose.ameijeiras@usc.es (Jose Ameijeiras-Alonso)

into account periodicity. Several applications related to orientations or periodic phenomena can be found in Ley & Verdebout (2018).

The complicated features that data tend to exhibit on the circle (Mardia, 1972, Section 1.4), such as skewness or varying peakedness (i.e., the curvature of the density function) around the modal direction, lead to the exploration of new more flexible models in this context. Ley & Verdebout (2017, Section 2.1.1) determined four kinds of flexible models for linear data: the skew symmetric distributions (such as the Umbach & Jammalamadaka, 2009, proposal), the transformation of variables distributions (Jones & Pewsey, 2012; Abe et al., 2013), the mixture distributions (Mardia & Sutton, 1975) and the two–piece distributions (see, e.g., Arellano-Valle et al., 2005; Cassart et al., 2008; Wallis, 2014; Gijbels et al., 2019, for the linear case). Up to our knowledge, none of the proposed models for circular data directly followed this last approach. The only remarkable exception constitute the circular densities, constructed from linear ones, using the wrapping approach or projections (see, e.g. Chaubey & Karmaker, 2021). The aim of this paper is to provide a genuine new flexible two–piece circular density. Therefore, we propose a method to introduce asymmetry and varying peakedness around the modal direction. We refer to this issue as the concept of peakedness–free model. We hereby start from any symmetric circular density f_0 (the base density), which can be seen as the main submodel. This new model is "mode–based" in the sense that a unimodal distribution is proposed preserving the modal direction of the base density f_0.

Although even multimodal distributions can be obtained with our approach, our focus is on unimodal densities. As then, each parameter of the model has a clear interpretation: modal direction, concentration, peakedness at the left and at the right of the modal direction. When the objective is modeling multimodal phenomena, densities with more than one local maximum can still be obtained with a mixture of these unimodal components. Another important advantage of the new proposed methodology is that when using the von Mises or the cardioid as the base density, the symmetric submodels (obtained when both peakedness parameters are equal) match the well–known peakedness–free models of Batschelet and Papakonstantinou. Therefore, some results in this paper also complement the findings in Abe et al. (2009); Pewsey et al. (2011) and Abe et al. (2013).

Several examples where the modal direction is of special relevance and a flexible circular distribution is needed can be found in the literature. Some applications include: modeling the daily time of gun crimes (Gill & Hangartner, 2010), analyzing the yearly time of wildfire occurrences (Ameijeiras-Alonso et al., 2019), modeling the wind orientation (Agostinelli, 2007), studying the angular positions of cracks in the cement mantle in a hip implant (Mann et al., 2003), analyzing the hourly temperature cycle changes (Oliveira et al., 2013), or studying the flight orientation of migrating raptors (Cabrera-Cruz & Villegas-Patraca, 2016b). This last example is revisited here to complement the findings of Cabrera-Cruz & Villegas-Patraca (2016b).

The paper is organized as follows. Section 2 introduces some useful terminology and includes a summary of circular models that are of special relevance for this paper. Section 3 provides the basic formulation of the new asymmetric and peakedness–free model; and its associated properties, in terms of modality, symmetry, peakedness and trigonometric moments; together with an algorithm for generating random data from the model. In Section 4 parameter estimation is studied. In particular, computation and asymptotic properties of maximum likelihood (ML) estimators are studied in detail. In Section 5, we illustrate the application of the new proposal in an example in the ecology field. Section 6 summarizes relevant points of discussion. The Appendix summarizes some basic circular terminology and it contains proofs of the main theoretical results provided in Sections 3 and 4. Additional results are included as Supplementary Material: (i) proofs of the other theoretical results of the paper; (ii) a study of the shape of the densities and some shape measures of the proposed family for different choices of the base and weight functions; (iii) a study of the shape measures related to the trigonometric moments for the generalized Papakonstantinou model; (iv) some further details in the generation of random numbers and in the computation of the ML estimators; and (v) the finite–sample behavior of the ML estimates for the generalized Batschelet distribution.

2. Basic models for circular data

In the following, we describe some relevant circular models that are employed throughout this paper. A recent review of different circular models can be found in Pewsey et al. (2013, Section 4.3) and Ley & Verdebout (2017, Section 2.2). Our main focus is on unimodal models, i.e., distributions for which the density has a unique local maximum at the modal direction m, and a minimum at the antimodal direction, in the interval $[-\pi, \pi)$. The modal direction can be employed as a location parameter, and a circular density f_m belongs to the location family of a base density f_0, if $f_m(\theta) = f_0(\theta - m)$, for all $\theta \in [-\pi, \pi)$. Throughout this section, we assume that the reader is familiarized with the circular statistics terminology. Alternatively, we refer to Appendix A where some important definitions, related to the circular random variables Θ, are summarized.

Besides the location parameter a second shape parameter indicating how "concentrated" the data are towards this center, called concentration c, is included in most of the classical circular distributions (see, e.g., Mardia & Jupp, 2000, Ch. 3). Table 1 lists some circular densities depending on a concentration parameter c. For some densities, such as the cardioid, c coincides with the mean resultant length, but for others, such as the von Mises, it is just a shape parameter controlling the spread of the distribution. A general mechanism for constructing circular densities is the wrapping approach, provided in Table 1. Besides the classic circular distributions, Table 1 also includes the wrapped Laplace, studied by Jammalamadaka & Kozubowski (2003). The main drawback of the wrapped densities is that they do not always simplify to a closed form, as it occurs for the wrapped normal.

The classical circular densities given in Table 1 share the limitation of being symmetric around 0. Throughout this paper symmetry refers to reflective

Name	Expression of the density	Parameter	NSRC
	standard circular densities f_Θ		
von Mises	$f_{0,c}(\theta) = \exp\left(c\,\cos(\theta)\right) / \left(2\pi\,\mathcal{I}_0(c)\right)$	$c > 0$	–
cardioid	$f_{0,c}(\theta) = (1 + 2c\,\cos(\theta)) / (2\pi)$	$c \in (0, 0.5)$	–
	wrapped circular densities		
$f_\Theta(\theta) = \sum_{s=1}^{\infty} f_X(\theta + 2s\pi)$	f_X density of a linear random variable X		
wrapped Cauchy	$f_{\Theta;0,c}(\theta) = (1 - c^2) / \big(2\pi(1 + c^2 - 2c\,\cos(\theta))\big)$	$c \in (0, 1)$	–
wrapped Laplace	$f_{X;0,c}(x) = c\exp(-c\lvert x\rvert)/2$	$c > 0$	(A6), (A7)
wrapped normal	$f_{X;0,c}$ a normal $N(0, \sigma^2)$ density, $c = e^{-\sigma^2/2}$	$c \in (0, 1)$	–

Table 1: Some known circular densities, depending on a concentration parameter c. The column NSRC includes the non–satisfied regularity conditions, when using that density as a base density of (3). $\mathcal{I}_n(\cdot)$ is the modified Bessel function of the first kind of order n.

symmetry. Some general approaches for constructing asymmetric circular distributions are reviewed in Table 2. Note that, for the k–sine–skewed and the densities in Table 3, the location parameter does not designate the modal direction. Therefore the parameter μ instead of m is employed. The main drawback of the k–sine–skewed distributions is that unimodality does not always hold. For example, the 1–sine–skewed von Mises distribution is bimodal when the absolute value of s is "large" for values of $c > 3$ (see Abe & Pewsey, 2011, Fig. 2). The main advantage of the k–sine–skewed densities is that the normalizing constant is the same as that of the base density f_0. The inverse 2–sine–skewed distribution has the main advantage of being always unimodal, with the same modal and antimodal directions as the base density f_0. Thus, it shares this property with the family proposed in this paper. A main inconvenience of the inverse 2–sine–skewed distribution is that the density needs to be computed numerically.

The asymmetry of a circular distribution may be measured by the skewness coefficient $\mathfrak{s}$ (see Appendix A). According to this coefficient, left– (with respect to the center of symmetry) skewed may refer to the case where $\mathfrak{s} < 0$, and

Name/Reference	Expression of the density	Parameters
Umbach & Jammalamadaka (2009)	$g_\mu(\theta) = 2f_0(\theta-\mu)H_0(w(\theta-\mu))$	
f_0 and h_0 circular densities symmetric around 0, $H_0(\theta) = \int_{-\pi}^{\theta} h_0(\psi)\mathrm{d}\psi$.		
$w : \mathbb{R} \to \mathbb{R}$, an odd and periodic weight function, satisfying $\lvert w(\theta)\rvert \leq \pi$		
k–sine–skewed densities	taking $H_0 = (\pi+\theta)/(2\pi)$, $w(\theta) = s\pi\sin(k\theta)$	$k \in \mathbb{Z}^+$
(Abe & Pewsey, 2011)	$g_{\mu,s}(\theta) = f_0(\theta-\mu)(1 + s\sin(k(\theta-\mu)))$	$s \in [-1,1]$
Abe et al. (2021)	$g_m(\theta) = f_0\left(\eta^{-1}(\theta)\right)$ $\eta(\theta) = 2\int_{-\pi}^{\theta} H_0(w(\psi))\mathrm{d}\psi - \pi$	
inverse 2–sine–skewed density	taking $\eta(\theta) = \theta + s\sin^2(\theta)$	$s \in [-1,1]$

Table 2: Some known skewed circular densities, depending on a skewing parameter s.

right–skewed, when $\mathfrak{s} > 0$. A drawback of using this terminology is that $\bar{\beta}_2 = 0$ does not always imply reflective symmetry (see Section S3 of the Supplementary Material). Testing for reflective symmetry around a point is thus not equivalent to testing for $\bar{\beta}_2 = 0$. See e.g. Pewsey (2002) for a test for testing circular symmetry based on an estimator for $\bar{\beta}_2$. Given all these considerations we use the definitions of left/right–skewed, with respect to θ_0, in terms of the density shape. Given a subset $\mathcal{A} \subsetneq (0,\pi)$ and its relative complement $\mathcal{A}^c = (0,\pi)\backslash\mathcal{A}$, we define symmetry/skewness as follows.

$$
\begin{aligned}
f \text{ is } \left\{\begin{matrix}\text{left–skewed}\\ \text{symmetric}\\ \text{right–skewed}\end{matrix}\right\} \text{with respect to } \theta_0, \text{ if } \quad f(\theta_0-\theta) &\left\{\begin{matrix} > \\ = \\ < \end{matrix}\right\} f(\theta_0+\theta), \text{ for all } \theta \in \mathcal{A}^c, \\
\text{and } f(\theta_0-\theta) &= f(\theta_0+\theta), \text{ for all } \theta \in \mathcal{A}.
\end{aligned} \tag{1}
$$

The peakedness–free term indicates when a density can be more flat–topped or more sharply peaked than a base density. In the circular literature *peakedness* may be measured with the kurtosis coefficient (see Appendix A) or with the curvature around the modal direction (see, e.g., Abe et al., 2013). Thus, in this paper, peakedness (curvature around the modal direction) is defined as,

$$|f''(\theta)|/[1+(f'(\theta))^2]^{3/2}, \text{ where } \theta \text{ is in a neighbourhood of } m. \tag{2}$$

Using that curvature concept, three existing peakedness–free symmetric models are described in Section 3; the Batschelet, the Papakonstantinou and the Abe et al. (2013, Section 2) distributions.

Up to our knowledge, the two main competitors that allow for flexible accommodation of skewness and peakedness, are: the inverse Batschelet distribution (Jones & Pewsey, 2012), which according to Pewsey et al. (2013, Section 4.3) was "the most flexible circular model to date"; and the Kato & Jones (2015) density, which is termed "a very flexible unimodal distribution" in Ley & Verdebout (2017, Section 2.2). The density functions of these existing four parameter distributions are provided in Table 3. When referring to the inverse Batschelet distribution, following Jones & Pewsey (2012), we employ the von Mises as the base density to provide a circular density belonging to their family of distributions. The inverse Batschelet distribution is already implemented in the code provided by Pewsey et al. (2013, Section 4.3.13).

The main advantage of the Kato & Jones (2015) distribution is its analytic expression of the normalizing constant and a clear interpretation of the parameters in terms of the classic characteristics coefficients (see Appendix A). More specifically: $\mu = \mu_1$, $c = \rho_1$, $s = \bar{\beta}_2$ and $p = \bar{\alpha}_2$. A disadvantage of this model is that the parameter space of (p, s) depends on the parameter configuration (see Table 3). For example, if $c = 0.25$, $p \in [-0.125, 0.25)$ when $s = 0$, and when $s \neq 0$ the range of the support of p decreases. Conversely, if $c = 0.5$, $p \in [0, 0.5)$ when $s = 0$, and its support range decreases when $s \neq 0$. The main submodels of the Kato & Jones (2015) distribution are: the cardioid (when $s = p = 0$) and the wrapped Cauchy (when $s = 0$ and $p = c^2$).

Regarding the inverse Batschelet family, one important advantage is that the modal and antimodal directions have closed forms, $(\mu - 2s)$ and $(\mu \pm \pi)$. Furthermore, there is the orthogonality between some parameters for the sub-

Name/Reference	Expression of the density	Parameter values
Inverse Batschelet	$g_{\mu,c,s,p}(\theta) = C_{c,p}^{-1} f_{\mu,c}(\eta_{2,p}(\eta_{1,s}^{-1}(\theta)))$	location: $\mu \in [-\pi, \pi)$
family of densities	$C_{c,p}$ normalizing constant	concentration: $c \in \mathbb{R}$
(Jones & Pewsey,	$\eta_{1,s}(\theta) = \theta - s(1+\cos(\theta))$	skewness: $-1 \leq s \leq 1$
2012)	$\eta_{3,p}(\theta) = \theta - (1+p)\sin(\theta)/2$	
	$\eta_{2,p}(\theta) = (1-p)\theta/(1+p) + 2p\eta_{3,p}^{-1}(\theta)/(1+p)$	if peakedness: $p \in (-1, 1]$
	$\eta_{2,p}(\theta) = \theta - \sin(\theta)$	if peakedness: $p = -1$
Inverse Batschelet	$f_{\mu,c}$ being the von Mises density	concentration: $c > 0$
distribution		
Kato & Jones	$g_{\mu,c,s,p}(\theta) = c^2(c\cos(\theta-\mu) - p)$	location: $\mu \in [-\pi, \pi)$
(2015)	$((c^2+p^2+s^2-2c(p\cos(\theta-\mu)$	concentration: $0 \leq c < 1$
	$+ s\sin(\theta-\mu)))\pi)^{-1} + \frac{1}{2\pi}$	skewness: s
		peakedness: p
	$(p-c^2)^2 + s^2 \leq c^2(1-c)^2$	
	$(p,s) \neq (c,0)$	

Table 3: Some known four–parameter flexible circular densities.

model with $p = 0$: the elements of the Fisher information matrix $\mathfrak{I}$, satisfy $\mathfrak{i}_{\mu c} = \mathfrak{i}_{cs} = 0$. A disadvantage is that the normalizing constant as well as the inverse functions have no analytic expressions and need to be computed numerically. Regarding the parameter interpretation, Jones & Pewsey (2012) claim that (when $s = 0$) the most flat–topped scenarios for their density occur when $p = -1$ and the most sharply peaked when $p = 1$. Regarding the circular skewness, studying the values of $\bar{\beta}_2$ for different configurations (when $p = 0$), they found that $\mathfrak{s}$ is equal to zero, for $s = 0$, negative for $s > 0$, and decreasing as a function of s. The main submodel of the inverse Batschelet family of densities is the base density $f_{\mu,c}$, when $s = p = 0$.

3. The two–piece circular distributions and their properties

The main objective of this section is to provide the basic formulation of the new asymmetric and peakedness–free models and their general properties.

Using any symmetric and unimodal density as a basis, with two parameters, modal direction m ($-\pi \leq m < \pi$) and concentration c ($c \geq 0$), this new model depends on also two extra parameters: peakedness at left ($p_L \in \mathbb{R}$) and at right ($p_R \in \mathbb{R}$) of the modal direction m. This model is constructed with the objective to keep unimodality with the modal direction at m, independently of whether it is symmetrical or not.

Given a circular symmetric and unimodal density (with a modal direction at 0), denoted by $f_{0,c}$, and a weight function $w : \mathbb{R} \to \mathbb{R}$, the new density, in a point $\theta \in [-\pi, \pi)$, is defined as

$$g_{m,c,p_L,p_R}(\theta) = \frac{1}{C_{c,p_L,p_R}} \begin{cases} f_{0,c}[(\theta - m) + p_L w(\theta - m)] & \text{if } \theta \in I_{m,1}, \\ f_{0,c}[(\theta - m) + p_R w(\theta - m)] & \text{if } \theta \in I_{m,2}, \end{cases} \tag{3}$$

where $C_{c,p_L,p_R} = C_{c,p_L} + C_{c,p_R}$ is the normalizing constant, with $C_{c,p_L} = \int_{I_{0,1}} f_{0,c}[\theta + p_L w(\theta)]d\theta$, $C_{c,p_R} = \int_{I_{0,2}} f_{0,c}[\theta + p_R w(\theta)]d\theta$, and the support $I_{m,2} = [-\pi, \pi) \backslash I_{m,1}$, with $I_{m,1}$ defined as

$$I_{m,1} = \begin{cases} [-\pi + m, m) & \text{if } m \geq 0, \\ [-\pi, m) \cup [\pi + m, \pi) & \text{if } m < 0. \end{cases}$$

Note that a little abuse of notation was made as both g_{m,c,p_L,p_R} and C_{c,p_L,p_R} also depend on $f_{0,c}$ and w. In order to ensure that g is a density and that some properties for this family hold, some conditions on the symmetric base density $f_{0,c}$ and on the weight function w are needed. These conditions could be relaxed if the only objective is to obtain a new circular distribution, in which case the density (3) is defined outside $[-\pi, \pi)$, satisfying $g_{m,c,p_L,p_R}(\theta) = g_{m,c,p_L,p_R}(\theta + 2k\pi)$, for any integer k.

The simplest sufficient conditions to guarantee that g is a circular density function are: (i) $f_{0,c}$ is a circular density function; (ii) w is periodic, with period 2π; (iii) $0 < C_{c,p_L,p_R} < \infty$. The latter can be obtained, e.g., if $f_{0,c}$ is positive and bounded. In what follows, we give some extra conditions, needed to establish some properties of the two-piece distributions. Below, we denote $f' \equiv df/d\theta$.

Regularity conditions on the base density.

(A1) $f_{0,c}$ belongs to a location family, i.e., $f_{m,c}(\theta) = f_{0,c}(\theta + m)$, for any θ and m.

(A2) $f_{0,c}$ is periodic with period 2π, i.e., for all integers k, $f_{0,c}(\theta) = f_{0,c}(\theta+2k\pi)$; and $\int_{-\pi}^{\pi} f_{0,c}(\theta)d\theta = 1$.

(A3) $f_{0,c}(\theta) > 0$, for all $\theta \in [-\pi, \pi)$.

(A4) $f_{0,c}$ is a bounded function with a unique maximum at 0 and a minimum at $-\pi$, in the interval $[-\pi, \pi)$.

(A5) $f_{0,c}$ is an even function, i.e., $f_{0,c}(-\theta) = f_{0,c}(\theta)$, for all θ.

(A6) $f_{0,c}$ has a continuous derivative satisfying $f'_{0,c}(\theta) > 0$ if $\theta \in I_{0,1}$ and $f'_{0,c}(\theta) < 0$ if $\theta \in I_{0,2}$.

(A7) $f_{0,c}$ has a continuous second derivative in a neighborhood of 0.

Note that in the definition of density (3) we employed Condition (A1). Another way of formulating this model without this condition is using $f_{m,c}$ instead of $f_{0,c}$ in density (3). Most of the classical circular distributions satisfy Conditions (A1)–(A7) (see Table 1).

Regularity conditions on the weight function.

(B1) w is periodic, with period 2π.

(B2) w is a non–constant odd function.

(B3) w is a bounded function with continuous derivative, satisfying, for all $\theta \in [-\pi, \pi)$ and for some $\mathfrak{l} > 0$: (i) $\mathfrak{l}|w(\theta)| \leq |\theta|$; (ii) $\mathfrak{l}|w(\theta)| \leq \pi - |\theta|$; (iii) $\mathfrak{l}|w'(\theta)| < 1$, if $\theta \neq k\pi$, for any integer k. The quantity $l > 0$ denotes the largest value such that the previous conditions are satisfied for all $\mathfrak{l} \in [0, l]$.

(B4) w has a continuous second derivative in a neighborhood of 0.

Some of these conditions are similar to those required on the weight function in Umbach & Jammalamadaka (2009). In their case, the non–constant condition (part of Condition (B2)) and Conditions (B3) and (B4) are replaced by $|w(\theta)| \leq \pi$. The weight function $w(\theta) = \sin(k\theta)$, with k being a integer different from zero, is the main example of a function satisfying both the conditions in Umbach & Jammalamadaka (2009) and Conditions (B1)–(B4). Other examples of weight functions are provided in Table 4, and discussed below. As before, some conditions can be removed in our setting when the objective is only to obtain a circular density that may not be continuous or unimodal (see Section 3.2).

3.1. Main submodels

When considering the weight $w(\theta) = \sin(\theta)$ and $p = p_L = p_R$, some symmetric peakedness–free models available in the literature are obtained as particular cases of the general model in (3). For example, taking $f_{0,c}$ the cardioid density, one obtains the Papakonstantinou model (see, e.g., Abe et al., 2009). Taking $f_{0,c}$ the von Mises density, density (3) results into the Batschelet density (see, e.g., Pewsey et al., 2011). If $f_{0,c}(\theta) = h(cos(\theta))$, where h denotes a symmetric circular density which is a function of $\cos(\theta)$, then the model provided in Section 2 of Abe et al. (2013) is obtained as a special case. Thus, the general formulation in (3) provides a flexible way of creating asymmetric alternatives of these densities. In the following, we focus on some of them.

Generalized Batschelet. Concerning the Batschelet distribution, an asymmetric generalization can be obtained by taking density (3) with $w(\theta) = \sin(\theta)$ and $f_{0,c}$ being the von Mises density. This leads to the density

$$g_{m,c,p_L,p_R}(\theta) = \frac{1}{2\pi\mathcal{I}_0(c)C_{c,p_L,p_R}} \begin{cases} \exp[c\cos((\theta-m)+p_L\sin(\theta-m))] & \text{if } \theta\in I_{m,1}, \\ \exp[c\cos((\theta-m)+p_R\sin(\theta-m))] & \text{if } \theta\in I_{m,2}. \end{cases} \tag{4}$$

In Figure 1 this density is depicted for different values of the parameters, with the objective of showing the flexibility and shape of this density. Note that

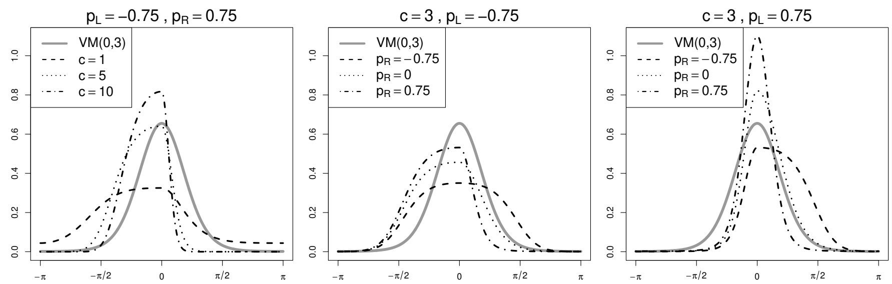

Figure 1: Generalized Batschelet density with $m = 0$. Solid grey line: von Mises (VM) density when $c = 3$ ($p_L = p_R = 0$). Broken black lines: parameters indicated in the graphics. Left: effect of the concentration parameter. Center and right: effect of peakedness parameter at right, when peakedness at left is negative ($p_L = -0.75$, center) and positive ($p_L = 0.75$, right).

this density satisfies all the previous regularity conditions. Thus, one can use Propositions 1–4, to see that this density has a continuous derivative, a unique modal direction at m, and that the parameters p_L and p_R control the symmetry and peakedness behaviour.

Generalized Papakonstantinou. This model is obtained when considering as base density $f_{0,c}$ the cardioid density and $w(\theta) = \sin(\theta)$. For this particular case, the normalizing constant has an explicit analytic expression that can be obtained as an immediate consequence of having a symmetric density $g_{0,c,p,p}$ for which the normalizing constant is $1 - cJ_1(p)$ (see Abe et al., 2009). Then, we obtain the density

$$g_{m,c,p_L,p_R}(\theta) = \frac{1}{2\pi - \pi c(J_1(p_L) + J_1(p_R))} \begin{cases} 1 + c\cos[(\theta - m) + p_L \sin(\theta - m)] & \text{if } \theta \in I_{m,1}, \\ 1 + c\cos[(\theta - m) + p_R \sin(\theta - m)] & \text{if } \theta \in I_{m,2}. \end{cases} \tag{5}$$

As is explained in Section 3.3, this density is not very suitable for modeling “concentrated” data. This disadvantage is similar to the one already observed for the cardioid base model. Figure 2 (left) presents density (5) for different parameter configurations.

k–sine–weighted submodels. A simplification of the model is obtained when considering $w(\theta) = \sin(k\theta)$, with $k \in \mathbb{Z}$ and $k \neq 0$. When $k = 1$ and the base density is $f_0(\theta) = h(\cos(\theta))$, with h a symmetric circular density, the generalized Abe et al. (2013, Section 2) distribution is obtained. Figure 2 shows

the sine–weighted submodels for different base densities satisfying $\rho_1 = 0.45$, in its symmetric version. From Figure 2 it is seen that, while respecting the shape of the base model, the sine–weighted submodels allow for wider ranges of skewness and peakedness. In Section S2 of the Supplementary Material we further illustrate the flexibility of this model, comparing its appropriateness using three base densities (the cardioid, the von Mises and the wrapped Cauchy densities) when data are generated from some other circular distribution. In general, the sine–weighted submodels provide a good approximation to many existing circular unimodal densities.

The main cases where the previous statement is not true is when for a circular unimodal density the modal and antimodal direction are pronounced and "close" to each other (see Section 3.2). The ranges of the shape measures in (3) may be very wide depending on the base density. The generalized Papakonstantinou density does not allow for "highly" concentrated distributions ($\rho_1 > 0.79$, see Section 3.3). In Table S1 of Section S2 we study the ranges of the shape measures related to the trigonometric moments, $\mathbb{s}$ and $\mathbb{k}$, of the sine–weighted submodels, when using the cardioid, the von Mises or the wrapped Cauchy as base densities, for different parameter configurations. One of the findings there is that the proposed model allows for wider ranges of the skewness and kurtosis coefficients as ρ_1 increases. Thus, the generalized Batschelet and the sine–weighted wrapped Cauchy allow for wider ranges of skewness and kurtosis than the generalized Papakonstantinou. When ρ_1 is large ($0.7 < \rho_1 \leq 0.9$), the wrapped Cauchy allows for wider ranges of the shape measures than the von Mises base density. This effect can be also seen in the different plots of Section S2. The wrapped Cauchy base density seems to be more appropriate to obtain a good model for data generated from highly concentrated densities.

In Table S1, the generalized Batschelet, the generalized Papakonstantinou and the sine–weighted wrapped Cauchy are also compared with the other two very flexible unimodal circular distributions described in Table 3, in terms of $\mathbb{s}$ and $\mathbb{k}$. In general, the inverse Batschelet distribution is the distribution obtaining the wider ranges of $\mathbb{s}$ and $\mathbb{k}$. For the different studied scenarios, independent

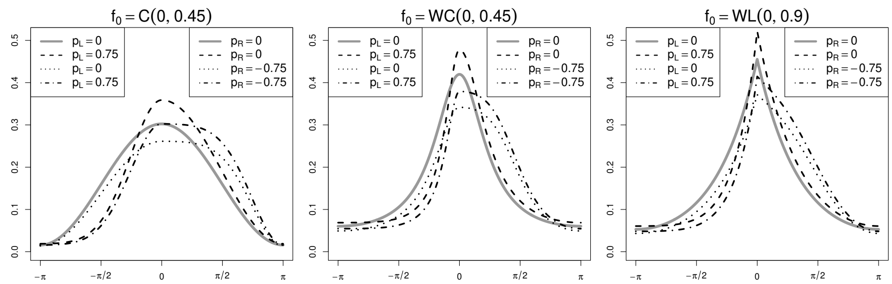

Figure 2: Density (3) with $m = 0$, $w(\theta) = \sin(\theta)$ and the base models: cardioid (left), wrapped Cauchy (center) and wrapped Laplace (right). Solid grey line: symmetric submodels, with $\rho_1 = 0.45$. Broken lines: asymmetric submodels, with different peakedness parameters.

of ρ_1 and whether $\mathbb{s}$ is "small" or "large", both the generalized Batschelet and the inverse Batschelet always contain a submodel with the same kurtosis as the wrapped normal ($\mathbb{k} = 0$), while the same does not hold for the Kato & Jones (2015) distribution. In the studied scenarios, the inverse Batschelet distribution is the one allowing for larger and smaller values of $\mathbb{k}$, independently on the configurations of ρ_1 and $\mathbb{s}$. But, by comparing the results with the three base densities, it seems that the proposed model could have larger values of $\mathbb{k}$ if a more highly–peaked distribution is used as base density. For "small" absolute values of $\mathbb{k}$, the generalized Batschelet or the Kato & Jones (2015) distributions (depending on ρ_1) are the ones having wider ranges of $\mathbb{s}$. For "large" absolute values of $\mathbb{k}$, the inverse Batschelet is the distribution that provides wider ranges of $\mathbb{s}$.

The weights $\sin(k\theta)$, when $|k| > 1$, can be useful to model data coming from a density that presents "shoulders", i.e., a density with almost flat parts outside the modal and antimodal directions (see Figure 3, left). One advantage of the newly proposed density when comparing with the k–sine–skewed densities (Abe & Pewsey, 2011) is that unimodality always holds if $-1/|k| \leq p_L, p_R \leq 1/|k|$ (see Proposition 2).

Weight function	NSRC	Control of left/right skewness by p_L and p_R?
$w(\theta) = \sin(k\,\theta) \quad k \in \mathbb{Z},\, k \neq 0$	–	if $\lvert k \rvert = 1$
$w(\theta)=$ a linear combination of $\sin(k\theta)$ functions	–	depends
$w(\theta) = \mathrm{Re}\,\{i\mathrm{Li}_k(\exp(-i\theta))\}$, $k > 1$ $\mathrm{Li}_k(z)$ polylogarithm function of order k, i.e. $\mathrm{Li}_k(z) = \sum_{q=1}^{\infty} \frac{z^q}{q^k}$	–	yes
$w(\theta) = \sin(2\pi H(\theta))$ H cumulative distribution function (CDF) of a circular density symmetric around 0, with $H(0) = 0$	depends on H	yes
H being the CDF of $\mathrm{VMM}(\mu, \kappa) = (\mathrm{VM}(\mu, \kappa) + \mathrm{VM}(-\mu, \kappa))/2$	–	yes
triangle wave weight function $w(\theta) = \begin{cases} \pi + \theta & \text{if} \quad -\pi \leq \theta \leq -\pi/2 \\ \lvert\theta\rvert & \text{if} \quad -\pi/2 \leq \theta \leq \pi/2 \\ \pi - \theta & \text{if} \quad \pi/2 \leq \theta < \pi \end{cases}$ $w(\theta) = w(\theta + 2k\pi)$, $\forall k \in \mathbb{N}$	(B3), (B4)	yes

Table 4: Examples of weight functions w. The column NSRC includes the non–satisfied regularity conditions, when using w as a weight function of (3).

Other weighting functions. Weight functions are not limited to k–sine functions. Some useful weighting functions are summarized in Table 4. The effect of these weighting functions, when the von Mises is employed as base density, is shown in Figure 3. Different weight functions lead to more flexibility and the shape of the proposed density moves away from the original von Mises shape.

3.2. Basic properties

As can already be anticipated from Figure 1, the general regularity conditions on w and $f_{0,c}$ lead to some interesting properties of density model (3). In particular, it is shown that under appropriate conditions: g_{m,c,p_L,p_R} is a circular density with a continuous derivative; is unimodal with modal direction at m when $-l \leq p_L, p_R \leq l$, where l is the positive value in (B3); the density is

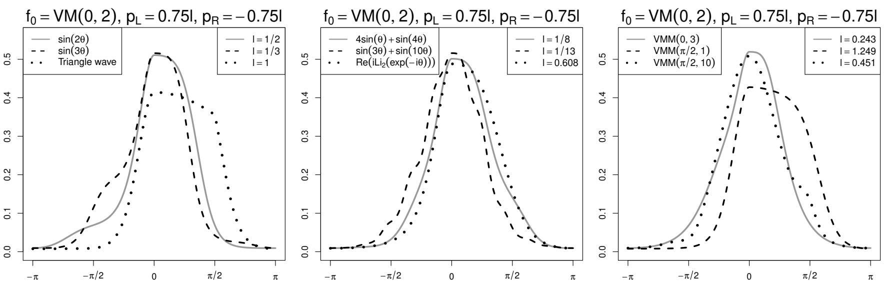

Figure 3: Density (3), $g_{0,2,3l/4,-3l/4}$, when employing the von Mises as base density, with the weights of Table 4 and where the values of l are given in the right–hand legend.

symmetric if and only if $p_L = p_R$. These results are stated formally in Propositions 1–4, the proofs of which are provided in Appendix B.1 or in Section S1 of the Supplementary Material.

In Appendix B.1 and Section S1, we also comment on how to relax the conditions on $f_{0,c}$ and w, while still obtaining some of the properties below. Among the simplest examples of base densities are the wrapped unimodal two–piece densities, that have a continuous derivative, except at the modal direction, as for the wrapped Laplace distribution (Table 1 and Figure 2, right). From the proofs, it is seen that with the wrapped Laplace as a base density, the density g_{m,c,p_L,p_R} in fact has all properties below, except for having a continuous derivative at the modal direction. A similar remark can be made for the triangle wave function (see Table 4, and Figure 3, left panel, dotted line), with which g_{m,c,p_L,p_R} satisfies all properties below (when l approaches to one from the left), with exception of the continuity of the derivative at the points $m \pm \pi/2$.

Proposition 1. *If $f_{0,c}$ satisfies Conditions (A1)–(A4) and w verifies (B1), then g_{m,c,p_L,p_R} is a circular density. If, in addition, Conditions (A6), (B2) and (B3) hold, then g_{m,c,p_L,p_R} has a continuous derivative.*

Proposition 2. *If $f_{0,c}$ satisfies Conditions (A1)–(A4) and (A6), w verifies (B1)–(B3), and $-l \leq p_L, p_R \leq l$, where l is the positive value in Condition (B3); then g_{m,c,p_L,p_R} is a unimodal density, in $[-\pi, \pi)$, with modal direction at*

m and antimodal direction at $(m \pm \pi)$.

In particular, when considering the sine as the weight function, unimodality always holds considering $l = 1$. So, the generalized Papakonstantinou and Batschelet densities are always unimodal if $-1 \leq p_L, p_R \leq 1$. The same result is indicated by Abe et al. (2009) and Pewsey et al. (2011) for the particular case when $p_L = p_R$. In the case of considering the sine weight function, Proposition 2 is an if and only if result. From the proof in Section S1, it is easy to see that if $|p_L| > 1$ or $|p_R| > 1$, there would be at least one extra modal direction, respectively, in the arc $I_{m,1}$ or $I_{m,2}$.

If $l \neq 1$, alternatively, for a general w function, the "normalized" weight function $w(\theta)/l$ may be considered, so g_{m,c,p_L,p_R} is always unimodal if $-1 \leq p_L, p_R \leq 1$.

Proposition 2 also highlights the main limitation of density (3): under the mentioned hypotheses, one important constraint is that the modal and the antimodal directions are antipodal. Thus, a distribution belonging to the proposed family should not be employed for modeling data for which the modal and antimodal direction are pronounced and "close" to each other. An example of this "bad" fitting can be seen, e.g., in Figure S3 of the Supplementary Material, for data from the Kato & Jones (2015) distribution.

Proposition 3. *If $f_{0,c}$ satisfies (A1)–(A6) and w verifies (B1)–(B3), then g_{m,c,p_L,p_R} is reflective symmetric around m if and only if $p_L = p_R$. Also, when $p_L \neq p_R$, if $(p_L - p_R)w(\theta) \geq 0$, for all $\theta \in (-\pi, 0)$, then g_{m,c,p_L,p_R} is skewed to the left (see Equation 1), while if $(p_L - p_R)w(\theta) \leq 0$, for all $\theta \in (-\pi, 0)$, then g_{m,c,p_L,p_R} is right–skewed.*

From Proposition 3 we can see that the peakedness parameters may also control when the proposed distribution is left– or right–skewed. Table 4 indicates which weight functions allow for this feature. For the other weight functions, the skewness behavior can be analyzed only, locally, in a neighborhood of the modal direction.

Proposition 4. *Suppose that $f_{0,c}$ satisfies (A1)–(A7) and w verifies (B1)–(B4). Then, considering the values $-l \leq p_{L_1}, p_{L_2}, p_{R_1}, p_{R_2} \leq l$,*

$$sgn\left(\lim_{\theta\to m^-} g''_{m,c,p_{L_1},p_{R_1}}(\theta) - \lim_{\theta\to m^-} g''_{m,c,p_{L_2},p_{R_1}}(\theta)\right) = -sgn\left(w'(0)\right)\cdot sgn\left(p_{L_1}-p_{L_2}\right),$$
$$sgn\left(\lim_{\theta\to m^+} g''_{m,c,p_{L_1},p_{R_1}}(\theta) - \lim_{\theta\to m^+} g''_{m,c,p_{L_1},p_{R_2}}(\theta)\right) = -sgn\left(w'(0)\right)\cdot sgn\left(p_{R_1}-p_{R_2}\right).$$

Considering that $g'_{m,c,p_L,p_R}(m) = 0$ and $w'(0) \neq 0$, we obtain, from Proposition 4, that the parameter p_L (respectively p_R) controls the peakedness at the left (respectively at the right) of the modal direction (see Equation 2).

3.3. Trigonometric moments

Given density (3), the expression of the trigonometric moments can be obtained as follows,

$$\begin{aligned}\alpha_r &= \frac{1}{C_{c,p_L,p_R}}\left(\int_{I_{m,1}} \cos(r\theta) f_{0,c}[(\theta-m)+p_L w(\theta-m)]d\theta\right.\\ &\qquad \left.+\int_{I_{m,2}} \cos(r\theta) f_{0,c}[(\theta-m)+p_R w(\theta-m)]d\theta\right), \qquad (6)\\ \beta_r &= \frac{1}{C_{c,p_L,p_R}}\left(\int_{I_{m,1}} \sin(r\theta) f_{0,c}[(\theta-m)+p_L w(\theta-m)]d\theta\right.\\ &\qquad \left.+\int_{I_{m,2}} \sin(r\theta) f_{0,c}[(\theta-m)+p_R w(\theta-m)]d\theta\right). \qquad (7)\end{aligned}$$

Even for the simple case with $p_L = p_R$ and $f_{0,c}$ being the von Mises density, Pewsey et al. (2011) claim that there is no known analytical expression for these quantities and hence they must be calculated numerically. The same occurs when computing the trigonometric moments about the mean direction, with the extra difficulty that, in general, $m \neq \mu_1$, except when $p_L = p_R$. In that case, g_{m,c,p_L,p_R} is symmetric and the modal direction coincides with the mean direction. When $p_L = p_R$, because of the symmetry, the value of the rth sine moment about μ_1 is $\bar{\beta}_r = 0$, for all $r \in \mathbb{Z}$ (see Mardia & Jupp, 2000, Section 3.4.4). This allows us to provide the mean resultant length, $0 \leq \rho_1 \leq 1$, of the symmetric version of the new model when the weight is the sine function, $w(\theta) = \sin(\theta)$. The result is stated in the following proposition.

Proposition 5. *If $f_{0,c}$ satisfies Conditions (A1)–(A5), $w(\theta) = \sin(\theta)$ and $p_L = p_R = p \neq 0$, then $\rho_1 = \bar{\alpha}_1 = (1 - 2C_{c,p})/(2pC_{c,p})$, where $C_{c,p} = C_{c,p,p}/2$.*

This proposition gives an analytical expression of the mean resultant length from the values of the normalizing constant $C_{c,p}$. Thus, it can be useful when the objective is to directly compute the mean resultant length of the Pewsey et al. (2011) or Abe et al. (2013, Section 2) models from their value of $C_{c,p}$. It also provides an idea about how concentrated the density is at the left and at the right of the modal direction in the proposed model.

Generalized Papakonstantinou. As mentioned before, in general, explicit expressions for the different integrals in (6) and (7) cannot be provided. However for the generalized Papakonstantinou density presented in (5) in Section 3.1 we can provide such explicit expressions. In that case, if $p_L, p_R \neq 0$, the rth cosine and sine moments are equal to

$$\alpha_r = \frac{c}{2\pi - \pi c(J_1(p_L) + J_1(p_R))} \begin{cases} \sum_{s=1}^{\infty} \frac{4r\sin(rm)}{r^2-4s^2}(J'_{2s}(p_L) - J'_{2s}(p_R)) \\ +\pi r\left(\frac{J_r(p_L)}{p_L} + \frac{J_r(p_R)}{p_R}\right)\cos(rm) & \text{, if } r \text{ is odd,} \\ \sum_{s=1}^{\infty} \frac{4r(2s-1)\sin(rm)}{r^2-(2s-1)^2}\left(\frac{J_{2s-1}(p_L)}{p_L} - \frac{J_{2s-1}(p_R)}{p_R}\right) \\ +\pi(J'_r(p_L) + J'_r(p_R))\cos(rm) & \text{, if } r \text{ is even,} \end{cases}$$

$$\beta_r = \frac{c}{2\pi - \pi c(J_1(p_L) + J_1(p_R))} \begin{cases} \sum_{s=1}^{\infty} \frac{4r\cos(rm)}{r^2-4s^2}(J'_{2s}(p_R) - J'_{2s}(p_L)) \\ +\pi r\left(\frac{J_r(p_L)}{p_L} + \frac{J_r(p_R)}{p_R}\right)\sin(rm) & \text{, if } r \text{ is odd,} \\ \sum_{s=1}^{\infty} \frac{4r(2s-1)\cos(rm)}{r^2-(2s-1)^2}\left(\frac{J_{2s-1}(p_R)}{p_R} - \frac{J_{2s-1}(p_L)}{p_L}\right) \\ +\pi(J'_r(p_L) + J'_r(p_R))\sin(rm) & \text{, if } r \text{ is even.} \end{cases}$$

The terms $J_r(p_L)/p_L$, if $p_L = 0$, or $J_r(p_R)/p_R$, if $p_R = 0$, are replaced by $1/2$ when $r = 1$ and by 0 otherwise. The derivation of these results is given in Section S1.5 of the Supplementary Material. The values of α_r and β_r can be computed, in practice, by approximating the infinite sums by a finite number of terms, where the committed error can be controlled from $|J_s(p)| < 0.675s^{-1/3}$ (see Landau, 2000). Note also that for the Papakonstantinou model (i.e., when $p_L = p_R$), the infinite sum disappears and the same results as in Abe et al. (2009) are obtained using that $2pJ_r(p)/p = J_{r-1}(p) + J_{r+1}(p)$ and $2J'_r(p) = J_{r-1}(p) - J_{r+1}(p)$.

From the previous expressions, we can see that the circular mean direction, μ_1, only depends on m, p_L and p_R, being a monotonically increasing function with respect to p_L and monotonically decreasing with respect to p_R. Regarding the mean resultant length value ρ_1, we obtain that $\rho_1 \leq J_1(1)/(1 - J_1(1)) \approx 0.7859$, thus the generalized Papakonstantinou is not a suitable distribution for modeling "concentrated" data. The previous result is obtained by noting that the derivative of ρ_1 with respect to c is positive, and, numerically, we observed that the largest value of ρ_1 is obtained when considering $\lim c \to 0.5^-$ and $p_L = p_R = 1$. These and other shape measures of this distribution are analyzed in Section S3 of the Supplementary Material.

3.4. Simulation of random numbers

Assuming that we know how to generate random numbers from the base density $f_{0,c}$, we propose to employ an adaptive acceptance–rejection method to generate random numbers from the circular random variable Θ with density (3). Let h be a density function satisfying $g_{m,c,p_L,p_R}(\theta) \leq Mh(\theta)$, for all $\theta \in [-\pi, \pi)$ and for some $M \geq 1$. In general terms, the rejection sampling algorithm consists in generating two random values, one from the uniform distribution $U(0, 1)$, denoted as U_i, and another from the distribution associated with $h(\theta)$, and denoted by Ψ_i. The random value Ψ_i is accepted (as a random value drawn from g_{m,c,p_L,p_R}) if $U_i < g_{m,c,p_L,p_R}(\Psi_i)/(Mh(\Psi_i))$.

The key point for obtaining a computationally fast algorithm is then to get a close envelope bounding of the target density g_{m,c,p_L,p_R}. With that objective in mind we propose to employ the following auxiliary function,

$$h_{1;m,c,b}(\theta) = \begin{cases} f_{0,c}(\theta - m + b)/C_{c,p_L,p_R} & \text{if } -\pi \leq \theta - m < -b, \\ f_{0,c}(0)/C_{c,p_L,p_R} & \text{if } -b \leq \theta - m \leq b, \\ f_{0,c}(\theta - m - b)/C_{c,p_L,p_R} & \text{if } b < \theta - m < \pi, \end{cases} \tag{8}$$

where $b = \max_\theta(l|w(\theta)|) \leq \pi/2$, and with $b = 1$ when $w(\theta) = \sin(\theta)$. Under the assumptions of Proposition 3, the inequality $g_{m,c,p_L,p_R}(\theta) \leq h_{1;m,c,1}(\theta)/C_{c,p_L,p_R}$ always holds, as $f_{0,c}[(\theta - m) + p_L w(\theta - m)] \leq f_{0,c}(\theta - m + b)$ when $\theta - m \in [-\pi, -b)$ and the same applies in the other part of the support. In Figure 4 (left), this bounding envelope (dashed line) is shown for the generalized Batschelet density (grey solid line).

Note that if $p_L = 0$ or $p_R = 0$, we can use the original base density as the bounding function in $I_{m,1}$ or $I_{m,2}$. Even more, since $f_{0,c}(\theta)$ is increasing in $[-\pi, 0)$, then, if $p_L w(\theta) \leq 0$ when $\theta \in [-\pi, 0)$, it holds that $g_{m,c,p_L,p_R}(\theta) \leq f_{0,c}(\theta)/C_{c,p_L,p_R}$, for these values of θ, and the same applies for $\theta \in [0, \pi)$ if $p_R w(\theta) \geq 0$. Thus, a computationally faster algorithm can be obtained considering the following auxiliary function:

$$h_{2;m,c,b,p_L,p_R}(\theta) = \begin{cases} f_{0,c}(\theta - m)/C_{c,p_L,p_R} & \text{if } \theta \in I_{m,1}, \text{ and } p_L w(\theta) \leq 0, \\ f_{0,c}(\theta - m)/C_{c,p_L,p_R} & \text{if } \theta \in I_{m,2} \text{ and } p_R w(\theta) \geq 0, \\ h_{1;m,c,b}(\theta) & \text{otherwise.} \end{cases} \tag{9}$$

For illustrative purposes, we depict function (9) in Figure 4 for the generalized Batschelet. Section S4 of the Supplementary Material provides more details of the proposed algorithm for generating random numbers from (3).

Remark 1. *Note that this algorithm also provides a computationally faster method to generate data from the different particular cases of the density (3). For the Batschelet distribution, the described algorithm is faster than the one suggested in Section 2.2 of Pewsey et al. (2011). In the acceptance–rejection method, the average number of random values required to generate a datum point is equal to M. Considering $p_L = p_R = p \geq 0$, this quantity is equal*

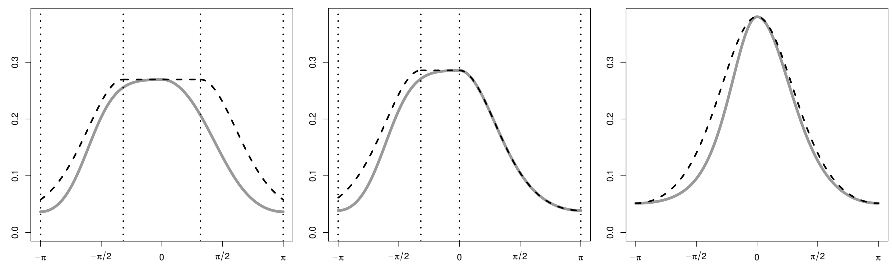

Figure 4: Solid grey line: generalized Batschelet density with $m = 0$ and $c = 1$. Dashed line: bounding function employed to generate random values. Vertical dotted lines: separating pieces in which the support is divided for obtaining the auxiliary functions. Left: $p_L = -0.8$ and $p_R = -0.3$ ($h_{1;m,c,1}$). Center: $p_L = -0.8$ and $p_R = 0$ ($h_{2;m,c,b,p_L,p_R}$). Right: $p_L = 0.4$ and $p_R = 0.1$ ($h_{2;m,c,b,p_L,p_R}$).

to $1/C_{c,p,p}$*, for the algorithm proposed in this section; while* M *is equal to* $\exp(c)/(\mathcal{I}_0(c)C_{c,p,p})$*, for the Pewsey et al. (2011) algorithm. Thus, for example, when* $c = 10$*, the average number of random values required to generate a datum point is divided by* $\exp(10)/\mathcal{I}_0(10) \approx 7.82$ *with respect to the algorithm proposed by Pewsey et al. (2011). The higher the* c *value, the better the performance. When* $c \to 0$*, both algorithms accept almost all values, and thus show a similar performance.*

4. Parameter estimation

Let $\boldsymbol{\Theta} = (\Theta_1, \ldots, \Theta_n)$ denote a random sample of angles obtained from a distribution with density (3). Given a base density $f_{0,c}$ and a weight function w, the objective of this section is to discuss maximum likelihood estimation of the parameters m, c, p_L and p_R, to establish the asymptotic behaviour of the ML estimates, and to discuss the construction of the confidence intervals (CIs) for the parameters.

We focus on the maximum likelihood estimation procedure, as the method of moments may provide multiple solutions or no solution. For example, for the symmetric version of the Papakonstantinou distribution (i.e., with $p_L =$

$p_R = p$), Abe et al. (2009) propose to use, as method of moments estimators, the sample mean direction $\hat{\mu}_1 = \mathrm{Arg}(\sum_{i=1}^n \exp(i\Theta_i))$ to estimate m, and $\bar{a}_1$ and $\bar{a}_2$, with $\bar{a}_r = n^{-1}\sum_{i=1}^n \cos r(\Theta_i - \hat{\mu}_1)$, to estimate c and p, through ρ_1 and $\rho_2 = |\mathbb{E}[Z_2]|$. In their paper, they showed that there is no solution when $\bar{a}_2/\bar{a}_1 > J_2'(1)/J_1(1) \approx 0.478$. Since even for one of the most simple subcases, the method of moments may not provide a solution we do not recommended to use this method for obtaining parameter estimators. Therefore the focus is entirely on maximum likelihood estimation.

4.1. Maximum likelihood estimation

The log–likelihood function for the full vector of parameters (m, c, p_L, p_R) of density (3) is

$$\ell(m, c, p_L, p_R) = -n \ln(C_{c,p_L,p_R}) + \sum_{i=1}^{n} \mathbb{I}(\Theta_i \in I_{m,1}) \ln[f_{0,c}((\Theta_i - m) + p_L w(\Theta_i - m))]$$
$$+ \sum_{i=1}^{n} \mathbb{I}(\Theta_i \in I_{m,2}) \ln[f_{0,c}((\Theta_i - m) + p_R w(\Theta_i - m))], \tag{10}$$

where $\mathbb{I}$ denotes the indicator function. If S_c denotes the support for c for the base density, i.e., $c \in S_c$, then the ML estimator of $\boldsymbol{\lambda} = (m, c, p_L, p_R)^T$ is the solution of $\hat{\boldsymbol{\lambda}}_n = (\hat{m}_n, \hat{c}_n, \hat{p}_{L_n}, \hat{p}_{R_n})^T = \mathrm{argmax}_{\boldsymbol{\lambda}\in\boldsymbol{\Lambda}} \ell(m, c, p_L, p_R)$, where $\boldsymbol{\Lambda} = [-\pi, \pi) \times S_c \times [-l, l] \times [-l, l]$ is the parameter space of $\boldsymbol{\lambda}$ and the superscript T denotes the transpose of the vector. For instance, the parameter space of the generalized Batschelet density (4) is $\boldsymbol{\Lambda} = [-\pi, \pi) \times \mathbb{R}^+ \times [-1, 1] \times [-1, 1]$.

When maximizing the log–likelihood function (10) the main issue is that, in general, this function is not differentiable with respect to the parameter m at the points $m = \Theta_i + k\pi$, with $i \in \{1, \ldots, n\}$ and k being an integer. Under some assumptions (see Proposition 1), the log–likelihood has a continuous derivative in almost every point. In Section S5 of the Supplementary Material, we discuss the algorithm that is employed for searching the parameters maximizing the log–likelihood, using the box–constraints provided by $\boldsymbol{\Lambda}$. The use of different initial values is generally recommended when performing the optimization to avoid identifying a local maximum, rather than the global maximum of the log–likelihood. Concerning the computational efficiency, we experienced in our simulation studies for the generalized Batschelet density that when using as starting values in the algorithm the ML estimators of the base density $f_{0,c}$ and the peakedness parameters p_L and p_R equal to zero, the true global maximum of the log–likelihood is correctly identified, with high probability. See Sections S5 and S6 in the Supplementary Material.

4.2. Asymptotic behavior of the maximum likelihood estimators

We next study the asymptotic behavior of the ML estimator $\hat{\boldsymbol{\lambda}}_n$. Note that in our context we cannot apply the classical asymptotic theory as the log–likelihood function (10) is not differentiable in some points of the support. In what follows, we show that $\hat{\boldsymbol{\lambda}}_n$ is a consistent estimator of the true parameter $\boldsymbol{\lambda}^0 = (m^0, c^0, p_L^0, p_R^0)^T$ and derive its asymptotic distribution under certain specified assumptions. The proofs of these asymptotic results are given in Appendix B or in Section S1.6 of the Supplementary Material.

Assumptions for the asymptotic results.

(C1) Let $\boldsymbol{\Lambda}_R$ be a compact subset of $\boldsymbol{\Lambda}$ and assume that $\boldsymbol{\lambda}^0$ is in the interior of $\boldsymbol{\Lambda}_R$.

(C2) If $c_1 > c_2$, the base density satisfies that $f_{0,c_1}(0) > f_{0,c_2}(0)$ and $f_{0,c_1}(-\pi) < f_{0,c_2}(-\pi)$.

(C3) The base density $f_{0,c}$ is differentiable with respect to c and the following functions have a bounded integral with respect to θ in the interval $[-\pi, \pi)$: (i) $f'_{0,c}(\theta)$, (ii) $(\partial/\partial c) f_{0,c}(\theta)$ and (iii) $w(\theta) f'_{0,c}(\theta + p w(\theta))$, for any $-l \leq p \leq l$.

(C4) For any $-l \leq p_L, p_R \leq l$, the quantities:

$$\begin{aligned} D^{k_1,k_2,k_3,k_4;-}_{c,p_L,p_R} &= \int_{-\pi}^{0} [f_{0,c}(\theta + p_L w(\theta))]^{-1} [w(\theta)]^{k_1} [w'(\theta)]^{k_2} \\ &\times [f'_{0,c}(\theta + p_L w(\theta))]^{k_3} \left[\frac{\partial}{\partial c} f_{0,c}(\theta + p_L w(\theta))\right]^{k_4} d\theta, \\ D^{k_1,k_2,k_3,k_4;+}_{c,p_L,p_R} &= \int_{0}^{\pi} [f_{0,c}(\theta + p_R w(\theta))]^{-1} [w(\theta)]^{k_1} [w'(\theta)]^{k_2} \\ &\times [f'_{0,c}(\theta + p_R w(\theta))]^{k_3} \left[\frac{\partial}{\partial c} f_{0,c}(\theta + p_R w(\theta))\right]^{k_4} d\theta, \\ D^{k_1,k_2,k_3,k_4;\pm}_{c,p_L,p_R} &= D^{k_1,k_2,k_3,k_4;-}_{c,p_L,p_R} + D^{k_1,k_2,k_3,k_4;+}_{c,p_L,p_R}, \end{aligned}$$

have a finite value when the following elements are considered in the vector (k_1, k_2, k_3, k_4): $(0, 0, 0, 2)$, $(0, 0, 2, 0)$, $(0, 1, 2, 0)$, $(0, 2, 2, 0)$, $(1, 0, 1, 1)$, $(1, 0, 2, 0)$, $(1, 1, 2, 0)$, $(2, 0, 2, 0)$.

(C5) $f_{0,c}$ and w have a bounded continuous second order derivative with respect to θ and also with respect to c (in the case of $f_{0,c}$).

(C6) $(\partial/\partial c) f_{0,c}$ is an even function.

Note that $(\partial/\partial c) f_{0,c}(\theta + pw(\theta))$ denotes the partial derivative of $f_{0,c}$ with respect to c evaluated at the point $(\theta + pw(\theta))$. Assumptions (C3) and (C4) are always satisfied if $f'_{0,c}(\theta)$ and $(\partial/\partial c) f_{0,c}(\theta)$ are bounded in $[-\pi, \pi)$ (so they are a consequence of Assumption (C5)) and $\inf_\theta f_{0,c}(\theta) > 0$. Assumption (C2) should always be satisfied using the traditional concept of concentration. In particular, Conditions (C2)–(C5) are satisfied for the following circular densities: von Mises, cardioid, wrapped Cauchy. Condition (C6) is only needed to simplify the expression of the Fisher information matrix.

Excluded here are the cases where the parameters are at the boundary of the parameter space. For the reader interested in parameter estimation when parameters are at the boundary of the parameter space, we refer to, for example, Self & Liang (1987). We also exclude base and weight densities that do not have a continuous derivative. Following the proofs in Appendix B and Section S1.6

one can see that the previous conditions may be relaxed. This is the case for $f_{0,c}$ being the wrapped Laplace or w being the triangle wave. In general, the following results remain true when the base and weight densities have a bounded continuous second order derivative with respect to θ and c in almost every point.

Theorem 1. *Suppose that the base density $f_{0,c}$ satisfies (A1)–(A4) and (A6) and the weight function w verifies (B1)–(B3). Then under Assumptions (C1) and (C2) the ML estimator $\hat{\boldsymbol{\lambda}}_n$ of $\boldsymbol{\lambda}^0$ is weakly consistent, i.e., $\hat{\boldsymbol{\lambda}}_n \xrightarrow{P} \boldsymbol{\lambda}^0$, as $n \to \infty$.*

In Proposition 6, we establish results for the Fisher information matrix, of which the elements depend on the base density $f_{0,c}$ and on the weight function w. We also prove that the expected value, with respect to the true underlying distribution, of the score vector of Θ is zero.

Proposition 6. *Suppose that the base density $f_{0,c}$ satisfies (A1)–(A6) and the weight function w verifies (B1)–(B3). Then under Assumption (C3) the expected value of the derivative of $\ln g_{\boldsymbol{\lambda}}(\Theta)$ with respect to each parameter is zero, i.e.,*

$$\mathbb{E}\left[\frac{\partial}{\partial \boldsymbol{\lambda}} \ln g_{\boldsymbol{\lambda}}(\Theta)\right] = \mathbf{0}.$$

If in addition Assumption (C4) holds, then all elements of the Fisher information matrix $\mathfrak{I}$ are finite. If, furthermore, Assumption (C6) holds, then element $(1,2)$, denoted $\mathfrak{i}_{mc}$, and the element $(2,1)$, denoted $\mathfrak{i}_{cm}$, of the Fisher information matrix are equal to zero, i.e.,

$$\mathbb{E}\left[\left(\frac{\partial}{\partial m} \ln g_{\boldsymbol{\lambda}}(\Theta)\right)\left(\frac{\partial}{\partial c} \ln g_{\boldsymbol{\lambda}}(\Theta)\right)\right] = 0.$$

When considering the symmetric submodel with $p = p_L = p_R$, the element $(1,3)$, denoted $\mathfrak{i}_{mp}$, and the element $(3,1)$, denoted $\mathfrak{i}_{pm}$, of the Fisher information matrix associated with the parameters (m, c, p) are also equal to zero, i.e.,

$$\mathbb{E}\left[\left(\frac{\partial}{\partial m} \ln g_{\boldsymbol{\lambda}}(\Theta)\right)\left(\frac{\partial}{\partial p} \ln g_{\boldsymbol{\lambda}}(\Theta)\right)\right] = 0.$$

Under the mentioned assumptions, Proposition 6 reveals that the pair of parameters (m, c) is always orthogonal. Therefore, the ML estimator of m is asymptotically independent of that for c. This proposition also shows that the estimator of m is asymptotically independent of the remaining parameters when considering the symmetric submodel with $p = p_L = p_R$.

Denoting by $C^{\lambda_i}_{c,p_L,p_R}$ the partial derivative of C_{c,p_L,p_R} with respect to the parameter λ_i, the elements of the symmetric Fisher Information matrix are (see Section S1.6 of the Supplementary Material for derivations)

$$
\begin{aligned}
\mathfrak{i}_{mm} &= \frac{1}{C_{c,p_L,p_R}}\left(D^{0,0,2,0;\pm}_{c,p_L,p_R} + p_L^2 D^{0,2,2,0;-}_{c,p_L,p_R} + 2p_L D^{0,1,2,0;-}_{c,p_L,p_R} + p_R^2 D^{0,2,2,0;+}_{c,p_L,p_R}\right.\\
&+ \left. 2p_R D^{0,1,2,0;+}_{c,p_L,p_R}\right),\\
\mathfrak{i}_{mc} &= \frac{1}{C_{c,p_L,p_R}}\left(-\int_{-\pi}^{\pi}\frac{f'_{0,c}(\theta)\left(\frac{\partial}{\partial c}f_{0,c}(\theta)\right)}{f_{0,c}(\theta)}d\theta\right) (= 0 \text{ if Assumption (C6) holds}),\\
\mathfrak{i}_{mp_L} &= \frac{1}{C_{c,p_L,p_R}}\left(-D^{1,0,2,0;-}_{c,p_L,p_R} - p_L D^{1,1,2,0;-}_{c,p_L,p_R}\right),\\
\mathfrak{i}_{mp_R} &= \frac{1}{C_{c,p_L,p_R}}\left(-D^{1,0,2,0;+}_{c,p_L,p_R} - p_R D^{1,1,2,0;+}_{c,p_L,p_R}\right),\\
\mathfrak{i}_{cc} &= \frac{1}{C_{c,p_L,p_R}}\left(-\frac{(C^{c}_{c,p_L,p_R})^2}{C_{c,p_L,p_R}} + D^{0,0,0,2;\pm}_{c,p_L,p_R}\right),\\
\mathfrak{i}_{cp_L} &= \frac{1}{C_{c,p_L,p_R}}\left(-\frac{(C^{c}_{c,p_L,p_R})(C^{p_L}_{c,p_L,p_R})}{C_{c,p_L,p_R}} + D^{1,0,1,1;-}_{c,p_L,p_R}\right),\\
\mathfrak{i}_{cp_R} &= \frac{1}{C_{c,p_L,p_R}}\left(-\frac{(C^{c}_{c,p_L,p_R})(C^{p_R}_{c,p_L,p_R})}{C_{c,p_L,p_R}} + D^{1,0,1,1;+}_{c,p_L,p_R}\right),\\
\mathfrak{i}_{p_Lp_L} &= \frac{1}{C_{c,p_L,p_R}}\left(-\frac{(C^{p_L}_{c,p_L,p_R})^2}{C_{c,p_L,p_R}} + D^{2,0,2,0;-}_{c,p_L,p_R}\right),\\
\mathfrak{i}_{p_Lp_R} &= \frac{1}{C_{c,p_L,p_R}}\left(-\frac{(C^{p_L}_{c,p_L,p_R})(C^{p_R}_{c,p_L,p_R})}{C_{c,p_L,p_R}}\right),\\
\mathfrak{i}_{p_Rp_R} &= \frac{1}{C_{c,p_L,p_R}}\left(-\frac{(C^{p_R}_{c,p_L,p_R})^2}{C_{c,p_L,p_R}} + D^{2,0,2,0;+}_{c,p_L,p_R}\right).
\end{aligned}
$$

Theorem 2. *Suppose the base density $f_{0,c}$ satisfies (A1)–(A6) and the weight function w verifies (B1)–(B3). Then under Assumptions (C1), (C2) and (C5) if the determinant of the Fisher information matrix is not null (i.e., $\det(\mathfrak{I}) \neq 0$), the ML estimator $\hat{\boldsymbol{\lambda}}_n$ of $\boldsymbol{\lambda}^0$ is asymptotically normally distributed with mean $\boldsymbol{\lambda}^0$ and variance–covariance matrix the inverse of $n\mathfrak{I}$, i.e.,*

$$\sqrt{n}(\hat{\boldsymbol{\lambda}}_n - \boldsymbol{\lambda}^0) \xrightarrow{d} N\left(\mathbf{0}, \mathfrak{I}^{-1}\right), \text{ as } n \to \infty.$$

With $tr\,\mathcal{A}$ denoting the trace of a matrix $\mathcal{A}$ and $\boldsymbol{I}_4$ the 4 by 4 identity matrix, the variance–covariance matrix $\mathfrak{I}^{-1}$ equals

$$\frac{1}{\det(\mathfrak{I})}\left[\frac{1}{6}\left((tr\,\mathfrak{I})^3 - 3tr(\mathfrak{I})tr\left(\mathfrak{I}^2\right) + 2tr\left(\mathfrak{I}^3\right)\right)\boldsymbol{I}_4 - \frac{1}{2}\left((tr\,\mathfrak{I})^2 - tr\,\mathfrak{I}^2\right)\mathfrak{I} + (tr\,\mathfrak{I})\mathfrak{I}^2 - \mathfrak{I}^3\right].$$

Note that for establishing Theorem 2, the assumption $\det(\mathfrak{I}) \neq 0$ is needed. Numerically, we have observed that $\det(\mathfrak{I}) \neq 0$ for the sine–weighted submodels when the von Mises, cardioid or the wrapped Cauchy are employed as base density. This is not the case in the vicinity of symmetry for the 1–sine–skewed von Mises distribution (i.e., in a vicinity of $s = 0$, see Table 2), due to the collinearity of the scores for location and skewness. This is an important issue as, in that case, locally and asymptotically optimal tests–in the Le Cam sense–against asymmetric alternatives of the 1–sine–skewed von Mises distribution cannot be derived (see Ameijeiras-Alonso et al., 2021).

A similar result to that provided by Proposition 6 and Theorem 2 can be derived for the symmetric submodel with $p = p_L = p_R$. When $\mathfrak{i}_{cc}\mathfrak{i}_{pp} \neq \mathfrak{i}_{cp}^2$, a proof similar to that employed for Theorem 2 yields that, under the same assumptions, the asymptotic distribution of $(\hat{m}_n, \hat{c}_n, \hat{p}_n)^T$, as $n \to \infty$, is

$$\sqrt{n}\left((\hat{m}_n, \hat{c}_n, \hat{p}_n)^T - (m^0, c^0, p^0)^T\right) \xrightarrow{d} N\left(\mathbf{0}, \begin{bmatrix} \frac{1}{\mathfrak{i}_{mm}} & 0 & 0 \\ 0 & \frac{\mathfrak{i}_{pp}}{\mathfrak{i}_{cc}\mathfrak{i}_{pp}-\mathfrak{i}_{cp}^2} & -\frac{\mathfrak{i}_{cp}}{\mathfrak{i}_{cc}\mathfrak{i}_{pp}-\mathfrak{i}_{cp}^2} \\ 0 & -\frac{\mathfrak{i}_{cp}}{\mathfrak{i}_{cc}\mathfrak{i}_{pp}-\mathfrak{i}_{cp}^2} & \frac{\mathfrak{i}_{cc}}{\mathfrak{i}_{cc}\mathfrak{i}_{pp}-\mathfrak{i}_{cp}^2} \end{bmatrix}\right). \tag{11}$$

Remark 2. *From (11) and Proposition 6, the asymptotic behavior of the model proposed by Abe et al. (2013, Section 2) is obtained as a special case. For the particular case of the Batschelet and Papakonstantinou distributions, taking w the sine function and $f_{0,c}$ respectively the von Mises and the cardioid distribution, the asymptotic normality in (11) coincides with the results in Pewsey et al. (2011, Section 3.2 and Appendix 3) and Abe et al. (2009, Section 3.2).*

In Section S6 of the Supplementary Material we present a simulation study to investigate the finite-sample performance of the ML estimates when considering the generalized Batschelet density (4).

4.3. Confidence intervals

Confidence intervals for the parameters of density (3) can be constructed in two ways: using the asymptotic theory or bootstrap methods. Given a significance level α, the $(1-\alpha)100\%$ asymptotic CIs or confidence regions are obtained directly from Theorem 2, using the Gaussian distribution $N(\hat{\boldsymbol{\lambda}}_n, (n\hat{\mathfrak{I}})^{-1})$, where $\hat{\mathfrak{I}}$ is the Fisher information matrix obtained by replacing, in the expression of $\mathfrak{I}$, the unknown parameters by their ML estimates. The approximate limits for the confidence interval (CI) for the parameter $\boldsymbol{\lambda}_j$ are $\hat{\boldsymbol{\lambda}}_{jn} \pm z_{\alpha/2}((\hat{\mathfrak{I}}^{-1})_{jj}/n)^{1/2}$, where $(\hat{\mathfrak{I}}^{-1})_{jj}$ is the (j,j)th component of $\hat{\mathfrak{I}}^{-1}$. Alternatively one can apply parametric bootstrap to obtain approximate CIs, according to the following resampling strategy. (i) Compute the ML estimators of the parameters, $\hat{\boldsymbol{\lambda}}$, from the original sample $\boldsymbol{\Theta} = (\Theta_1, \ldots, \Theta_n)$. (ii) Generate B parametric bootstrap resamples of size n from the distribution associated with $g_{\hat{\boldsymbol{\lambda}}}$ (see Section 3.4), and denote these bootstrap resamples by $\boldsymbol{\Theta}^{*b}$, with $b \in \{1, \ldots, B\}$. (iii) For each bootstrap resample $\boldsymbol{\Theta}^{*b}$ compute the ML estimator, $\hat{\boldsymbol{\lambda}}^{*b}$. Given a significance level α, compute the $\alpha/2$ and $(1-\alpha/2)$ sample quantiles of $\hat{\boldsymbol{\lambda}}_j^{*b}$, with $b \in \{1, \ldots, B\}$, for each parameter $\boldsymbol{\lambda}_j$.

5. Real data application

In ecology, one can find many applications where the use of circular statistics is necessary. In particular, our objective is modeling the flight orientation of migrating raptors in response to an increasing number of wind farms. The data, available in Cabrera-Cruz & Villegas-Patraca (2016a), consist of 3169 flight bearings of migrating raptors recorded in an area located on an important migratory corridor in southern Mexico (6 km radius of radar detection around the centroid 16.590° North latitude, −94.822° West longitude). Data were collected during the autumn migration seasons (from mid–September until early November) from 2009 to 2014, with the number of observations per year: 789, 228, 166, 894, 827 and 265. Data were obtained with a marine radar and hawk–watch monitoring stations, a full description of the employed tools and the data are provided in Cabrera-Cruz & Villegas-Patraca (2016b). This region and period were chosen as the number of wind farms increased from one (period 2009–2011, located at the centroid of the studied area) to three (period 2012–2014). The two new wind farms being located to the east and northeast of the first wind farm. One of the main objectives of the study in Cabrera-Cruz & Villegas-Patraca (2016b) was to analyze if migrating raptors adjusted their main flight orientations to avoid new wind farms. Thus, for modeling these data, the proposed mode–based family of distributions can be very useful as $\hat{m}_n$ already provides an estimator of the preferred orientation of the migrating raptors.

The histograms provided by Cabrera-Cruz & Villegas-Patraca (2016b) are reproduced on the real line in Figure 5 (gray rectangles). From these plots, a unimodal pattern is observed in the first period (2009–2011), with an almost flat–zero part in the region being at a distance π from the histogram peak(s). Thus this dataset concerns an example where the newly introduced model can be useful. The estimated densities when employing the generalized Batschelet density on the flight orientation data, for the first period, are displayed in Figure 5 (continuous black lines). Table 5 gives the estimated parameters and the corresponding 95% asymptotic and bootstrap CIs.

Since for the second period (2012–2014) it seems that a bimodal pattern is obtained, a two–component mixture of generalized Batschelet densities is employed for modeling the data during these years 2012–2014. Hence the considered density is

$$h_{\boldsymbol{m},\boldsymbol{c},\boldsymbol{p}_L,\boldsymbol{p}_R,\gamma}(\theta) = \gamma g_{m_1,c_1,p_{L1},p_{R1}}(\theta) + (1-\gamma) g_{m_2,c_2,p_{L2},p_{R2}}(\theta), \qquad (12)$$

where $\gamma \in [0,1]$ and g is the generalized Batschelet density function. The parameters of the mixture model are estimated by maximum likelihood, using the same algorithm as the one described in Section S5, considering the nine parameters and including the constraint $\gamma \in [0,1]$. The estimated densities $h_{\hat{\boldsymbol{m}},\hat{\boldsymbol{c}},\hat{\boldsymbol{p}}_L,\hat{\boldsymbol{p}}_R,\hat{\gamma}}(\theta)$, for the second period, are shown in Figure 5 (continuous black lines). Table 5 contains the estimated parameters and the corresponding 95%

bootstrap CIs. Note that the component label of the bootstrap estimated parameters was assigned according to the distance between the bootstrap modal directions, $\hat{m}^{*b}_{1,n}$ and $\hat{m}^{*b}_{2,n}$, and the original modal directions, $\hat{m}_{1,n}$ and $\hat{m}_{2,n}$.

We first discuss the results in Table 5, referring to the period where the studied area had just one wind farm (2009–2011). First note that almost the same behavior is observed when comparing the asymptotic and the bootstrap CIs, with the only exception of c in 2010 and 2011. The latter is probably due to the "small" sample size and the "large" estimated value of the concentration parameter (see Section S6). Looking at the CIs for m, it is clear that the migrating raptors always kept the same peak orientation during these years, all the CIs contain the arc (134°, 137°). Note that the modal direction estimator already provides more insight into the results in Cabrera-Cruz & Villegas-Patraca (2016b). They studied the mean direction (around 141.8°) and could not conclude that the main flight orientation did not change during these years. The point estimate of the concentration changed during the studied years. Data were more concentrated towards the modal direction in 2010 ($\hat{c}_n = 6.546$) than in 2009 ($\hat{c}_n = 2.227$). Looking at the CIs for c, all of them include the range 2.068–2.446, for the three years. Studying the point estimates of the peakedness parameters, a right–skewed distribution, $(\hat{p}_{L_n} - \hat{p}_{R_n}) > 0$, is always obtained. The estimated density is more peaked at the left than at the right of the modal direction; being at the left always more peaked than the corresponding von Mises density, $\hat{p}_{L_n} > 0$. Regarding the CIs for p_L it is noted that, for the year 2009, the lower CI limit for p_L is larger than the upper CI limit for p_R. Thus the symmetric Batschelet density is not an appropriate density to describe these data. The point 0 is always contained in the CIs for p_L and p_R for the years 2010 and 2011. For that reason we investigate below (see Table 6) whether a simpler von Mises distribution could be an appropriate model.

As mentioned before the two–component mixture (12) is employed to model the data in the second period (2012–2014). From Figure 5 as well as Table 5 it is clear that the flight orientation of the migrating raptors changed in the second period. For both components the modal direction in 2013 (lower 95%

confidence limits are 86° and 169°) was different from the modal direction in 2012 (upper 95% confidence limits are 82° and 168°). Just in the year 2014, we could assume that a group of birds had the same peak flight orientation as that followed by the migrating raptors in 2009–2011. The modal direction estimators for the first period are contained in the 95% CI of m_1. But according to the 95% CI of γ, at most 51% of the birds followed the direction associated with that first component.

Looking a the CIs for all nine parameters, we can observe that the flight orientation was similar in the years 2012 and 2014. The year 2013 exhibits a different behavior, with respect to the other two years, with a first component less concentrated and a second component more concentrated around the modal direction. Also, the probability to belong to the first component is larger in the year 2013 (lower 95% confidence limit is 0.7) than in 2012 or 2014 (upper 95% confidence limits are 0.541 and 0.51). Regarding the point estimates for the peakedness parameters we always obtained a right–skewed first component and a left–skewed second component. According to the 95% CIs, for the first component in 2012 and 2013, the estimated density is more peaked at the left and more flat–topped at the right than the von Mises density. For the second component in 2012, the opposite behavior is observed. Thus, for 2012 (both components) and 2013 (first component), the asymmetric version of the generalized Batschelet distribution is needed to model these data. Using the 95% CIs of p_{L1}, p_{L2}, p_{R1} and p_{R2}, it could be assumed that the von Mises mixture is a "good distribution" to model the flight orientation data in 2014.

To investigate further the appropriateness of model (3), and possible other models, we also fitted the two very–flexible four–parameter models: the inverse Batschelet distribution (Jones & Pewsey, 2012) and the Kato & Jones (2015) distribution; and its main submodel, the von Mises distribution. As mentioned before a bimodal pattern is observed in the second period and the estimate of at least one of the peakedness parameters in the generalized Batschelet distribution is at the boundary. These facts motivated the relaxation of the peakedness parameter restrictions when using a one–component distribution in the second

Year		m		c		p_L		p_R	
2009	Est	2.310(132°)		2.227		0.966		0.374	
	ACL	2.227(128°)	2.393(137°)	2.008	2.446	0.639	1	0.167	0.580
	BCL	2.253(129°)	2.393(137°)	2.068	2.504	0.651	1	0.165	0.593
2010	Est	2.450(140°)		6.546		0.622		-0.119	
	ACL	2.332(134°)	2.568(147°)	0.212	12.88	-0.378	1	-0.730	0.493
	BCL	2.308(132°)	2.570(147°)	4.425	36.115	-0.394	1	-0.740	0.268
2011	Est	2.261(130°)		3.970		0.831		-0.000	
	ACL	2.103(121°)	2.420(139°)	2.048	5.891	-0.012	1	-0.449	0.449
	BCL	2.112(121°)	2.448(140°)	3.127	14.226	-0.149	1	-0.638	0.397
Year		m_1		c_1		p_{L_1}		p_{R_1}	
2012	Est	1.342(77°)		33.368		1		-0.627	
	BCL	1.273(73°)	1.431(82°)	18.759	164.736	0.250	1	-0.916	-0.381
2013	Est	1.56(89°)		2.545		1		-0.132	
	BCL	1.496(86°)	1.682(96°)	2.322	3.061	0.534	1	-0.382	0.161
2014	Est	1.380(79°)		9.477		0.885		0.779	
	BCL	1.227(70°)	1.533(88°)	6.007	104.663	-0.529	1	-0.567	1
Year		m_2		c_2		p_{L_2}		p_{R_2}	
2012	Est	2.852(163°)		2.896		-0.589		1	
	BCL	2.641(151°)	2.935(168°)	2.529	4.197	-0.930	-0.050	0.285	1
2013	Est	3.030(174°)		18.649		0.330		1	
	BCL	2.952(169°)	3.099(178°)	18.602	67.847	-0.409	0.945	-0.065	1
2014	Est	2.723(156°)		2.181		-0.360		0.627	
	BCL	2.054(118°)	2.964(170°)	1.750	3.445	-1	1	-0.548	1
	Year	2012		2013		2014			
γ	Est	0.436		0.753		0.314			
	BCL	0.351	0.541	0.700	0.796	0.199	0.510		

Table 5: Flight orientation data: parameter estimates, in the first block (2009–2011) for the generalized Batschelet distribution; in the second block (2012–2014) for the two–component mixture of generalized Batschelet distributions. Est indicates the point estimates, ACL refers to the asymptotic and BCL to the bootstrap 95% confidence limits.

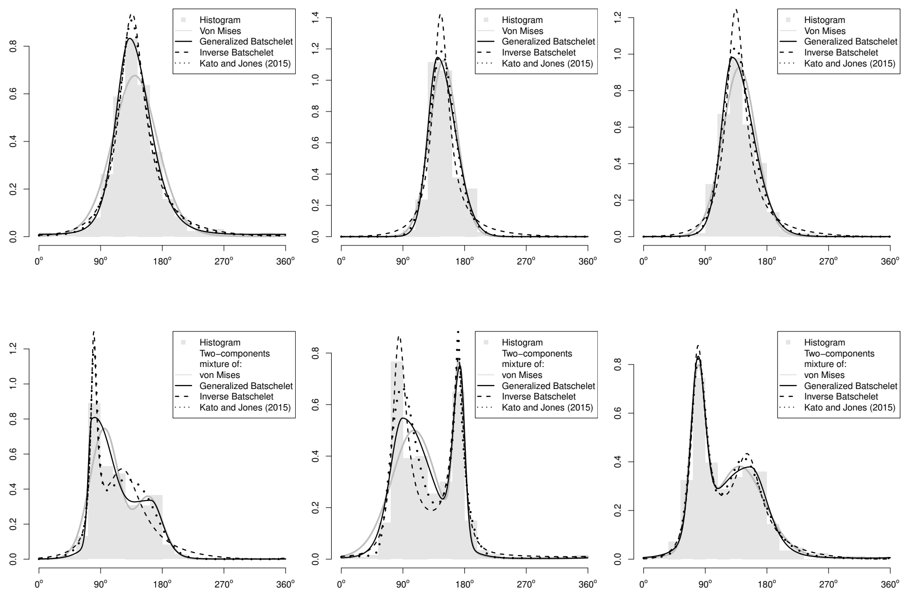

Figure 5: Flight orientation data. Histogram and maximum likelihood fits of different distributions, as indicated on the legends. Top: with one wind farm; from left to right, years 2009–2011. Bottom: with three wind farms; from left to right, years 2012–2014.

period. The results in Table 6, for all the years in the period 2012–2014, are obtained with $\hat{p}_{L_n} > 1$. Note that even if Proposition 2 does not hold, $\hat{m}_n$ is still the point at which $g_{\hat{m}_n,\hat{c}_n,\hat{p}_{L_n},\hat{p}_{R_n}}$ achieves its global maximum. Also, for the second period, the two–component mixture of the previous distributions are employed to determine which model provided the highest estimated log–likelihood value and the lowest Akaike Information Criterion (AIC). The achieved results are given in Table 6. Note that for the mixture models the number of parameters is twice the number of parameters of each component plus one parameter for the mixing probability, i.e, it is always 9, except for the mixture of von Mises, where this number is equal to 5. The estimated densities are shown in Figure 5.

Among the studied distributions, the generalized Batschelet is always the "best" one for the first period according to the log–likelihood and AIC, except in 2011, where it is the second best one and a slightly better performance is

	Year	2009	2010	2011	2012	2013	2014
Log–likelihood	Generalized Batschelet	**-714.927**	**-87.534**	-96.040	**-803.692**	**-854.823**	**-296.753**
	Kato–Jones	-724.157	-104.998	-101.888	-888.831	-979.946	-310.360
	Inverse Batschelet	-717.602	-89.311	**-95.231**	-837.414	-918.090	-299.529
	Von Mises	-751.599	-91.257	-99.931	-889.829	-959.163	-320.127
	Two–component mixtures						
	Generalized Batschelet	—	—	—	**-759.168**	-806.165	**-283.902**
	Kato–Jones	—	—	—	-810.811	-799.289	-284.017
	Inverse Batschelet	—	—	—	-772.575	**-792.025**	-284.323
	Von Mises	—	—	—	-807.873	-848.354	-285.541
AIC	Generalized Batschelet	**1437.854**	**183.068**	200.080	**1615.384**	**1717.647**	**601.505**
	Kato–Jones	1456.314	217.995	211.777	1785.661	1967.893	628.719
	Inverse Batschelet	1443.204	186.622	**198.462**	1682.827	1844.180	607.058
	Von Mises	1507.199	186.514	203.862	1783.658	1922.325	644.254
	Two–component mixtures						
	Generalized Batschelet	—	—	—	**1536.337**	1630.329	585.804
	Kato–Jones	—	—	—	1639.622	1616.578	586.035
	Inverse Batschelet	—	—	—	1563.150	**1602.050**	586.646
	Von Mises	—	—	—	1625.746	1706.707	**581.082**

Table 6: Flight orientation data. Estimated maximal log–likelihood and AIC values for the generalized Batschelet, the inverse Batschelet, the Kato & Jones (2015) and the von Mises distributions. For the second period (2012–2014), the two–component mixture of these distributions is also employed. In bold the best obtained value for the one– and two–component distributions.

obtained with the inverse Batschelet. In the second period (years 2012–2014), there is always at least one mixture model that provided a better model fit than the one–component generalized Batschelet density. Note that in 2012 the generalized Batschelet density provides a better fit than the other two-component mixtures. The generalized Batschelet mixture gives the best performance in terms of log–likelihood in the years 2012 and 2014, and in terms of AIC in 2012. For 2013, the inverse Batschelet mixture, and for 2014 the von Mises distribution (due to fewer parameters) provided a better model fit (than the generalized

	$\sin(\theta)$			$\sin(2\theta)$			$\sin(3\theta)$		
Year	VM	WC	WN	VM	WC	WN	VM	WC	WN
2009	**-714.93**	-718.27	-724.77	-727.85	-718.92	-747.16	-749.26	-732.85	-804.27
2010	-87.53	-90.26	-88.08	-86.37	-92.94	**-86.36**	-91.03	-103.27	-90.36
2011	-96.04	-99.98	**-95.73**	-98.42	-101.72	-98.23	-99.66	-107.75	-99.91
2012	-803.69	-820.83	-809.36	-830.79	**-780.67**	-854.47	-835.69	-915.12	-861.58
2013	-854.82	-821.72	-877.05	-872.64	**-805.08**	-915.36	-861.77	-894.09	-869.76
2014	-296.75	**-292.49**	-297.92	-295.12	-298.62	-300.79	-307.08	-326.30	-306.98

Table 7: Flight orientation data. Estimated maximal log–likelihood values for the two-pieces distributions. When employing as base functions: the von Mises (VM), wrapped Cauchy (WC), and wrapped normal (WN) densities. The used weight functions are the $\sin(k\theta)$, with $k \in \{1, 2, 3\}$. For each year, in bold is highlighted the best–obtained log–likelihood value.

Batschelet density) in terms of AIC.

In this section, among the possible models belonging to the proposed family of two-piece distributions, we decided, for simplicity, to only explore the generalized Batschelet. The reason being that its two- and three-parameters submodels correspond to two of the most well-known circular distributions. Alternatively, if one is interested in obtaining the "best" fitting, in terms of the estimated log–likelihood, one could explore different combinations of the base density function in Table 1, with the weight functions in Table 4. Table 7 reports the estimated maximal log–likelihood for model (3), when combining different base functions (von Mises, wrapped Cauchy and wrapped normal) with the weight functions $\sin(k\theta)$, $k \in \{1, 2, 3\}$. Note that, as before, for the first period, we constrained the peakedness parameters, so the estimated distribution is unimodal, while for the second period this requirement was lifted. In Table 7, we can see that in the years 2010–2014, other configurations would provide a better fitting than the generalized Batschelet, in terms of the estimated log–likelihood.

6. Conclusions

In this paper, a new way of constructing two–piece densities for the circular case was proposed. Starting with a two parameters base model (location and concentration), which can be thought as the main submodel, these new four parameters distributions can model wider ranges of peakedness (at the left and at the right of the modal direction) and asymmetry. From the applied point of view, besides their flexibility and the clear parameter interpretation, these new distributions have the advantage of preserving the modal direction.

We established the general properties of the proposed model, together with the asymptotic normality of the ML estimators. Since the newly proposed densities also provide an extension of two of the most well–known peakedness–free models in the circular literature (the Papakonstantinou and Batschelet models), our findings complement previous results provided for both models.

Drawbacks are that there is generally no closed–form expression for computing the normalizing constant or the trigonometric moments and that the modal and antimodal directions are antipodal. The computational disadvantage of having to perform numerical approximations, in practice, is shared by most of the circular flexible and non–flexible models, such as the inverse Batschelet or the wrapped normal. In this paper, we also provide an easy way of simulating data from the model with density (3) and a way of computing the parameter estimates.

In comparison with the other four parameter flexible models, obtaining the ML estimators is more involved than in the Kato & Jones (2015) case, due to the normalizing constant. When analysing the data in Section 5 and in generated random samples, we found out that fewer random initial points are needed to compute the ML estimators with the generalized Batschelet. This makes the "computational expenses" comparable for both distributions. Another advantage of both, the proposed family and the inverse Batschelet, with respect to the Kato & Jones (2015) distribution, is that the main submodel is chosen by the practitioner. Thus, the von Mises or any other circular distribution can be chosen as the main submodel of the proposed distribution. This flexibility also allows, in practice, for wider ranges of symmetry and peakedness.

The proposed density provides a clear parameter interpretation in terms of the density shape. Alternatively to the proposed parametrization, Proposition 3 also suggests employing, as the third parameter, a skewness parameter $(p_L - p_R)$, which is only equal to zero when the distribution is symmetric. In that case, one possible fourth parameter candidate would be $(p_L + p_R)/2$, which could be coined as the (two-sided) peakedness parameter. Note that, for the symmetric subdistributions considered in this paper, the fourth parameter would coincide with their peakedness parameter. Thus, that parametrization would allow mimicking the classical "location-scale-skewness-kurtosis" paradigm. Finally, for the reader interested in the parameter interpretation in terms of the shape measures related to the trigonometric moments, we refer to the Kato & Jones (2015) distribution.

We applied the proposed generalized Batschelet distribution on data from the ecological field. The objective was to provide a further insight in the analysis of the flight orientation of migration raptors. This example shows the need for having a mode–based model since the estimated parameters provide information on the preferred orientation of the birds. By fitting various plausible densities, and using model selection type criteria, it was illustrated that the proposed general family of distributions can lead to a useful model in a practical setting.

The use of this model in a semiparametric regression approach is part of current research. Future research will include the study of the extension of a similar model to the multivariate (toroidal) setting.

Acknowledgements

The authors thank an Associate Editor and two anonymous reviewers for their valuable comments that led to an improvement of the paper.

Supported by the FWO research project G.0826.15N (Flemish Science Foundation). The first and second author were/are supported by GOA/12/014 and

C16/20/002 projects (Research Fund KU Leuven). The first author acknowledges support by Project MTM2016–76969–P from the Spanish State Research Agency (AEI) co–funded by the European Regional Development Fund (ERDF) and the Competitive Reference Groups 2017–2020 (ED431C 2017/38) from the Xunta de Galicia through the ERDF. This research was done while the first author was a postdoctoral researcher at the KU Leuven.

Appendix A. Basic circular terminology

The first objective of this section is to define the main terminology, related with a circular random variable Θ. With that objective, let us denote the rth order complex exponential with $Z_r = \exp(ir\Theta)$, $r \in \mathbb{Z}$ and i the imaginary unit. Then, the cosine and sine trigonometric moments are, respectively, defined as $\alpha_r = \mathrm{Re}\,(\mathbb{E}[Z_r])$ and $\beta_r = \mathrm{Im}\,(\mathbb{E}[Z_r])$, where $\mathrm{Re}(z)$ denotes the real part and $\mathrm{Im}(z)$ the imaginary part of $z \in \mathbb{C}$. The circular mean direction μ_1 is equal to the argument of $\mathbb{E}[Z_1]$, i.e., $\mu_1 = \mathrm{Arg}\,(\mathbb{E}[Z_1])$. The mean resultant length ρ_1 is the modulus of $\mathbb{E}[Z_1]$, $\rho_1 = |\mathbb{E}[Z_1]|$. The cosine and sine trigonometric moments about the mean direction are, respectively, $\bar{\alpha}_r = \mathrm{Re}\,(\mathbb{E}[Z_r - \mu_1])$ and $\bar{\beta}_r = \mathrm{Im}\,(\mathbb{E}[Z_r - \mu_1])$. Using the previous notation, in general, in the circular literature (see, e.g. Mardia & Jupp, 2000, Section 3.4), the skewness coefficient is defined as $\mathbb{s} = \bar{\beta}_2/(1-\rho_1)^{3/2}$, while the kurtosis coefficient is $\mathbb{k} = (\bar{\alpha}_2 - \rho_1^4)/(1-\rho_1)^2$.

Appendix B. Proofs of the main theoretical results

Appendix B.1. Proof of Proposition 1

For simplicity, in order to use Equation (3), we consider that Condition (A1) is also satisfied. But note that this condition is not necessary for deriving this result. Under Conditions (A3) and (A4), it is easy to see that $0 < C_{c,p_L,p_R} < \infty$. The first inequality is obtained from the fact that the integral of a strictly positive-definite function in a non-zero measure set is always positive, $f_{0,c}[\theta - m + p_L w(\theta - m)] > 0$ and $f_{0,c}[\theta - m + p_R w(\theta - m)] > 0$. The second inequality is a consequence of $f_{0,c}[\theta - m + p_L w(\theta - m)]$ and $f_{0,c}[\theta - m + p_R w(\theta - m)]$ being bounded functions which are integrated over a bounded set.

Now, if $0 < C_{c,p_L,p_R} < \infty$, to show that g_{m,c,p_L,p_R} is a circular density, Conditions (A2), (A3) and (B1) are enough. The function g_{m,c,p_L,p_R} is a circular density (see, e.g., Mardia & Jupp, 2000, Section 3.2) if:

(i) $g_{m,c,p_L,p_R}(\theta) \geq 0$, almost everywhere on $(-\infty, \infty)$.

(ii) $\int_{-\pi}^{\pi} g_{m,c,p_L,p_R}(\theta) d\theta = 1$.

(iii) $g_{m,c,p_L,p_R}(\theta) = g_{m,c,p_L,p_R}(\theta + 2k\pi)$, almost everywhere on $(-\infty, \infty)$, for any integer k.

The first two parts, (i) and (ii), are immediate consequences of having a strictly positive and finite value of C_{c,p_L,p_R} and Conditions (A2) and (A3). Now, in order to prove (iii), by Conditions (A2) and (B1), for any integer k, we have that

$$f_{0,c}[\theta - m + p_L w(\theta - m)] = f_{0,c}[\theta - m + p_L w(\theta - m) + 2k\pi] = f_{0,c}[(\theta + 2k\pi - m) + p_L w(\theta + 2k\pi - m)],$$

and the same is true replacing p_L by p_R, which leads to Condition (iii).

Remark 3. *Note that if the only objective is to provide a circular density from (3), some of the previous conditions could be relaxed in order to obtain a more broad family. In particular, providing that the normalizing constant has a finite non–zero value, the strict inequality in (A3) could be replaced by a non–strict one, i.e.,* $g_{m,c,p_L,p_R}(\theta) = 0$ *for some values of* θ*. Then, just Conditions (A2) on* $f_{0,c}$ *and (B1) on* w *are needed in order to obtain a circular density.*

Finally, for seeing that g_{m,c,p_L,p_R} has a continuous derivative, using Conditions (A2), (A6), (B1), (B2) and (B3), we have that the composite of continuous mappings is continuous, thus the derivative of g_{m,c,p_L,p_R} is continuous in the interior of the subsets $I_{m,1}$ and $I_{m,2}$. Now, referring to the two remaining points, m and $(m - \pi)$ or $(m + \pi)$ (depending on the value of m), using (A2) and (A6) we obtain that $f'_{0,c}(0) = 0$ and $f'_{0,c}(\pi) = 0$. Using (B1) and (B2), we can see also that $w(0) = w(-\pi) = 0$. Finally, for the derivative we find that

$$\begin{aligned}
\lim_{\theta \to m^-} g'_{m,c,p_L,p_R}(\theta) &= \lim_{\theta \to m^-} C^{-1}_{c,p_L,p_R}[1 + p_L w'(\theta - m)] f'_{0,c}[(\theta - m) + p_L w(\theta - m)] = 0, \\
\lim_{\theta \to m^+} g'_{m,c,p_L,p_R}(\theta) &= \lim_{\theta \to m^+} C^{-1}_{c,p_L,p_R}[1 + p_R w'(\theta - m)] f'_{0,c}[(\theta - m) + p_R w(\theta - m)] = 0.
\end{aligned}$$

In an analogous way, the same result is derived for $(m \pm \pi)$. These same arguments could be employed to see the continuity of g_{m,c,p_L,p_R} in these two points. Thus, if just the continuity of g_{m,c,p_L,p_R} is required, Conditions (A6) and (B3) could be relaxed, replacing them by the continuity of $f_{0,c}$ and w.

Appendix B.2. Proof of Theorem 1

For obtaining the consistency of the ML estimators, sufficient conditions are given in Theorem 2.1 of Newey & McFadden (1994). Now, if the density $g_{\boldsymbol{\lambda}}$ satisfies the following assumptions, the sufficient conditions of Theorem 2.1 are obtained using Lemmas 2.2 and 2.4 of Newey & McFadden (1994).

- D.1 $\boldsymbol{\Lambda}_R$ is compact.
- D.2 $\boldsymbol{\lambda}^0$ is identified, i.e., if $\boldsymbol{\lambda} \neq \boldsymbol{\lambda}^0$ and $\boldsymbol{\lambda}^0, \boldsymbol{\lambda} \in \boldsymbol{\Lambda}_R$, then $g_{\boldsymbol{\lambda}}(\cdot) \neq g_{\boldsymbol{\lambda}^0}(\cdot)$.
- D.3 $\mathbb{E}[|\ln g_{\boldsymbol{\lambda}}(\Theta)|] < \infty$, for all $\boldsymbol{\lambda} \in \boldsymbol{\Lambda}_R$.
- D.4 $\ln g_{\boldsymbol{\lambda}}(\Theta_i)$, with $i = 1\ldots, n$; is continuous at each $\boldsymbol{\lambda} \in \boldsymbol{\Lambda}_R$ with probability one.

Condition D.1 is satisfied by considering Assumption (C1). For obtaining Condition D.2, we assume that $g_{\boldsymbol{\lambda}}(\cdot) = g_{\boldsymbol{\lambda}^0}(\cdot)$ for some $\boldsymbol{\lambda} \neq \boldsymbol{\lambda}^0$, being $\boldsymbol{\lambda} = (m, c, p_L, p_R)^T \in \boldsymbol{\Lambda}_R$. Since both densities are equal, this implies that both have the same modal direction, thus $m = m^0$.

Referring to the parameter c, first note that, since $g_{\boldsymbol{\lambda}}(m) = g_{\boldsymbol{\lambda}^0}(m^0)$, the following equality is obtained $f_{0,c}(0)/C_{c,p_L,p_R} = f_{0,c^0}(0)/C_{c,p_L^0,p_R^0}$. Now, assume that $c \neq c^0$, then if $c > c^0$, by Assumption (C2), $f_{0,c}(0) > f_{0,c^0}(0)$. Thus, $C_{c,p_L,p_R} > C_{c^0,p_L^0,p_R^0}$. Applying the same arguments when evaluating g in the antimodal direction $m \pm \pi$ we obtain that $C_{c,p_L,p_R} < C_{c^0,p_L^0,p_R^0}$, which leads to a contradiction. The same arguments applies when considering $c < c^0$, thus, the only possibility is $c = c^0$. In that case, we also obtain that $C_{c,p_L,p_R} = C_{c^0,p_L^0,p_R^0}$.

We see next that $p_L = p_L^0$ and $p_R = p_R^0$. Using the regularity Condition (B2) on w, there exists a point $\theta_0 < 0$ for which $w(\theta_0) \neq 0$. Condition (B3) implies that there exists a neighborhood of θ_0 for which $\mathrm{sign}(w(\theta)) = \mathrm{sign}(w(\psi)) \neq 0$, for all the points $\theta, \psi \in (\theta_0-\delta, \theta_0+\delta)$, with $\delta > 0$. Now if $p_L \neq p_L^0$, $p_L w(\theta) < p_L^0 w(\theta)$ or $p_L w(\theta) > p_L^0 w(\theta)$, for all $\theta \in (\theta_0 - \delta, \theta_0 + \delta)$. Assume the first case, then, by Condition (B3), $(\theta - m) + p_L w(\theta - m) < (\theta - m) + p_L^0 w(\theta - m) < 0$. Since, by Condition (A3), the base density is strictly increasing in $(-\pi, 0)$, $f_{0,c}[(\theta-m)+p_L w(\theta-m)] < f_{0,c^0}[(\theta-m)+p_L^0 w(\theta-m)]$, for all $\theta \in (\theta_0-\delta, \theta_0+\delta)$ as $c = c^0$. Now, combining this last inequality with $g_{\boldsymbol{\lambda}}(\cdot) = g_{\boldsymbol{\lambda}^0}(\cdot)$, we obtain that the only possibility is $C_{c,p_L,p_R} > C_{c,p_L^0,p_R^0}$. Arguing in the same way, if $p_L w(\theta) > p_L^0 w(\theta)$, $C_{c,p_L,p_R} < C_{c,p_L^0,p_R^0}$. As we saw in the previous paragraph that $C_{c,p_L,p_R} = C_{c^0,p_L^0,p_L^0}$, the only possibility is $p_L = p_L^0$. The same arguments can be employed to obtain that $p_R = p_R^0$.

For deriving Condition D.3, we first consider the expression for $\ln g_{\boldsymbol{\lambda}}(\theta)$,

$$\begin{aligned} \ln g_{\boldsymbol{\lambda}}(\theta) &= -\ln(C_{c,p_L,p_R}) + \mathbb{I}(\theta \in I_{m,1}) \ln[f_{0,c}((\theta - m) + p_L w(\theta - m))] \\ &+ \mathbb{I}(\theta \in I_{m,2}) \ln[f_{0,c}((\theta - m) + p_R w(\theta - m))]. \end{aligned} \tag{B.1}$$

Then, for any $\boldsymbol{\lambda} \in \boldsymbol{\Lambda}_R$, we obtain that

$$\begin{aligned} \mathbb{E}[|\ln g_{\boldsymbol{\lambda}}(\Theta)|] \leq |\ln(C_{c,p_L,p_R})| &+ \mathbb{E}\left[\mathbb{I}(\Theta \in I_{m,1}) |\ln(f_{0,c}((\Theta - m) + p_L w(\Theta - m)))|\right] \\ &+ \mathbb{E}\left[\mathbb{I}(\Theta \in I_{m,2}) |\ln(f_{0,c}((\Theta - m) + p_R w(\Theta - m)))|\right]. \end{aligned}$$

Now, since $0 < C_{c,p_L,p_R} < \infty$ (see Section Appendix B.1) and $0 < f_{0,c}(\theta) < \infty$ by Conditions (A3) and (A4), both $|\ln(C_{c,p_L,p_R})|$ and $|\ln(f_{0,c}(\cdot)|$ are bounded. Thus, using that the supports $I_{m,1}$ and $I_{m,2}$ are also bounded, we obtain that the expected value in Condition D.3 is bounded.

Finally, Condition D.4 is an immediate consequence of (A6) and the continuity of the composition of continuous functions.

Appendix B.3. Proof of Theorem 2

Since we have that the likelihood function is non–differentiable, the proof of Theorem 2 is obtained following Huber (1967). For doing so, first, let us introduce some notation,

$$\boldsymbol{\Phi}(\theta, \boldsymbol{\lambda}) = \begin{bmatrix} \frac{\partial}{\partial m} \ln g_{\boldsymbol{\lambda}}(\theta) \\ \frac{\partial}{\partial c} \ln g_{\boldsymbol{\lambda}}(\theta) \\ \frac{\partial}{\partial p_L} \ln g_{\boldsymbol{\lambda}}(\theta) \\ \frac{\partial}{\partial p_R} \ln g_{\boldsymbol{\lambda}}(\theta) \end{bmatrix}. \tag{B.2}$$

From Equation (S1.5, see Section S1.6 of the Supplementary Material), if the base density satisfies Condition (A6) and the weight function (B3), we obtain that $\boldsymbol{\Phi}(\theta, \boldsymbol{\lambda})$ is a continuous function but, even if the base density $f_{0,c}$ has a continuous second derivative, it may be non differentiable at the points $\theta = m + k\pi$ with k an integer. The quantity $\boldsymbol{\varpi}(\boldsymbol{\lambda})$ represents the expected value of $\boldsymbol{\Phi}(\Theta, \boldsymbol{\lambda})$, i.e.,

$$\boldsymbol{\varpi}(\boldsymbol{\lambda}) = \mathbb{E}[\boldsymbol{\Phi}(\Theta, \boldsymbol{\lambda})].$$

Denoting by $||\cdot||$ to the Euclidean norm, the function u is defined as follows,

$$u(\theta, \boldsymbol{\lambda}, \delta) = \sup_{||\boldsymbol{\vartheta} - \boldsymbol{\lambda}|| \leq \delta} ||\boldsymbol{\Phi}(\theta, \boldsymbol{\vartheta}) - \boldsymbol{\Phi}(\theta, \boldsymbol{\lambda})||.$$

Using these definitions, if the Fisher information matrix is invertible and the conditions below are satisfied, from the corollary of Theorem 3 of Huber (1967), we obtain the asymptotic normal distribution of $\sqrt{n}(\hat{\boldsymbol{\lambda}}_n - \boldsymbol{\lambda}^0)$ with mean zero and asymptotic variance–covariance matrix upper bounded by $\mathfrak{I}^{-1}$. The variance–covariance matrix is equal to this upper bound when the Fisher information matrix is continuous at $\boldsymbol{\lambda}^0$, which is the case under Assumption (C5).

E.1 For each $\boldsymbol{\lambda} \in \boldsymbol{\Lambda}_R$, $\boldsymbol{\Phi}(\theta, \boldsymbol{\lambda})$ is $\mathcal{F}$–measurable and separable (in the sense of Doob), where the σ–algebra $\mathcal{F}$ is a collection of all the possible events.

E.2 There is a $\boldsymbol{\lambda}^0 \in \boldsymbol{\Lambda}_R$ such that $\boldsymbol{\varpi}(\boldsymbol{\lambda}^0) = 0$.

E.3 There exist strictly positive numbers a, b, c, δ_0 such that

(a) $||\boldsymbol{\varpi}(\boldsymbol{\lambda})|| \geq a||\boldsymbol{\lambda} - \boldsymbol{\lambda}^0||$, for $||\boldsymbol{\lambda} - \boldsymbol{\lambda}^0|| \leq \delta_0$,

(b) $\mathbb{E}[u(\Theta, \boldsymbol{\lambda}, \delta)] \leq b\delta$, for $||\boldsymbol{\lambda} - \boldsymbol{\lambda}^0|| + \delta \leq \delta_0$, with $\delta \geq 0$;

(c) $\mathbb{E}[u(\Theta, \boldsymbol{\lambda}, \delta)^2] \leq c\delta$, for $||\boldsymbol{\lambda} - \boldsymbol{\lambda}^0|| + \delta \leq \delta_0$, with $\delta \geq 0$.

E.4 The expectation $\mathbb{E}[||\boldsymbol{\Phi}(\Theta, \boldsymbol{\lambda})||^2]$ is finite.

E.5 $(1/\sqrt{n}) \sum_{i=1}^{n} \boldsymbol{\Phi}(\Theta_i, \hat{\boldsymbol{\lambda}}_n) \to \mathbf{0}$ in probability.

Condition E.1 can be obtained taking into account that $\mathbf{\Phi}$ is a continuous function and $\mathbf{\Lambda}_R$ is a compact set by Assumption (C1). Conditions E.2 and E.4 are immediate consequences of Proposition 6. Condition E.5 is a consequence of the ML estimates being weakly consistent (see Theorem 1).

Then, the only condition that remains to be proven is Condition E.3. Conditions E.3(b) and E.3(c) are a consequence of g and $u(\theta, \boldsymbol{\lambda}, \delta)$ being continuous and bounded on the compact set $\mathbf{\Lambda}_R$. Now, to prove Condition E.3(a), first assume that $\boldsymbol{\lambda} \neq \boldsymbol{\lambda}^0$ (if they are equal, this condition is trivially satisfied). Using the theorem's assumptions, $||\boldsymbol{\varpi}(\boldsymbol{\lambda}^0)|| = 0$ and the total derivative of $\boldsymbol{\varpi}(\boldsymbol{\lambda})$ exists at $\boldsymbol{\lambda}^0$ and it is equal to $-\mathfrak{I}$. Since the determinant of the Fisher information matrix is non–null, its norm is also different to zero. Using the reverse triangular inequality and the sub–multiplicativity of the norm, for any value of ϵ satisfying $0 < \epsilon < ||\mathfrak{I}||$, there exists a δ_0 such that if $||\boldsymbol{\lambda} - \boldsymbol{\lambda}^0|| \leq \delta_0$, then

$$\begin{aligned} -\epsilon||\boldsymbol{\lambda} - \boldsymbol{\lambda}^0|| &\leq -||\boldsymbol{\varpi}(\boldsymbol{\lambda}) - \boldsymbol{\varpi}(\boldsymbol{\lambda}^0) + \mathfrak{I}(\boldsymbol{\lambda} - \boldsymbol{\lambda}^0)|| \\ &\leq ||\boldsymbol{\varpi}(\boldsymbol{\lambda})|| - ||\boldsymbol{\varpi}(\boldsymbol{\lambda}^0)|| - ||\mathfrak{I}|| \cdot ||(\boldsymbol{\lambda} - \boldsymbol{\lambda}^0)||. \end{aligned}$$

Thus, considering $a = ||\mathfrak{I}|| - \epsilon > 0$, the proof of Condition E.3(a) is finished.

References

Abe, T., Miyata, Y., & Shiohama, T. (2021). Bayesian estimation for mode and anti-mode preserving circular distributions. *Econometrics and Statistics*, *to appear*.

Abe, T., & Pewsey, A. (2011). Sine–skewed circular distributions. *Statistical Papers*, *52*, 683–707.

Abe, T., Pewsey, A., & Shimizu, K. (2009). On Papakonstantinou's extension of the cardioid distribution. *Statistics & Probability Letters*, *79*, 2138–2147.

Abe, T., Pewsey, A., & Shimizu, K. (2013). Extending circular distributions through transformation of argument. *Annals of the Institute of Statistical Mathematics*, *65*, 833–858.

Agostinelli, C. (2007). Robust estimation for circular data. *Computational Statistics & Data Analysis*, *51*, 5867–5875.

Ameijeiras-Alonso, J., Lagona, F., Ranalli, M., & Crujeiras, R. M. (2019). A circular nonhomogeneous hidden Markov field for the spatial segmentation of wildfire occurrences. *Environmetrics*, *30*, 1–19.

Ameijeiras-Alonso, J., Ley, C., Pewsey, A., & Verdebout, T. (2021). On optimal tests for circular reflective symmetry about an unknown central direction. *Statistical Papers*, *62*, 1651–1674.

Arellano-Valle, R. B., Gómez, H. W., & Quintana, F. A. (2005). Statistical inference for a general class of asymmetric distributions. *Journal of Statistical Planning and Inference*, *128*, 427–443.

Cabrera-Cruz, S. A., & Villegas-Patraca, R. (2016a). Data from: Response of migrating raptors to an increasing number of wind farms. Figshare, `https://dx.doi.org/10.6084/m9.figshare.3123100.v2`. Accessed: 2020–03–16.

Cabrera-Cruz, S. A., & Villegas-Patraca, R. (2016b). Response of migrating raptors to an increasing number of wind farms. *Journal of Applied Ecology*, *53*, 1667–1675.

Cassart, D., Hallin, M., & Paindaveine, D. (2008). Optimal detection of fechner-asymmetry. *Journal of Statistical Planning and Inference*, *138*, 2499–2525.

Chaubey, Y. P., & Karmaker, S. C. (2021). On some circular distributions induced by inverse stereographic projection. *Sankhya B*, *83*, 319–341.

Gijbels, I., Karim, R., & Verhasselt, A. (2019). On quantile-based asymmetric family of distributions: Properties and inference. *International Statistical Review*, *87*, 471–504.

Gill, J., & Hangartner, D. (2010). Circular data in political science and how to handle it. *Political Analysis*, *18*, 316–336.

Huber, P. J. (1967). The behavior of maximum likelihood estimates under nonstandard conditions. In *Proceedings of the fifth Berkeley symposium on mathematical statistics and probability* (pp. 221–233). University of California Press volume 1.

Jammalamadaka, S. R., & Kozubowski, T. (2003). A new family of circular models: The wrapped Laplace distributions. *Advances and Applications in Statistics*, *3*, 77–103.

Jones, M., & Pewsey, A. (2012). Inverse Batschelet distributions for circular data. *Biometrics*, *68*, 183–193.

Kato, S., & Jones, M. (2015). A tractable and interpretable four–parameter family of unimodal distributions on the circle. *Biometrika*, *102*, 181–190.

Landau, L. (2000). Bessel functions: monotonicity and bounds. *Journal of the London Mathematical Society*, *61*, 197–215.

Ley, C., & Verdebout, T. (2017). *Modern Directional Statistics*. Boca Raton, Florida: CRC Press.

Ley, C., & Verdebout, T. (2018). *Applied Directional Statistics: Modern Methods and Case Studies*. Boca Raton, Florida: CRC Press.

Mann, K. A., Gupta, S., Race, A., Miller, M. A., & Cleary, R. J. (2003). Application of circular statistics in the study of crack distribution around cemented femoral components. *Journal of Biomechanics*, *36*, 1231–1234.

Mardia, K. V. (1972). *Statistics of Directional Data*. London: Academic Press.

Mardia, K. V., & Jupp, P. E. (2000). *Directional Statistics*. Chichester: Wiley.

Mardia, K. V., & Sutton, T. W. (1975). On the modes of a mixture of two von Mises distributions. *Biometrika*, *62*, 699–701.

Newey, W. K., & McFadden, D. (1994). Large sample estimation and hypothesis testing. In J. J. Heckman, & E. E. (Eds.), *Handbook of Econometrics* (pp. 2111–2245). Amsterdam, Netherlands: North Holland.

Oliveira, M., Crujeiras, R. M., & Rodríguez-Casal, A. (2013). Nonparametric circular methods for exploring environmental data. *Environmental and Ecological Statistics*, *20*, 1–17.

Pewsey, A. (2002). Testing circular symmetry. *Canadian Journal of Statistics*, *30*, 591–600.

Pewsey, A., Neuhäuser, M., & Ruxton, G. D. (2013). *Circular Statistics in R*. Oxford, United Kingdom: Oxford University Press.

Pewsey, A., Shimizu, K., & de la Cruz, R. (2011). On an extension of the von Mises distribution due to Batschelet. *Journal of Applied Statistics*, *38*, 1073–1085.

Self, S. G., & Liang, K.-Y. (1987). Asymptotic properties of maximum likelihood estimators and likelihood ratio tests under nonstandard conditions. *Journal of the American Statistical Association*, *82*, 605–610.

Umbach, D., & Jammalamadaka, S. R. (2009). Building asymmetry into circular distributions. *Statistics & Probability Letters*, *79*, 659–663.

Wallis, K. F. (2014). The two-piece normal, binormal, or double gaussian distribution: its origin and rediscoveries. *Statistical Science*, *29*, 106–112.