Test
DOI 10.1007/s11749-008-0111-0

ORIGINAL PAPER

Two nested families of skew-symmetric circular distributions

William J. Reed · Arthur Pewsey

Received: 10 July 2007 / Accepted: 20 May 2008

Abstract In this paper, we introduce two new flexible families of unimodal circular distributions obtained by wrapping onto the unit circle two recently explored heavy-tailed distributions defined on the real line. The first, the four-parameter wrapped normal–Laplace distribution, is nested within the second, the five-parameter wrapped generalized normal–Laplace distribution. Both families contain the wrapped normal and wrapped Laplace and generalized Laplace distributions as special cases. Stochastic models for the genesis of these new distributions, which may be useful in identifying situations in which they are likely to occur, are developed. The basic properties of the new distributions are derived and model fitting by maximum likelihood discussed. Examples which illustrate fitting the two distributions to exact and grouped data are presented.

Keywords Circular Brownian motion · Generalized normal–Laplace distribution · Wrapped distributions

Mathematics Subject Classification (2000) MSC 60E10 · MSC 60E07

The research of W. J. Reed was supported by NSERC grant OGP 7252, and partly carried out at the University of Waikato, New Zealand, whose support is gratefully acknowledged. The research of Arthur Pewsey was partially supported by project MTM2007-61470 of the Spanish Ministry of Education and Science.

W.J. Reed (✉)
Department of Mathematics and Statistics, University of Victoria, PO Box 3045, Victoria, BC V8W 3P4, Canada
e-mail: reed@math.uvic.ca

A. Pewsey
Departamento de Matemáticas, Escuela Politécnica, Universidad de Extremadura, 10071 Cáceres, Spain
e-mail: apewsey@unex.es

1 Introduction

The standard parametric models used in the analysis of circular data, i.e., the von Mises, cardioid, wrapped normal, and wrapped Cauchy distributions (see Mardia and Jupp 2000, Sect. 3.5), are all symmetric, despite the fact that directional data seldom exhibit symmetry. In recent years, several new unimodal circular distributions capable of modeling symmetry as well as asymmetry have been proposed in the literature. These include the wrapped versions of the skew-normal (Pewsey 2000, 2006); Laplace (Jammalamadaka and Kozubowski 2003, 2004); and stable (Gatto and Jammalamadaka 2003; Pewsey 2008) distributions. An alternative approach to modeling asymmetric unimodal circular data is to use finite mixtures of symmetric unimodal circular distributions; see, for example, Fisher (1993, Sect. 4.6); Jammalamadaka and SenGupta (2001, Sect. 4.3); and Mardia and Jupp (2000, Sect. 5.5). Finite mixtures with symmetric components can involve relatively large numbers of parameters when compared with the wrapped skew-symmetric models mentioned above and those that we propose here. Nevertheless, their interpretation is often simple. Moreover, finite mixtures with unimodal (symmetric or asymmetric) component distributions can be used as a means of modeling multimodal circular data.

In this paper we propose two new flexible families of wrapped unimodal distributions: the four-parameter *wrapped normal–Laplace* and the five-parameter *wrapped generalized normal–Laplace* distributions. Both families of distributions include the wrapped normal, wrapped Laplace, and wrapped generalized Laplace distributions and can be used to model asymmetry, varying levels of kurtosis and heavy-tailedness. In Sect. 2 we define the two distributions and derive them from a simple stochastic model involving Brownian motion on the circle. This model, which has several variants, can provide analysts of circular data with a theoretical justification for the use of these new distributions. Properties of the wrapped generalized normal–Laplace and wrapped normal–Laplace distributions are presented in Sect. 3. Maximum likelihood estimation is discussed in Sect. 4, and the two families are fitted to real data and their fits compared with those of other candidate distributions in Sect. 5. The paper ends, in Sect. 6, with some concluding remarks.

2 The wrapped normal–Laplace and generalized normal–Laplace distributions—definitions and genesis

The *generalized normal–Laplace* (GNL) distribution defined on the real line was introduced by Reed (2007) and has characteristic function

$$\phi_{\mathrm{GNL}}(s) = \left[\frac{\exp(\mathrm{i}\eta s - \tau^2 s^2/2)}{(1-\mathrm{i}as)(1+\mathrm{i}bs)}\right]^{\zeta}, \tag{1}$$

where $\eta \in \mathcal{R}$ and a, b, ζ, and τ^2 are non-negative reals (i.e., $\in \bar{\mathcal{R}}_{-}$). We will use the notation $X \sim \mathrm{GNL}(\eta, \tau^2, a, b, \zeta)$ to indicate that the random variable (rv) X has the above distribution. In general, closed-form expressions for the probability density and distribution functions of the GNL distribution are not available.

For the special case when $\zeta = 1$, the distribution reduces to the (ordinary) *normal–Laplace* (NL) distribution for which closed forms for the density and distribution function have been established (Reed and Jorgensen 2004). The reason for the name "normal–Laplace" is because the distribution arises as that of a convolution of Gaussian ($\mathrm{N}(\eta, \tau^2)$) and (skew-) Laplace (with density $f(x) = (a+b)^{-1}\mathrm{e}^{x/b}$ for $x < 0$ and $= (a+b)^{-1}\mathrm{e}^{-x/a}$ for $x > 0$) components. It occurs when the state of a Brownian motion (with normally distributed starting state) is observed after an exponentially distributed time (Reed and Jorgensen 2004).

The GNL distribution can be represented as a convolution of independent Gaussian and *generalized Laplace* components (Kotz et al. 2001). Since the generalized Laplace distribution can be represented as the difference between two independent gamma rv's with the same shape parameter, one can represent a GNL rv X as $X \overset{d}{=} \eta\zeta + \tau\sqrt{\zeta}Z + aV_1 - bV_2$, where Z, V_1 and V_2 are independent, $Z \sim \mathrm{N}(0, 1)$ and V_1 and V_2 are identically distributed gamma rv's with shape parameter ζ and scale parameter 1 (Reed 2007). This representation provides the easiest way to generate pseudo-random variates from a GNL distribution (and, thus, from its wrapped version).

The family of GNL distributions is quite rich. It contains, as special cases, the normal distribution (when $a = b = 0$); the generalized Laplace distribution (when $\eta = \tau = 0$); and the (skew-) Laplace distribution (when $\eta = \tau = 0$ and $\zeta = 1$).

It is clear from (1) that: (i) both the NL and GNL distributions are infinitely divisible; (ii) sums of independent NL rv's with the same a and b will follow a GNL distribution; (iii) the GNL distribution is closed under summation, in the sense that sums of independent GNL rv's with the same a and b will also follow a GNL distribution.

If $X \sim \mathrm{GNL}(\eta, \tau^2, a, b, \zeta)$ then we will say that the circular rv $\Theta = X \pmod{2\pi} \in [0, 2\pi)$ follows the *wrapped generalized normal–Laplace* (WGNL) distribution, and denote this by $\Theta \sim \mathrm{WGNL}(\eta, \tau^2, a, b, \zeta)$. It then follows (Mardia and Jupp 2000, p. 48) that the characteristic function of Θ has complex Fourier coefficients

$$\phi_p = \phi_{\mathrm{GNL}}(p) = \left[\frac{\exp(\mathrm{i}\eta p - \tau^2 p^2/2)}{(1 - \mathrm{i}ap)(1 + \mathrm{i}bp)}\right]^{\zeta}, \tag{2}$$

for $p = 0, \pm 1, \pm 2, \ldots$.

When $\zeta = 1$, the distribution of Θ is *wrapped normal–Laplace* (WNL). Using an obvious notation, we will denote this by $\Theta \sim \mathrm{WNL}(\eta, \tau^2, a, b)$. Given the relationships between the GNL, NL, and normal distributions, clearly $\mathrm{WGNL}(\eta, \tau^2, 0, 0, 1) \equiv \mathrm{WNL}(\eta, \tau^2, 0, 0) \equiv \mathrm{WN}(\eta, \tau^2)$, where WN is an abbreviation for 'wrapped normal'. Thus, both the WGNL and WNL families contain the wrapped normal distribution as a special case. The ability of the wrapped normal distribution to closely approximate the von Mises distribution is discussed at length in Pewsey and Jones (2005). In a similar fashion, the WNL and WGNL distributions contain the wrapped Laplace distribution and the wrapped generalized Laplace distributions as special cases ($\mathrm{WGNL}(0, 0, a, b, 1) \equiv \mathrm{WNL}(0, 0, a, b)$ and $\mathrm{WGNL}(0, 0, a, b, \zeta)$, respectively).

Stephens (1963) showed that the WN distribution can arise as the state of a Brownian motion on the circle after evolving for a fixed length of time. The wrapped versions of the ordinary normal–Laplace and the generalized normal–Laplace distributions have related interpretations, being the state of Brownian motion on the circle after certain random lengths of time. To see this, consider a particle following a Brownian motion on the circle with infinitesimal mean drift $\mu\,\mathrm{d}t$, infinitesimal variance $\sigma^2\,\mathrm{d}t$ and initial direction θ_0. The direction of the particle at time t has a wrapped normal distribution, with characteristic function

$$\phi_p = e^{i\theta_0 p}\exp\bigl(i\mu t p - (\sigma^2/2)tp^2\bigr). \tag{3}$$

Now suppose that the time for which the Brownian motion has been evolving is a rv. Specifically, assume that this time, T, is a constant plus a gamma-distributed rv. Thus,

$$T \stackrel{d}{=} t_0 + \frac{1}{\lambda}G, \tag{4}$$

where t_0 is a constant and G has a gamma distribution with unit scale parameter and shape parameter ζ, i.e., G has pdf $f_G(x) = x^{\zeta-1}e^{-x}/\Gamma(\zeta)$, where $x,\zeta > 0$. The characteristic function of the direction of the particle after the random time T can be found by integrating the characteristic function (3) with respect to the density $\lambda f_G(\lambda(t-t_0))$ over $t \in (x_0,\infty)$. After making the change of variable $s = t - t_0$ and integrating with respect to s over $(0,\infty)$ this yields the characteristic function

$$\phi_p = \exp\bigl(i(\theta_0 + \mu t_0)p - (\sigma^2/2)t_0 p^2\bigr)\left(\frac{\lambda}{\lambda - i\mu p + (\sigma^2/2)p^2}\right)^{\zeta},$$

which is exactly of the form (2) with

$$\eta = \frac{\theta_0 + \mu t_0}{\zeta} \pmod{2\pi}, \qquad \tau^2 = \frac{\sigma^2 t_0}{\zeta},$$

$$a = \sqrt{\left(\frac{\mu}{2\lambda}\right)^2 + \frac{\sigma^2}{2\lambda}} + \frac{\mu}{2\lambda}, \qquad b = \sqrt{\left(\frac{\mu}{2\lambda}\right)^2 + \frac{\sigma^2}{2\lambda}} - \frac{\mu}{2\lambda}.$$

Thus, the direction of the particle after the random time T follows a WGNL distribution.

We note that an alternative model to (4) for the evolution time of the Brownian motion, which produces similar results, is when a Brownian motion on the circle with initial direction following a wrapped normal distribution, $\mathrm{WN}(\mu_0,\sigma_0^2)$ say, evolves for a gamma-distributed (or exponentially-distributed) random time. This results in the same WGNL (or WNL) distribution (with μt_0 and $\sigma^2 t_0$ replaced by μ_0 and σ_0^2, respectively).

A number of special cases are of interest:

- $\zeta = 1$, i.e., the random component of the evolution time is *exponentially* distributed. In this case the state after time T follows an ordinary WNL distribution.
- $\mu = 0$, i.e., there is no mean drift in the circular Brownian motion. In this case the WGNL or WNL distribution is symmetric ($a = b = \sigma/\sqrt{2\lambda}$).

- $t_0 = 0$. In this case the time T follows a gamma distribution and the state after the random time T follows the wrapped generalized Laplace distribution (Kotz et al. 2001).
- $\zeta = 1$ and $t_0 = 0$. In this case the state at time T follows a wrapped Laplace distribution.

These representations of the WGNL and WNL distributions may well be useful in statistical modeling. For example, consider a particle or organism which when created has a fixed orientation, and then subsequently suffers random perturbations to its orientation (in a circular random walk or Brownian motion). Suppose this continues for a fixed time t_0 and possibly beyond that time until 'death' occurs, with hazard rate λ (i.e., the probability of death in an infinitesimal interval of length $\mathrm{d}t$ is $\lambda\,\mathrm{d}t + o(\mathrm{d}t)$). At the time of death, the particle's orientation will follow a WNL distribution. If instead of a constant hazard rate, λ, the hazard rate is that of a gamma distribution, the resulting distribution of the particle's orientation at death will be WGNL.

A related process is for a growing population of particles or organisms each of which has an orientation which evolves randomly over time. More precisely, suppose that the number of particles grows as a homogeneous birth process (i.e., assume any particle alive at t can give birth to a new particle in $(t, t + \mathrm{d}t)$ with probability $\lambda\,\mathrm{d}t + o(\mathrm{d}t)$) and that when particles are born their orientation follows a wrapped normal distribution. Assume that subsequent to birth the orientation of a particle evolves following a circular Brownian motion. As the time since the start of the homogeneous birth process tends to infinity, so the distribution of the time in existence of any randomly chosen particle from the current population tends to an exponential distribution (this follows from the properties of order statistic processes—see, e.g., Reed and Hughes (2007). Thus, the distribution of the orientation of particles over the whole population will tend to a WNL distribution with a and b related to the Brownian motion parameters and the birth intensity λ; and η and τ^2 the mean and variance parameters for the wrapped normal distribution of the initial orientation.

3 Properties of WGNL and WNL distributions

As stated in Sect. 2, the rv $\Theta \sim \mathrm{WGNL}(\eta, \tau^2, a, b, \zeta)$ has a characteristic function defined by the complex Fourier coefficients $\phi_p = \alpha_p + \mathrm{i}\beta_p = \phi_{\mathrm{GNL}}(p)$, for $p = 0, \pm 1, \pm 2, \ldots$, given in (2). The coefficients α_p and β_p are the pth cosine and sine moments, respectively, of Θ, i.e., $\alpha_p + \mathrm{i}\beta_p = \mathrm{E}[\cos(p\Theta)] + \mathrm{i}\mathrm{E}[\sin(p\Theta)]$. The coefficient ϕ_p can be represented in polar form as $\phi_p = \rho_p \mathrm{e}^{\mathrm{i}\mu_p}$, where $\rho_p \in [0, 1]$ is known as the pth *mean resultant length*, and μ_p as the pth *mean direction*. Thus,

$$\alpha_p = \rho_p \cos \mu_p, \qquad \beta_p = \rho_p \sin \mu_p. \tag{5}$$

In particular, $\mu = \mu_1$ is known as the *mean direction* and $\rho = \rho_1$ as the *mean resultant length*.

The probability density function (pdf) of a circular distribution can be represented in terms of its Fourier coefficients as

$$f(\theta) = \frac{1}{2\pi}\left[1 + 2\sum_{p=1}^{\infty}\left\{\alpha_p \cos(p\theta) + \beta_p \sin(p\theta)\right\}\right], \tag{6}$$

or, alternatively, using (5), as

$$f(\theta) = \frac{1}{2\pi}\left[1 + 2\sum_{p=1}^{\infty}\rho_p \cos(p\theta - \mu_p)\right]. \tag{7}$$

For the special case of the WNL distribution (WGNL with $\zeta = 1$),

$$\alpha_p = \frac{e^{-\tau^2 p^2/2}}{(1+a^2p^2)(1+b^2p^2)}\left[\left(1+abp^2\right)\cos(\eta p) + (b-a)p\sin(\eta p)\right] = \alpha_p^1, \quad \text{say,} \tag{8}$$

and

$$\beta_p = \frac{e^{-\tau^2 p^2/2}}{(1+a^2p^2)(1+b^2p^2)}\left[\left(1+abp^2\right)\sin(\eta p) - (b-a)p\cos(\eta p)\right] = \beta_p^1, \quad \text{say.} \tag{9}$$

For the more general WGNL distribution, one can write

$$\alpha_p = \left[\frac{e^{-\tau^2 p^2}}{(1+a^2p^2)(1+b^2p^2)}\right]^{\zeta/2}\cos\left(\zeta\mu_p^1\right), \tag{10}$$

$$\beta_p = \left[\frac{e^{-\tau^2 p^2}}{(1+a^2p^2)(1+b^2p^2)}\right]^{\zeta/2}\sin\left(\zeta\mu_p^1\right), \tag{11}$$

where μ_p^1 is the argument of $\alpha_p^1 + i\beta_p^1$. It follows that the pth mean resultant length of the WGNL distribution is

$$\rho_p = \left[\frac{e^{-\tau^2 p^2}}{(1+a^2p^2)(1+b^2p^2)}\right]^{\zeta/2} \tag{12}$$

and the pth mean direction is $\mu_p = \zeta\mu_p^1 \pmod{2\pi}$. The mean direction, μ, is given by

$$\mu = \mu_1 = \zeta\left[\eta + \tan^{-1}\left(\frac{a-b}{1+ab}\right)\right] \pmod{2\pi} \tag{13}$$

and the circular variance, ν, by

$$\nu = 1 - \rho = 1 - \left[\frac{e^{-\tau^2}}{(1+a^2)(1+b^2)}\right]^{\zeta/2}. \tag{14}$$

The circular standard deviation, defined as $\sigma = \{-2\log\rho_1\}^{1/2}$, reduces to $\sigma = \{\zeta[\tau^2 + \log(1+a^2) + \log(1+b^2)]\}^{1/2}$. Like the variance, it is increasing in ζ, τ^2, a, and b.

The pth cosine and sine moments about the mean direction μ, i.e., $\bar{\alpha}_p = \mathrm{E}[\cos(p(\Theta-\mu))]$ and $\bar{\beta}_p = \mathrm{E}[\sin(p(\Theta-\mu))]$, can be obtained by replacing η with $\tan^{-1}(\frac{a-b}{1+ab})$ in (8)–(11) above. The coefficients of circular skewness, s, and kurtosis, k, defined as $s = \bar{\beta}_2/(1-\rho)^{3/2}$ and $k = (\bar{\alpha}_2 - \rho^4)/(1-\rho)^2$, then follow directly. Note that if $a = b$, $\tan^{-1}(\frac{a-b}{1+ab}) = 0$, so that $s = 0$ and the distribution is symmetric.

To compute the pdf's of the WNL or WGNL distributions using (6) or (7) one must approximate their infinite sums in some way. The easiest seems to be to use a finite sum approximation, using sufficient terms so that the additional contribution of any omitted term is less than some specified tolerance. We note, from (12), that ρ_p decreases as p increases. So, in practice, one would continue summing up to the Nth term if the mean resultant length ρ_N was greater than the specified tolerance value but ρ_{N+1} was not. In practice, with the tolerance level set at 1×10^{-12}, the number of terms in the finite series approximation seldom exceeds 100, and is usually considerably less. The cases requiring the largest number of terms are those for which σ^2 and ζ are both small and the resulting pdf extremely peaked.

Figure 1 displays the shape of the WNL distribution (top row) and WGNL distribution with $\zeta = 0.7$ (second row). In both rows the left-hand and center panels show symmetric cases ($a = b$); for the left-hand panels $a = b = 0.1$ for various values of τ^2, and in the center panels τ^2 is fixed and $a = b$ varied. The right-hand panel shows the effect of changing a while b is kept fixed at zero. The second row is the same as the first except that ζ is set at 0.7. Note how decreasing ζ leads to greater peakedness and thinner flanks. Note also, in the right-hand panels, how a greater difference between a and b leads to greater skewness and at the same time a flatter distribution.

4 Maximum likelihood estimation

4.1 Independent exact observations

For a given set of independent exactly observed directions $\theta^T = (\theta_1, \theta_2, \ldots, \theta_n)$ from the WNGL distribution with parameters $(\eta, \tau^2, a, b, \zeta)$, one can compute the log-likelihood numerically to a given level of precision as $\ell(\eta, \tau^2, a, b, \zeta) = \sum_{i=1}^n \log f(\theta_i)$ using a finite sum approximation to (6) or (7). If ζ is set to unity, the log-likelihood $\ell(\eta, \tau^2, a, b)$ of the WNL distribution is produced. Numerical methods of optimization must then be used to maximize the log-likelihood over $(\eta, \tau^2, a, b) \in [0, 2\pi) \times \bar{\mathcal{R}}_-^3$ to obtain estimates for the WNL distribution; or over $(\eta, \tau^2, a, b, \zeta) \in [0, 2\pi) \times \bar{\mathcal{R}}_-^4$ for the WGNL distribution.

However, it is preferable to re-parameterize in terms of other quantities that are easier to estimate, and so as to avoid some of the high correlations that exist between the estimates of the original parameters. One possible re-parameterization results from using the mean direction, μ, and mean resultant length, ρ, given in (13) and (14), instead of η and τ^2. This simply involves substituting $\frac{\mu}{\zeta} - \tan^{-1}(\frac{a-b}{1+ab})$ (mod 2π)

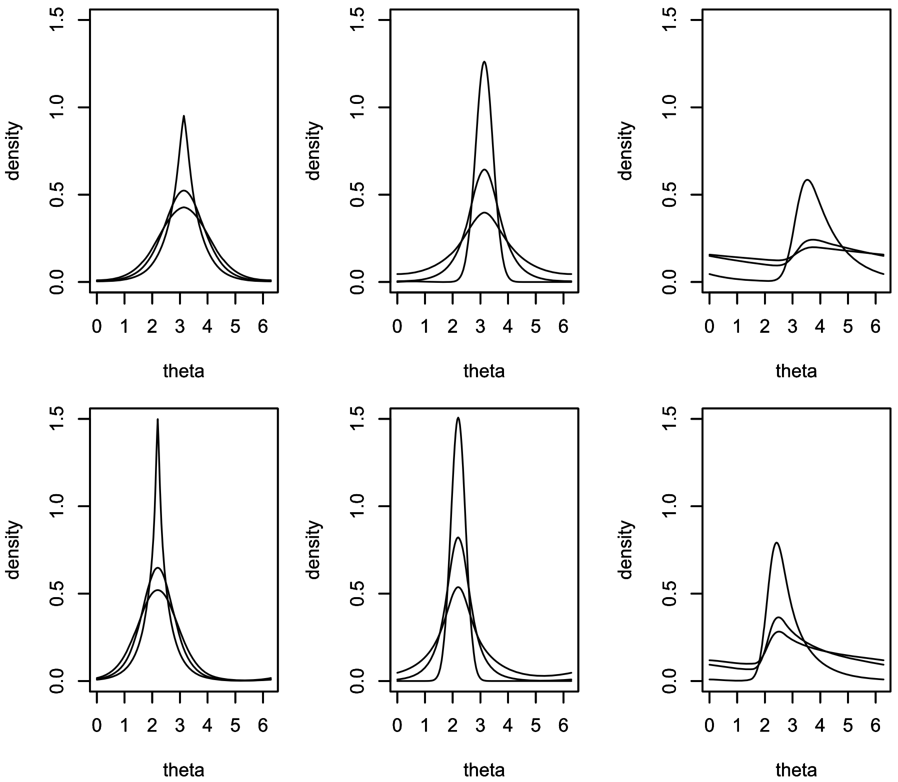

Fig. 1 Densities of wrapped normal–Laplace (WNL) (*top row*) and wrapped generalized normal–Laplace (WGNL) (*bottom row*) distributions for various parameter values. *In the top row* $\zeta = 1$, and *in second row* $\zeta = 0.7$. *The left-hand panels* show the effect of changing the parameter τ^2, with the other parameters kept constant ($a = b = 0.5$ and $\eta = \pi$). For *the three curves* (moving downwards) $\tau^2 = 0, 0.25$ and 0.5, respectively. The case $\tau^2 = 0$ corresponds to a wrapped Laplace distribution (*top row*) and wrapped generalized Laplace distribution (*second row*). *The center panels* show the effect of changing a and b while keeping them equal. *The three curves* (moving downward) correspond to $a = b = 0$ (wrapped normal), $a = b = 0.5$ and 1.0 (with $\eta = \pi$, $\tau^2 = 0.1$). Note how increasing a and b has the effect of flattening the distribution. *The right-hand panels* show the effect of increasing the difference between a and b. *The three curves* (moving downwards) correspond to $a = 1, b = 0$, $a = 5, b = 0$ and $a = 10, b = 0$ (with $\eta = \pi$, $\tau^2 = 0.1$). Note how the skewness and flatness of the distributions increase with the increased difference between a and b. In comparing *the top row* ($\zeta = 1$) and *bottom row* ($\zeta = 0.7$) note how the smaller value of ζ leads to taller peaks and thinner flanks, and also a move to the left of the mode

and $-\frac{2}{\zeta}\log\rho - \log(1 + a^2) - \log(1 + b^2)$ for η and τ^2, respectively, in the log-likelihood function. To accommodate the fact that μ is a circular variable, one can replace it by $\mu^* \in \mathcal{R}$ in the log-likelihood and then maximize over $(\mu^*, \rho, a, b, \zeta) \in \mathcal{R} \times [0, 1] \times \bar{\mathcal{R}}^3_-$. The MLE of μ, $\hat{\mu}$, is then given by $\hat{\mu}^*$ (mod 2π). The numerical optimization can be performed using standard routines such as those in optim in the R stats package, or the S-Plus routine nlminb. We have found the Nelder–Mead simplex algorithm (the default in optim) to be very dependable for this problem. Nevertheless, many other approaches are available which could be used to carry out the numerical optimization.

Before fitting the five-parameter WGNL distribution, we suggest firstly fitting the four-parameter WNL distribution, using the sample mean direction, $\bar{\theta}$, and sample mean resultant length, $\bar{R}$, as starting values for μ and ρ, and arbitrary starting values for a and b. Indeed to accommodate the possibility of multiple local maxima, we recommend the use of a variety of starting values for a and b. Following this, the five-parameter WGNL distribution can be fitted using the maximum likelihood estimates (MLE's) of μ, ρ, a and b for the fitted WNL distribution, and $\zeta = 1$, as starting values.

Asymptotic standard errors and correlation coefficients for the MLE's of μ, ρ, a, b, and ζ can be obtained in the usual way from the observed information matrix. The latter can be calculated using the inverse of the Hessian matrix (produced by optim when the option "hessian = TRUE" is specified) evaluated at the maximum.

4.2 Grouped data

If the data consist of independent observations grouped into N cells given by the half-open intervals $[0, \theta_{(1)})$, $[\theta_{(1)}, \theta_{(2)}), \ldots, [\theta_{(N-1)}, 2\pi)$ with observed frequencies $f_1, f_2, \ldots, f_N$, the log-likelihood is of multinomial form $\ell = \sum_{j=1}^{N} f_j \log \phi_j$, where $\phi_j = F(\theta_{(j)}) - F(\theta_{(j-1)})$ and F is the cumulative distribution function (cdf) of the WGNL distribution. The latter is given by

$$F(\theta) = \frac{1}{2\pi}\left[\theta + 2\sum_{p=1}^{\infty} \frac{\rho_p}{p}\left\{\sin\left[p(\theta - \eta) - \mu_p\right] + \sin[p\eta + \mu_p]\right\}\right],$$

and can be computed to any given tolerance level using a finite sum approximation in the same way as described above for computing the pdf, $f(\theta)$. Again, it proves worthwhile to re-parameterize as described in Sect. 4.1 and then maximize the log-likelihood over the transformed parameters.

5 Examples

5.1 Exact observations

As an example of independent exact observations we use the data, reported in Bruderer and Jenni (1990), of $n = 1827$ flight headings, a "heading" being the direction of a bird's body measured clockwise from north. A histogram of the headings is presented in Fig. 2. The large-sample test of Pewsey (2002) emphatically rejects underlying reflective symmetry, with a p-value of 0.000. The sample mean direction and mean resultant length are $\bar{\theta} = 3.923$ (radians) and $\bar{R} = 0.737$, respectively.

Pewsey (2008) fitted several parametric models to these data: the wrapped stable distribution, a two component mixture of von Mises distributions, and a mixture with circular uniform and wrapped skew-normal components. Here we fit the WNL and WGNL distributions and compare their fits with those obtained by Pewsey.

Fitting the four-parameter WNL distribution using the (μ, ρ, a, b) parameterization, $\bar{\theta} = 3.923$ and $\bar{R} = 0.737$ as starting values for μ and ρ and 1 as the

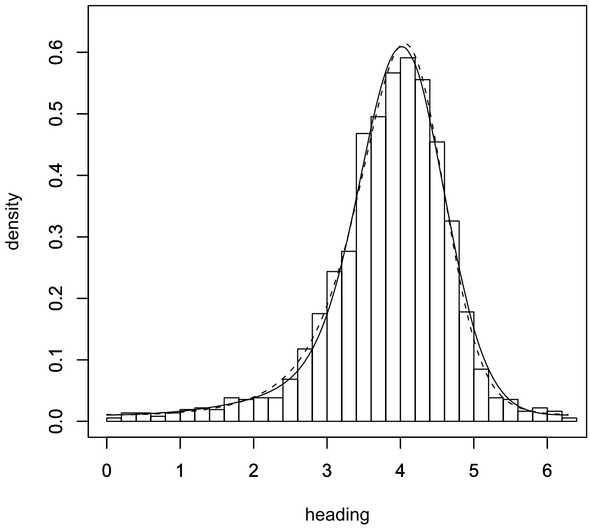

Fig. 2 Histogram of 1827 bird-flight "headings" (direction measured clockwise from north, in radians) of migrating birds. Also shown are the densities of the maximum likelihood fits for the WGNL (*solid line*) and wrapped stable (*broken line*) distributions

starting value for both a and b, leads to the MLE's $\hat{\mu} = 3.913$ (0.0177), $\hat{\rho} = 0.734$ (0.0091), $\hat{a} = 0.369$ (0.0319), $\hat{b} = 0.696$ (0.0326), with a log-likelihood of -2137.08. The figures in brackets are asymptotic standard errors computed from the observed information matrix. The asymptotic correlations between the estimates are all quite small except for that between $\hat{\rho}$ and $\hat{b}$ (-0.73). The Pearson chi-square goodness-of-fit statistic, calculated using 18 class intervals each of width 20 degrees, has a value of 30.56 on 13 degrees of freedom (p-value = 0.004), indicating an inadequate fit.

Next, fitting the WGNL distribution using the (μ, ρ, a, b, ζ) parameterization and the MLE's for the WNL fit and $\zeta = 1$ as starting values, leads to the MLE's $\hat{\mu} = 3.921$ (0.0172), $\hat{\rho} = 0.736$ (0.0095), $\hat{a} = 0.487$ (0.1849), $\hat{b} = 2.200$ (1.2080), $\hat{\eta} = 0.166$ (0.0913). All the asymptotic correlations involving $\hat{\mu}$ and $\hat{\rho}$ are small. Those between $\hat{a}, \hat{b}$ and $\hat{\zeta}$ are larger, with that between $\hat{b}$ and $\hat{\zeta}$ being particularly strong and negative (-0.98). The value of the log-likelihood for this fit is -2129.83 and the Pearson chi-square goodness-of-fit statistic has a value of 13.40 on 12 degrees of freedom (p-value = 0.341). This suggests a very good fit, particularly considering the sample size involved.

For these data, the fit of the WGNL distribution is not as good (as measured by maximized log-likelihood) as that of the four-parameter wrapped stable distribution (max $\ell = -2127.73$), nor that of a four-parameter mixture with circular uniform and skew-normal components (max $\ell = -2128.03$). However it is better than the fit of a five-parameter mixture of two von Mises distributions (max $\ell = -2130.10$) (see Pewsey 2008). Given these log-likelihood results and the number of parameters of the various models, the wrapped stable distribution would be identified as the superior model using any of the standard information criteria such as AIC, BIC, etc. The pdf of the WGNL fit is shown superimposed on the histogram of the data, together with the density of the wrapped stable fit, in Fig. 2.

Computing time was not found to be a major concern. Using the R routine optim with the tolerance for computing the pdf using a finite sum (Sect. 3) set at 1×10^{-12},

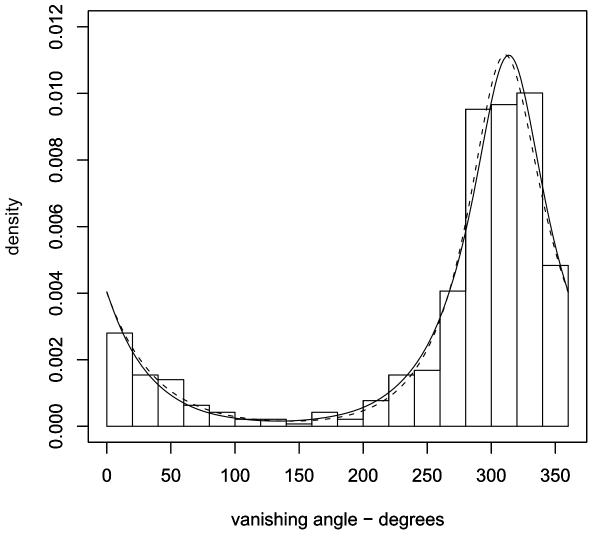

Fig. 3 Histogram of the "vanishing angles" of 714 mallard ducks (grouped in twenty degree segments). Also shown are the densities of the maximum likelihood fits for the WNL (*broken line*) and symmetric WNL (*solid line*) distributions

about two minutes of computing time on a desktop PC with a 2.66 GHz processor was required to obtain the MLEs for the five-parameter WGNL model.

5.2 Grouped data

As an example of fitting our new models to grouped data, we consider the data on the 714 vanishing angles of mallard ducks presented in Mardia and Jupp (2000, p. 3; histogram, p. 4). The MLE's for the four-parameter WNL distribution, using the (μ, ρ, a, b) parameterization, are $\hat{\mu} = 5.486$ (0.0290), $\hat{\rho} = 0.722$ (0.0154), $\hat{a} = 0.651$ (0.0484), $\hat{b} = 0.532$ (0.0453), with a log-likelihood value of -1618.96. The largest asymptotic correlation between the MLE's is that between $\hat{\rho}$ and $\hat{a}$ (0.62). For these data $\bar{\theta} = 5.481$ (radians) and $\bar{R} = 0.726$, which agree closely with the MLE's $\hat{\mu}$ and $\hat{\rho}$ for the WNL fit. The Pearson chi-square goodness-of-fit statistic has a value of 20.49 on 13 degrees of freedom (p-value $= 0.084$).

Fitting the five-parameter WGNL distribution leads to a maximized log-likelihood of -1617.99, with convergence being obtained in under a minute. There appears to be some 'flatness' of the log-likelihood around the maximum, since different starting values lead to convergence to different points, but all with similar log-likelihood values. Furthermore the likelihood ratio test of $\zeta = 1$, using the usual asymptotic chi-square approximation, leads to a p-value of 0.16. This suggests that the WGNL is an over-parameterized model for these data and that the WNL distribution provides an adequate fit.

Given the similarity of the estimates for a and b (when considered together with their asymptotic standard errors) it is of interest to investigate the hypothesis of an underlying symmetric WNL population (i.e., with $a = b$). Fitting such a distribution to these data results in the estimates $\hat{\mu} = 5.474$ (0.0281), $\hat{\rho} = 0.759$ (0.0260), $\hat{a} = \hat{b} = 0.593$ (0.0290), and a log-likelihood value of -1620.25. All of the asymptotic correlations between the estimates are small except that between $\hat{\rho}$ and $\hat{a}$ (0.86). The likelihood ratio test of $a = b$ has a p-value of 0.108. This compares with a p-value of 0.124 for the large-sample test of underlying reflective symmetry of Pewsey (2002) (without any allowance for grouping). These results provide support for the hypothesis of underlying symmetry, and indicate that the most parsimonious fit for these data is that of the symmetric WNL distribution. However, the Pearson chi-square goodness-of-fit statistic for this model is 22.55 on 14 degrees of freedom (p-value = 0.068), suggesting a barely adequate fit. The ambiguity as to which of the 4-parameter WNL model or the 3-parameter symmetric WNL model is superior, is reflected in the AIC values (3245.92 and 3246.50) and the BIC values (3264.20 and 3260.21) for the two models. The more stringent BIC criterion supports the superiority of the symmetric WNL model whereas AIC identifies the four-parameter WNL model as being marginally better. The pdf's of the maximum likelihood fits for the WNL and symmetric WNL distributions are shown superimposed on the histogram of the data in Fig. 3.

6 Conclusions

In this article we have introduced two new families of unimodal circular distributions obtained by wrapping the normal–Laplace and generalized normal–Laplace distributions onto the unit circle. In Sect. 2, a simple stochastic model which can lead to the WNL and WGNL distributions was outlined. One version of this involves a growing population of objects each of which has a property of orientation (corresponding to an angle between 0 and 2π radians) which evolves in a random way. Under certain assumptions, it can be shown that the distribution of orientations over the population after a long time has passed should be approximately WNL. The assumptions are that: (i) the population grows exponentially (in expectation) as a homogeneous birth process; (ii) when new objects are born their initial orientation follows a wrapped normal distribution; (iii) from birth onwards, orientations of objects evolve independently as Brownian motion on the circle. Whether this construction will prove useful in physics, biology, or any other area of application remains to be seen. We, nevertheless, hope that the two new families of distributions will find use as models for circular data, whether produced by this or a related process or otherwise, that the existing distributions proposed in the literature are incapable of modeling.

References

Bruderer B, Jenni L (1990) Migration across the Alps. In: Gwinner E (ed) Bird migration: physiology and ecophysiology. Springer, Berlin, pp 60–70

Fisher NI (1993) Statistical analysis of circular data. Cambridge University Press, Cambridge

Gatto R, Jammalamadaka SR (2003) Inference for wrapped symmetric alpha-stable circular models. Sankhya 65:333–355

Jammalamadaka SR, SenGupta A (2001) Topics in circular statistics. World Scientific, Singapore

Jammalamadaka SR, Kozubowski TJ (2003) A new family of circular models: the wrapped Laplace distributions. Adv Appl Stat 3:77–103

Jammalamadaka SR, Kozubowski TJ (2004) New families of wrapped models for modeling skew circular data. Commun Stat Theory Methods 33:2059–2074

Kotz S, Kozubowski TJ, Podgórski K (2001) The Laplace distribution and generalizations. Birkhäuser, Boston

Mardia KV, Jupp PE (2000) Directional statistics. Wiley, Chichester

Pewsey A (2000) The wrapped skew-normal distribution on the circle. Commun Stat Theory Methods 29:2459–2472

Pewsey A (2002) Testing circular symmetry. Can J Stat 30:591–600

Pewsey A (2006) Modelling asymmetrically distributed circular data using the wrapped skew-normal distribution. Environ Ecol Stat 13:257–269

Pewsey A (2008) The wrapped stable family of distributions as a flexible model for circular data. Comput Stat Data Anal 52:1516–1523

Pewsey A, Jones MC (2005) Discrimination between the von Mises and wrapped normal distributions: just how big does the sample size have to be? Statistics 39:81–89

Reed WJ (2007) Brownian–Laplace motion and its application in financial modeling. Commun Stat Theory Methods 36:473–484

Reed WJ, Hughes BD (2007) Theoretical size distribution of fossil taxa: analysis of a null model. Theor Biol Med Model 4:12

Reed WJ, Jorgensen M (2004) The double Pareto–lognormal distribution—a new parametric model for size distribution. Commun Stat Theory Methods 33:1733–1753

Stephens MA (1963) Random walk on a circle. Biometrika 50:385–390