Geodesic Normal distribution on the circle

Jean-François Coeurjolly[1,2] and Nicolas Le Bihan[2,3]
[1] Laboratoire Jean Kuntzmann, Grenoble University, France.
[2] GIPSA-lab, Grenoble University, France, [3] CNRS, France.

December 21, 2010

Abstract

This paper is concerned with the study of a circular random distribution called geodesic Normal distribution recently proposed for general manifolds. This distribution, parameterized by two real numbers associated to some specific location and dispersion concepts, looks like a standard Gaussian on the real line except that the support of this variable is $[0, 2\pi)$ and that the Euclidean distance is replaced by the geodesic distance on the circle. Some properties are studied and comparisons with the von Mises distribution in terms of intrinsic and extrinsic means and variances are provided. Finally, the problem of estimating the parameters through the maximum likelihood method is investigated and illustrated with some simulations.

Circular statistics deal with random variables taking values on hyperspheres and can be included in the broader field of directional statistics. Applications of circular statistics are numerous and can be found, for example, in fields such as climatology (wind direction data [MJ00]), biology (pigeons homing performances [Wat83]) or earth science (earthquake locations occurence and other data types, see [MJ00] for examples) among others.

A circular distribution is a probability distribution function (pdf) which mass is concentrated on the circumference of a unit circle. The support of a random variable θ representing an angle measured in radians may be taken to $[0, 2\pi)$ or $[-\pi, \pi)$. We will focus here on continuous circular distributions, that is on absolutely continuous (w.r.t. the Lebesgue measure on the circumference) distributions. A pdf of a circular random variable has to fulfill the following axioms

(i) $f(\theta) \geq 0$.

(ii) $\int_0^{2\pi} f(\theta) d\theta = 1$

(iii) $f(\theta) = f(\theta + 2k\pi)$ for any integer k (i.e. f is periodic).

Among many models of circular data, the von Mises distribution plays a central role (essentially due to the similarities shared with the Normal distribution on the real line). A circular random variable (for short r.v.) θ is said to have a von Mises distribution, denoted by $vM(\mu, \kappa)$, if it has the density function

$$f(\theta; \mu, \kappa) = \frac{1}{2\pi I_0(\kappa)} e^{\kappa \cos(\theta - \mu)},$$

where $\mu \in [0, 2\pi)$ and $\kappa \geq 0$ are parameters and where $I_0(\kappa)$ is the modified Bessel function of order 0. The aim of this paper is to review some properties of another circular distribution introduced by [Pen06] which also shares similarities with the Normal distribution on the real line. For some $\mu \in [0, 2\pi)$ and for some parameter $\gamma \geq 0$, a r.v. θ is said to have a geodesic Normal distribution denoted in the following $gN(\mu, \gamma)$, if it has the density function

$$f(\theta; \mu, \gamma) = k^{-1}(\gamma) e^{-\frac{\gamma}{2} d_G(\mu, \theta)^2},$$

where $d_G(\cdot, \cdot)$ is the geodesic distance on the circle and where $k(\gamma)$ is the normalizing constant defined by

$$k(\gamma) := \sqrt{\frac{2\pi}{\gamma}} \left(\Phi(\pi\sqrt{\gamma}) - \Phi(-\pi\sqrt{\gamma}\right) = \sqrt{\frac{2\pi}{\gamma}} erf\left(\pi\sqrt{\frac{\gamma}{2}}\right),$$

where Φ is the cumulative distribution function of a standard Gaussian random variable and where $erf(\cdot)$ is the error function.
Let us underline that Pennec [Pen06] introduced the geodesic Normal distribution for general Riemannian manifolds. We focus here on a special (and simple) manifold, the circle, in order to highlight its basic properties and compare them with the most classical circular distribution, namely the $vM(\mu, \kappa)$ distribution. That is, we present here a new study of the gN distribution in the framework of circular statistics and provide results in terms of estimation and asymptotic behaviour.

One of the main conclusions of this paper may be summarized as follows. While the von Mises distribution has strong relations with the notion of extrinsic moments (that is with trigonometric moments), we will emphasize, in this paper, that the geodesic Normal distribution has strong relations with intrinsic moments that is with the Fréchet mean (defined as the angle α minimizing the expectation of $d_G(\alpha, \theta)^2$) and the geodesic variance. vM and gN distributions definition are closely related to respectively extrinsic and intrinsic moments, and we present their similarities together with dissimilarities.

After introducing the gN distribution in Section 1, we present a brief review on intrinsic and extrinsic quantities that allow characterization of distributions on the circle in Section 2. In Section 3, we present extrinsic and intrinsic properties of vM and gN distributions. Then, in Section 4, we rapidly explain how to simulate gN random variables on the circle. Finally, in Section 5, we present the maximum likelihood estimators for the gN distributions and study their asymptotic behaviour. Numerical simulations illustrate the presented results.

1 Geodesic Normal distribution through the tangent space

As introduced in [Pen06] the geodesic Normal distribution is defined for random variables taking values on Riemannian manifolds and is based on the "geodesic distance" concept [BP03, BP05, Pen06]. On a Riemannian manifold $\mathcal{M}$, one can define at each point $x \in \mathcal{M}$ a scalar product $\langle .,. \rangle_x$ in the tangent plane $T_x\mathcal{M}$ attached to the manifold at x. On $\mathcal{M}$, among the possible smooth curves between two points x and y, the curve of minimum length is called a *geodesic*. The length of the curve is understood as integration of the norm of its instantaneous velocity along the path, and with the norm at position x on $\mathcal{M}$ taken as: $\|.\| = \sqrt{\langle .,. \rangle_x}$. It is well-known that, given a point $x \in \mathcal{M}$ and a vector $\overrightarrow{v} \in T_x\mathcal{M}$, there exists only one geodesic $\gamma(t)$ with $\gamma(t=0) = x$ and with tangent vector $\overrightarrow{x}$.

Through the exponential map, to each vector $\in T_x\mathcal{M}$ is associated to a point $y \in \mathcal{M}$ reached in unit time, *i.e.* $\gamma(1) = y$. Using the notation adopted in [Pen06], the vector defined in $T_x\mathcal{M}$ associated to the geodesic that passes through x at time $t = 0$ and reaches y at time $t = 1$ is denoted $\overrightarrow{xy}$. Thus, the exponential map (at point x) maps a vector of T_xM to a point y, *i.e.* $y = \exp_x(\overrightarrow{xy})$. Now the *geodesic distance*, denoted $d_G(x,y)$, between $x \in \mathcal{M}$ and $y \in \mathcal{M}$ is:

$$d_G(x,y) = \sqrt{\langle \overrightarrow{xy}, \overrightarrow{xy} \rangle_x}$$

The Log map is the inverse map that associates to a point $y \in \mathcal{M}$ in the neighbourhood of x a vector $\overrightarrow{xy} \in T_x\mathcal{M}$, *i.e.* $\overrightarrow{xy} = Log_x(y)$.

A random variable Y taking values in $\mathcal{M}$, with density function $f(y;\mu,\Gamma)$ is said to have a geodesic Normal distribution with parameters $\mu \in \mathcal{M}$ and Γ, a (N,N) matrix, denoted by $gN(\mu,\Gamma)$, if:

$$f(y;\mu,\Gamma) = k^{-1} \exp\left(-\frac{\overrightarrow{\mu y}^T.\Gamma.\overrightarrow{\mu y}}{2}\right).$$

The parameter μ is related to some specific location concept. Namely, [Pen06] has proved that μ corresponds to the intrinsic or Fréchet mean of the random variable Y (see Section 2.2 for more details). The normalizing constant is given by:

$$k = \int_{\mathcal{M}} \exp\left(-\frac{\overrightarrow{\mu y}^T.\Gamma.\overrightarrow{\mu y}}{2}\right) d\mathcal{M}(y)$$

where $d\mathcal{M}(y)$ is the Riemannian measure (induced by the Riemannian metric). The matrix Γ is called the concentration matrix and is related to the covariance matrix of the vector $\overrightarrow{\mu Y}$ given for random variables on a Riemannian manifold by

$$\Sigma := E[\overrightarrow{\mu Y}\overrightarrow{\mu Y}^T] = k \int_{\mathcal{M}} \overrightarrow{\mu y}.\overrightarrow{\mu y}^T \exp\left(-\frac{\overrightarrow{\mu y}^T.\Gamma.\overrightarrow{\mu y}}{2}\right) d\mathcal{M}(y).$$

Note that in the case where $\mathcal{M} = \mathbb{R}^d$, the manifold is flat and the geodesic distance is nothing more than the Euclidian distance. In this case, we retrieve the classical definition of a Gaussian variable in $\mathbb{R}^d$ with μ and Γ corresponding respectively to the classical expectation and to the inverse of the covariance matrix Σ^{-1}.

The case of the circle We now present, as done in [Pen06], the case where $\mathcal{M}$ is the unit circle $\mathcal{S}^1$. The exponential chart is the angle $\theta \in]-\pi;\pi[$. Note that this chart is "local" and is thus defined at a point on the manifold. This must be kept in mind especially for explicit calculation. Here, as the tangent plane to $\mathcal{S}^1$ is the real line, θ takes values on the segment between $-\pi$ and π. Note that π and $-\pi$ are omitted as they are in the cut locus of the development point (point on the manifold where the tangent plane is attached). $d\mathcal{M}$ is simply $d\theta$ here.

As stated before, the gN distribution on the circle has density function given for $\theta \in (\mu-\pi, \mu+\pi)$ by:

$$f(\theta;\mu,\gamma) = k^{-1}(\gamma)e^{-\frac{\gamma}{2}d_G(\mu,\theta)^2},$$

where γ is a nonnegative real number. Note that $d_G(\mu,\theta)$ is the arc length between μ and θ. The normalization $k(\gamma)$ is:

$$k(\gamma) = \int_{\mu-\pi}^{\mu+\pi} e^{-\frac{\gamma}{2}d_G(\mu,\theta)^2}d\theta = \sqrt{\frac{2\pi}{\gamma}}erf\left(\pi\sqrt{\frac{\gamma}{2}}\right),$$

with the development made around μ.

In order to consider a gN distribution as a classical circular distribution (defined by axioms (i)-(iii) in the introduction), one must extend the support of this distribution from $(\mu-\pi,\mu+\pi)$ to $\mathbb{R}$. The way to achieve this is to make the geodesic distance periodic. Let us consider the distance $\tilde{d}_G$ for an angle $\alpha \in \mathbb{R}\setminus\{\mu+k\pi, k\in\mathbb{Z}\}$ defined by $\tilde{d}_G(\mu,\alpha) = d_G(\mu,\tilde{\alpha})$ with $\tilde{\alpha} = (\alpha-\mu+\pi)(mod\, 2\pi)$. Let $\tilde{f}$ the density f where d_G is replaced by $\tilde{d}_G$. This new density defines a circular disribution satisfying axioms (i)-(iii). In particular,

$$\int_0^{2\pi}\tilde{f}(\theta,\mu,\gamma)d\theta = \int_0^{2\pi}k^{-1}(\gamma)e^{-\frac{\gamma}{2}\tilde{d}_G(\mu,\theta)^2}d\theta = \int_{\mu-\pi}^{\mu+\pi}k^{-1}(\gamma)e^{-\frac{\gamma}{2}d_G(\mu,\theta)^2}d\theta = 1.$$

For the sake of simplicity, d_G will be understood as the distance $\tilde{d}_G$ in the rest of the paper. And therefore, the density of a gN distribution is considered periodic, defined on $\mathbb{R}\setminus\{\mu+k\pi, k\in\mathbb{Z}\}$ and with values on $\mathbb{R}$. In the case of a vM distribution (and actually for most of circular distributions) no such considerations are needed; the periodic nature being included through the $\cos(\cdot)$ function.

2 Measures of location and dispersion for circular random variables

We briefly present the concepts of *extrinsic* and *intrinsic* moments for random variables on the circle. While the former are well-known in circular statistics (*e.g.* [MJ00]), the later are based on the geodesic distance on the circle.

2.1 Extrinsic moments

In circular statistics, it is well established that trigonometric moments give access to measures of *mean direction* and *circular variances*. Considering a cicular random variable θ, its $p-th$ order trigonometric moment is defined as:

$$\varphi_p = E[e^{ip\theta}] = \alpha_p + i\beta_p$$

where $\alpha_p = E[\cos p\theta]$ and $\beta_p = E[\sin p\theta]$. These later quantities are *extrinsic* by definition[1]. The first order trigonometric moment is thus:

$$\varphi_1 = E[e^{i\theta}] = \rho e^{i\mu^E}$$

where ρ is called the *mean resultant length* ($0 \leq \rho \leq 1$) and μ^E is the *mean direction*. In the following, we refer to μ^E as the *extrinsic mean*. The *extrinsic variance* σ_E^2 is indeed the circular variance defined as:

$$\sigma_E^2 = 1 - \rho. \tag{1}$$

In the sequel, *extrinsic moments* will be used in place of *trigonometric moments*, keeping in mind that they are the same quantities.

For more details on trigonometric moments see [MJ00, JAS01].

2.2 Intrinsic moments

Another way to consider moments for distributions of random variables on the circle is to use the fact that $\mathcal{S}^1$ is a Riemannian manifold and thus the geodesic distance can be used to define *intrinsic moments*, namely the intrinsic mean μ^I and variance σ_I^2. Given a random variable θ with values on $\mathcal{S}^1$, its intrinsic or Fréchet mean μ^I is:

$$\mu^I = \text{argmin}_{\tilde{\mu} \in \mathcal{S}^1} E[d_G(\tilde{\mu}, \theta)^2], \tag{2}$$

with $d_G(.,.)$ the geodesic distance on the circle, *i.e.* the arc length. From the defnition, one understands that the intrinsic mean is the counterpart of the mean in Euclidian space, except that the distance used here is no more the Euclidian distance, but the geodesic one.

The *intrinsic variance* σ_I^2 is also naturally defined for a random variable on the circle using the geodesic distance in the following way:

$$\sigma_I^2 = E\left[d_G(\mu^I, \theta)^2\right], \tag{3}$$

with μ^I the *intrinsic mean* defined above. Note that, once again, this definition is consistent with the “classical” definition of the variance of a random variable on the real, with the difference that the natural distance to use on the circle is the geodesic distance. For more details and a thourough study of intrinsic statistics for random variables on Riemannian manifolds, see [Pen06] or [BP03, BP05].

In the next section, we will concentrate on first and second order statistics (intrinsic and extrinsic) of the gN and vM distributions.

[1] We denote by *extrinsic* any quantity that is not computed explicitly on the circle.

3 Basic properties of the $gN(\mu,\gamma)$ and $vM(\mu,\kappa)$ distributions

Symmetry property

First, let us say that like the vM distribution, the gN distribution has a mode for $\theta = \mu$, a symmetry around $\theta = \mu$. The vM distribution has an anti-mode for $\theta = \mu \pm \pi$. For a gN distribution, the density is not defined at these points. However, the shared behaviour is the decreasing of both densities on each interval $(\mu - \pi, \mu]$ and $[\mu, \mu + \pi)$

Measures of location and dispersion

The following proposition gives simple conditions on the density of a circular distribution ensuring the existence and unicity of the Fréchet mean.

Proposition 1 (Kaziska and Srivastava [KS08])**.** *Let f be the density of a circular pdf satisfying $f(\mu+\theta) = f(\mu-\theta)$ for $\theta \in [0,\pi)$ and such that $f(\theta)$ is differentiable for $\theta = \mu$ and $f'(\theta) > 0$ for $\theta \in (\mu-\pi,\mu)$ and $f'(\theta) < 0$ for $\theta \in (\mu, \mu+\pi)$, then the pdf has a unique Fréchet mean at μ.*

In particular, and as a consequence of this result, it is easily verified that the Fréchet mean of a $vM(\mu,\kappa)$ or a $gN(\mu,\gamma)$ is unique and corresponds to the parameter μ. Of course, this was already established by [Pen06] for the $gN(\mu,\gamma)$ by an entropy argument (see the concerned paragraph above, p.8). As reviewed in Section 2, the classical way to characterize a circular random variable is the computation of its trigonometric moments. These are stated below for both distributions.

Proposition 2. *The p-th trigonometric moment ($p \in \mathbb{N}^*$) of a circular random variable θ, denoted by φ_p is defined by $\varphi_p := E\left[e^{ip\theta}\right]$.*
(i) For $\theta \sim vM(\mu,\kappa)$, then

$$\varphi_p = e^{ip\mu}\,\frac{I_p(\kappa)}{I_0(\kappa)},$$

where $I_p(\cdot)$ is the modified Bessel function of the first kind of order p.
(ii) For $\theta \sim gN(\mu,\gamma)$, then

$$\varphi_p = e^{ip\mu}\,e^{-\frac{p^2}{2\gamma}}\,\frac{Re\left(erf\left(\pi\sqrt{\frac{\gamma}{2}} - i\frac{p}{\sqrt{2\gamma}}\right)\right)}{erf\left(\pi\sqrt{\frac{\gamma}{2}}\right)},$$

where erf is the error function defined for any complex number by $erf(z) = \frac{2}{\sqrt{\pi}}\int_{-\infty}^{z} e^{-t^2}\,dt$.

Proof. (i) See *e.g.* [MJ00]. (ii) Let $p \geq 1$,

$$E[e^{ip\theta}] = k^{-1}(\gamma)\int_0^{2\pi} e^{ip\theta}e^{-\frac{\gamma}{2}d_G(\theta,\mu)^2}d\theta = k^{-1}(\gamma)\int_{\mu-\pi}^{\mu+\pi} e^{ip\theta}e^{-\frac{\gamma}{2}d_G(\theta,\mu)^2}d\theta.$$

Since for $\theta \in (\mu - \pi, \mu + \pi)$, $d_G(\theta, \mu) = |\theta - \mu|$,

$$\begin{aligned} E[e^{ip\theta}] &= k^{-1}(\gamma) \int_{-\pi}^{\pi} e^{ip(\theta+\mu)} e^{-\frac{\gamma}{2}\theta^2} d\theta \\ &= e^{ip\mu} k^{-1}(\gamma) \int_{-\pi}^{\pi} e^{-\left(\left(\theta\sqrt{\frac{\gamma}{2}} - i\frac{p}{\sqrt{2\gamma}}\right)^2 - \left(i\frac{p}{\sqrt{2\gamma}}\right)^2\right)} d\theta \\ &= e^{ip\mu} e^{-\frac{p^2}{2\gamma}} k^{-1}(\gamma) \sqrt{\frac{2\pi}{\gamma}} \frac{1}{2} \left(erf\left(\pi\sqrt{\frac{\gamma}{2}} - i\frac{p}{\sqrt{2\gamma}}\right) - erf\left(-\pi\sqrt{\frac{\gamma}{2}} - i\frac{p}{\sqrt{2\gamma}}\right)\right) \\ &= e^{ip\mu}\, e^{-\frac{p^2}{2\gamma}} \frac{Re\left(erf\left(\pi\sqrt{\frac{\gamma}{2}} - i\frac{p}{\sqrt{2\gamma}}\right)\right)}{erf\left(\pi\sqrt{\frac{\gamma}{2}}\right)}, \end{aligned}$$

since for any complex number z, $erf(z) - erf(-\overline{z}) = erf(z) - \overline{erf(-z)} = erf(z) + \overline{erf(z)} = 2Re(erf(z))$. ■

A consequence of Proposition 2 is that the extrinsic mean of a $gN(\mu, \gamma)$ is μ and, according to the definition 1, the extrinsic variance equals $1 - e^{-\frac{1}{2\gamma}} \frac{Re\left(erf\left(\pi\sqrt{\frac{\gamma}{2}} - i\frac{1}{\sqrt{2\gamma}}\right)\right)}{erf\left(\pi\sqrt{\frac{\gamma}{2}}\right)}$. To complete the comparison, let us compute the intrinsic variances of a gN and vM distribution.

Proposition 3. *(i) For $\theta \sim gN(\mu, \gamma)$,*

$$\begin{aligned} \sigma_I^2(\gamma) &= \frac{1}{\gamma}\left(1 - 2\pi k^{-1}(\gamma) e^{-\frac{\gamma\pi^2}{2}}\right) = -2\,\frac{k'(\gamma)}{k(\gamma)} \\ &= \frac{1}{\gamma} - \frac{\frac{2\pi}{\sqrt{\gamma}}\varphi(\pi\sqrt{\gamma})}{2\Phi(\pi\sqrt{\gamma}) - 1}, \end{aligned}$$

where we recall that $k(\gamma) = \sqrt{\frac{2\pi}{\gamma}}\left(\Phi(\pi\sqrt{\gamma}) - \Phi(-\pi\sqrt{\gamma})\right) = \sqrt{\frac{2\pi}{\gamma}}\, erf\left(\pi\sqrt{\frac{\gamma}{2}}\right)$.
(ii) For $\theta \sim vM(\mu, \kappa)$,

$$\sigma_I^2(\kappa) = \frac{1}{2\pi I_0(\kappa)} \int_{-\pi}^{\pi} \alpha^2 e^{\kappa\cos(\alpha)} d\alpha.$$

Proof. The case of a geodesic Normal distribution has already been considered by [Pen06]. For a $vM(\mu, \kappa)$, since the Fréchet mean is μ, we have to compute

$$E[d_G(\theta, \mu)^2] = \frac{1}{2\pi I_0(\kappa)} \int_{\mu-\pi}^{\mu+\pi} (\alpha - \mu)^2 e^{\kappa\cos(\alpha-\mu)} d\alpha,$$

which is the stated result. ■

Table 3 summarizes extrinsic and intrinsic means and variances for both distributions and Figure 3 shows the evolutions of the extrinsic and intrinsic variances in terms of the concentration parameter κ for the vM and γ for the gN. It is interesting to notice that the von Mises distribution and the geodesic Normal distributions have intrinsic variance equal to $\frac{\pi^2}{3}$ when γ or κ equals zero, corresponding to the variance of the uniform distribution on the circle. Note also that intrinsic and extrinsic variances tend to zero as γ or κ tend to infinity.

Distribution	μ^I	μ^E	σ_I^2	σ_E^2
$vM(\mu,\kappa)$	μ	μ	$\frac{1}{2\pi I_0(\kappa)}\int_{-\pi}^{\pi}\alpha^2 e^{\kappa\cos(\alpha)}d\alpha$	$1-\frac{I_1(\kappa)}{I_0(\kappa)}$
$gN(\mu,\gamma)$	μ	μ	$\frac{1}{\gamma}\left(1-2\pi k^{-1}(\gamma)e^{-\frac{\gamma\pi^2}{2}}\right)$	$1-e^{-\frac{1}{2\gamma}}\frac{Re\left(erf\left(\pi\sqrt{\frac{\gamma}{2}}-i\frac{1}{\sqrt{2\gamma}}\right)\right)}{erf\left(\pi\sqrt{\frac{\gamma}{2}}\right)}$

Table 1: Summary of extrinsic and intrinsic means and variances for the von Mises and geodesic Normal distributions.

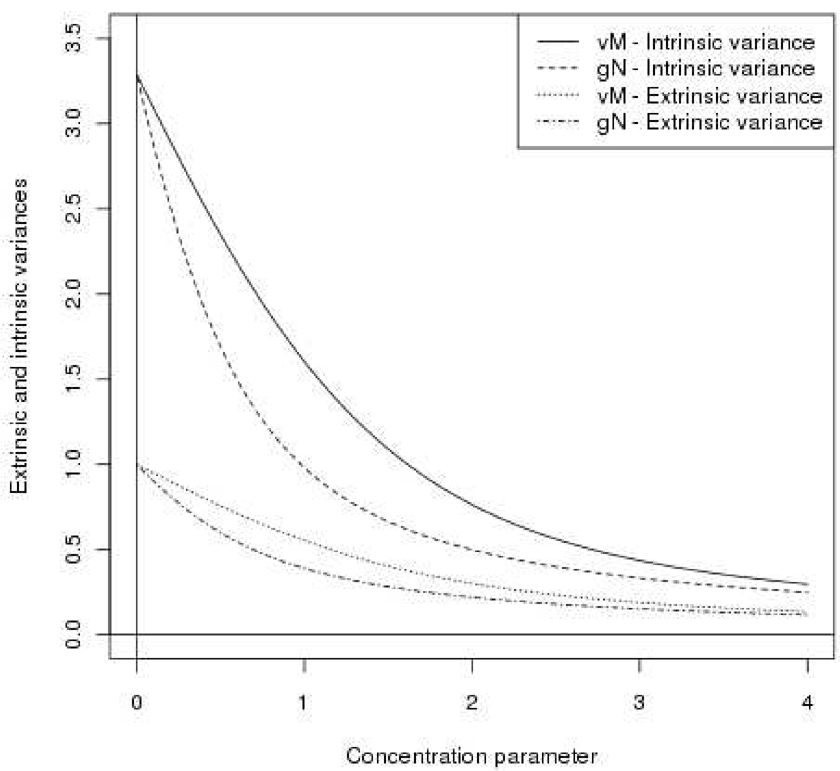

Figure 1: Evolutions of extrinsic and intrinsic variances in terms of the concentration parameter, γ for the gN distribution and κ for the vM one.

Entropy property

gN and vM have both a peculiar position respectively amongst distributions defined on Riemannian manifolds and circular statistics distributions: both maximize a certain definition of the entropy.

As explained in [JAS01] (characterization due to Mardia [Mar72]), the circular distribution that maximizes the entropy defined with respect to the angular random variable θ (see [MJ00] for exact definition), subject to the constraint that the first trigonometric moment is fixed, *i.e.* for μ^E and σ_E^2 fixed, is a vM(μ^E,κ), where $\sigma_E^2 = 1 - I_1(\kappa)/I_0(\kappa)$.

In a similar way, as demonstrated in [Pen06], the distribution defined using the geodesic distance on $\mathcal{S}^1$ which maximizes the entropy, when it is defined in the tangent plane (recall that exponential and Log maps allow this), and subject to the constraints that μ^I and σ_I^2 are fixed, is the $gN(\mu^I,\sigma_I^2)$.

One can thus conclude that vM and gN distributions play "similar" roles in that they maximize the entropy with respect to either extrinsic or intrinsic moments of the distribution.

Linear approximation of $\overline{\mu\theta}$

Recall that the random variable $\overline{\mu\theta}$ represents the algebraic measure of the vector $\overrightarrow{\mu\theta}$. The support of $\overline{\mu\theta}$ is $(-\pi,\pi)$. Its cdf is given for $t \in (-\pi,\pi)$ by

$$F_{\overline{\mu\theta}}(t) = k^{-1}(\gamma)\int_{\mathcal{S}^1} \mathbf{1}_{[-\pi,t]}(\overline{\mu y})e^{-\frac{\gamma}{2}\overline{\mu y}^2}d\mathcal{M}(y) = k^{-1}(\gamma)\int_{-\pi}^{t} e^{-\frac{\gamma}{2}\theta^2}d\theta.$$

This reduces to $F_{\overline{\mu\theta}}(t) = \frac{\Phi(t\sqrt{\gamma})-\Phi(-\pi\sqrt{\gamma})}{\Phi(\pi\sqrt{\gamma})-\Phi(-\pi\sqrt{\gamma})}$. In other words, $\overline{\mu\theta}$ is nothing else than a truncated Gaussian random variable with support $(-\pi,\pi)$ with mean 0 and scale parameter $1/\sqrt{\gamma}$, that is

$$\overline{\mu\theta} \stackrel{d}{=} Z\big|\ |Z| \leq \pi, \text{ where } Z \sim \mathcal{N}(0, 1/\sqrt{\gamma}). \tag{4}$$

For large concentration parameters γ, the geodesic Normal distribution can be "approximated" by a linear normal distribution in the following sense.

Proposition 4. *Let $\theta \sim gN(\mu,\gamma)$, then as $\gamma \to +\infty$, $\sqrt{\gamma}\,\overline{\mu\theta} \stackrel{d}{\to} \mathcal{N}(0,1)$.*

Proof. Let us denote by $h(\cdot)$ the moment generating function of the random variable $\sqrt{\gamma}\,\overline{\mu\theta}$, then for $t \in \mathbb{R}$

$$\begin{aligned} h(t) &= k^{-1}(\gamma)\int_{-\pi}^{\pi} e^{\sqrt{\gamma}t\theta}e^{-\frac{\gamma}{2}\theta^2}d\theta \\ &= e^{\frac{t^2}{2}}\ k^{-1}(\gamma)\sqrt{\gamma}\int_{-\pi\sqrt{\gamma}}^{\pi\sqrt{\gamma}} e^{-\frac{1}{2}(\theta-t)^2}d\theta \\ &= e^{\frac{t^2}{2}}\ \frac{\Phi\left(\pi\sqrt{\gamma}-t\right)-\Phi\left(-\pi\sqrt{\gamma}-t\right)}{\Phi\left(\pi\sqrt{\gamma}\right)-\Phi\left(-\pi\sqrt{\gamma}\right)}. \end{aligned}$$

Therefore, as $\gamma \to +\infty$, for fixed t, $h(t)$ converges towards $e^{t^2/2}$ which is the moment generating function of a standard Gaussian random variable. ∎

Such a result also holds for the von Mises distribution: as $\kappa \to +\infty$, $\sqrt{\kappa}(\theta-\mu) \stackrel{d}{\to} \mathcal{N}(0,1)$, see *e.g.* Proposition 2.2 in [JAS01].

4 Simulation of a geodesic Normal distribution and examples

The generation of a gN distribution with support on $(0, 2\pi)$ is extremely simple following (4). It consists in two steps.

1. Generate $Z\big|\ |Z| \leq \pi,$ where $Z \sim \mathcal{N}(0, 1/\sqrt{\gamma})$
2. Set $\theta = \mu + Z (\text{mod } 2\pi)$.

Figure 4 presents some examples for different values of location parameter μ and concentration parameter γ.

5 Maximum Likelihood Estimation

5.1 Preliminary and notation

Let us consider now the identification problem of estimating the parameters of a gN distribution from the n observations $\theta_1, \ldots, \theta_n \in (0, 2\pi)$. In this section, we will denote by $\mu^\star$ and $\gamma^\star$ the unknown parameters to estimate. We assume that $\mu^\star \in [0, 2\pi)$ and $\gamma^\star > 0$. Also, we propose to denote by $\widehat{\mu}^I$ and $\widehat{\mu}^E$ the empirical intrinsic and extrinsic means defined by

$$\widehat{\mu}^I \quad := \quad \text{argmin}_{\mu \in \mathcal{S}^1} \frac{1}{n} \sum_{i=1}^{n} d_G(\mu, \theta_i)^2. \tag{5}$$

$$\widehat{\mu}^E \quad := \quad \text{Arg}\left(\widehat{\varphi}_1\right), \text{ with } \widehat{\varphi}_1 := \frac{1}{n} \sum_j \cos(\theta_j) + i \frac{1}{n} \sum_j \sin(\theta_j). \tag{6}$$

The natural intrinsic and extrinsic variances are then denoted by $\widehat{\sigma}_I^2$ and $\widehat{\sigma}_E^2$ and given by

$$\widehat{\sigma}_I^2 = \frac{1}{n} \sum_{i=1}^{n} d_G(\widehat{\mu}^I, \theta_i)^2 \quad \text{and} \quad \widehat{\sigma}_E^2 = 1 - |\widehat{\varphi}_1|. \tag{7}$$

In the following, we will need the following Lemma and notation.

Lemma 5. *For a random variable $\theta_\gamma \sim gN(\mu^\star, \gamma)$, let $\sigma_I^2(\mu, \gamma) := E[d_G(\mu, \theta_\gamma)^2]$ for $\mu \in (\mu^\star - \pi, \mu^\star + \pi)$ and $\delta = \mu^\star - \mu$, then*

$$\sigma_I^2(\mu, \gamma) = g(\delta) \mathbf{1}_{[0,\pi)}(\delta) + g(-\delta) \mathbf{1}_{(-\pi,0]}(\delta),$$

where

$$g(\delta) := \int_{-\pi}^{\pi} (\delta + \alpha)^2 f(\alpha) d\alpha + 4\pi \int_{\pi-\delta}^{\pi} (2\pi - (\delta + \alpha)) f(\alpha) d\alpha$$

and $f(\alpha) = k^{-1}(\gamma) e^{-\frac{\gamma}{2}\alpha^2}$.

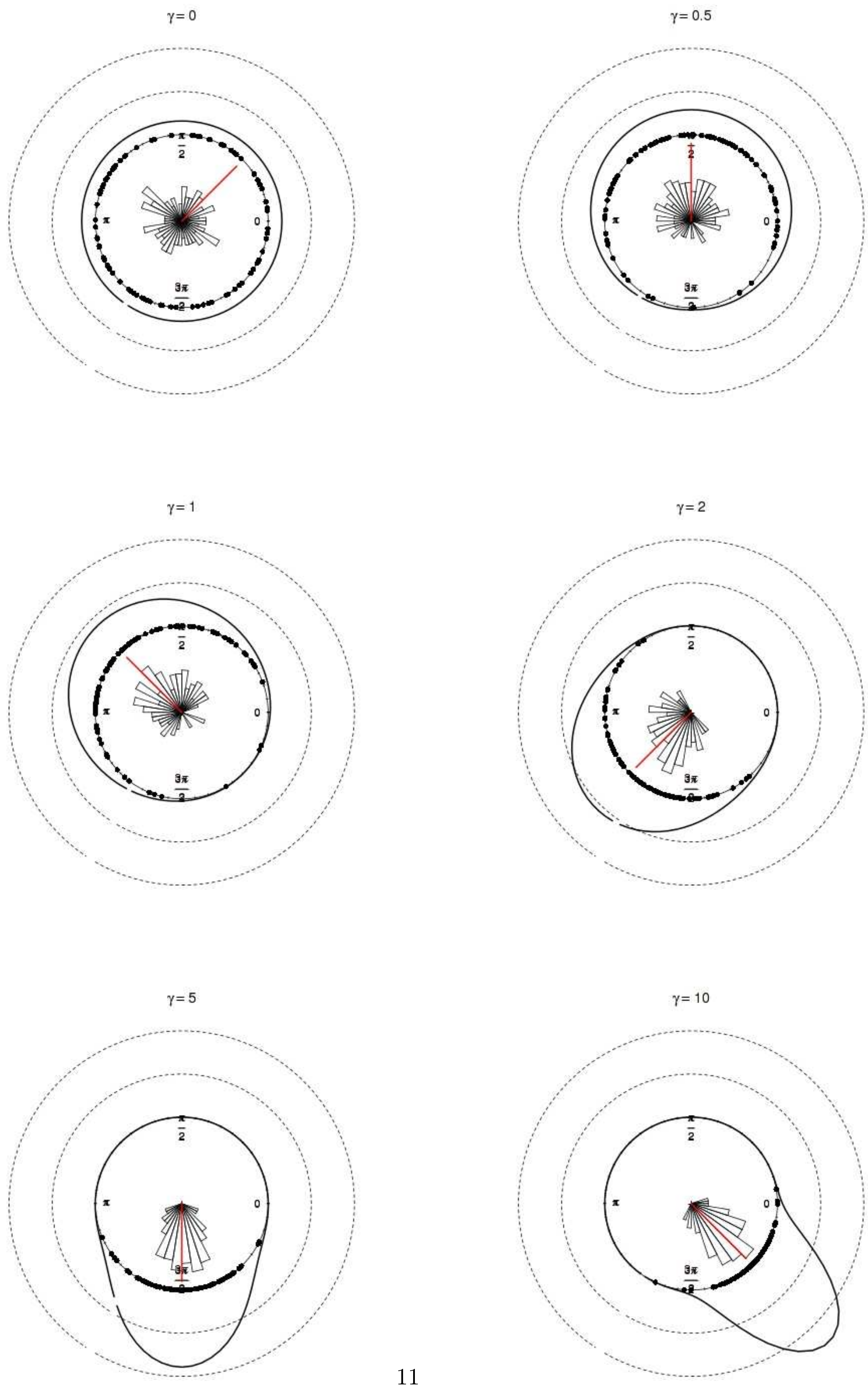

Figure 2: Example of $n = 100$ $gN(\mu, \gamma)$ realizations for different parameters. The red line indicated the value of μ. These different plots have been produced using the R packages circular, CircStats maintained by C. Agostinelli, and related to the book [JAS01].

Proof. Let us fix $\mu \in (\mu^\star - \pi, \mu^\star]$, then

$$\overline{\mu\alpha} = \begin{cases} \overline{\mu\mu^\star} + \overline{\mu^\star\alpha} & \text{when } \alpha \in (\mu^\star - \pi, \mu^\star + \pi - (\mu^\star - \mu)) \\ \overline{\mu\mu^\star} + \overline{\mu^\star\alpha} - 2\pi & \text{when } \alpha \in (\mu^\star + \pi - (\mu^\star - \mu), \mu^\star + \pi) \end{cases}$$

Denoting $\delta := \mu^\star - \mu$, this expansion allows us to derive

$$\begin{aligned} \sigma_I^2(\mu, \gamma) &= E[d_G(\mu, \theta_\gamma)^2] \\ &= \int_{-\pi}^{\pi-\delta} (\delta + \alpha)^2 f(\alpha) d\alpha + \int_{\pi-\delta}^{\pi} (2\pi - \delta - \alpha)^2 f(\alpha) d\alpha \\ &= \int_{-\pi}^{\pi} (\delta + \alpha)^2 f(\alpha) d\alpha + 4\pi \int_{\pi-\delta}^{\pi} (2\pi - (\delta + \alpha)) f(\alpha) d\alpha. \end{aligned}$$

Now, let $\mu \in [\mu^\star, \mu^\star + \pi)$, then

$$\overline{\mu\alpha} = \begin{cases} \overline{\mu\mu^\star} + \overline{\mu^\star\alpha} & \text{when } \alpha \in (\mu^\star - \pi + (\mu - \mu^\star), \mu^\star + \pi) \\ \overline{\mu\mu^\star} + \overline{\mu^\star\alpha} + 2\pi & \text{when } \alpha \in (\mu^\star - \pi, \mu^\star - \pi + (\mu - \mu^\star)), \end{cases}$$

which leads to

$$\begin{aligned} \sigma_I^2(\mu, \gamma) &= \int_{-\pi-\delta}^{\pi} (\delta + \alpha)^2 f(\alpha) d\alpha + \int_{-\pi}^{-\pi-\delta} (2\pi + \delta + \alpha)^2 f(\alpha) d\alpha \\ &= \int_{-\pi}^{\pi} (\delta + \alpha)^2 f(\alpha) d\alpha + 4\pi \int_{-\pi}^{-\pi-\delta} (2\pi + (\delta + \alpha)) f(\alpha) d\alpha \\ &= \int_{-\pi}^{\pi} (\delta + \alpha)^2 f(\alpha) d\alpha + 4\pi \int_{\pi+\delta}^{\pi} (2\pi + (\delta - \alpha)) f(\alpha) d\alpha \\ &= g(-\delta). \end{aligned}$$

■

Note that from Proposition 3, $\sigma_I^2(\mu^\star, \gamma)$, which corresponds to the intrinsic variance of θ_γ, does not depend on $\mu^\star$ and will therefore be simplified to $\sigma_I^2(\gamma)$.

5.2 Maximum Likelihood Estimate

The log-likelihood expressed for n i.i.d. gN distributions is given by :

$$\ell(\mu, \gamma) = -n \log(k(\gamma)) - \frac{\gamma}{2} \sum_{i=1}^{n} d_G(\mu, \theta_i)^2.$$

Let $\Omega = \{(\mu, \gamma) : \mu \in [0, 2\pi) \setminus \{(\mu^\star \pm \pi)\ mod(2\pi)\}, \gamma > 0\}$ and assume that the true parameter $(\mu^\star, \gamma^\star)$ belongs to the interior of Ω. The MLE estimates and asymptotic results are given by the following proposition.

Proposition 6.
(i) The MLE estimates of $\mu^\star$ *and* $\gamma^\star$ *are uniquely given by* $\widehat{\mu}^{MLE} := \widehat{\mu}^I$ *and* $\sigma_I^2\left(\widehat{\gamma}^{MLE}\right) = \widehat{\sigma}_I^2$.
(ii) As $n \to +\infty$*, the MLE estimates are consistent and satisfy the following central limit theorem*

$$\sqrt{n}\left(\widehat{\mu}^{MLE} - \mu^\star, \widehat{\gamma}^{MLE} - \gamma^\star\right)^T \xrightarrow{d} \mathcal{N}\left(0, J^{-1}(\gamma^\star)\right),$$

where $J(\gamma^\star)$ *is the Fisher information matrix given by* $J(\gamma^\star) = \begin{pmatrix} J_1(\gamma^\star) & 0 \\ 0 & J_2(\gamma^\star) \end{pmatrix}$ *with* $J_1(\gamma^\star) := \gamma^\star\left(1 - 2\pi k^{-1}(\gamma^\star) e^{-\frac{\gamma^\star}{2}\pi^2}\right)$ *and* $J_2(\gamma^\star) := \frac{k''(\gamma^\star)}{k(\gamma^\star)} - \left(\frac{k'(\gamma^\star)}{k(\gamma^\star)}\right)^2$.

As Mardia and Jupp note for the von Mises distribution ([MJ00], Section 5.3 p. 86) the asymptotic normality means that $\widehat{\mu}^{MLE}$ is regarded as unwrapped onto the line.

As it is for a Gaussian distribution on the real line or for a vM distribution on the circle, the Fisher information matrix of a gN distribution does not depend on the true location parameter $\mu^\star$ and the two estimates of $\mu^\star$ and $\gamma^\star$ are asymptotically independent. Let us also note that the geodesic moment estimates, that is the estimates of $\mu^\star$ and $\gamma^\star$ based on the two first geodesic moments equations μ^I and $\sigma_I^2(\gamma)$ exactly fit to the maximum likelihood estimates. Here is again another analogy with the vM distribution, since the MLE of a vM distribution correspond to the estimates of μ and κ based on the extrinsic moments (see [MJ00] for further details).

Proof. (i) Since the minimum of $\sum_{i=1}^n d_G(\mu, \theta_i)^2$ defines the intrinsic empirical mean, the MLE of μ correponds to the intrinsic mean estimate of $\mu^\star$. Now, the partial derivative of ℓ with respect to γ is given by

$$\frac{\partial \ell}{\partial \gamma}(\mu, \gamma) = -\frac{k'(\gamma)}{k(\gamma)} - \frac{1}{2}\sum_{i=1}^n d_G(\mu, \theta_i)^2.$$

Let us note that

$$k'(\gamma) = -\frac{1}{2}\int_{-\pi}^{\pi} \theta^2 e^{-\frac{\gamma}{2}\theta^2} d\theta = -\frac{k(\gamma)}{2} E[d_G(\mu^\star, \theta_\gamma)^2] = -\frac{k(\gamma)}{2}\sigma_I^2(\gamma). \tag{8}$$

Replacing μ by its MLE estimate and taking the derivative of ℓ w.r.t. γ equal to zero implies that the MLE estimate of γ is defined by the following equation:

$$\sigma_I^2(\widehat{\gamma}^{MLE}) = \frac{1}{n}\sum_{i=1}^n d_G(\widehat{\mu}^{MLE}, \theta_i)^2 =: \widehat{\sigma}_I^2.$$

The proof is ended by showing that $\sigma_I^2(\gamma)$ is a striclty decreasing function on $\mathbb{R}^+$. Similarly to (8), we notice that $k''(\gamma) = \frac{k(\gamma)}{4} E[d_G(\mu^\star, \theta_\gamma)^4]$. Now,

$$\begin{aligned}
(\sigma_I^2)'(\gamma) &= -2\left(\frac{k''(\gamma)}{k(\gamma)} - \left(\frac{k'(\gamma)}{k(\gamma)}\right)^2\right) \\
&= -2\left(\frac{1}{4}E[d_G(\mu^\star, \theta_\gamma)^4] - E[d_G(\mu^\star, \theta_\gamma)^2/2]^2\right) \\
&= -\frac{1}{2}Var[d_G(\mu^\star, \theta_\gamma)^2] < 0. \qquad (9)
\end{aligned}$$

(*ii*) Standard theory of maximumn likelihood estimators (Theorem 5.1 p. 463 of [LC98]) shows that asymptotic normality result holds. The verification of the assumptions (A-D) of [LC98], p.462-463 are omitted; we just focus, here, on the computation of the Fisher information matrix. The antidiagonal term is given by

$$J_{12} := \frac{1}{2} E\left[\frac{\partial}{\partial \mu} d_G(\mu, \theta_\gamma)^2\right]\bigg|_{\mu=\mu^\star, \gamma=\gamma^\star} = \frac{1}{2}\frac{\partial \sigma_I^2}{\partial \mu}(\mu^\star, \gamma^\star) = 0,$$

since $\mu^\star$ corresponds to the intrinsic mean and thus minimizes the geodesic variance. The asymptotic variance of $\sqrt{n}\widehat{\gamma}^{MLE}$ is obviously given by the inverse of

$$J_2(\gamma^\star) = \frac{k''(\gamma^\star)k(\gamma^\star) - k'(\gamma^\star)^2}{k(\gamma^\star)^2}.$$

Recall that from (9), this constant is positive. Now, the last term to compute is the asymptotic variance of $\sqrt{n}\widehat{\mu}^{MLE}$ given by the inverse of

$$J_1(\gamma^\star) := \frac{\gamma}{2} E\left[\frac{\partial^2}{\partial \mu^2} d_G(\mu, \theta_\gamma)^2\right]\bigg|_{\mu=\mu^\star, \gamma=\gamma^\star} = \frac{\gamma^\star}{2}\frac{\partial^2(\sigma_I^2)}{\partial \mu^2}(\mu^\star, \gamma^\star).$$

From Lemma 5, $\sigma_I^2(\mu, \gamma)$ is a function of $\delta = \mu^\star - \mu$. Without loss of generality, assume $\delta \geq 0$ (the other case leads to the same conclusion), then the function g (in Lemma 5) is twice continuous differentiable on $[0, \pi)$ and

$$g'(\delta) = 2\delta - 4\pi\left(F(\pi) - F(\pi - \delta)\right) \quad \text{and} \quad g''(\delta) = 2 - 4\pi f(\pi - \delta).$$

Setting $\delta = 0$ in the last equation leads to the stated result. ■

5.3 Simulation study

We have investigated the efficiency of the maximum likelihood estimates in a simulation study. A part of the results are presented in Table 5.3. As expected, the empirical MSE of both estimates of the parameters $\mu^\star$ and $\gamma^\star$ converge towards zero as the sample size grows. We also notice that it's more complicated to estimate the intrinsic mean when the concentration parameter is low. Unlike this, the concentration parameter is better estimated for low values of $\gamma^\star$. These facts are confirmed by Figure 5.3 plotting the constants of the asymptotic variances (for both estimates), i.e. $1/J_1(\gamma^\star)$ and $1/J_2(\gamma^\star)$, in terms of $\gamma^\star$. Finally, Figure 5.3 illustrates the central limit theorem satisfied by the MLE estimates.

Location parameter $\mu^\star$					
Simulation	Sample size				
Parameters	$n=10$	$n=20$	$n=50$	$n=100$	$n=500$
$\mu^\star=\frac{\pi}{4},\gamma^\star=.5$	3.3668	1.6515	0.2218	0.0313	0.0048
$\mu^\star=\frac{3\pi}{4},\gamma^\star=1$	0.1088	0.0537	0.0206	0.0103	0.0020
$\mu^\star=\frac{5\pi}{4},\gamma^\star=5$	0.0199	0.0102	0.0040	0.0019	0.0004
$\mu^\star=\frac{7\pi}{4},\gamma^\star=10$	0.0099	0.0050	0.0020	0.0010	0.0002

Concentration parameter $\gamma^\star$					
Simulation	Sample size				
Parameters	$n=10$	$n=20$	$n=50$	$n=100$	$n=500$
$\mu^\star=\frac{\pi}{4},\gamma^\star=.5$	0.2893	0.0748	0.0255	0.0141	0.0065
$\mu^\star=\frac{3\pi}{4},\gamma^\star=1$	1.0559	0.2268	0.0565	0.0252	0.0046
$\mu^\star=\frac{5\pi}{4},\gamma^\star=5$	22.8649	5.4092	1.3365	0.5757	0.1037
$\mu^\star=\frac{7\pi}{4},\gamma^\star=10$	88.5093	20.1554	5.3607	2.2723	0.4179

Table 2: Empirical Mean Squared Error (MSE) of MLE estimates of the location parameter $\mu^\star$ (top) and the concentration parameter $\gamma^\star$ (bottom) based on $m = 5000$ replications of gN distributions for differents choices of parameters and different sample sizes.

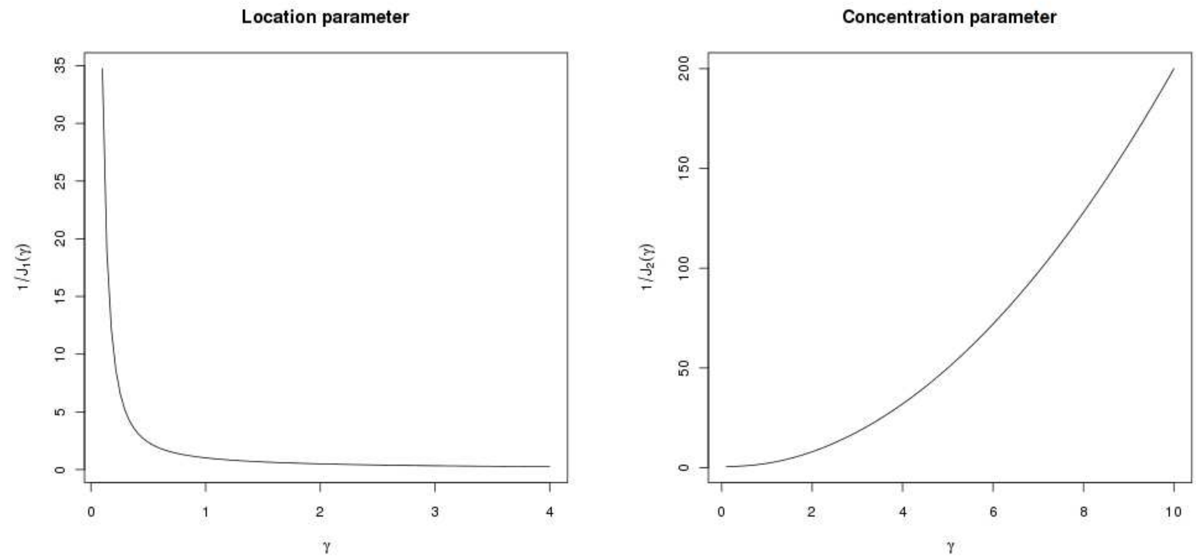

Figure 3: Plots of the constants of the asymptotic variances of $\widehat{\mu}^{MLE}$ (left) and $\widehat{\gamma}^{MLE}$ (right), i.e. the constants $1/J_1(\gamma^\star)$ and $1/J_2(\gamma^\star)$ given in Proposition 6, in terms of $\gamma^\star$.

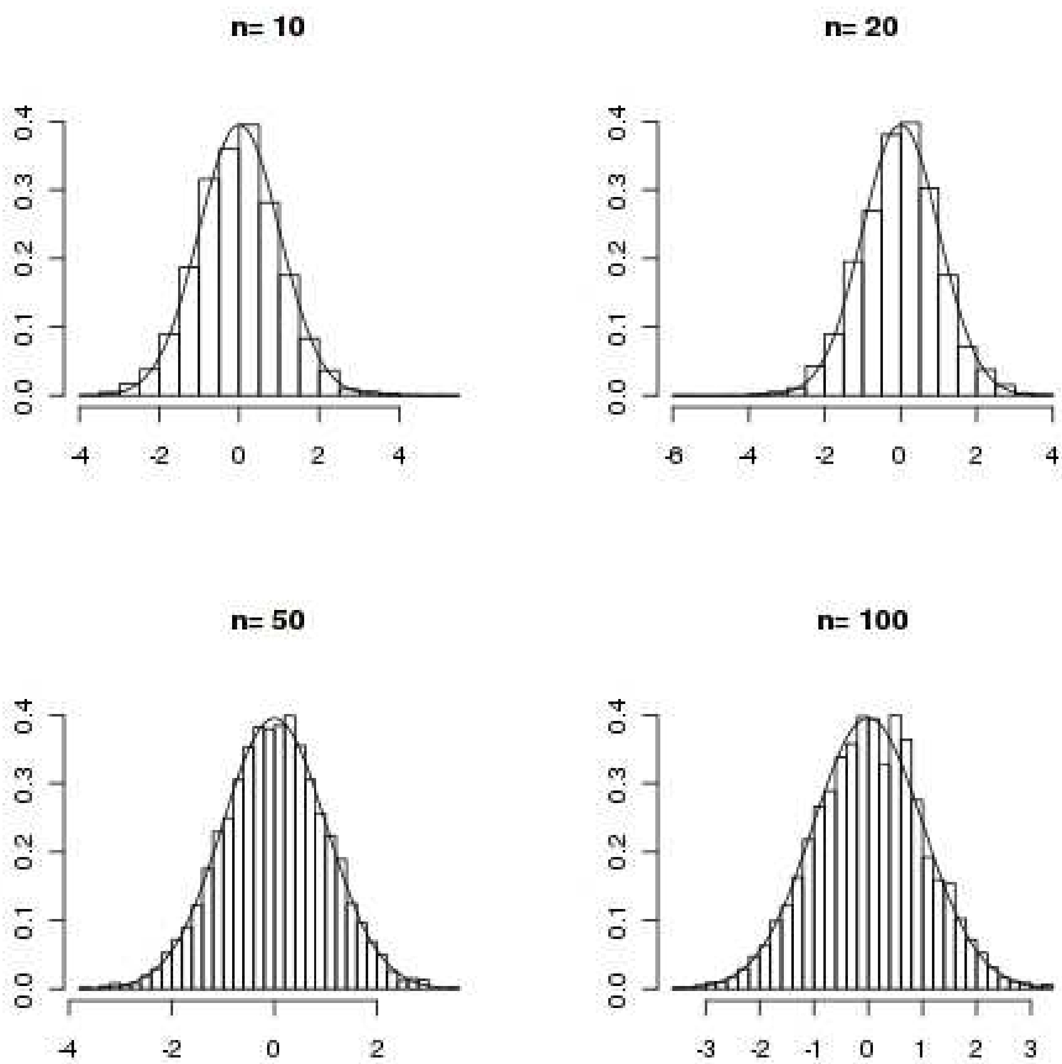

Figure 4: Histograms of $\left(\sqrt{n}(\widehat{\mu}_j^{MLE} - \mu^\star)\right)_{j=1,\ldots,5000}$ based on $m = 5000$ replications of a $gN(3\pi/4, 1)$ distribution for different sample sizes. The curve corresponds to the density of a Gaussian random variable with mean zero and variance $1/J_1(1)$ given in Proposition 6.

References

[BP03] R. Bhattacharya and V. Patrangenaru. Large Sample Theory of Intrinsic and Extrinsic Sample Means on Manifolds: I. *Annals of statistics*, 31(1):1–29, 2003.

[BP05] R. Bhattacharya and V. Patrangenaru. Large Sample Theory of Intrinsic and Extrinsic Sample Means on Manifolds: II. *Annals of statistics*, 33:1225–1259, 2005.

[JAS01] S.R. Jammalamadaka and A. A. Sengupta. *Topics in circular statistics*. World Scientific Pub Co Inc, 2001.

[KS08] D. Kaziska and A. Srivastava. The Karcher mean of a class of symmetric distributions on the circle. *Statistics & Probability Letters*, 78(11):1314–1316, 2008.

[LC98] E.L. Lehmann and G. Casella. *Theory of point estimation*. Springer Verlag, 1998.

[Mar72] K.V. Mardia. *Statistics of directional data*. Academic Press, London, 1972.

[MJ00] K.V. Mardia and P.E. Jupp. *Directional statistics*. Wiley Chichester, 2000.

[Pen06] X. Pennec. Intrinsic statistics on Riemannian manifolds: Basic tools for geometric measurements. *Journal of Mathematical Imaging and Vision*, 25(1):127–154, 2006.

[Wat83] G.S. Watson. *Statistics on spheres*. Wiley-Interscience, 1983.