APPENDIX
Directional Statistics with the Spherical Normal Distribution
— Supplementary Material —

Søren Hauberg
Department of Mathematics and Computer Science
Technical University of Denmark
Kgs. Lyngby, Denmark
sohau@dtu.dk

The isotropic spherical normal distribution has density

\[ SN(x \mid \mu, \lambda) = \frac{1}{Z_1(\lambda)} \exp\left(-\frac{\lambda}{2} \text{Log}_\mu(x)\text{Log}_\mu(x)^\top\right), \quad \|\mu\| = 1, \quad \lambda > 0. \] (1)

In the main paper, we claim that the normalization constant is

\[ Z_1(\text{even}) = \frac{A_{D-2}}{2^{D-2}} \left(\frac{D-2}{D/2-1}\right) \sqrt{\frac{\pi}{2\lambda}} \text{erf}\left(\frac{\pi}{\sqrt{2}} \frac{\lambda}{2}\right) + \frac{A_{D-2}}{2^{D-3}} \frac{2\pi}{\lambda} \left(\frac{D-2}{2}\right) \sum_{k=0}^{D/2-2} (-1)^k \left(\frac{D-2}{k}\right) \exp\left(-\frac{(D-2-2k)^2}{2\lambda}\right) \] (2a)

\[ Z_1(\text{odd}) = \frac{A_{D-2}}{2^{D-3}} \left(\frac{D-3}{D/2-1}\right) \sqrt{\frac{\pi}{2\lambda}} \sum_{k=0}^{(D-3)/2} (-1)^k \left(\frac{D-2}{k}\right) \exp\left(-\frac{(D-2-2k)^2}{2\lambda}\right) \] (2b)

\[ \cdot \left\{ \text{Im}\left[\text{erf}\left(\frac{(D-2-2k)i}{\sqrt{2\lambda}}\right)\right] + \text{Im}\left[\text{erf}\left(\frac{\pi\lambda-(D-2-2k)i}{\sqrt{2\lambda}}\right)\right]\right\} \]

when \( D \) is even and odd. Here \( \text{erf} \) is the imaginary error function, while \( \text{Re}\{\cdot\} \) and \( \text{Im}\{\cdot\} \) takes the real and imaginary parts of a complex number, respectively. In this section, we provide the derivation of this constant.

By definition, we have

\[ Z_1(\lambda) = \int_{S^{D-1}} \exp\left(-\frac{\lambda}{2} \text{Log}_\mu(x)\text{Log}_\mu(x)^\top\right) dx. \] (3)

We express this integral in the tangent space of the sphere at \( \mu \), i.e. we perform the substitution

\[ v = \text{Log}_\mu(x). \] (4)

The appropriate Jacobian is

\[ \det(J) = \left(\frac{\sin(||v||)}{||v||}\right)^{D-2} \] (5)

and the integral becomes

\[ Z_1(\lambda) = \int_{||v||<\pi} \exp\left(-\frac{\lambda||v||^2}{2}\right) \left(\frac{\sin(||v||)}{||v||}\right)^{D-2} dv. \] (6)

This project has received funding from the European Research Council (ERC) under the European Union’s Horizon 2020 research and innovation programme (grant agreement n° 757360). SH was supported by a research grant (15334) from VILLUM FONDEN.
We then write this in hyper-spherical coordinates
\[ Z_1(\lambda) = \int_{r=0}^{\pi} \int_{\phi_1=0}^{\pi} \cdots \int_{\phi_D=0}^{\pi} \sin(\phi_1)^D \sin(\phi_2)^D \cdots \sin(\phi_{D-3}) d\phi_{D-2} d\phi_{D-3} \cdots d\phi_2 d\phi_1 dr. \] (7)

Since the radius \( r \) and the angles \( \phi \) are never mixed in the integrand, we can split this into the product of two integrals
\[ Z_1(\lambda) = \int_{r=0}^{\pi} \exp\left(-\frac{\lambda r^2}{2}\right) \sin^{D-2}(r) dr \] (8)
\[ \int_{\phi_1=0}^{\pi} \cdots \int_{\phi_{D-3}=0}^{\pi} \sin(\phi_1)^D \sin(\phi_2)^D \cdots \sin(\phi_{D-3}) d\phi_{D-2} d\phi_{D-3} \cdots d\phi_2 d\phi_1. \]

The second term is merely the surface area of the \( D - 2 \) dimensional unit sphere
\[ A_{D-2} = \frac{2\pi^{(D-1)/2}}{\Gamma((D-1)/2)}. \] (9)
where \( \Gamma \) is the usual Gamma function. The integral (9) then reduces to
\[ Z_1(\lambda) = A_{D-2} \int_{r=0}^{\pi} \exp\left(-\frac{\lambda r^2}{2}\right) \sin^{D-2}(r) dr. \] (10)

To evaluate this expression we need the trigonometric power formulas [2]
\[ \sin^{2n}(x) = \frac{1}{2^n} \binom{2n}{n} + (-1)^n \sum_{k=0}^{n-1} (-1)^k \binom{2n}{k} \cos(2(n-k)x), \] (11)
\[ \sin^{2n+1}(x) = (-1)^n \sum_{k=0}^{n} (-1)^k \binom{2n+1}{k} \sin((2n+1-2k)x). \] (12)

We are now ready evaluate the normalization constant. First we consider the case were \( D - 2 \) is even.
\[ Z_1^{(\text{even})}(\lambda) = A_{D-2} \int_{r=0}^{\pi} \exp\left(-\frac{\lambda r^2}{2}\right) \sin^{D-2}(r) dr \] (13)
\[ = A_{D-2} \frac{D-2}{2^{D-2}} \frac{D/2-1}{\Gamma(D/2-1)} \int_{r=0}^{\pi} \exp\left(-\frac{\lambda r^2}{2}\right) dr \]
\[ + A_{D-2} \frac{(-1)^{D/2-1}}{2^{D-3}} \sum_{k=0}^{D/2-2} (-1)^k \binom{D-2}{k} \int_{r=0}^{\pi} \exp\left(-\frac{\lambda r^2}{2}\right) \cos((D-2-2k)r) dr. \] (14)

To evaluate this we need two simple integrals, which we evaluate using Maple,
\[ \int_{r=0}^{\pi} \exp\left(-\frac{\lambda r^2}{2}\right) dr = \frac{\sqrt{\pi}}{\sqrt{2\lambda}} \erf\left(\frac{\pi \sqrt{\lambda}}{\sqrt{2}}\right) \] (15)
\[ \int_{r=0}^{\pi} \exp\left(-\frac{\lambda r^2}{2}\right) \cos(a \cdot r) dr = \frac{\sqrt{2\pi}}{4\sqrt{\lambda}} \exp\left(-\frac{a^2}{2\lambda}\right) \left( \erf\left(\frac{\pi \lambda - ai}{\sqrt{2\lambda}}\right) + \erf\left(\frac{\pi \lambda + ai}{\sqrt{2\lambda}}\right) \right), \] (16)
where \( i \) denotes the complex unit. Inserting these expressions into Eq. 14 and simplifying expressions gives the desired result (2a). Note that in the simple special case \( D = 2 \), the spherical normal is a distribution over the unit circle. Here, the normalization constant reduce to
\[ Z_1^{(D=2)}(\lambda) = \frac{\sqrt{2\pi}}{\sqrt{\lambda}} \erf\left(\frac{\pi \sqrt{\lambda}}{\sqrt{2}}\right). \] (17)

1Using the convention of https://en.wikipedia.org/wiki/N-sphere#Spherical_coordinates
We now consider the case where \( D - 2 \) is odd. Akin to the previous derivation, we insert Eq. 12 into Eq. 10 and get

\[
Z_{1}^{(\text{odd})}(\lambda) = A_{D-2} \int_{r=0}^{\pi} \exp \left( -\frac{\lambda r^2}{2} \right) \sin^{D-2}(r) dr
\]

\[
= A_{D-2} \left(\frac{-1}{2^{D-3}} \right) \sum_{k=0}^{(D-3)/2} (-1)^k \binom{D-2}{k} \int_{r=0}^{\pi} \exp \left( -\frac{\lambda r^2}{2} \right) \sin((D-2-2k)r) dr. \tag{18}
\]

To evaluate this, we need to evaluate the integral

\[
\int_{r=0}^{\pi} \exp \left( -\frac{\lambda r^2}{2} \right) \sin(a \cdot r) dr
\]

\[
= -\frac{i\sqrt{2\pi}}{4\lambda} \exp \left( \frac{-a^2}{2\lambda} \right) \left( 2\text{erf} \left( \frac{ia}{\sqrt{2\lambda}} \right) + \text{erf} \left( \frac{\pi\lambda - ia}{\sqrt{2\lambda}} \right) - \text{erf} \left( \frac{\pi\lambda + ia}{\sqrt{2\lambda}} \right) \right), \tag{20}
\]

which we have evaluated using Maple. Inserting this expression into Eq. 19 and simplifying gives the result in Eq. 2b. In the important special-case \( D = 3 \), the normalization constant reduce to

\[
Z_{1}^{(D=3)}(\lambda) = \frac{-i\pi^{3/2}}{\sqrt{2\lambda}} \exp \left( \frac{-1}{2\lambda} \right) \left\{ 2\text{erf} \left( \frac{i}{\sqrt{2\lambda}} \right) + \text{erf} \left( \frac{\pi\lambda - i}{\sqrt{2\lambda}} \right) - \text{erf} \left( \frac{\pi\lambda + i}{\sqrt{2\lambda}} \right) \right\}. \tag{21}
\]

This concludes the derivation.

A. Quality of the approximation

We approximate the inverse normalization constant \( 1/Z_{1} \) with a straight line in order to derive an expression for the anisotropic normalization constant. Figure 1 (center) show the inverse normalization constant for varying values of \( \lambda \). The right panel of the figure show the difference between the inverse normalization and a fitted straight line. From this we draw two conclusions: 1) the inverse normalization is indeed not a straight line; 2) a straight line is, however, a good approximation. While using one globally fitted straight line gives a fairly accurate estimate of the normalization constant, we find that accuracy can be slightly improved by fitting the line locally. We make this local fit through \( 1/Z_{1}(\lambda_{2}) \) and \( 1/Z_{1}(\lambda_{2} + \alpha(\lambda_{1} - \lambda_{2})) \). By extensive numerical optimization we have found that \( \alpha^{-1} = 0.46 \lambda_{2} + 1.55 \) minimizes the worst-case approximation error of the integral.

At times it may be easier to interpret a variance parameter rather than a concentration parameter. The variance of the spherical normal distribution is defined as [1]

\[
\text{Var}[x] = \int_{S^{D-1}} \text{arccos}^{2}(x^{T}\mu) \ SN(x \mid \mu, \Lambda) dx. \tag{22}
\]

When the distribution is isotropic, this expression can be evaluated for \( S^{2} \) similarly to the proof of proposition 1 to give

\[
\text{Var}[x] = -\frac{\pi}{\lambda^{3/2}} \left( -\frac{i\sqrt{\pi}(\lambda - 1)}{\sqrt{2}} \exp(-1/2\lambda) \text{erf} \left( \frac{i - \pi\lambda}{\sqrt{2\lambda}} \right) \right.
\]

\[
- \frac{i\sqrt{\pi}(\lambda - 1)}{\sqrt{2}} \exp(-1/2\lambda) \text{erf} \left( \frac{i + \pi\lambda}{\sqrt{2\lambda}} \right) + i(\lambda - 1)\sqrt{2\pi} \exp(-1/2\lambda) \text{erf} \left( \frac{i}{\sqrt{2\lambda}} \right)
\]

\[
- 2\sqrt{\lambda} \exp(-\pi^{2}/2) - 2\sqrt{\lambda} \right). \tag{23}
\]

The left panel of Fig. 1 show how the variance change as a function of \( \lambda \). Notice that the curve is roughly shaped as \( 1/\lambda \).
Fig. 1. **Left:** the variance as a function of the concentration. **Center:** the inverse normalization constant for the isotropic distribution. **Right:** The deviation between the inverse normalization constant and a single linear approximation.

### REFERENCES
