Suplementary Documents

Pablo Fernandez-Gonzalez, Concha Bielza, Pedro Larrañaga

Department of Artificial Intelligence, Universidad Politécnica de Madrid

1 APPENDIX. Proofs of the results

1.1 Proof of Lemma 1

1. If $f_{tvM}(\theta;\mu,\kappa,a,b)$ is strictly decreasing $\Leftrightarrow$ a,b satisfy $a,b\in\mathbb{O}_{\mu,\mu+\pi}$.The proof follows simply from noting that the von Mises distribution decreases from its maximum to its minimum value in $\mathbb{O}_{\mu,\mu+\pi}$. If $\mathbb{O}_{a,b}\subset\mathbb{O}_{\mu,\mu+\pi}$, the resulting truncated von Mises distribution exhibits a monotonic decreasing behavior.

2. Analogously, if $f_{tvM}(\theta;\mu,\kappa,a,b)$ is strictly increasing $\Leftrightarrow$ truncation parameters a,b satisfy $\mathbb{O}_{a,b}\subset\mathbb{O}_{\mu-\pi,\mu}$.

3. If $f_{tvM}(\theta;\mu,\kappa,a,b)$ increases and decreases reaching a single maximum $\Leftrightarrow$ truncation parameters a,b satisfy $\mu\in\mathbb{O}_{a,b}$ and $\mu+\pi,\notin\mathbb{O}_{a,b}$

4. If $f_{tvM}(\theta;\mu,\kappa,a,b)$ increases and decreases reaching a single minimum $\Leftrightarrow$ truncation parameters a,b satisfy $\mu+\pi\in\mathbb{O}_{a,b}$ and $\mu,\mu+2\pi\notin\mathbb{O}_{a,b}$.

5. If $f_{tvM}(\theta;\mu,\kappa,a,b)$ increases and decreases with both single maximum and single minimum $\Leftrightarrow$ the truncation parameters a,b satisfy either $\mu,\mu+\pi\in\mathbb{O}_{a,b}$ or $\mu,\mu-\pi\in\mathbb{O}_{a,b}$

1.2 Proof of Lemma 2.

We have, by means of the power series expansion of the $e^{(\cdot)}$ function,

$$I(\theta;\mu,\kappa)=\int f_{uvM}(\theta;\mu,\kappa)d\theta=\int e^{\kappa\cos(\theta-\mu)}d\theta=\int\sum_{n=0}^{\infty}\frac{(\kappa\cos(\theta-\mu))^n}{n!}d\theta,$$

where $f_{uvM}(\theta;\mu,\kappa)$ is the unnormalized von Mises distribution, and $I(\theta;\mu,\kappa)$ is its distribution function. Therefore, $\int_a^b f_{uvM}(\theta;\mu,\kappa)=I(b;\mu,\kappa)-I(a;\mu,\kappa)$.

Considering that $\sum_{n=0}^{\infty}\frac{|(\kappa\cos(\theta-\mu))^n|}{n!}$ is a solely positive continuous bounded function in $[1,e^{\kappa}]$, and, therefore, for any finite integral coefficients $i_1,i_2\in\mathbb{R}$, it satisfies $\int_{i_1}^{i_2}\sum_{n=0}^{\infty}\frac{(\kappa\cos(\theta-\mu))^n}{n!}d\theta<$

∞, we can conclude that it satisfies the Fubini-Tonelli theorem conditions for integral summation exchange.

We then follow with the procedure for the indefinite integral:

$$\begin{aligned} I(\theta;\mu,\kappa) &= \int \sum_{n=0}^{\infty} \frac{(\kappa\cos(\theta-\mu))^n}{n!} d\theta \\ &= \sum_{n=0}^{\infty} \int \frac{(\kappa\cos(\theta-\mu))^n}{n!} d\theta \\ &= \sum_{n=0}^{\infty} \frac{\kappa^n}{n!} \int \cos^n(\theta-\mu) d\theta. \end{aligned} \tag{1}$$

The above integral is defined in a recursive way as

$$\int \cos^n(\theta-\mu)d\theta = \frac{\sin(\theta-\mu)\cos^{n-1}(\theta-\mu)}{n} + \frac{n-1}{n}\int \cos^{n-2}(\theta-\mu)d\theta.$$

And it can be calculated by the procedure of integration by parts. In this appendix, however, we give a non-recursive expression:

$$\int \cos^n(\theta-\mu)d\theta = \sin(\theta-\mu)\left(\sum_{i=0}^{\lfloor \frac{n}{2} \rfloor + \bmod \frac{n}{2} - 1} \left(\cos^{n-2i-1}(\theta-\mu)\frac{\prod_{j=0}^{2i}(n-j)}{\prod_{j=0}^{i}(n-2j)^2}\right)\right) \forall n \text{ such that } n = 2m+1$$

with $m \in \mathbb{N}$. This materializes out of the observation of the numerical regularities that appear when "unfolding" the recursive expression:

$$\begin{aligned} \int \cos^n(\theta-\mu)d\theta &= \frac{\sin(\theta-\mu)\cos^{n-1}(\theta-\mu)}{n} + \frac{n-1}{n}\int \cos^{n-2}(\theta-\mu)d\theta \\ &= \frac{\sin(\theta-\mu)\cos^{n-1}(\theta-\mu)}{n} + \frac{n-1}{n}\left(\frac{\sin(\theta-\mu)\cos^{n-3}(\theta-\mu)}{n-2} + \frac{n-3}{n-2}\int \cos^{n-4}(\theta-\mu)d\theta\right) \\ &= \frac{1}{n}\sin(\theta-\mu)\cos^{n-1}(\theta-\mu) + \frac{n-1}{n(n-2)}\sin(\theta-\mu)\cos^{n-3}(\theta-\mu) \\ &\quad + \frac{(n-1)(n-3)}{n(n-2)(n-4)}\sin(\theta-\mu)\cos^{n-5}(\theta-\mu) + \frac{(n-1)(n-3)(n-5)}{n(n-2)(n-4)}\int \cos^{n-6}(\theta-\mu)d\theta \end{aligned}$$

They can be primary generalized using the expression

$$\sin(\theta-\mu)\left(\sum_{i=0}^{\lfloor \frac{n}{2} \rfloor + \bmod \frac{n}{2} - 1} \left(\cos^{n-2i-1}(\theta-\mu)\frac{\prod_{j=0}^{2i}(n-j)}{\prod_{j=0}^{i}(n-2j)^2}\right)\right)$$

However, while this first expression does suffice for odd n, an extra term appears if n is even as we reach the point at which the term $\int \cos^0(\theta-\mu)d\theta$ is computed. This can be reflected properly by adding an addend that takes into account the parity of the formula. In our case, it has the form:

$$g(n,x) = \frac{(-1)^n h(x) + h(x)}{2} = \frac{((-1)^n + 1)h(x)}{2},$$

where $\forall n \in \mathbb{N}$ such that $n = 2m$ and $m \in \mathbb{N}, g(n,x) = h(x)$ and 0 otherwise.

In a shorter notation and adding the parity term, the expression becomes

$$\int \cos^n(\theta-\mu)d\theta = \sin(\theta-\mu)\left(\sum_{i=0}^{\lfloor \frac{n}{2} \rfloor + \bmod \frac{n}{2} - 1}\left(\cos^{n-2i-1}(\theta-\mu)\prod_{j=0}^{2i}(n-j)^{-(-1)^j}\right) + \frac{((-1)^n+1)\prod_{j=0}^{\lfloor \frac{n}{2} \rfloor + \bmod \frac{n}{2} - 1}(n-j)^{-(-1)^j}(\theta-\mu)}{2}\right).$$

Thus, substituting in (A.1) we obtain the final expression for $\int e^{\kappa\cos(\theta-\mu)}d\theta$.

1.3 Proof of Theorem 1

The theorem is entirely derived by means of the trigonometrical equality:

$$\begin{aligned}&\kappa_2\cos(x) + c_2\sin(x)\\ &= \left[\kappa_2\cos\left(\arctan\left(\frac{c_2}{\kappa_2}\right)\right) + c_2\sin\left(\arctan\left(\frac{c_2}{\kappa_2}\right)\right)\right]\cos\left(x - \arctan\left(\frac{c_2}{\kappa_2}\right)\right).\end{aligned} \tag{2}$$

From the equality we can express the exponent of the conditional distribution in (11) using a formula of the type $\kappa'\cos(x-\mu')$. Now if we consider that

$$\kappa_2\cos\left(\arctan\left(\frac{c_2}{\kappa_2}\right)\right) + c_2\sin\left(\arctan\left(\frac{c_2}{\kappa_2}\right)\right) = \frac{\kappa_2 + \frac{c_2^2}{\kappa_2}}{\sqrt{1+\left(\frac{c_2}{\kappa_2}\right)^2}} = \sqrt{\kappa_2^2 + c_2^2},$$

then (A.1) becomes

$$\kappa_2\cos(x) + c_2\sin(x) = \sqrt{\kappa_2^2+c_2^2}\cos\left(x - \arctan\left(\frac{c_2}{\kappa_2}\right)\right). \tag{3}$$

Thus, we can adapt the truncated conditional distribution to the univariate truncated von Mises exponent by properly selecting:

$$\begin{aligned}\kappa' &= \sqrt{\kappa_2^2+c_2^2}\\ \mu' &= \mu_2 + \arctan\left(\frac{c_2}{\kappa_2}\right),\end{aligned}$$

where $c_2 = \lambda\sin(\theta_1 - \mu_1)$.

1.4 Proof of Theorem 2

We consider

$$f_{umtvM}(\theta_{1'}) = e^{\kappa_1 \cos(\theta_{1'})} \int_{a_2}^{b_2} e^{\kappa_2 \cos(\theta_2 - \mu_2) + \lambda \sin(\theta_{1'}) \sin(\theta_2 - \mu_2)} d\theta_2 \tag{4}$$

to be the unnormalized marginal truncated von Mises distribution. For simplicity's sake, the proof is developed in a linear context (using classical intervals [x,y], with their associated constraints, instead of circular intervals $\mathbb{O}_{x,y}$), whose extension to the circle is deemed as known and trivial at this point. Also, unless otherwise specified, $\lambda > 0$ is assumed and a_2, b_2 truncation parameters are referred to simply as the truncation parameters. The proof is as follows:

(a) Determination of the derivative expression and the $T(\cdot,\cdot,\cdot,\cdot,\cdot,\cdot)$ function

(b) Analysis of the marginal expression with focus on the case of symmetrical truncation parameters in order to prove cases 1 and 2

(c) Further analysis for the case of non-symmetrical truncation parameters, determining all distinctive behaviors of the integral subterm of the marginal expression

(d) Monotony study divided by cases of the circular distance of the truncation parameters w.r.t. μ_2 and subintervals of the $\theta_{1'} \in [-\pi, \pi]$ interval in order to prove case 3. Case 4 is proven by ruling out every other possible outcome.

In (a), $T(\cdot,\cdot,\cdot,\cdot,\cdot,\cdot)$ is derived from a particularization of the second derivative of the marginal function. The meaning of the value of the $T(\cdot,\cdot,\cdot,\cdot,\cdot,\cdot)$ function is clarified for the symmetrical truncation parameters. In (b) and (c), the analysis aims to characterize the behavior of the integral term of the marginal distribution. In (b), the analysis will first observe the particularities of the integral term, especially, how $\theta_{1'}$ modifies the location and concentration parameters of the von Mises distribution inside the integral, and then derive from it some properties and insights will also be used for the proof of case 3. We then prove how thsese variations affect the area under the curve and their relationships to the truncation parameters. Finally, partial and total analyses of the derivate of the integral term are performed, concluding the proof of the first two cases of the theorem. In (c), an analysis of the derivate of the integral term for non-symmetrical truncation parameters w.r.t. μ_2 is performed. Using the previous insights, the analysis first determines the cases where, according to the truncation parameter values, the marginal integral term follows a unimodal distribution. The analysis then focuses on the remaining cases in order to prove that the global maximum of the integral term necessarily appears at the associated point of the truncation parameter ($-\frac{\pi}{2}$ for a_2 and $\frac{\pi}{2}$ for b_2), which has the largest circular distance w.r.t. μ_2. Also, in the bi-modal case for non-symmetrical truncation parameters, we analyze how the minimum comprehended between the modes appears in the $\frac{\pi}{2}$−length interval with 0 as an extrema associated with the truncation parameter that has the smallest circular distance w.r.t. μ_2 ($[-\frac{\pi}{2}, 0]$ for a_2 and $[0, \frac{\pi}{2}]$ for b_2), and its relationship with the minimum that appears in $[-\pi, -\frac{\pi}{2}]$ for the associated interval $[-\frac{\pi}{2}, 0]$ or in $[\frac{\pi}{2}, \pi]$ for the associated interval $[0, \frac{\pi}{2}]$. In (d), the monotony study identifies all different behaviors and the subinterval in which more than one critical point can occur, thus enabling us to detect bi-modality with different valued maxima with the proposed criteria.

(a) By differentiating $f_{umtvM}(\theta_{1'})$ w.r.t. θ_1 we obtain:

$$\begin{aligned} f'_{umtvM}(\theta_{1'}) &= -\kappa_1 \sin(\theta_{1'}) e^{\kappa_1 \cos(\theta_{1'})} \int_{a_2}^{b_2} e^{\kappa_2 \cos(\theta_2-\mu_2)+\lambda \sin(\theta_{1'}) \sin(\theta_2-\mu_2)} d\theta_2 \\ &\quad +\lambda \cos(\theta_{1'}) e^{\kappa_1 \cos(\theta_{1'})} \int_{a_2}^{b_2} \sin(\theta_2-\mu_2) e^{\kappa_2 \cos(\theta_2-\mu_2)+\lambda \sin(\theta_{1'}) \sin(\theta_2-\mu_2)} d\theta_2 \\ &= e^{\kappa_1 \cos(\theta_{1'})} \Big(-\kappa_1 \sin(\theta_{1'}) \int_{a_2}^{b_2} e^{\kappa_2 \cos(\theta_2-\mu_2)+\lambda \sin(\theta_{1'}) \sin(\theta_2-\mu_2)} d\theta_2 \\ &\quad +\lambda \cos(\theta_{1'}) \int_{a_2}^{b_2} \sin(\theta_2-\mu_2) e^{\kappa_2 \cos(\theta_2-\mu_2)+\lambda \sin(\theta_{1'}) \sin(\theta_2-\mu_2)} d\theta_2 \Big). \end{aligned} \tag{5}$$

We observe that

$$\begin{aligned} f'_{umtvM}(0) &= \lambda e^{\kappa_1} \left(\int_{a_2}^{b_2} \sin(\theta_2-\mu_2) e^{\kappa_2 \cos(\theta_2-\mu_2)} d\theta_2 \right) \\ &= \frac{\lambda}{\kappa_2} e^{\kappa_1} \left(e^{\kappa_2 \cos(a_2-\mu_2)} - e^{\kappa_2 \cos(b_2-\mu_2)} \right). \end{aligned} \tag{6}$$

If and only if $\cos(b_2-\mu_2) = \cos(a_2-\mu_2)$, it follows that $f_{umtvM}(\theta_{1'})$ has a critical point at μ_1.

Solving and assessing the equation $f''_{umtvM}(\theta_{1'}) = 0$ in order to obtain information about the curvature for $\theta_{1'} = 0$ results in

$$-\frac{\kappa_1}{\lambda^2} + \frac{\int_{a_2}^{b_2} \sin^2(\theta_2-\mu_2) e^{\kappa_2 \cos(\theta_2-\mu_2)} d\theta_2}{\int_{a_2}^{b_2} e^{\kappa_2 \cos(\theta_2-\mu_2)} d\theta_2} = 0,$$

from which we can define the $T(\cdot,\cdot,\cdot,\cdot,\cdot,\cdot)$ function as

$$T(\lambda, \mu_2, \kappa_1, \kappa_2, a_2, b_2) = -\frac{\kappa_1}{\lambda^2} + \frac{\int_{a_2}^{b_2} \sin^2(\theta_2-\mu_2) e^{\kappa_2 \cos(\theta_2-\mu_2)} d\theta_2}{\int_{a_2}^{b_2} e^{\kappa_2 \cos(\theta_2-\mu_2)} d\theta_2}. \tag{7}$$

However, we still need to understand whether Equation (A.7) is sufficient to distinguish between cases 1 and 2 established in the theorem.

(b) In order to understand the truncated marginal behavior, if we rewrite the integral term in $f_{umtvM}(\theta_{1'})$ by means of Equation (A.3) we have

$$f_{umtvM}(\theta_{1'}) = e^{\kappa_1 \cos(\theta_{1'})} \int_{a_2}^{b_2} e^{\sqrt{\kappa_2^2 + (\lambda \sin(\theta_{1'}))^2} \cos\left(x_2 - \mu_2 - \arctan\left(\frac{\lambda \sin(\theta_{1'})}{\kappa_2} \right) \right)} d\theta_2.$$

It is apparent that the integral term computes the area of location-concentration varying von Mises distributions as $\int_{a_2}^{b_2} f_{tvM}\left(\theta_2; \mu_2 + \arctan\left(\frac{\lambda \sin(\theta_{1'})}{\kappa_2}\right), \sqrt{\kappa_2^2 + (\lambda \sin(\theta_{1'}))^2}\right) d\theta_2$. If we consider the location variations over $[-\pi, \pi]$ by means of the $\sin(\theta_{1'})$ function, the distribution in the integrand is displaced over the interval $[-\arctan\left(\frac{\lambda}{\kappa_2}\right), 0]$ when $\sin(\theta_{1'}) < 0$ (from displacement 0 to displacement $-\arctan\left(\frac{\lambda}{\kappa_2}\right)$ when $\theta_{1'} \in [-\pi, -\frac{\pi}{2}]$ and from displacement $-\arctan\left(\frac{\lambda}{\kappa_2}\right)$ to displacement 0 when $\theta_{1'} \in [-\frac{\pi}{2}, 0]$), and over the interval $[0, \arctan\left(\frac{\lambda}{\kappa_2}\right)]$ when $\sin(\theta_{1'}) > 0$

(similary for $\theta_{1'} \in [0, \frac{\pi}{2}]$ and $\theta_{1'} \in [\frac{\pi}{2}, \pi]$). If we consider concentration variations, we can regard the source of bi-modality of the integral term as the $\sqrt{\kappa_2^2 + (\lambda \sin(\theta_{1'}))^2}$ subterm, given that $\sin^2(\theta_{1'})$ is a π−periodic solely positive function. Additionally, from $\theta_{1'} = 0$ to $\theta_{1'} = \frac{\pi}{2}$ and from $\theta_{1'} = -\pi$ to $\theta_{1'} = -\frac{\pi}{2}$, the concentration parameter grows from its minimum value κ_2 to its maximum value $\sqrt{\kappa_2^2 + \lambda^2}$, while it decreases from its maximum to its minimum value in the cases of $\theta_{1'}$ from $-\frac{\pi}{2}$ to 0 and from $\frac{\pi}{2}$ to π.

The proof then follows trivially by noting that, truncation parameters aside, the function's behavior in $[\mu_2 - \pi, \mu_2]$ can be considered symmetrical w.r.t. μ_2 to the function's behavior in $[\mu_2, \mu_2 + \pi]$. The symmetry w.r.t. μ_2 in the truncation parameters selects two subintervals of symmetrical behavior w.r.t. μ_1, thus producing a function that is symmetrical w.r.t. μ_1.

Further analyzing the integral term we look to determine the critical points and understand how the selection of truncation parameters affects the integral term behaviour. We take

$$\begin{aligned}
v_1(\theta_{1'}) &= \int_{a_2}^{b_2} e^{\kappa_2 \cos(\theta_2 - \mu_2) + \lambda \sin(\theta_{1'}) \sin(\theta_2 - \mu_2)} d\theta_2 \\
v_2(\theta_{1'}) &= \int_{a_2}^{b_2} \sin(\theta_2 - \mu_2) e^{\kappa_2 \cos(\theta_2 - \mu_2) + \lambda \sin(\theta_{1'}) \sin(\theta_2 - \mu_2)} d\theta_2,
\end{aligned}$$

where

$$\lambda \cos(\theta_{1'}) v_2(\theta_{1'}) = v_1'(\theta_{1'}).$$

We now want to analyze $v_2(\theta_{1'})$ as it is part of the derivate expression of $v_1(\theta_{1'})$. Taking the integrand of $v_2(\theta_{1'})$ to be

$$f_{v_2}(\theta_2; \theta_{1'}) = \sin(\theta_2 - \mu_2) e^{\kappa_2 \cos(\theta_2 - \mu_2) + \lambda \sin(\theta_{1'}) \sin(\theta_2 - \mu_2)}$$

Note that, in $f_{v_2}(\theta_2; \theta_{1'})$, the argument is θ_2 since it creates the area that is to be computed in $v_2(\theta_{1'})$. $\theta_{1'}$ can be considered here as a modifying parameter. The $f_{v_2}(\theta_2; \theta_{1'})$ function comprises the product of a strictly positive function $e^{(\cdot)}$ and a $\sin(\cdot)$ function. Therefore, the sign of $f_{v_2}(\theta_2; \theta_{1'})$ is solely determined by the sign of the $\sin(\cdot)$ function. To be precise, if $\theta_2 \in [\mu_2 - \pi, \mu_2]$ then $f_{v_2}(\theta_2; \theta_{1'}) \le 0$ and if $\theta_2 \in [\mu_2, \mu_2 + \pi]$ then $f_{v_2}(\theta_2; \theta_{1'}) \ge 0$. Therefore, we can subdivide $v_2(\theta_{1'})$ as

$$v_2(\theta_{1'}) = \int_{a_2}^{\mu_2} f_{v_2}(\theta_2; \theta_{1'}) d\theta_2 + \int_{\mu_2}^{b_2} f_{v_2}(\theta_2; \theta_{1'}) d\theta_2,$$

where the first addend is a solely negative term and the second addend is a solely positive term provided that $\mu_2 \in (a_2, b_2)$. In the symmetry case, if $\theta_{1'} = 0$ we have

$$-\int_{a_2}^{\mu_2} f_{v_2}(\theta_2; 0) d\theta_2 = \int_{\mu_2}^{b_2} f_{v_2}(\theta_2; 0) d\theta_2; \tag{8}$$

for $\theta_{1'} \in (0, \pi)$ we have

$$-\int_{a_2}^{\mu_2} f_{v_2}(\theta_2; \theta_{1'}) d\theta_2 < \int_{\mu_2}^{b_2} f_{v_2}(\theta_2; \theta_{1'}) d\theta_2; \tag{9}$$

and for $\theta_{1'} \in (-\pi, 0)$ we have

$$-\int_{a_2}^{\mu_2} f_{v_2}(\theta_2; \theta_{1'}) d\theta_2 > \int_{\mu_2}^{b_2} f_{v_2}(\theta_2; \theta_{1'}) d\theta_2 \tag{10}$$

Intuitively, the displaced exponential w.r.t. the μ_2 term increases all the values of either the negative or the positive curve of the $\sin(\theta_2 - \mu_2)$ function and reduces the curve of the opposite sign in less amount, therefore defining the sign and the value of $v_2(\theta_{1'})$. Formally, this to hold, we need to prove that $\forall \theta_{1'} \in (-\pi, 0)$ $f_{v_2}(\theta_2; 0) - f_{v_2}(\theta_2; \theta_{1'}) > 0$ if $\theta_2 \in (\mu_2 - \pi, \mu_2)$ and $\forall \theta_{1'} \in (-\pi, 0)$ $f_{v_2}(\theta_2; 0) - f_{v_2}(\theta_2; \theta_{1'}) < 0$ if $\theta_2 \in (\mu_2, \mu_2 + \pi)$ for the negative displacement, and an analogous statement for $\theta_{1'} \in (0, \pi)$ positive displacement. For the negative displacement case, it follows that

$$\begin{aligned} \sin(\theta_2 - \mu_2) e^{\kappa_2 \cos(\theta_2 - \mu_2)} - \sin(\theta_2 - \mu_2) e^{\kappa_2 \cos(\theta_2 - \mu_2) + \lambda \sin(\theta_{1'}) \sin(\theta_2 - \mu_2)} &> 0 \\ \sin(\theta_2 - \mu_2) \left(e^{\kappa_2 \cos(\theta_2 - \mu_2)} - e^{\kappa_2 \cos(\theta_2 - \mu_2) + \lambda \sin(\theta_{1'}) \sin(\theta_2 - \mu_2)} \right) &> 0. \end{aligned}$$

As $\sin(\theta_2 - \mu_2) < 0$ in $\theta_2 \in [\mu_2 - \pi, \mu_2]$ it suffices if

$$e^{\kappa_2 \cos(\theta_2 - \mu_2)} - e^{\kappa_2 \cos(\theta_2 - \mu_2) + \lambda \sin(\theta_{1'}) \sin(\theta_2 - \mu_2)} < 0$$

in $\theta_2 \in [\mu_2 - \pi, \mu_2]$. We proceed as follows:

$$\begin{aligned} e^{\kappa_2 \cos(\theta_2 - \mu_2)} - e^{\kappa_2 \cos(\theta_2 - \mu_2) + \lambda \sin(\theta_{1'}) \sin(\theta_2 - \mu_2)} &< 0 \\ e^{-\lambda \sin(\theta_{1'}) \sin(\theta_2 - \mu_2)} &< 1 \\ -\lambda \sin(\theta_{1'}) \sin(\theta_2 - \mu_2) &< 0 \end{aligned}$$

and, since we have specified $\theta_{1'} \in (-\pi, 0)$ and then $\sin(\theta_{1'}) < 0$, we have $-\lambda \sin(\theta_{1'}) > 0$. Therefore, the sign of $-\lambda \sin(\theta_{1'}) \sin(\theta_2 - \mu_2)$ follows from that of $\sin(\theta_2 - \mu_2)$. This proves the statement for both θ_2 intervals in the case of negative displacement. The proof for positive displacement is analogous.

This result implies that the selection of truncation parameters that are symmetrical w.r.t. μ_2 does not change the monotony of $v_1(\theta_{1'})$. More generally, this result implies that no selection of truncation parameters changes the monotonicity of $v_2(\theta_{1'})$, that is, increasing in $[-\frac{\pi}{2}, \frac{\pi}{2}]$ and decreasing otherwise.

Since (A.8), (A.9) and (A.10) hold, we can now perform the sign and critical points analysis of $\lambda \cos(\theta_{1'}) v_2(\theta_{1'})$ to obtain that $v_1(\theta_{1'})$ follows the monotony of $\sin^2(\theta_{1'})$ for any a_2, b_2 such that $\cos(b_2 - \mu_2) = \cos(a_2 - \mu_2)$, with critical points $\{-\frac{\pi}{2}, 0, \frac{\pi}{2}\}$. Therefore, in Equation (A.4), unimodal/bimodal observed distributions are "decided" for this case by the product of $v_1(\theta_{1'})$ with $e^{\kappa_1 \cos(\theta_{1'})}$.

Therefore, if $T(\lambda, \mu_2, \kappa_1, \kappa_2, a_2, b_2) > 0$ then $f_{umtvM}(\theta_{1'})$ presents a minimum critical point at μ_1 and the distribution has two equal symmetrical maxima in $[-\frac{\pi}{2}, \frac{\pi}{2}]$ (the maxima location interval can be proven as a result of monotony and sign comparisons between $v_1(\theta_{1'})$ and $e^{\kappa_1 \cos(\theta_{1'})}$). Respectively, if $T(\lambda, \mu_2, \kappa_1, \kappa_2, a_2, b_2) < 0$ then $f_{umtvM}(\theta_{1'})$ presents a maximum critical point and the distribution is unimodal. This result generalizes the outcome for the non-truncated case to symmetrical parameters other than a_2, b_2 such that $b_2 - a_2 = 2\pi$ (Singh (2002)). This suffices to prove cases 1 and 2 of the theorem.

(c) For case 3 we want to observe the behavior of the marginal distribution for different cases of circular distances of a_2, b_2 truncation parameters w.r.t. μ_2. Thus, we need knowledge about the subterm $v_2(\theta_{1'})$ when a_2, b_2 truncation parameters are not symmetrical w.r.t. μ_2 in order to reach useful results. We will address this point first.

If we now observe $\lambda\cos(\theta_{1'})v_2(\theta_{1'}) = 0$ for non-symmetrical parameters we can as before, isolate two critical points:

$$\begin{aligned}\theta_{1'} &= -\frac{\pi}{2},\\ \theta_{1'} &= \frac{\pi}{2}\end{aligned}$$

and a third critical point at some $\theta_{1'}$ such that $-\int_{a_2}^{\mu_2} f_{v_2}(\theta_2;\theta_{1'}) + \int_{\mu_2}^{b_2} f_{v_2}(\theta_2;\theta_{1'}) = 0$ if a_2, b_2 are not truncation parameters that satisfy any of the following conditions:

(i) $a_2, b_2 \in [\mu_2, \mu_2 + \pi]$ as then $v_2(\theta_{1'}) > 0\ \forall\theta_{1'} \in [-\pi, \pi]$

(ii) $a_2, b_2 \in [\mu_2 - \pi, \mu_2]$ as then $v_2(\theta_{1'}) < 0\ \forall\theta_{1'} \in [-\pi, \pi]$

(iii) $\mu_2 \in (a_2, b_2)$ such as $-\int_{a_2}^{\mu_2} f_{v_{2'}}(\theta_2; -\frac{\pi}{2})d\theta_2 \le \int_{\mu_2}^{b_2} f_{v_{2'}}(\theta_2; -\frac{\pi}{2})d\theta_2$ as then $v_2(\theta_{1'}) > 0\ \forall\theta_{1'} \in [-\pi, \pi]$

(iv) $\mu_2 \in (a_2, b_2)$ such as $\int_{\mu_2}^{b_2} f_{v_{2'}}(\theta_2; \frac{\pi}{2})d\theta_2 \le -\int_{a_2}^{\mu_2} f_{v_{2'}}(\theta_2; \frac{\pi}{2})d\theta_2$ as then $v_2(\theta_{1'}) < 0\ \forall\theta_{1'} \in [-\pi, \pi]$.

Notice that from the viewpoint of truncation parameters, cases (iii) and (iv) can be considered opposite. Also, as highlighted by the previous analysis, it is clear that case (iii) implies $\cos(b_2 - \mu_2) < \cos(a_2 - \mu_2)$ (more intuitively, $\cos(b_2 - \mu_2) \ll \cos(a_2 - \mu_2)$) and case (iv) $\cos(b_2 - \mu_2) > \cos(a_2 - \mu_2)$ (more intuitively, $\cos(b_2 - \mu_2) \gg \cos(a_2 - \mu_2)$). We will refer to cases (iii) and (iv) as the strong lower parameter cases.

Therefore, by manipulating a_2, b_2 truncation parameters, it is possible to reshape $v_1(\theta_{1'})$ to exhibit a minimum in $-\frac{\pi}{2}$ and a maximum in $\frac{\pi}{2}$ if case (i) or (iii) applies or to exhibit a maximum in $-\frac{\pi}{2}$ and a minimum in $\frac{\pi}{2}$ if case (ii) or (iv) applies. In these cases, $v_1(\theta_{1'})$ is an integral term with unimodal behavior.

It follows that any other case for non-symmetrical truncation parameters implies $\mu_2 \in (a_2, b_2)$, and $v_1(\theta_{1'})$ exhibits two differentiated maxima in $-\frac{\pi}{2}$ and $\frac{\pi}{2}$. Also, $v_2(-\frac{\pi}{2}) < 0$ and $v_2(\frac{\pi}{2}) > 0$. If we examine the case of $\theta_{1'} = 0$ for truncation parameters a_2, b_2 such that $\cos(b_2 - \mu_2) > \cos(a_2 - \mu_2)$ then $-\int_{a_2}^{\mu_2} f_{v_2}(\theta_2; 0)d\theta_2 > \int_{\mu_2}^{b_2} f_{v_2}(\theta_2; 0)d\theta_2$ and therefore $v_2(\theta_{1'}) = 0$ for some $\theta_{1'}^* \in [0, \frac{\pi}{2}]$ such that $v_2(\theta_{1'}) < 0$ if $\theta_{1'} \in [0, \theta_{1'}^*)$ and $v_2(\theta_{1'}) > 0$ if $\theta_{1'} \in (\theta_{1'}^*, \frac{\pi}{2}]$. It follows that this also impliesthe existence of another minimum in $[\frac{\pi}{2}, \pi]$ as $v_2(\theta_{1'}) > 0\ \forall\theta_{1'} \in [\frac{\pi}{2}, \pi - \theta_{1'}^*)$ and $v_2(\theta_{1'}) < 0\ \forall\theta_{1'} \in (\pi - \theta_{1'}^*, \pi]$. Similarly, if $\cos(b_2 - \mu_2) < \cos(a_2 - \mu_2)$ then $-\int_{a_2}^{\mu_2} f_{v_2}(\theta_2; 0)d\theta_2 < \int_{\mu_2}^{b_2} f_{v_2}(\theta_2; 0)d\theta_2$ and therefore $v_2(\theta_{1'}) = 0$ for some $\theta_{1'}^* \in [-\frac{\pi}{2}, 0]$ and $-\pi - \theta_{1'}^* \in [-\pi, -\frac{\pi}{2}]$, that is, the minimum of $v_1(\theta_{1'})$ that appears in $[-\frac{\pi}{2}, \frac{\pi}{2}]$ is more precisely located in the $\frac{\pi}{2}$–length interval associated with the truncation parameter that presents the smallest circular distance w.r.t. μ_2 and imples an additional minimum located in the contiguous $\frac{\pi}{2}$–length interval more distant from $\theta_{1'} = 0$.

Additionally, the global maximum of the two differentiated maxima is that of the $\frac{\pi}{2}$–length interval associated with the truncation parameter that has the largest circular distance w.r.t. μ_2. We can prove this by comparing both maxima as follows:

$$v_1\left(-\frac{\pi}{2}\right) - v_1\left(\frac{\pi}{2}\right) > 0 \text{ if } \cos(b_2 - \mu_2) > \cos(a_2 - \mu_2).$$

Thus if we take $\kappa' = \sqrt{\kappa_2^2 + (\lambda)^2}$ we have

$$\int_{a_2}^{b_2} e^{\kappa' \cos\left(\theta_2 - \mu_2 - \arctan\left(-\frac{\lambda}{\kappa_2}\right)\right)} d\theta_2 - \int_{a_2}^{b_2} e^{\kappa' \cos\left(\theta_2 - \mu_2 - \arctan\left(\frac{\lambda}{\kappa_2}\right)\right)} d\theta_2 > 0.$$

Expressing this by means of the distribution function we obtain

$$\left[I(\theta, -\mu_2 - \arctan\left(-\frac{\lambda}{\kappa_2}\right), \kappa')\right]_{a_2}^{b_2} - \left[I(\theta, -\mu_2 - \arctan\left(\frac{\lambda}{\kappa_2}\right), \kappa')\right]_{a_2}^{b_2} > 0. \quad (11)$$

Clearly, $I(\theta, \mu, \kappa)$ is strictly increasing and $e^{\kappa' \cos\left(\theta_2 - \mu_2 - \arctan\left(-\frac{\lambda}{\kappa_2}\right)\right)}$ is symmetrical to $e^{\kappa' \cos\left(\theta_2 - \mu_2 - \arctan\left(\frac{\lambda}{\kappa_2}\right)\right)}$ w.r.t. μ_2. Therefore

1. $$\left[I(\theta, -\mu_2 - \arctan\left(\frac{-\lambda}{\kappa_2}\right), \kappa')\right]_{2\mu_2 - b_2}^{\mu_2} = \left[I(\theta, -\mu_2 - \arctan\left(\frac{\lambda}{\kappa_2}\right), \kappa')\right]_{\mu_2}^{b_2}$$

2. $$\left[I(\theta, -\mu_2 - \arctan\left(\frac{-\lambda}{\kappa_2}\right), \kappa')\right]_{\mu_2}^{2\mu_2 - a_2} = \left[I(\theta, -\mu_2 - \arctan\left(\frac{\lambda}{\kappa_2}\right), \kappa')\right]_{a_2}^{\mu_2}$$

taking

$$\left[I(\theta, -\mu_2 - \arctan\left(\frac{-\lambda}{\kappa_2}\right), \kappa')\right] = Ie_1(\theta)$$
$$\left[I(\theta, -\mu_2 - \arctan\left(\frac{\lambda}{\kappa_2}\right), \kappa')\right] = Ie_2(\theta),$$

we can rewrite inequation (A.11) as

$$[Ie_1(\theta)]_{a_2}^{\mu_2} + [Ie_1(\theta)]_{\mu_2}^{b_2} - [Ie_2(\theta)]_{a_2}^{\mu_2} - [Ie_2(\theta)]_{\mu_2}^{b_2} > 0,$$

substituting,

$$\begin{aligned} [Ie_1(\theta)]_{\mu_2}^{a_2} + [Ie_1(\theta)]_{b_2}^{\mu_2} - [Ie_1(\theta)]_{\mu_2}^{2\mu_2 - a_2} - [Ie_1(\theta)]_{2\mu_2 - b_2}^{\mu_2} &> 0 \\ -Ie_1(a_2) + Ie_1(b_2) - Ie_1(2\mu_2 - a_2) + Ie_1(2\mu_2 - b_2) &> 0 \\ [Ie_1(\theta)]_{a_2}^{2\mu_2 - b_2} - [Ie_1(\theta)]_{b_2}^{2\mu_2 - a_2} &> 0, \end{aligned}$$

that is, the inequation reduces to the comparison between the area in two subintervals of equal length that are symmetrical w.r.t. μ_2. By this symmetry and by the fact that the mode is in $(-\frac{\pi}{2}, 0)$ and the anti-mode in $(\frac{\pi}{2}, \pi)$ in $e^{\kappa' \cos\left(\theta_2-\mu_2-\arctan\left(-\frac{\lambda}{\kappa_2}\right)\right)}$, we can safely conclude that the inequation holds thus proving the statement. Therefore, for any marginal truncated distribution, the global maximum in the integral term is located in $\theta_{1'} = \frac{\pi}{2}$ if $\cos(a_2 - \mu_2) > \cos(b_2 - \mu_2)$ and in $\theta_{1'} = -\frac{\pi}{2}$ if $\cos(a_2 - \mu_2) < \cos(b_2 - \mu_2)$.

At this point all behaviors for critical points and monotony of $v_1(\theta_{1'})$ have been characterized.
Analogously to the non-truncated case, the effect of the $e^{\kappa_1 \cos(\theta_{1'})}$ subterm has to be taken into consideration in order to determine the shape of the distribution. To do this, we perform a monotony study that incorporates all previous developments.

(d) After conducting the study on $v_2(\theta_{1'})$ and $v_1(\theta_{1'})$, we proceed by equating function (A.5) to zero, resulting in

$$-\kappa_1 \sin(\theta_{1'})v_1(\theta_{1'}) + \lambda \cos(\theta_{1'})v_2(\theta_{1'}) = 0.$$

If we consider the cases where $a_2, b_2 \in [\mu_2, \mu_2 + \pi]$ or a_2 is a strong lower parameter w.r.t b_2 we have:

1. $v_2(\theta_{1'}) > 0 \ \forall \theta_{1'} \in [-\pi, \pi]$.

2. If $\theta_{1'} \in [-\pi, -\frac{\pi}{2}]$, then $\sin(\theta_{1'}) \leq 0$ and $\cos(\theta_{1'}) \leq 0$. In this case, at least a minimum and a critical point of $f_{umtvM}(\theta_{1'})$ can be found in the examined interval as shown by:

$$\begin{aligned} f'_{umtvM}(-\pi) &= e^{-\kappa_1}\left(-\lambda \int_{a_2}^{b_2} \sin(\theta_2 - \mu_2) e^{\kappa_2 \cos(\theta_2-\mu_2)} d\theta_2\right) \\ f'_{umtvM}\left(-\frac{\pi}{2}\right) &= \kappa_1 \int_{a_2}^{b_2} e^{\kappa_2 \cos(\theta_2-\mu_2)-\lambda \sin(\theta_2-\mu_2)} d\theta_2 > 0, \end{aligned}$$

   where $f'_{umtvM}(-\pi) < 0$. Notice that if $a_2, b_2 \in [\mu_2, \mu_2 + \pi]$ the critical point neccesarily exists regardless of the effect of the other parameters.

3. If $\theta_{1'} \in [-\frac{\pi}{2}, 0]$, then $\sin(\theta_{1'}) \leq 0$ and $\cos(\theta_{1'}) \geq 0$. $f_{umtvM}(\theta_{1'})$ exhibits a monotonic increasing behavior, as all terms involved in the expression are positive.

4. If $\theta_{1'} \in [0, \frac{\pi}{2}]$, then $\sin(\theta_{1'}) \geq 0$ and $\cos(\theta_{1'}) \geq 0$. Here, at least a maximum and a critical point can be found in the interval by considering Equation (A.6), where $f'_{umtvM}(0) > 0$, and
$$f'_{umtvM}\left(\frac{\pi}{2}\right) = -\kappa_1 \int_{a_2}^{b_2} e^{\kappa_2 \cos(\theta_2-\mu_2)-\lambda \sin(\theta_2-\mu_2)} d\theta_2 < 0.$$

5. If $\theta_{1'} \in [\frac{\pi}{2}, \pi]$, then $\sin(\theta_{1'}) \geq 0$ and $\cos(\theta_{1'}) \leq 0$. $f_{umtvM}(\theta_{1'})$ exhibits a monotonic decreasing behavior, as all terms involved in the expression are negative.

Therefore, for this case, the distribution exhibits critical points in two non-contiguous intervals. By the previous developments, such a distribution of critical points would only correspond to the unimodal case and also, as the contribution of $e^{\kappa_1 \cos(\theta_{1'})}$ is symmetrical w.r.t. μ_1 or $\theta_{1'} = 0$, the marginal function could only have one global maximum in $\theta_{1'} \in [0, \frac{\pi}{2}]$ interval and one global minimum in $\theta_{1'} \in [-\pi, -\frac{\pi}{2}]$.

The case where $a_2, b_2 \in [\mu_2 - \pi, \mu_2]$ or b_2 is a strong lower parameter w.r.t a_2 can be understood as "symmetric behavior w.r.t μ_1", since the results for $\theta_{1'} \in [-\pi, -\frac{\pi}{2}]$ now hold for $\theta_{1'} \in [\frac{\pi}{2}, \pi]$ and the results for $\theta_{1'} \in [-\frac{\pi}{2}, 0]$ now hold for $\theta_{1'} \in [0, \frac{\pi}{2}]$. This property, general to the $[-\pi, \pi]$ interval, guarantees that in our case, it suffices to determine the behavior for one of the two remaining cases to completely determine the behavior of the marginal function.

We now consider the remaining parameter configurations that satisfy $\cos(b_2 - \mu_2) > \cos(a_2 - \mu_2)$.

1. If $\theta_{1'} \in [-\pi, -\frac{\pi}{2}]$, then $v_2(\theta_{1'}) < 0$, thus resulting in $f_{umtvM}(\theta_{1'})$, which exhibits a strictly increasing behavior, as all terms involved in the expression are now positive.

2. If $\theta_{1'} \in [-\frac{\pi}{2}, 0]$, then $v_2(\theta_{1'}) < 0$. In this case, after performing sign comparisons on the extrema, there is at least one critical point and one maximum in the interval.

3. If $\theta_{1'} \in [0, \frac{\pi}{2}]$, $v_2(\theta_{1'}) < 0 \ \forall \theta_{1'} \in [0, \theta_{1'}^*)$ and $v_2(\theta_{1'}) > 0 \ \forall \theta_{1'} \in [\theta_{1'}^*, \frac{\pi}{2})$. Therefore, no critical point exists in $[0, \theta_{1'}^*)$, since $f_{umtvM}(\theta_{1'})$ exhibits a decreasing behavior and all terms involved in the expression are negative. In $[\theta_{1'}^*, \frac{\pi}{2})$, no, one or two critical points can occur as both sign and monotony comparisons were not conclusive.

4. If $\theta_{1'} \in [\frac{\pi}{2}, \pi]$, then $v_2(\theta_{1'}) > 0 \ \forall \theta_{1'} \in [\frac{\pi}{2}, \pi - \theta_{1'}^*)$ and $v_2(\theta_{1'}) < 0 \ \forall \theta_{1'} \in (\pi - \theta_{1'}^*, \pi]$. Therefore, no critical point exists in $[\frac{\pi}{2}, \pi - \theta_{1'}^*)$ since $f_{umtvM}(\theta_{1'})$ exhibits a decreasing behavior as all terms involved in the expression are negative. In $(\pi - \theta_{1'}^*, \pi]$, after performing sign comparisons on the extrema, at least one critical point can occur. Therefore, for this case, the distribution has three contiguous intervals containing critical points. Since clearly no more than two critical points are allowed in a $\frac{\pi}{2}$–length interval, the case with two possible critical points in $[\theta_{1'}^*, \frac{\pi}{2})$ is the case of bi-maximality (differentiated maxima) with a minimum and a maximum in $\theta_{1'} \in [\theta_{1'}^*, \frac{\pi}{2})$ and a maximum in $\theta_{1'} \in [-\frac{\pi}{2}, 0]$. Complementarily, this distribution of critical points "corresponds" to the bi-maximal (differentiated maxima) behavior of $v_1(\theta_{1'})$, and, therefore, the critical point in $\theta_{1'} \in [-\frac{\pi}{2}, 0]$ is necessarily a maximum, and the critical point in $[\frac{\pi}{2}, \pi]$ is necessarily a minimum. Thus, it can be concuded that in the case of bimodality, the interval associated with the truncation parameter that has the shortest circular distance w.r.t. μ_2 contains the two critical points, whereas the interval associated with the truncation parameter that has the largest circular distance w.r.t. μ_2 contains the global maximum.

If $\lambda < 0$, the proof follows trivially by noting that the displacement caused by the $\sin(\cdot)$ function in the exponent that appears in the $v_1(\theta_{1'})$ subterm is the opposite. This in turn causes the distribution to have an opposite symmetrical behaviour w.r.t. μ_1. This suffices to prove case 3 of the theorem. Case 4 can also be proven with the developed theory. However, it can additionally be proven by ruling out any other possible outcome, considering the three previously developed cases.

References

Singh, H. (2002), “Probabilistic model for two dependent circular variables,” *Biometrika*, 89(3), 719–723.