Toroidal Information Fusion
Based on the Bivariate von Mises Distribution

Gerhard Kurz* and Uwe D. Hanebeck*

***Abstract*— Fusion of toroidal information, such as correlated angles, is a problem that arises in many fields ranging from robotics and signal processing to meteorology and bioinformatics. For this purpose, we propose a novel fusion method based on the bivariate von Mises distribution. Unlike most literature on the bivariate von Mises distribution, we consider the full version with matrix-valued parameter rather than a simplified version. By doing so. we are able to derive the exact analytical computation of the fusion operation. We also propose an efficient approximation of the normalization constant including an error bound and present a parameter estimation algorithm based on a maximum likelihood approach. The presented algorithms are illustrated through examples.**

I. Introduction

Many applications involve the consideration of correlated angles or other periodic quantities. For example, the direction a person's head is facing and the direction the torso is facing are highly correlated. Other examples are the wind direction at two different measurement locations or at two different times, the orientation of two joints of a robotic arm, or the phase of a signal received over multiple paths.

Motivated by the large number of highly relevant problems in a variety of areas, we propose a novel method for the fusion of measurements of correlated angles in this paper. Because each angle can be represented by a point on the unit circle, we consider the Cartesian product of two unit circles, which is given by the torus.

As most commonly used distributions such as the Gaussian distribution are defined on $\mathbb{R}^n$, they are not suitable to represent uncertainty on the torus (although they can still be used to locally approximate a toroidal distribution). Consequently, it is important to consider probability distributions defined on the torus. We resort to the field of directional statistics [1], a subfield of statistics that deals with distributions on certain manifolds, the torus being one example. One suitable toroidal distribution is the bivariate von Mises distribution (see Fig. 1), which was proposed by Mardia in [2] and [3] (see also [1, Sec. 3.7.1]). This distribution constitutes a generalization of the univariate von Mises distribution [4], a commonly used distribution on the unit circle. Special cases of the bivariate von Mises distribution were considered by Rivest [5], Singh et al. [6], and Mardia et al. [7], [8]. A possible generalization to higher dimensions was discussed in [9].

* The authors are with the Intelligent Sensor-Actuator-Systems Laboratory (ISAS), Institute for Anthropomatics and Robotics, Karlsruhe Institute of Technology (KIT), Germany. E-mail: gerhard.kurz@kit.edu, uwe.hanebeck@ieee.org

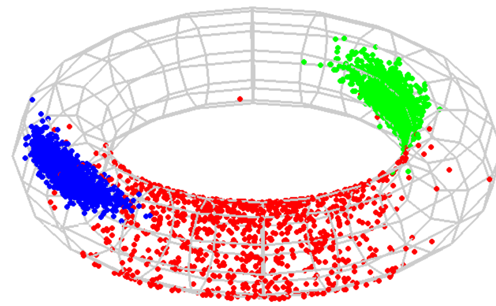


Fig. 1: Samples on the torus drawn from three bivariate von Mises distributions with different parameters.

The bivariate von Mises distribution is not the only toroidal probability distribution capable of representing correlations. An interesting alternative is the bivariate wrapped normal distribution (see, for example, [10]). We previously investigated the applicability of this distribution for recursive filtering in [11]. Although the bivariate wrapped normal has a number of nice properties and is well-suited for propagation of information, it is not closed under multiplication, which makes the fusion operation difficult and computationally expensive. Some other distributions can be found in literature, such as a bivariate distribution with von Mises marginals[1] [12], a model based on Fourier series [13], and a distribution induced by a Möbius transformation [14].

The contributions of this paper can be summarized as follows. First, we present a series representation for the normalization constant of the bivariate von Mises distribution that can be used to efficiently approximate its value with fairly low computational cost. There has been some previous work on the normalization constant for certain special cases, but our method is to the authors' knowledge the first result for the general bivariate von Mises. Second, we propose a parameter estimation scheme based on maximum likelihood estimation of the distribution's parameters. Third, we prove that the general form of the bivariate von Mises distribution is closed under multiplication and present an efficient algorithm to analytically calculate the product of two bivariate von Mises distributions. This method allows efficient Bayesian fusion on the torus.

[1]The bivariate von Mises distribution counterintuitively does not have von Mises marginals in general.

II. Bivariate von Mises Distribution

Before we introduce the bivariate von Mises distribution, we briefly review the properties of the univariate von Mises distribution [4]. Its probability density function (pdf) is given by

$$f(x;\mu,\kappa) = \frac{1}{2\pi I_0(\kappa)} \exp(\kappa\cos(x-\mu)) \ ,$$

where $x \in [0, 2\pi)$, $\mu \in [0, 2\pi)$ and $\kappa \geq 0$. The normalization constant contains $I_0(\kappa)$, the modified Bessel function of the first kind [15, Sec. 9.6]. The parameter μ determines the location of the mode of the distribution, whereas κ is a concentration parameter that represents the uncertainty.

The bivariate von Mises (BVM) distribution is defined by the probability density function

$$\begin{aligned} f(\underline{x}; \underline{\mu}, \underline{\kappa}, \mathbf{A}) &= C(\underline{\kappa}, \mathbf{A}) \cdot \exp\Bigg(\kappa_1 \cos(x_1-\mu_1) + \kappa_2\cos(x_2-\mu_2) \\ &\quad + \begin{bmatrix} \cos(x_1-\mu_1) \\ \sin(x_1-\mu_1) \end{bmatrix}^T \mathbf{A} \begin{bmatrix} \cos(x_2-\mu_2) \\ \sin(x_2-\mu_2) \end{bmatrix} \Bigg) \ , \end{aligned} \tag{1}$$

where $\underline{x} \in [0,2\pi)^2$, $\underline{\mu} \in [0,2\pi)^2$,

$$\mathbf{A} = \begin{bmatrix} a_{1,1} & a_{1,2} \\ a_{2,1} & a_{2,2} \end{bmatrix} \in \mathbb{R}^{2\times 2} \ ,$$

and $C(\underline{\kappa}, \mathbf{A}) > 0$ is the normalization constant[2]. We discuss the calculation of $C(\underline{\kappa}, \mathbf{A})$ in detail in Sec. III. The parameter $\underline{\mu}$ once again controls the location of the distribution, whereas κ controls the concentration. Additionally, there is the parameter $\mathbf{A}$ that influences the correlation of the two angles. It is important to emphasize that $\mathbf{A}$ is an arbitrary 2×2 matrix, i.e., it does not need to be symmetric, positive definite, or invertible. To illustrate the influence of the matrix $\mathbf{A}$, we depict several examples of the resulting pdf in Fig. 2. As can be seen, the pdf can become bimodal for certain choices of $\mathbf{A}$, which is typically not intended[3]. Also, it is noteworthy that there are four parameters, the entries of $\mathbf{A}$, parameterizing the correlation. This may seem surprising as intuitively a single scalar parameter should be sufficient (similar to the bivariate wrapped normal distribution [11]).

The marginals of the BVM distribution with matrix parameter can be obtained according to

$$\begin{aligned} f(x_1) &= \int_0^{2\pi} f(x_1, x_2) dx_2 \\ &= C \cdot \exp(\kappa_1 \cos(x_1 - \mu_1)) \\ &\quad \cdot \int_0^{2\pi} \exp(\alpha\cos(x_2)) \exp(\beta \sin(x_2)) dx_2 \\ &= C \cdot \exp(\kappa_1\cos(x_1-\mu_1)) \cdot 2\pi I_0(\sqrt{\alpha^2+\beta^2}) \end{aligned}$$

where

$$\begin{aligned} \alpha &= \kappa_2 + \cos(x_1-\mu_1)a_{11} + \sin(x_1-\mu_1)a_{21} \ , \\ \beta &= \sin(x_1-\mu_1)a_{22} + \cos(x_1-\mu_1)a_{12} \ . \end{aligned}$$

Note that the marginals are VM distributed if $\mathbf{A}$ is a zero-matrix, but not in general. This result constitutes a generalization of the results by Singh et al. [6, eq. (2.2)] and Mardia et al. [7, eq. (5)].

Several special cases of the bivariate von Mises distribution have been considered in literature, in which $\mathbf{A}$ was restricted to fulfill certain properties. Jupp et al. have discussed a distribution, where $\mathbf{A}$ is a multiple of an orthogonal matrix [17, eq. (3·3)]. Another subclass was discussed by Rivest [5], in which $\mathbf{A}$ is restricted to be a diagonal matrix, i.e., $a_{1,2} = a_{2,1} = 0$. Later, further special cases of Rivest's model were considered by Singh et al. [6] and Mardia et al. [7]. Singh et al. proposed the sine model, where $a_{1,1} = a_{1,2} = a_{2,1} = 0$ and only $a_{2,2} \in \mathbb{R}$ can be chosen. Analogously, Mardia et al. proposed the cosine model, where $a_{1,2} = a_{2,1} = a_{2,2} = 0$ and only $a_{1,1} \in \mathbb{R}$ can be chosen. Both models are compared in [7]. Several interesting results have been published for these special cases, e.g., solutions for the normalization constant, bimodality conditions, marginals, parameter estimation schemes, etc.

However, very little is known about the general bivariate von Mises distribution, as it is more difficult to analyze than its special cases. In this paper, we seek to address this deficiency and present several results for the general bivariate von Mises distribution. The following sections will be devoted to deriving an approximation of the normalization constant, presenting a maximum likelihood estimator, and proposing a method for Bayesian fusion.

III. Normalization Constant

In this section, we derive a method for calculating the normalization constant of a bivariate von Mises distribution. Some authors such as Singh et al. [6], Rivest [5], and Jupp et al. [17] have previously derived the normalization constant for certain special cases of the bivariate von Mises distribution but their solutions do not apply to the more general distribution with matrix parameter $\mathbf{A}$.

A. Series Representation

Obviously, the normalization constant is independent of the location of the distribution, i.e., we do not need to consider the value of $\underline{\mu}$. By rewriting the exponential function using its power series representation [15, 4.2.1], we obtain

$$\begin{aligned} C(\kappa, \mathbf{A})^{-1} = \int_0^{2\pi}\int_0^{2\pi} \exp\Bigg(\kappa_1\cos(x_1) + \kappa_2\cos(x_2) \\ + \begin{bmatrix} \cos(x_1) \\ \sin(x_1) \end{bmatrix}^T \mathbf{A} \begin{bmatrix} \cos(x_2) \\ \sin(x_2) \end{bmatrix} \Bigg) dx_1 dx_2 \end{aligned}$$

[2] Some authors use a slightly different (but equivalent) parameterization, in which angles are represented by unit vectors, see e.g., [16, eq. (2.10)]

[3] The exact condition when the density is bimodal are not yet known for the general case of a bivariate von Mises distribution. Bimodality conditions for the sine and cosine models are given in [7, Theorem 2,3]

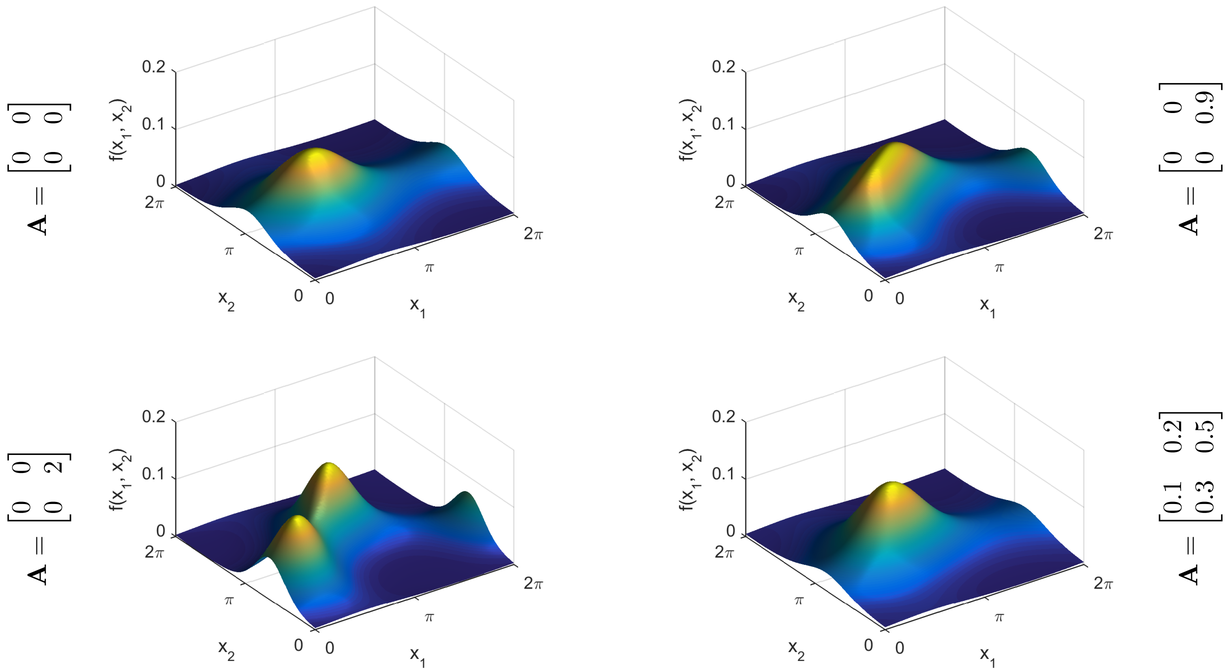


Fig. 2: Bivariate von Mises distributions with $\underline{\mu} = [2, 3]^T$, $\underline{\kappa} = [0.7, 1.3]^T$, for different correlation matrices $\mathbf{A}$. Note that both x_1 and x_2 are 2π-periodic.

$$= \sum_{n=0}^{\infty} \frac{1}{n!} \int_0^{2\pi} \int_0^{2\pi} \Bigg(\kappa_1 \cos(x_1) + \kappa_2 \cos(x_2) + \begin{bmatrix} \cos(x_1) \\ \sin(x_1) \end{bmatrix}^T \mathbf{A} \begin{bmatrix} \cos(x_2) \\ \sin(x_2) \end{bmatrix} \Bigg)^n dx_1 dx_2$$
$$= \sum_{n=0}^{\infty} \frac{s_n(\underline{\kappa}, \mathbf{A})}{n!} \ , \qquad (2)$$

where we define for $n \in \mathbb{N}_0$

$$s_n(\underline{\kappa}, \mathbf{A}) = \int_0^{2\pi} \int_0^{2\pi} \big(\kappa_1 \cos(x_1) + \kappa_2 \cos(x_2) + \cos(x_1) a_{1,1} \cos(x_2) + \cos(x_1) a_{1,2} \sin(x_2) + \sin(x_1) a_{2,1} \cos(x_2) + \sin(x_1) a_{2,2} \sin(x_2)\big)^n dx_1 dx_2 \ .$$

The term $s_n(\underline{\kappa}, \mathbf{A})$ can be calculated analytically for arbitrary n, but the solution gets complicated for large values of n. We give solutions for the first few terms in the Appendix. In the case of low concentrations, it is sufficient to consider only the first couple of terms of the power series, as the summands quickly converge to zero.

B. Error Bound

We can prove an error bound of this approximation by considering the inequality for the n-th term

$$|s_n(\underline{\kappa}, \mathbf{A})| \leq \int_0^{2\pi} \int_0^{2\pi} |\kappa_1 + \kappa_2 + a_{1,1} + a_{1,2} + a_{2,1} + a_{2,2}|^n dx_1 dx_2 = |4\pi^2(\kappa_1 + \kappa_2 + a_{1,1} + a_{1,2} + a_{2,1} + a_{2,2})|^n \ .$$

Note that this bound is not tight. To simplify the notation, we define the abbreviation

$$t = 4\pi^2(\kappa_1 + \kappa_2 + a_{1,1} + a_{1,2} + a_{2,1} + a_{2,2}).$$

Now, we only consider summands 0 to $N-1$ for calculation of the normalization constant. Then, the absolute value of the truncated part of the series can be bounded according to

$$\left| \sum_{n=N}^{\infty} \frac{s_n(\underline{\kappa}, \mathbf{A})}{n!} \right| \leq \sum_{n=N}^{\infty} \left| \frac{t^n}{n!} \right| = \sum_{n=0}^{\infty} \frac{|t|^{N+n}}{(N+n)!} = |t|^N \sum_{n=0}^{\infty} \frac{|t|^n}{(N+n)!} \leq \frac{|t|^N}{N!} \sum_{n=0}^{\infty} \frac{|t|^n}{n!} = \frac{|t|^N}{N!} \exp(t) \to_{N \to \infty} 0 \ .$$

Thus, the error decreases for fixed t when the number of summands N is chosen large enough. As can be seen from this bound, the proposed approximation is mostly suitable for situations where the concentration is small, i.e., the uncertainty is high. In cases where the concentration is high, other approximations may need to be found, e.g., approximations based on the similarity of von Mises distributions of high concentration to the Gaussian distributions.

IV. Parameter Estimation

For many practical problems, it is essential to be able to estimate the parameters of a distribution based on a set of samples. In this section, we propose an algorithm based on the concept of maximum likelihood estimation (MLE) to address this problem. The related problem of parameter estimation for the bivariate wrapped normal distribution is discussed in [18]

A. Maximum Likelihood Approach

We assume that n independent samples $\underline{x}^{(1)}, \dots, \underline{x}^{(n)} \in [0, 2\pi)^2$ on the torus are given. Then, the likelihood of obtaining these samples is given by

$$L(\underline{\mu}, \underline{\kappa}, \mathbf{A}) = \prod_{j=1}^{n} f(\underline{x}^{(j)}; \underline{\mu}, \underline{\kappa}, \mathbf{A}) \ ,$$

and we seek to obtain the parameters $\underline{\mu}, \underline{\kappa}, \mathbf{A}$ that maximize $L(\underline{\mu}, \underline{\kappa}, \mathbf{A})$. For this purpose, we consider the log-likelihood

$$\begin{aligned}
&\log(L(\underline{\mu}, \underline{\kappa}, \mathbf{A}) = \sum_{j=1}^{n} \log(f(\underline{x}^{(j)}; \underline{\mu}, \underline{\kappa}, \mathbf{A})) \\
&= n \log(C(\kappa, \mathbf{A})) + \sum_{j=1}^{n} \kappa_1 \cos(x_1^{(j)} - \mu_1) \\
&+ \kappa_2 \cos(x_2^{(j)} - \mu_2) + \begin{bmatrix} \cos(x_1^{(j)} - \mu_1) \\ \sin(x_1^{(j)} - \mu_1) \end{bmatrix}^T \mathbf{A} \begin{bmatrix} \cos(x_2^{(j)} - \mu_2) \\ \sin(x_2^{(j)} - \mu_2) \end{bmatrix}
\end{aligned}$$

and obtain the parameters according to

$$\begin{aligned}
\arg \max_{\underline{\mu}, \underline{\kappa}, \mathbf{A}} \ & \log(L(\underline{\mu}, \underline{\kappa}, \mathbf{A}) \\
\text{s. t. } & \kappa_1 > 0, \kappa_2 > 0
\end{aligned}$$

Unfortunately, a closed-form solution of this problem does not seem to exist. Therefore, we use a numerical optimization algorithm to find the best solution. As this is a non-convex optimization problem, we cannot guarantee to find the global optimum, but starting with a well-chosen initial value seems to be sufficient to obtain a good local optimum in practice.

In order to obtain the initial value for the optimization, we propose the following method. First, we observe that if $\mathbf{A}$ is a zero matrix, the marginals

$$\int_0^{2\pi} f(\underline{x}; \underline{\mu}, \underline{\kappa}, \mathbf{A}) dx_1 \ , \ \int_0^{2\pi} f(\underline{x}; \underline{\mu}, \underline{\kappa}, \mathbf{A}) dx_2$$

are von Mises distributions with parameters μ_1, κ_1 and μ_2, κ_2, respectively. Therefore, we fit von Mises distributions to the marginals of the samples using moment matching (see [19]) and use their parameters to obtain an initial estimate for $\underline{\mu}$ and $\underline{\kappa}$. This is equivalent to obtaining $\underline{\mu}$ from the first trigonometric moment of the samples (see [11]). The matrix $\mathbf{A}$ is then initialized with a zero matrix.

B. Derivatives

Some optimization algorithms allow the user to specify the derivatives of the value function in order speed up the optimization process. For this reason, we derive analytical expression for the derivatives in this section. The derivatives of the log-likelihood can be calculated as follows. Deriving with respect to μ_1 yields

$$\begin{aligned}
\frac{\partial}{\partial \mu_1} \log(L(\underline{\mu}, \underline{\kappa}, \mathbf{A}) = \sum_{j=1}^{n} \Bigg(& \kappa_1 \sin(x_1^{(j)} - \mu_1) \\
&+ \begin{bmatrix} \sin(x_1^{(j)} - \mu_1) \\ -\cos(x_1^{(j)} - \mu_1) \end{bmatrix}^T \mathbf{A} \begin{bmatrix} \cos(x_2^{(j)} - \mu_2) \\ \sin(x_2^{(j)} - \mu_2) \end{bmatrix} \Bigg)
\end{aligned}$$

and the derivative with respect to μ_2 can be obtained analogously. The derivative with respect to κ_l ($l = 1, 2$) is given by

$$\begin{aligned}
&\frac{\partial}{\partial \kappa_l} \log(L(\underline{\mu}, \underline{\kappa}, \mathbf{A}) \\
&= -nC(\underline{\kappa}, \mathbf{A}) \left(\frac{\partial}{\partial \kappa_l} \left(C(\underline{\kappa}, \mathbf{A})^{-1} \right) \right) + \sum_{j=1}^{n} \cos(x_l^{(j)} - \mu_l) \ ,
\end{aligned}$$

where the derivative of the inverse normalization constant is given by

$$\frac{\partial}{\partial \kappa_l} (C(\underline{\kappa}, \mathbf{A})^{-1}) = \sum_{n=0}^{\infty} \frac{\frac{\partial}{\partial \kappa_l} s_n(\underline{\kappa}, \mathbf{A})}{n!} \ .$$

The derivative of s_n can be calculated analytically because s_n is a polynomial for all n (see Appendix). If we differentiate with respect to $a_{1,1}$, we obtain

$$\begin{aligned}
&\frac{\partial}{\partial a_{1,1}} \log(L(\underline{\mu}, \underline{\kappa}, \mathbf{A}) \\
&= -n \frac{\frac{\partial}{\partial a_{1,1}} \left(C(\kappa, \mathbf{A})^{-1} \right)}{C(\kappa, \mathbf{A})^{-1}} + \sum_{j=1}^{n} \cos(x_1^{(j)} - \mu_1) \cos(x_2^{(j)} - \mu_2) \ ,
\end{aligned}$$

where

$$\frac{\partial}{\partial a_{1,1}} (C(\kappa, \mathbf{A})^{-1}) = \sum_{n=0}^{\infty} \frac{\frac{\partial}{\partial a_{1,1}} s_n(\underline{\kappa}, \mathbf{A})}{n!} \ .$$

Once again, the derivative of s_n can be calculated analytically. The other entries of $\mathbf{A}$ can be obtained in a similar way.

V. Bayesian Fusion

Now that we have shown how to estimate the parameters of a single bivariate wrapped normal distribution, we address the questions of how to fuse several of these densities. In order to perform Bayesian fusion of multiple densities, we need to derive a way to calculate the product of two bivariate von Mises densities. The reason is that according to Bayes' theorem, we have

$$f(\underline{x}|\underline{z}) = \frac{f(\underline{z}|\underline{x}) \cdot f(\underline{x})}{f(\underline{z})} \propto f(\underline{z}|\underline{x}) \cdot f(\underline{x}) \ ,$$

$$\mathbf{M}(\underline{\mu}) = \begin{bmatrix} \cos(\mu_1)\cos(\mu_2) & -\sin(\mu_1)\cos(\mu_2) & -\cos(\mu_1)\sin(\mu_2) & \sin(\mu_1)\sin(\mu_2) \\ \sin(\mu_1)\cos(\mu_2) & \cos(\mu_1)\cos(\mu_2) & -\sin(\mu_1)\sin(\mu_2) & -\cos(\mu_1)\sin(\mu_2) \\ \cos(\mu_1)\sin(\mu_2) & -\sin(\mu_1)\sin(\mu_2) & \cos(\mu_1)\cos(\mu_2) & -\sin(\mu_1)\cos(\mu_2) \\ \sin(\mu_1)\sin(\mu_2) & \cos(\mu_1)\sin(\mu_2) & \sin(\mu_1)\cos(\mu_2) & \cos(\mu_1)\cos(\mu_2) \end{bmatrix}$$

Fig. 3: Matrix required for Bayesian fusion of bivariate von Mises densities.

i.e., if we assume $f(\underline{z}|\underline{x})$ and $f(\underline{x})$ to be distributed according to bivariate von Mises distributions, we can obtain $f(\underline{x}|\underline{z})$ as the (renormalized) product of the two.

Before we calculate the product, we first consider the probability density function (1) and rewrite the matrix multiplication, which yields

$$\begin{aligned} f(\underline{x};\underline{\mu},\underline{\kappa},\mathbf{A}) = C \cdot \exp(&\kappa_1 \cos(x_1-\mu_1) + \kappa_2 \cos(x_2-\mu_2) \\ &+ \cos(x_1-\mu_1) a_{11} \cos(x_2-\mu_2) \\ &+ \cos(x_1-\mu_1) a_{12} \sin(x_2-\mu_2) \\ &+ \sin(x_1-\mu_1) a_{21} \cos(x_2-\mu_2) \\ &+ \sin(x_1-\mu_1) a_{22} \sin(x_2-\mu_2)) \ . \end{aligned}$$

By using the addition theorems for sine and cosine, we can reformulate the exponent as

$$\begin{aligned} &\kappa_1 \cos(\mu_1)\cos(x_1) + \kappa_2 \cos(\mu_2)\cos(x_2) \\ &+ \kappa_1 \sin(\mu_1)\sin(x_1) + \kappa_2 \sin(\mu_2)\sin(x_2) \\ &+ \cos(x_1)\cos(x_2) \cdot \mathbf{M}(\underline{\mu})_{1,:} \cdot \underline{a} \\ &+ \sin(x_1)\cos(x_2) \cdot \mathbf{M}(\underline{\mu})_{2,:} \cdot \underline{a} \\ &+ \cos(x_1)\sin(x_2) \cdot \mathbf{M}(\underline{\mu})_{3,:} \cdot \underline{a} \\ &+ \sin(x_1)\sin(x_2) \cdot \mathbf{M}(\underline{\mu})_{4,:} \cdot \underline{a} \ , \end{aligned}$$

where $\underline{a} = [a_{1,1},\ a_{2,1},\ a_{1,2},\ a_{2,2}]^T$ is the vectorized form of $\mathbf{A}$ and $\mathbf{M}(\underline{\mu})_{j,:}$ refers to the j-th row of $\mathbf{M}(\underline{\mu})$, which is given in Fig. 3.

Now we consider two bivariate von Mises distributions $f(\underline{x};\underline{\mu}^X,\underline{\kappa}^X,\mathbf{A}^X)$ and $f(\underline{x};\underline{\mu}^Y,\underline{\kappa}^Y,\mathbf{A}^Y)$ with parameters $\underline{\mu}^X,\underline{\kappa}^X,\mathbf{A}^X$ and $\underline{\mu}^Y,\underline{\kappa}^Y,\mathbf{A}^Y$, respectively. We seek to obtain the parameters $\underline{\mu}^Z,\underline{\kappa}^Z,\mathbf{A}^Z$ of a third bivariate von Mises distribution $f(\underline{x};\underline{\mu}^Z,\underline{\kappa}^Z,\mathbf{A}^Z)$, such that

$$f(\underline{x};\underline{\mu}^X,\underline{\kappa}^X,\mathbf{A}^X) \cdot f(\underline{x};\underline{\mu}^Y,\underline{\kappa}^Y,\mathbf{A}^Y) \propto f(\underline{x};\underline{\mu}^Z,\underline{\kappa}^Z,\mathbf{A}^Z)$$

For this purpose, we consider the exponent of the product and use the method of equating the coefficients to obtain the system of eight equations

$$\kappa_1^X \cos(\mu_1^X) + \kappa_1^Y \cos(\mu_1^Y) = \kappa_1^Z \cos(\mu_1^Z) \ , \quad (3)$$
$$\kappa_1^X \sin(\mu_1^X) + \kappa_1^Y \sin(\mu_1^Y) = \kappa_1^Z \sin(\mu_1^Z) \ , \quad (4)$$
$$\kappa_2^X \cos(\mu_2^X) + \kappa_2^Y \cos(\mu_2^Y) = \kappa_2^Z \cos(\mu_2^Z) \ , \quad (5)$$
$$\kappa_2^X \sin(\mu_2^X) + \kappa_2^Y \sin(\mu_2^Y) = \kappa_2^Z \sin(\mu_2^Z) \ , \quad (6)$$
$$\mathbf{M}(\underline{\mu}^X)_{1,:}\underline{a}^X + \mathbf{M}(\underline{\mu}^Y)_{1,:}\underline{a}^Y = \mathbf{M}(\underline{\mu}^Z)_{1,:}\underline{a}^Z \ , \quad (7)$$
$$\mathbf{M}(\underline{\mu}^X)_{2,:}\underline{a}^X + \mathbf{M}(\underline{\mu}^Y)_{2,:}\underline{a}^Y = \mathbf{M}(\underline{\mu}^Z)_{2,:}\underline{a}^Z \ , \quad (8)$$
$$\mathbf{M}(\underline{\mu}^X)_{3,:}\underline{a}^X + \mathbf{M}(\underline{\mu}^Y)_{3,:}\underline{a}^Y = \mathbf{M}(\underline{\mu}^Z)_{3,:}\underline{a}^Z \ , \quad (9)$$
$$\mathbf{M}(\underline{\mu}^X)_{4,:}\underline{a}^X + \mathbf{M}(\underline{\mu}^Y)_{4,:}\underline{a}^Y = \mathbf{M}(\underline{\mu}^Z)_{4,:}\underline{a}^Z \ . \quad (10)$$

The equations (3)–(6) can be solved to obtain $\underline{\mu}^Z$ and $\underline{\kappa}^Z$, which yields

$$\begin{aligned} \kappa_1^Z &= \sqrt{(m_1^C)^2 + (m_1^S)^2} \ , \\ \kappa_2^Z &= \sqrt{(m_2^C)^2 + (m_2^S)^2} \ , \\ \mu_1^Z &= \operatorname{atan2}(m_1^S, m_1^C) \ , \\ \mu_2^Z &= \operatorname{atan2}(m_2^S, m_2^C) \ , \end{aligned}$$

where

$$\begin{aligned} m_1^C &= \kappa_1^X \cos(\mu_1^X) + \kappa_1^Y \cos(\mu_1^Y) \ , \\ m_1^S &= \kappa_1^X \sin(\mu_1^X) + \kappa_1^Y \sin(\mu_1^Y) \ , \\ m_2^C &= \kappa_2^X \cos(\mu_2^X) + \kappa_2^Y \cos(\mu_2^Y) \ , \\ m_2^S &= \kappa_2^X \sin(\mu_2^X) + \kappa_2^Y \sin(\mu_2^Y) \ . \end{aligned}$$

Effectively, this corresponds to the multiplication formula for univariate von Mises distributions in each dimension.

The last four equations can be simplified to

$$\mathbf{M}(\underline{\mu}^Z) \cdot \underline{a}^Z = \mathbf{M}(\underline{\mu}^X) \cdot \underline{a}^X + \mathbf{M}(\underline{\mu}^Y) \cdot \underline{a}^Y \ ,$$

which can be solved by inverting $\mathbf{M}(\underline{\mu}^Z)$

$$\underline{a}^Z = \mathbf{M}(\underline{\mu}^Z)^{-1}(\mathbf{M}(\underline{\mu}^X) \cdot \underline{a}^X + \mathbf{M}(\underline{\mu}^Y) \cdot \underline{a}^Y) \quad (11)$$

to obtain $\underline{a}^Z$, and thus $\mathbf{A}^Z$. It can be shown that $\det \mathbf{M}(\underline{\mu}) = 1$, irrespective of $\underline{\mu}$. Hence, $\mathbf{M}(\underline{\mu})$ is always invertible and (11) can always be solved.

It is worth mentioning that our result for the product of two bivariate von Mises densities also shows that the restricted versions discussed by Singh et al. [6] and Mardia et al. [7] are, in general, not closed under multiplication because an entry of $\mathbf{A}$ can be non-zero after multiplication, even if it was zero before, depending on $\mathbf{M}(\underline{\mu})$.

VI. Examples

In this section, we give some examples of how the proposed methods can be applied.

A. Parameter Estimation Example

The following example illustrates the parameter estimation method discussed in Sec. IV. For this purpose, we use the wind data set from the R package *circular* [20]. This data set was also discussed in [21]. It contains measurements of the wind direction measured by a meteorological station at Col De La Roa, Italy, from January 29th, 2001 to March 31st, 2001. For each day, we consider the measurements obtained at 3:00 am and at 3:15 am. In this way, we obtain 62 two-dimensional samples defined on the torus.

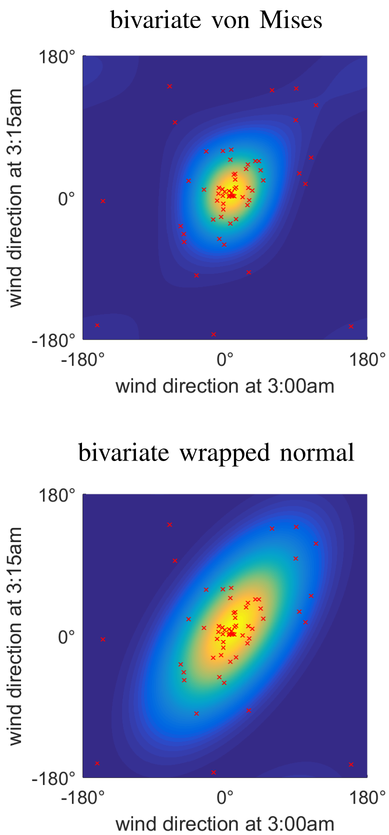


Fig. 4: Parameter estimation example. The samples are shown in red and the value of the probability density function is given by the color of the background. Note that both the x and y axes are 360°-periodic.

Then, the maximum likelihood estimation procedure proposed in Sec. IV was used to obtain the parameters of a bivariate von Mises distribution. For comparison, we also obtained the parameters of a bivariate wrapped normal distribution using numerical MLE [18], because this distribution is also a common model for toroidal data. The resulting densities are depicted in Fig. 4.

It can be seen that both distributions are able to represent the obvious correlation between the two measurements. However, the bivariate von Mises seems to reflect the true underlying density more closely because it is able to reflect the asymmetric distribution of the samples as a result of the additional degrees of freedom. This is illustrated by the fact that the log-likelihood of the BVM distribution is higher (-147.3259) than the log-likelihood of the BWN distribution (-164.5940). In future work, a rigorous goodness of fit test could be used to better quantify this result.

B. Fusion Example

In this example, we consider two bivariate von Mises distributions $f(\underline{x};\underline{\mu}^X,\underline{\kappa}^X,\mathbf{A}^X)$ and $f(\underline{x};\underline{\mu}^Y,\underline{\kappa}^Y,\mathbf{A}^Y)$ with parameters

$$\underline{\mu}^X = [2,4]^T,\ \underline{\kappa}^X = [0.7, 0.2]^T,\ \mathbf{A}^X = \begin{bmatrix} 0 & 0 \\ 0 & -0.3 \end{bmatrix}$$

and

$$\underline{\mu}^Y = [1.5,5]^T,\ \underline{\kappa}^Y = [0.1, 2]^T,\ \mathbf{A}^Y = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} ,$$

respectively. By applying the Bayesian fusion formulas given in Sec. V, we obtain the parameters of the product $f(\underline{x};\underline{\mu}^Z,\underline{\kappa}^Z,\mathbf{A}^Z)$, which are given by

$$\begin{aligned} \underline{\mu}^Z &= [1.9392, 4.9203]^T , \\ \underline{\kappa}^Z &= [0.7892, 2.1148]^T , \\ \mathbf{A}^Z &= \begin{bmatrix} 0.0145 & 0.0110 \\ -0.2383 & -0.1813 \end{bmatrix} . \end{aligned}$$

This example is visualized in Fig. 5. As shown in Sec. V, this result does not involve any approximation, i.e.,

$$f(\underline{x};\underline{\mu}^Z,\underline{\kappa}^Z,\mathbf{A}^Z) \propto f(\underline{x};\underline{\mu}^X,\underline{\kappa}^X,\mathbf{A}^X) \cdot f(\underline{x};\underline{\mu}^Y,\underline{\kappa}^Y,\mathbf{A}^Y) .$$

Observe that all entries of $\mathbf{A}^Z$ are non-zero, even though $\mathbf{A}^X$ had just one non-zero entry and $\mathbf{A}^Y$ was a zero matrix. This illustrates the fact that the sine model considered by Singh et al. [6] is not closed under multiplication.

VII. Conclusion

In this paper, we have presented a number of results on the bivariate von Mises distribution. In particular, we have proposed a series representation of the approximation constant that allows the efficient computation of an approximation. Furthermore, we have discussed parameter estimation based on a maximum likelihood approach. Finally, we have shown that the bivariate von Mises distribution is closed under multiplication and given an algorithm to calculate the parameters of the renormalized product in closed form. This allows efficient Bayesian fusion of bivariate von Mises densities, which facilitates their use in a variety of fusion problems on the torus.

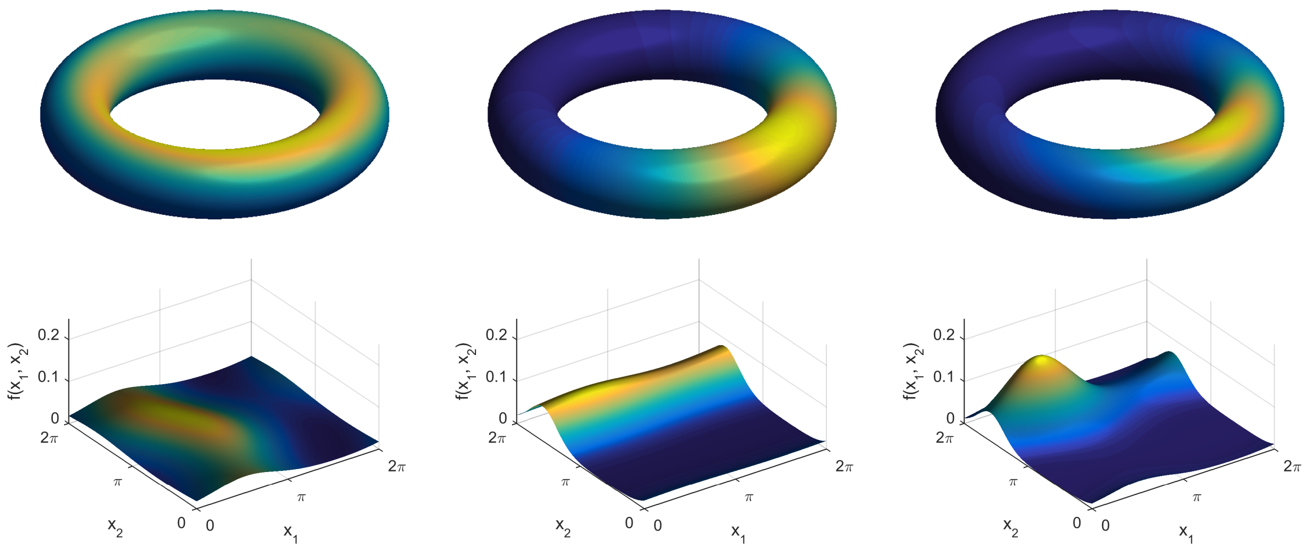
Future work may include an extension of the discussed methods to the n-torus (see also [9]), and the derivation of a convolution (or more general propagation) algorithm to create a recursive filter as was developed for the bivariate wrapped normal distribution [11]. Moreover, further investigation of the normalization constant may be of interest, as it may be possible to represent it as a series of Bessel functions, similar to the results for the special cases given in [5], [6], and [7].

An implementation of the algorithms proposed in this paper is available as part of the MATLAB library `libDirectional` [22], a library dedicated to directional statistics and estimation involving directional quantities.

Appendix

The first terms of $s_n(\kappa,\mathbf{A})$ in (2) are given by

$$\begin{aligned} s_0(\underline{\kappa},\mathbf{A}) &= 4\pi^2 , \\ s_1(\underline{\kappa},\mathbf{A}) &= 0 , \\ s_2(\underline{\kappa},\mathbf{A}) &= \pi^2(a_{1,1}^2 + a_{1,2}^2 + a_{2,1}^2 + a_{2,2}^2 + 2\kappa_1^2 + 2\kappa_2^2) , \\ s_3(\underline{\kappa},\mathbf{A}) &= 6\kappa_1\kappa_2 a_{1,1}\pi^2 , \end{aligned}$$



$f(\underline{x};\underline{\mu}^X,\underline{\kappa}^X,\mathbf{A}^X)$ $f(\underline{x};\underline{\mu}^Y,\underline{\kappa}^Y,\mathbf{A}^Y)$ $f(\underline{x};\underline{\mu}^Z,\underline{\kappa}^Z,\mathbf{A}^Z)$

Fig. 5: Fusion example. The density on the left is multiplied with the density in the middle to obtain the density on the right. On the top, the densities are visualized on the torus, whereas on the bottom, the same densities are visualized as a function on the plane.

$$\begin{aligned}
s_4(\underline{\kappa},\mathbf{A}) = \frac{3}{16}\pi^2(&3a_{1,1}^4 + 6a_{1,1}^2a_{1,2}^2 + 6a_{1,1}^2a_{2,1}^2 + 2a_{1,1}^2a_{2,2}^2 \\
&+24a_{1,1}^2\kappa_1^2 + 24a_{1,1}^2\kappa_2^2 + 8a_{1,1}a_{1,2}a_{2,1}a_{2,2} + 3a_{1,2}^4 \\
&+2a_{1,2}^2a_{2,1}^2 + 6a_{1,2}^2a_{2,2}^2 + 24a_{1,2}^2\kappa_1^2 + 8a_{1,2}^2\kappa_2^2 + 3a_{2,1}^4 \\
&+6a_{2,1}^2a_{2,2}^2 + 8a_{2,1}^2\kappa_1^2 + 24a_{2,1}^2\kappa_2^2 + 3a_{2,2}^4 + 8a_{2,2}^2\kappa_1^2 \\
&+8a_{2,2}^2\kappa_2^2 + 8\kappa_1^4 + 32\kappa_1^2\kappa_2^2 + 8\kappa_2^4)\ , \\
s_5(\underline{\kappa},\mathbf{A}) = \frac{15}{4}\pi^2\kappa_1\kappa_2(&3a_{1,1}^3 + 2a_{1,2}a_{2,1}a_{2,2} \\
&+a_{1,1}(3a_{1,2}^2 + 3a_{2,1}^2 + a_{2,2}^2 + 4\kappa_1^2 + 4\kappa_2^2))\ .
\end{aligned}$$

References

[1] K. V. Mardia and P. E. Jupp, *Directional Statistics*, 1st ed. Wiley, 1999.

[2] K. V. Mardia, “Characterizations of Directional Distributions,” in *A Modern Course on Statistical Distributions in Scientific Work*. Springer Netherlands, 1975, vol. 17, pp. 365–385.

[3] ——, “Statistics of Directional Data,” *Journal of the Royal Statistical Society. Series B (Methodological)*, vol. 37, no. 3, pp. 349–393, 1975.

[4] R. von Mises, “Über die ”Ganzzahligkeit” der Atomgewichte und verwandte Fragen,” *Physikalische Zeitschrift*, vol. XIX, pp. 490–500, 1918.

[5] L.-P. Rivest, “A Distribution for Dependent Unit Vectors,” *Communications in Statistics - Theory and Methods*, vol. 17, no. 2, pp. 461–483, 1988.

[6] H. Singh, V. Hnizdo, and E. Demchuk, “Probabilistic Model for Two Dependent Circular Variables,” *Biometrika*, vol. 89, no. 3, pp. 719–723, 2002.

[7] K. V. Mardia, C. C. Taylor, and G. K. Subramaniam, “Protein Bioinformatics and Mixtures of Bivariate von Mises Distributions for Angular Data,” *Biometrics*, vol. 63, no. 2, pp. 505–512, 2007.

[8] K. V. Mardia and J. Frellsen, “Statistics of Bivariate von Mises Distributions,” in *Bayesian Methods in Structural Bioinformatics*, ser. Statistics for Biology and Health, T. Hamelryck, K. Mardia, and J. Ferkinghoff-Borg, Eds. Springer Berlin Heidelberg, 2012, pp. 159–178.

[9] K. V. Mardia, G. Hughes, C. C. Taylor, and H. Singh, “A Multivariate von Mises Distribution with Applications to Bioinformatics,” *Canadian Journal of Statistics*, vol. 36, no. 1, pp. 99–109, 2008.

[10] R. A. Johnson and T. Wehrly, “Measures and Models for Angular Correlation and Angular-Linear Correlation,” *Journal of the Royal Statistical Society. Series B (Methodological)*, vol. 39, no. 2, pp. 222–229, 1977.

[11] G. Kurz, I. Gilitschenski, M. Dolgov, and U. D. Hanebeck, “Bivariate Angular Estimation Under Consideration of Dependencies Using Directional Statistics,” in *Proceedings of the 53rd IEEE Conference on Decision and Control (CDC 2014)*, Los Angeles, California, USA, Dec. 2014.

[12] G. S. Shieh and R. A. Johnson, “Inferences Based on a Bivariate Distribution with von Mises Marginals,” *Annals of the Institute of Statistical Mathematics*, vol. 57, no. 4, pp. 789–802, 2005.

[13] J. J. Fernández-Durán, “Models for Circular–Linear and Circular–Circular Data Constructed from Circular Distributions Based on Nonnegative Trigonometric Sums,” *Biometrics*, vol. 63, no. 2, pp. 579–585, 2007.

[14] S. Kato and A. Pewsey, “A Möbius Transformation-induced Distribution on the Torus,” *Biometrika*, 2015.

[15] M. Abramowitz and I. A. Stegun, *Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables*, 10th ed. New York: Dover, 1972.

[16] P. E. Jupp and K. V. Mardia, “A Unified View of the Theory of Directional Statistics, 1975-1988,” *International Statistical Review / Revue Internationale de Statistique*, vol. 57, no. 3, pp. 261–294, 1989.

[17] ——, “A General Correlation Coefficient for Directional Data and Related Regression Problems,” *Biometrika*, vol. 67, no. 1, pp. 163–173, 1980.

[18] G. Kurz and U. D. Hanebeck, “Parameter Estimation for the Bivariate Wrapped Normal Distribution (submitted),” in *Proceedings of the 54th IEEE Conference on Decision and Control (CDC 2015)*, Osaka, Japan, Dec. 2015.

[19] G. Kurz, I. Gilitschenski, and U. D. Hanebeck, “Recursive Nonlinear Filtering for Angular Data Based on Circular Distributions,” in *Proceedings of the 2013 American Control Conference (ACC 2013)*, Washington D. C., USA, Jun. 2013.

[20] U. Lund and C. Agostinelli, “circular: Circular Statistics,” Nov. 2013. [Online]. Available: http://cran.r-project.org/web/packages/circular/index.html

[21] C. Agostinelli, “Robust Estimation for Circular Data,” *Computational Statistics & Data Analysis*, vol. 51, no. 12, pp. 5867–5875, 2007.

[22] G. Kurz, I. Gilitschenski, F. Pfaff, and L. Drude, “libDirectional,” 2015. [Online]. Available: https://github.com/libDirectional