# Bivariate Angular Estimation Under Consideration of Dependencies Using Directional Statistics

Gerhard Kurz[1], Igor Gilitschenski[1], Maxim Dolgov[1] and Uwe D. Hanebeck[1]

*Abstract*— Estimation of angular quantities is a widespread issue, but standard approaches neglect the true topology of the problem and approximate directional with linear uncertainties. In recent years, novel approaches based on directional statistics have been proposed. However, these approaches have been unable to consider arbitrary circular correlations between multiple angles so far. For this reason, we propose a novel recursive filtering scheme that is capable of estimating multiple angles even if they are dependent, while correctly describing their circular correlation. The proposed approach is based on toroidal probability distributions and a circular correlation coefficient. We demonstrate the superiority to a standard approach based on the Kalman filter in simulations.

*Index Terms*— recursive filtering, wrapped normal, circular correlation coefficient, moment matching.

Fig. 1: A bivariate wrapped normal probability distribution on the torus shown as a heat map.

## I. INTRODUCTION

There are many applications that require estimation of angular quantities. These applications include, but are not limited to, robotics, augmented reality, and aviation, as well as biology, geology, and medicine. In many cases, not just one, but several angles have to be estimated. Furthermore, correlations may exist between those angles and have to be taken into account in the estimation algorithm. For example, there may be dependencies between the orientations of head and torso of a person. Another example is a robot arm with several rotary joints that are affected by correlated noise.

Traditional approaches for estimating correlated angles are typically based on Gaussian distributions and use classical filtering algorithms such as the Kalman filter [1] or nonlinear extensions thereof, e.g., the unscented Kalman filter [2]. However, Gaussian distributions are defined on $n$-dimensional vector spaces rather than the proper manifold, in this case a torus or hypertorus.

Some filtering algorithms that are particularly well-suited for angular estimation have been proposed. They rely on periodic probability distributions that stem from the field of directional statistics [3], [4]. For example, Azmani et al. proposed a filtering algorithm based on the von Mises distribution [5], [6]. In our prior work, we proposed an algorithm based on the wrapped normal distribution [7]. However, these approaches are one-dimensional and unable to take correlations between several angles into account.

To address this deficiency, we propose a new filtering algorithm for estimation of correlated angles in this paper. To our knowledge, this is the first work on recursive estimation based on toroidal probability distributions.

It should be noted that there are some directional filters that, in a sense, take correlation between angles into account. We have performed research on a filter based on the hyperspherical Bingham distribution [8], [9], which captures correlations when estimating rotations represented as quaternions. A very similar approach has independently been published by Glover et al. [10]. Furthermore, Feiten et al. have used mixtures of projected Gaussians to deal with 6D pose estimation while considering correlations between different angles as well as correlation between position and orientation [11]. However, all of these approaches are intended for describing 3D rotations, which have a different underlying topology, namely the group $SO(3)$ [12], rather than the torus. For this reason, they cannot be used to estimate arbitrary correlated angles.

In the field of directional statistics, some previous work on toroidal distributions can be found. In particular, multivariate generalizations of the von Mises distribution have been studied by several authors [13], [14]. The multivariate wrapped normal distribution has also been considered [4, Sec. 2.3.2], [15], which will be the foundation of the algorithm that we propose in this paper. Furthermore, various directional versions of the correlation coefficient have been suggested. We use the circular correlation coefficient as defined in [4, Sec. 8.2] and [15].

## II. TOROIDAL STATISTICS

In this section, we give an introduction to toroidal statistics. First of all, we define the necessary topological spaces. The *unit circle*

$$S^1 = \{x \in \mathbb{C} : ||x|| = 1\}$$

[1] 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`, `gilitschenski@kit.edu`, `maxim.dolgov@kit.edu`, `uwe.hanebeck@ieee.org`).

$$=\{\cos(\phi) + i\sin(\phi) : 0 \le \phi < 2\pi\}$$

is identified with the interval $S^1 \equiv [0, 2\pi)$, while keeping the topology in mind. The *torus* $T^2 = S^1 \times S^1$ is obtained as the Cartesian product of two circles. More generally, the $n$-torus

$$T^n = \underbrace{S^1 \times \cdots \times S^1}_{n \text{ times}} = (S^1)^n$$

is obtained by the $n$-fold Cartesian product of circles. We only consider $T^2$ in the remainder of this paper. Most of the presented techniques can be generalized to the $n$-torus.

### A. Toroidal Distributions

Before we look at the bivariate toroidal wrapped normal distribution, we introduce the circular univariate wrapped normal distribution to show how the toroidal distribution is a generalization of the circular case.

**Definition 1** (Wrapped Normal Distribution).
A univariate *wrapped normal* (WN) distribution is given by its probability density function (pdf)

$$f(x; \mu, \sigma) = \sum_{j=-\infty}^{\infty} \mathcal{N}(x; \mu + 2\pi j, \sigma)$$

with $x \in S^1$, location parameter $\mu \in S^1$, dispersion parameter $\sigma > 0$, and normal density $\mathcal{N}(x; \mu, \sigma)$.

We use the notation $X \sim \mathcal{WN}(\mu, \sigma)$ to indicate that a random variable $X$ is distributed according to a WN distribution with parameters $\mu$ and $\sigma$. A WN distribution is obtained by wrapping a normal distribution around the unit circle. The normalization constant is already included in the Gaussian distributions, so it does not need to be calculated separately. This is a significant advantage compared to other periodic probability distributions whose normalization constants can be difficult to calculate. The WN distribution can be generalized to the bivariate case as follows.

**Definition 2** (Toroidal Wrapped Normal Distribution).
The *toroidal* (or bivariate) *wrapped normal* (TWN) distribution is given by the pdf

$$f(\underline{x}; \underline{\mu}, \boldsymbol{\Sigma}) = \sum_{j=-\infty}^{\infty} \sum_{k=-\infty}^{\infty} \mathcal{N}\left(\underline{x}; \underline{\mu} + \begin{bmatrix} 2\pi j \\ 2\pi k \end{bmatrix}, \boldsymbol{\Sigma}\right)$$

with $\underline{x} = [x_1, x_2]^T \in S^1 \times S^1$, location parameter $\underline{\mu} = [\mu_1, \mu_2]^T \in S^1 \times S^1$, and symmetric parameter matrix

$$\boldsymbol{\Sigma} = \begin{bmatrix} \sigma_1^2 & \rho\sigma_1\sigma_2 \\ \cdot & \sigma_2^2 \end{bmatrix} \in \mathbb{R}^{2\times 2}$$

with correlation[1] parameter $-1 < \rho < 1$, linear standard deviations $\sigma_1, \sigma_2 > 0$, and multivariate normal distribution $\mathcal{N}(\underline{x}; \underline{\mu}, \boldsymbol{\Sigma})$.

The notation $X \sim \mathcal{TWN}(\underline{\mu}, \boldsymbol{\Sigma})$ is used to indicate that a random variable $X$ is distributed according to a TWN

---

[1] While the linear correlation coefficient $\rho$ can reach values $-1$ and $1$, we do not consider these cases because they lead to a positive semidefinite rather than a positive definite covariance matrix.

distribution with parameters $\underline{\mu}$ and $\boldsymbol{\Sigma}$. Note that the parameter matrix stems from the covariance of a bivariate normal distribution, yet its meaning is different in the toroidal context. An example of the TWN distribution is depicted in Fig. 2a. It can be seen how $x_1$ and $x_2$ are $2\pi$-periodic and the distribution wraps at these locations.

### B. Toroidal Moments

In analogy to the traditional linear moments, we introduce the circular moments and subsequently generalize them to toroidal moments.

**Definition 3** (Circular Moments).
In the univariate case, the $n$-th *circular moment* (sometimes also referred to as trigonometric or angular moment) of a random variable $x$ is given by

$$m_n = \mathbb{E}(e^{inx}) = \int_0^{2\pi} f(x)e^{inx}dx \in \mathbb{C} ,$$

where $i$ is the imaginary unit.

Note that the $n$-th circular moment is a complex number, i.e., it has two degrees of freedom. The argument of $m_1$ determines the location of the circular mean, whereas the absolute value of $m_1$ determines the concentration. For this reason, a WN distribution is uniquely determined by its first circular moment. We generalize circular moments to the bivariate case in the following Definition.

**Definition 4** (Toroidal Moments).
For a random variable $x$ distributed according to a toroidal distribution, the $n$-th *toroidal moment* is given by

$$\underline{m}_n = \mathbb{E}\left(\begin{bmatrix} e^{inx_1} \\ e^{inx_2} \end{bmatrix}\right) = \int_0^{2\pi}\int_0^{2\pi} f(x) \begin{bmatrix} e^{inx_1} \\ e^{inx_2} \end{bmatrix} dx_1 dx_2 \in \mathbb{C}^2 .$$

The $n$-th bivariate circular moment is a vector of two complex numbers, and thus, has four degrees of freedom.

**Lemma 1** (Moments of a TWN distribution).
The $n$-th moment of $\mathcal{TWN}(\underline{\mu}, \boldsymbol{\Sigma})$ is given by

$$\underline{m}_n = \begin{bmatrix} m_{n,1} \\ m_{n,2} \end{bmatrix} = \begin{bmatrix} \exp(in\mu_1 - n^2\sigma_1^2/2) \\ \exp(in\mu_2 - n^2\sigma_2^2/2) \end{bmatrix} ,$$

i.e., the componentwise circular moment of a $\mathcal{WN}(\mu_1, \sigma_1)$ and a $\mathcal{WN}(\mu_2, \sigma_2)$.

The proof is given in the appendix. Note that the $n$-th bivariate circular moment does not depend on the linear correlation coefficient $\rho$.

### C. Circular Correlation Coefficient

Several circular correlation coefficients have been defined (for example by Mardia [16], Johnson [17], Jupp [18], and Fisher [19]). We use the definition of Jammalamadaka et al. [15], [4], because it is intuitive, easy to work with and has a variety of nice properties (see [15, Theorem 2.1]).
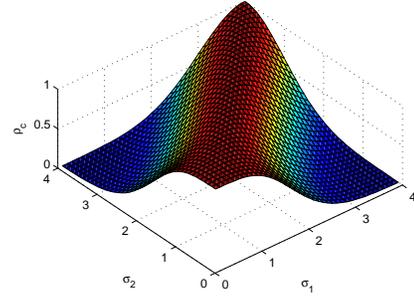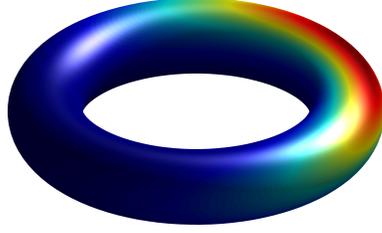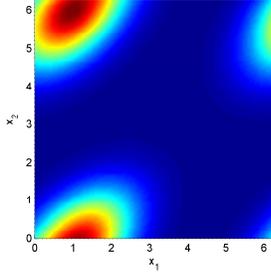
(a) A bivariate wrapped normal probability distribution with parameters $\underline{\mu} = [1, 6]^T$ and $\boldsymbol{\Sigma} = [1, 0.5; 0.5, 1]$ Keep in mind that $x_1$ and $x_2$ are $2\pi$-periodic.

(b) Circular correlation coefficient for a TWN distribution with given $\sigma_1, \sigma_2$ and $\rho \to 1$. As can be seen, the circular correlation coefficient $\rho_c$ is significantly lower than $\rho$ for large $\sigma_1, \sigma_2$.

Fig. 2: Toroidal wrapped normal distribution and circular correlation.

**Definition 5** (Circular Correlation Coefficient).
The circular correlation coefficient of random variables $\alpha, \beta$ with circular means $\mu, \nu$ is defined as

$$\rho_c(\alpha, \beta) = \frac{\mathbb{E}(\sin(\alpha - \mu)\sin(\beta - \nu))}{\sqrt{\operatorname{Var}(\sin(\alpha - \mu))\operatorname{Var}(\sin(\beta - \nu))}}$$

$$= \frac{\mathbb{E}(\sin(\alpha - \mu)\sin(\beta - \nu))}{\sqrt{\mathbb{E}(\sin^2(\alpha - \mu))\mathbb{E}(\sin^2(\beta - \nu))}} \in [-1, 1] .$$

Note that this definition of the circular correlation coefficient closely resembles the linear correlation coefficient

$$\rho(x, y) = \frac{\mathbb{E}((x - \mu)(y - \nu))}{\sqrt{\mathbb{E}((x - \mu)^2) \cdot \mathbb{E}((y - \nu)^2)}}$$

of real-valued random variables $x, y$ with means $\mu, \nu$. The circular correlation coefficient is obtained by applying the $\sin(\cdot)$ function to the considered quantities. Similar to the linear correlation coefficient, it holds that $-1 \le \rho_c \le 1$. For independent $\alpha$ and $\beta$, we have $\rho_c = 0$, but the converse is not necessarily true. On the other hand, we have $\rho_c = \pm 1$ if and only if $\alpha = \pm \beta + const \mod 2\pi$ (assuming $\alpha$ and $\beta$ have full support).

**Lemma 2** (Correlation Coefficient of TWN Distribution).
The circular correlation according to Definition 5 for $[x_1, x_2]^T \sim \mathcal{TWN}(\underline{\mu}, \boldsymbol{\Sigma})$ is given by

$$\rho_c(x_1, x_2) = \frac{\sinh(\sigma_1\sigma_2\rho)}{\sqrt{\sinh(\sigma_1^2)\sinh(\sigma_2^2)}} .$$

The derivation can be found in [15]. The TWN has five degrees of freedom, $(\mu_1, \mu_2, \sigma_1, \sigma_2, \rho)$, and it is uniquely determined by its first toroidal moment together with its circular correlation coefficient. The largest possible circular correlation coefficient for a TWN distribution is depicted in Fig. 2b.

*D. Moment Matching*

Moment matching is a well-known technique for a variety of problems and we can use a similar approach in the directional case. For univariate problems, it is natural to match the first circular moment, which captures both location and dispersion of the considered distribution (e.g., a WN distribution) [7]. In the bivariate case, however, we need to

match both the toroidal moments and the circular correlation to capture all five degrees of freedom of a TWN distribution.

An algorithm for estimating TWN parameters is given in [15, eq. (3.4), (3.5)], which is based on calculating the circular moments $m_{1,1} = \mathbb{E}(e^{ix_1})$ and $m_{1,2} = \mathbb{E}(e^{ix_2})$ as well as the product expectation value $\mathbb{E}(e^{i(x_1 - \mu_1)}e^{i(x_2 - \mu_2)})$ of some distribution and fitting a TWN distribution with identical moments. This is not always possible, because the resulting $\boldsymbol{\Sigma}$-matrix is not necessarily positive definite and thus, not a valid covariance matrix. Furthermore, this method does not maintain the circular correlation coefficient $\rho_c$.

We propose a better alternative obtained by matching the circular correlation coefficient.

**Lemma 3** (Estimation of TWN Parameters by Moment Matching).
For a given first moment $\underline{m}_1 \in \mathbb{C}^2$ and a given circular correlation coefficient $\rho_c \in (-1, 1)$, there exists a TWN distribution with identical first moment $\underline{m}_1$ and circular correlation coefficient $\rho_c$ if and only if $\rho \in (-1, 1)$, where

$$\underline{\mu} = \begin{bmatrix} \operatorname{atan2}(\operatorname{Im} m_{1,1}, \operatorname{Re} m_{1,1}) \\ \operatorname{atan2}(\operatorname{Im} m_{1,2}, \operatorname{Re} m_{1,2}) \end{bmatrix}$$

$$\sigma_1 = \sqrt{-2\log ||m_{1,1}||} , \quad \sigma_2 = \sqrt{-2\log ||m_{1,2}||}$$

$$\rho = \frac{1}{\sigma_1\sigma_2}\sinh^{-1}\left(\sqrt{\sinh(\sigma_1^2)\sinh(\sigma_2^2)} \cdot \rho_c\right) .$$

We give the proof in the appendix. The inverse of the hyperbolic sine can be calculated according to $\sinh^{-1}(x) = \log(x + \sqrt{x^2 + 1})$ [20, eq. (4.6.20)].

## III. OPERATIONS ON TWN DENSITIES

In order to create a toroidal filtering algorithm, we need to perform certain operations on TWN distributions, which we derive in this section.

*A. Convolution of TWN Densities*

The convolution of pdfs is required for prediction, as we will show in Sec. IV-A. Convolution of pdfs corresponds to the addition of random variables, in this case, addition modulo $2\pi$.

**Lemma 4** (Convolution of two TWN distributions).
For $A \sim \mathcal{TWN}(\underline{\mu}^a, \boldsymbol{\Sigma}^a)$ and $B \sim \mathcal{TWN}(\underline{\mu}^b, \boldsymbol{\Sigma}^b)$, we have

$A + B \sim \mathcal{TWN}(\underline{\mu}, \mathbf{\Sigma}) = \mathcal{TWN}(\underline{\mu}^a, \mathbf{\Sigma}^a) * \mathcal{TWN}(\underline{\mu}^b, \mathbf{\Sigma}^b)$, where $\underline{\mu} = \underline{\mu}^a + \underline{\mu}^b \mod 2\pi$, $\mathbf{\Sigma} = \mathbf{\Sigma}^a + \mathbf{\Sigma}^b$.

*Proof.* This result follows immediately from the convolution formula for Gaussian distributions. □

### B. Multiplication of TWN Densities

To perform a Bayesian update, we need to be able to calculate the product of two TWN probability density functions. Unfortunately, the resulting function is, in general, not the unnormalized pdf of a TWN, because TWN distributions are not closed under multiplication. This is not surprising since not even WN distributions are closed under multiplication [7].

*1) Problem Formulation:* We seek to derive an approximation, i.e., we try to fit a TWN distribution to the true product density. We propose to use moment matching in order to obtain the TWN distribution of the product.

We consider two TWN densities $\mathcal{TWN}(\underline{\mu}^a, \mathbf{\Sigma}^a)$ and $\mathcal{TWN}(\underline{\mu}^b, \mathbf{\Sigma}^b)$, and we try to obtain a TWN density $\mathcal{TWN}(\underline{\mu}, \mathbf{\Sigma})$, such that the moments and circular correlation of the renormalized true product density

$$\frac{\mathcal{TWN}(\underline{\mu}^a, \mathbf{\Sigma}^a) \cdot \mathcal{TWN}(\underline{\mu}^b, \mathbf{\Sigma}^b)}{\int_0^{2\pi} \int_0^{2\pi} \mathcal{TWN}(\underline{\mu}^a, \mathbf{\Sigma}^a) \cdot \mathcal{TWN}(\underline{\mu}^b, \mathbf{\Sigma}^b) dx_1 dx_2}$$

and the density of $\mathcal{TWN}(\underline{\mu}, \mathbf{\Sigma})$ match. The matrix $\mathbf{\Sigma}$ is composed of $\sigma_1, \sigma_2$, and $\rho$, as given in Definition 2.

*2) Obtaining $\underline{\mu}$ and $\sigma_1, \sigma_2$:* To determine $\underline{\mu}, \sigma_1, \sigma_2$, we need to calculate

$$c = \int_0^{2\pi} \int_0^{2\pi} f(\underline{x}; \underline{\mu}^a, \mathbf{\Sigma}^a) f(\underline{x}; \underline{\mu}^b, \mathbf{\Sigma}^b) dx_1 dx_2$$

$$m_{1,1} = \frac{1}{c} \mathbb{E}(\exp(ix_1))$$
$$= \frac{1}{c} \int_0^{2\pi} \int_0^{2\pi} e^{ix_1} f(\underline{x}; \underline{\mu}^a, \mathbf{\Sigma}^a) f(\underline{x}; \underline{\mu}^b, \mathbf{\Sigma}^b) dx_1 dx_2$$

$$m_{1,2} = \frac{1}{c} \mathbb{E}(\exp(ix_2))$$
$$= \frac{1}{c} \int_0^{2\pi} \int_0^{2\pi} e^{ix_2} f(\underline{x}; \underline{\mu}^a, \mathbf{\Sigma}^a) f(\underline{x}; \underline{\mu}^b, \mathbf{\Sigma}^b) dx_1 dx_2 .$$

These integrals are difficult to evaluate because even the integral over a two-dimensional Gaussian distribution can only be evaluated numerically [21]. For this reason we use numerical integration [22]. From $m_{1,1}$ and $m_{1,2}$, we obtain $\underline{\mu}, \sigma_1, \sigma_2$ as given in Lemma 3.

*3) Obtaining $\rho$:* To calculate the circular correlation coefficient, we need

$$\rho_c = \frac{s_{12}}{\sqrt{s_{11} s_{22}}}$$

where

$$s_{12} := \mathbb{E}(\sin(x_1 - \mu_1) \sin(x_2 - \mu_2))$$
$$= \int_0^{2\pi} \int_0^{2\pi} \sin(x_1 - \mu_1) \sin(x_2 - \mu_2)$$
$$\cdot f(\underline{x}; \underline{\mu}^a, \mathbf{\Sigma}^a) f(\underline{x}; \underline{\mu}^b, \mathbf{\Sigma}^b) dx_1 dx_2 ,$$
$$s_{11} := \mathbb{E}(\sin^2(x_1 - \mu_1))$$

$$= \int_0^{2\pi} \int_0^{2\pi} \sin^2(x_1 - \mu_1)$$
$$\cdot f(\underline{x}; \underline{\mu}^a, \mathbf{\Sigma}^a) f(\underline{x}; \underline{\mu}^b, \mathbf{\Sigma}^b) dx_1 dx_2 ,$$
$$s_{22} := \mathbb{E}(\sin^2(x_2 - \mu_2))$$
$$= \int_0^{2\pi} \int_0^{2\pi} \sin^2(x_2 - \mu_2)$$
$$\cdot f(\underline{x}; \underline{\mu}^a, \mathbf{\Sigma}^a) f(\underline{x}; \underline{\mu}^b, \mathbf{\Sigma}^b) dx_1 dx_2 .$$

We calculate the values $s_{12}, s_{11}$, and $s_{22}$ by numerical integration. Then, we apply Lemma 3 again to obtain $\rho$.

As mentioned in Lemma 3, a solution does not necessarily exist. There are different approaches to handle the cases where moment matching is not possible. For example, we can try to find the TWN distribution that, in some sense, is closest to the true distribution even though it does not have the same moments. Because this problem appears very rarely in practice, we handle it by ignoring the measurement in these cases. This approach is similar to the common solution of ignoring measurement that would cause the covariance matrix to loose the property of being positive definite in standard nonlinear filtering algorithms, such as the UKF.

### IV. TOROIDAL FILTERING ALGORITHM

We denote the system state at time step $k$ with $\underline{x}_k \in T^2$ and consider the system model

$$\underline{x}_{k+1} = \underline{x}_k + \underline{w}_k \mod 2\pi$$

with TWN distributed system noise $\underline{w}_k \sim \mathcal{TWN}(\underline{\mu}_k^w, \mathbf{\Sigma}_k^w)$. We interpret the modulo operator componentwise, i.e., $(a, b)^T \mod 2\pi = (a \mod 2\pi, b \mod 2\pi)^T$. The measurement $\hat{\underline{z}}_k \in T^2$ is disturbed by noise according to

$$\hat{\underline{z}}_k = \underline{x}_k + \underline{v}_k \mod 2\pi ,$$

where $\underline{v}_k$ is TWN distributed measurement noise with $\underline{v}_k \sim \mathcal{TWN}(\underline{\mu}_k^v, \mathbf{\Sigma}_k^v)$. The filtering algorithm consists of prediction and measurement update steps.

### A. Prediction Step

The prediction step propagates the current estimate through time. For this purpose, we assume the system model to be the identity. However, we allow for non-zero-mean system noise, which allows us to model system equations which add a known angle as well.

The predicted distribution consists of the convolution of the estimated distribution and the system noise distribution. This fact can be shown by calculating

$$f_{k+1}^p(\underline{x}_{k+1})$$
$$= \int_{T^2} f(\underline{x}_{k+1} | \underline{x}_k) f_k^e(\underline{x}_k) d\underline{x}_k$$
$$= \int_{T^2} \int_{T^2} f(\underline{x}_{k+1} | \underline{w}_k, \underline{x}_k) f_k^w(\underline{w}_k) d\underline{w}_k f_k^e(\underline{x}_k) d\underline{x}_k$$
$$= \int_{T^2} \int_{T^2} \delta(\underline{x}_{k+1} - \underline{w}_k - \underline{x}_k) f_k^w(\underline{w}_k) d\underline{w}_k f_k^e(\underline{x}_k) d\underline{x}_k$$
$$= \int_{T^2} f_k^w(\underline{x}_{k+1} - \underline{x}_k) f_k^e(\underline{x}_k) d\underline{x}_k$$
$$= (f_k^w * f_k^e)(\underline{x}_{k+1}) ,$$

where $*$ denotes convolution. Thus, we can directly apply the convolution as introduced in section III-A. The resulting algorithm is given in Algorithm 1.

---

**Algorithm 1:** Prediction of the proposed filter.

**Input**: estimate $\mathcal{TWN}(\underline{\mu}_k^e, \boldsymbol{\Sigma}_k^e)$, system noise $\mathcal{TWN}(\underline{\mu}_k^w, \boldsymbol{\Sigma}_k^w)$

**Output**: prediction $\mathcal{TWN}(\underline{\mu}_{k+1}^p, \boldsymbol{\Sigma}_{k+1}^p)$

$\underline{\mu}_{k+1}^p \leftarrow \underline{\mu}_k^e + \underline{\mu}_k^w \mod 2\pi$;
$\boldsymbol{\Sigma}_{k+1}^p \leftarrow \boldsymbol{\Sigma}_k^e + \boldsymbol{\Sigma}_k^w$;
**return** $\mathcal{TWN}(\underline{\mu}_{k+1}^p, \boldsymbol{\Sigma}_{k+1}^p)$;

---

### B. Measurement Update Step

In order to incorporate the information from a measurement into the state estimate, we perform a Bayesian measurement update step. In analogy to the derivation in [7], the estimated density is given by

$$f_k^e(\underline{x}_k) = f(\underline{x}_k | \hat{\underline{z}}_k) = c \cdot f(\hat{\underline{z}}_k | \underline{x}_k) f_k^p(\underline{x}_k)$$
$$= c \cdot f_k^v(\hat{\underline{z}}_k - \underline{x}_k) f_k^p(\underline{x}_k) \ ,$$

i.e., the renormalized product of the noise density $f_k^v(\hat{\underline{z}}_k - \underline{x}_k)$ and the prior density $f_k^p(\underline{x}_k)$. The density $f_k^v(\hat{\underline{z}}_k - \underline{x}_k)$ is given by $\mathcal{TWN}(\hat{\underline{z}}_k - \underline{\mu}_k^v, \boldsymbol{\Sigma}_k^v)$. The complete algorithm is given in Algorithm 2.

---

**Algorithm 2:** Measurement update of the proposed filter.

**Input**: prediction $\mathcal{TWN}(\underline{\mu}_k^p, \boldsymbol{\Sigma}_k^p)$, measurement noise $\mathcal{TWN}(\underline{\mu}_k^v, \boldsymbol{\Sigma}_k^v)$, measurement $\hat{\underline{z}}$

**Output**: estimate $\mathcal{TWN}(\underline{\mu}_k^e, \boldsymbol{\Sigma}_k^e)$

get $f_k^v(\hat{\underline{z}}_k - \underline{x}_k) = \mathcal{TWN}(\hat{\underline{z}}_k - \underline{\mu}_k^v, \boldsymbol{\Sigma}_k^v)$ ;
/* multiply $\mathcal{TWN}(\hat{\underline{z}}_k - \underline{\mu}_k^v, \boldsymbol{\Sigma}_k^v)$ and
$\mathcal{TWN}(\underline{\mu}_k^p, \boldsymbol{\Sigma}_k^p)$ */
get $c, m_{1,1}, m_{1,2}$ by numerical integration;
$\underline{\mu} \leftarrow \begin{bmatrix} \mathrm{atan2}(\mathrm{Im}\, m_{1,1}, \mathrm{Re}\, m_{1,1}) \\ \mathrm{atan2}(\mathrm{Im}\, m_{1,2}, \mathrm{Re}\, m_{1,2}) \end{bmatrix}$ ;
$\sigma_1 \leftarrow \sqrt{-2\log \|m_{1,1}\|}$ ;
$\sigma_2 \leftarrow \sqrt{-2\log \|m_{1,2}\|}$ ;
get $s_{12}, s_{11}, s_{22}$ by numerical integration;
$\rho_c \leftarrow s_{12}/\sqrt{s_{11}s_{22}}$;
$\rho \leftarrow \frac{1}{\sigma_1 \sigma_2} \sinh^{-1}\left( \sqrt{\sinh(\sigma_1^2)\sinh(\sigma_2^2)} \cdot \rho_c \right)$;
/* check result */
**if** $-1 < \rho < 1$ **then**
    $\underline{\mu}_k^e \leftarrow \underline{\mu}$ ;
    $\boldsymbol{\Sigma}_k^e \leftarrow \begin{bmatrix} \sigma_1^2 & \rho\sigma_1\sigma_2 \\ \rho\sigma_1\sigma_2 & \sigma_2^2 \end{bmatrix}$;
**else**
    $\underline{\mu}_k^e \leftarrow \underline{\mu}_k^p$;
    $\boldsymbol{\Sigma}_k^e \leftarrow \boldsymbol{\Sigma}_k^p$;
**end**
**return** $\mathcal{TWN}(\underline{\mu}_k^e, \boldsymbol{\Sigma}_k^e)$;

---

## V. EVALUATION

We evaluated the proposed approach in simulations. Unless specified otherwise, all angles are given in radians.

For comparison, we use a modified Kalman filter [1] with two-dimensional state vector. The unmodified Kalman filter fails completely once the periodic boundary is crossed. To perform a fair comparison, we modify the measurement update step of the regular Kalman filter by introducing a preprocessing of the measurement. Before the application of the Kalman filter update formulas, we reposition the measurement in such a way that its distance to the mean of the current estimate is at most $\pi$ in each dimension. The algorithm is given in Algorithm 3.

---

**Algorithm 3:** Measurement update for modified Kalman filter.

**Input**: prediction $\mathcal{N}(\underline{\mu}_k^p, \boldsymbol{\Sigma}_k^p)$, measurement noise $\mathcal{N}(\underline{\mu}_k^v, \boldsymbol{\Sigma}_k^v)$, measurement $\underline{z}$

**Output**: estimate $\mathcal{N}(\underline{\mu}_k^e, \boldsymbol{\Sigma}_k^e)$

/* preprocessing */
**for** $n = 1, 2$ **do**
    **if** $|(\underline{\mu}_k^p)_n - \underline{z}_n| > \pi$ **then**
        $\underline{z}_n \leftarrow \underline{z}_n + 2\pi \,\mathrm{sgn}\left( (\underline{\mu}_k^p)_n - \underline{z}_n \right)$;
    **end**
**end**
/* Kalman filter update */
$\mathbf{K} \leftarrow \boldsymbol{\Sigma}_k^p(\boldsymbol{\Sigma}_k^p + \boldsymbol{\Sigma}_k^v)^{-1}$;
$\underline{\mu}_k^e \leftarrow \underline{\mu}_k^p + \mathbf{K}(\underline{z} - \underline{\mu}_k^p) \mod 2\pi$ ;
$\boldsymbol{\Sigma}_k^e \leftarrow (\mathbf{I}_{2\times 2} - \mathbf{K})\boldsymbol{\Sigma}_k^p$;
**return** $\mathcal{N}(\underline{\mu}_k^e, \boldsymbol{\Sigma}_k^e)$;

---

| scenario | $\sigma_1^w$ | $\sigma_2^w$ | $\rho^w$ |
|---|---|---|---|
| 1n | 1 | 1 | 0 |
| 1c | 1 | 1 | 0.9 |
| 2n | 1 | 0.1 | 0 |
| 2c | 1 | 0.1 | 0.9 |

TABLE I: System noise parameters for the four scenarios.

We consider four scenarios (1n, 1c, 2n, 2c) with different system noise parameters, equal (1) or different (2) noise in the two dimensions, and uncorrelated (n) or correlated noise (c). The parameters are given in Table I. The mean of the system noise is zero in all cases. The zero-mean measurement noise has parameters $\sigma_1^v = 1$, $\sigma_2^v = 1$, $\rho_v = 0.5$.

In order to evaluate the performance, we use the angular RMSE (root mean square error) over $K$ time steps

$$e_j := \sqrt{\frac{1}{K} \sum_{k=1}^K \left( \min\left( (|\hat{x}_{k,j}\mu_{k,j}^e|, 2\pi - |\hat{x}_{k,j}\mu_{k,j}^e| \right) \right)^2}$$

in both dimensions $j = 1, 2$ separately. The angular RMSE considers the shorter of the two possible paths between two points on a circle.
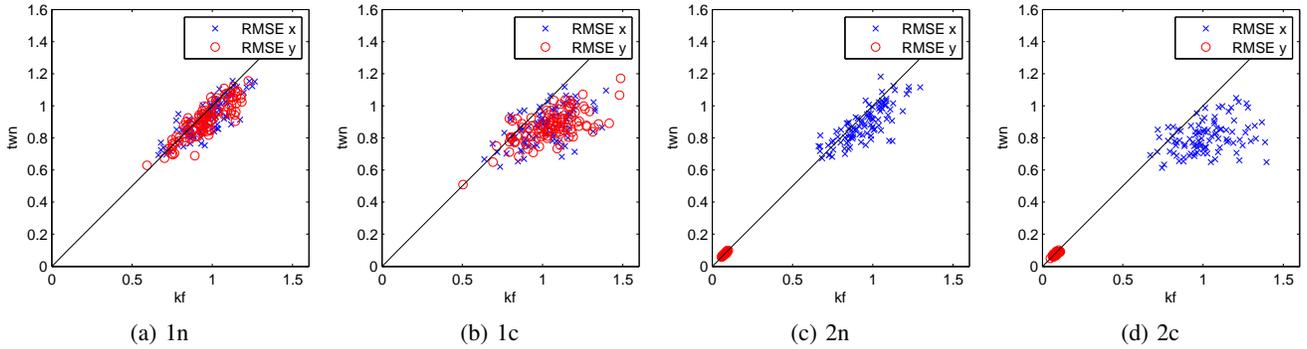
Fig. 3: Comparison of RMSE of Kalman filter (kf) and proposed filter (twn) over 100 Monte Carlo runs for uncorrelated system noise and correlated system noise. Values below the diagonal mean that the proposed filter is better, values above the diagonal mean that the Kalman filter is better. It can be seen that the proposed filter slightly outperforms the Kalman filter in the 1n and 2n scenarios and yields significantly better results in the 1c and 2c scenarios.

| scenario | # outperforming runs | | mean error quotient | |
|---|---|---|---|---|
| | $j = 1$ | $j = 2$ | $j = 1$ | $j = 2$ |
| 1n | 73 | 73 | 1.0465 | 1.0520 |
| 1c | 87 | 93 | 1.1601 | 1.2070 |
| 2n | 80 | 55 | 1.0667 | 1.0065 |
| 2c | 93 | 77 | 1.2618 | 1.0687 |

TABLE II: Results from 100 Monte Carlo runs. In each dimension $j = 1, 2$, we give the number of runs where the proposed filter outperforms the Kalman filter, as well as the mean of the quotient of the error of the Kalman filter and the error of the proposed filter (values larger than 1 indicate that the Kalman filter performs worse).

We performed 100 Monte Carlos runs with $K = 50$ time steps each. The results are depicted in Fig. 3. It can be seen that the proposed filter outperforms the Kalman filter significantly, in particular in the case of correlated system noise.

As far as runtime is concerned, numerical evaluation of the integrals for $c, m_{1,1}, m_{1,2}, s_{12}, s_{11}, s_{22}$ is by far the most costly step of the algorithm. Still, the algorithm is reasonably fast overall. On a standard computer with an Intel Core i7-4770 CPU, 16 GB RAM, and MATLAB 2013b, we can calculate all required integrals for one time step in less than $100 \, \text{ms}$.

## VI. CONCLUSION

In this paper, we presented a new method for estimating correlated angles using directional statistics. We derived a filter based on the toroidal wrapped normal distribution and evaluated its performance by comparing with a standard approach in multiple simulations. Our results suggest that the proposed approach outperforms standard approaches, particularly in cases of large noise and strong correlation. To the best of our knowledge, the presented algorithms constitute the first recursive filter on the torus that is based on directional statistics and correctly takes periodicity into account.

So far, the proposed filter is limited to two angular dimensions. Future work may include the generalization to $n$ correlated angles, i.e., estimation on the $n$-torus. Even though some of the methods used in this paper can easily be generalized to a higher number of dimensions, this generalization involves some additional challenges. Particularly, the moment-based approximation of the multiplication of partially wrapped normal densities may be difficult, as the numerical integration used for calculating the toroidal moments does not scale well for a higher number of dimensions. Therefore, other algorithms for approximating the product of multidimensional wrapped normal densities may be necessary.

We plan to extend the proposed filtering algorithm to nonlinear system and measurement equations by applying deterministic sampling techniques as have been used in [7] and [23]. Furthermore, we have published some results on a related circular-linear distribution [24], which might be used to generalize the toroidal filter to arbitrary partially periodic spaces.

## APPENDIX

### A. Proof of Lemma 1

We introduce the abbreviation

$$\underline{\mu}_{jk} := \underline{\mu} + [2\pi j, 2\pi k]^T \ ,$$

and calculate

$$\int_0^{2\pi}\int_0^{2\pi} f(x) \begin{bmatrix} e^{inx_1} \\ e^{inx_2} \end{bmatrix} dx_1 dx_2$$

$$= \int_0^{2\pi}\int_0^{2\pi} \sum_{j,k=-\infty}^{\infty} \mathcal{N}\left(\underline{x}; \underline{\mu}_{jk}, \mathbf{\Sigma}\right) \begin{bmatrix} e^{inx_1} \\ e^{inx_2} \end{bmatrix} dx_1 dx_2$$

$$\overset{(a)}{=} \begin{bmatrix} \int_0^{2\pi}\sum_{j=-\infty}^{\infty}\sum_{k=-\infty}^{\infty}\int_0^{2\pi}\mathcal{N}\left(\underline{x}; \underline{\mu}_{jk}, \mathbf{\Sigma}\right) dx_2 e^{inx_1} dx_1 \\ \int_0^{2\pi}\sum_{k=-\infty}^{\infty}\sum_{j=-\infty}^{\infty}\int_0^{2\pi}\mathcal{N}\left(\underline{x}; \underline{\mu}_{jk}, \mathbf{\Sigma}\right) dx_1 e^{inx_2} dx_2 \end{bmatrix}$$

$$\overset{(b)}{=} \begin{bmatrix} \int_0^{2\pi} \sum_{j=-\infty}^{\infty} \int_{-\infty}^{\infty} \mathcal{N}\left(\underline{x}; \underline{\mu} + \begin{bmatrix} 2\pi j \\ 0 \end{bmatrix}, \boldsymbol{\Sigma}\right) dx_2 e^{inx_1} dx_1 \\ \int_0^{2\pi} \sum_{k=-\infty}^{\infty} \int_{-\infty}^{\infty} \mathcal{N}\left(\underline{x}; \underline{\mu} + \begin{bmatrix} 0 \\ 2\pi k \end{bmatrix}, \boldsymbol{\Sigma}\right) dx_1 e^{inx_2} dx_2 \end{bmatrix}$$

$$\overset{(c)}{=} \begin{bmatrix} \int_0^{2\pi} \sum_{j=-\infty}^{\infty} \mathcal{N}\left(x_1; \mu_1 + 2\pi j, \sigma_1\right) e^{inx_1} dx_1 \\ \int_0^{2\pi} \sum_{k=-\infty}^{\infty} \mathcal{N}\left(x_2; \mu_2 + 2\pi k, \sigma_2\right) e^{inx_2} dx_2 \end{bmatrix}$$

$$= \begin{bmatrix} \exp(in\mu_1 - n\sigma_1^2/2) \\ \exp(in\mu_2 - n\sigma_2^2/2) \end{bmatrix} .$$

At (a), we use the dominated convergence theorem to interchange summation and integration. Step (b) is based on concatenation of integrals

$$\sum_{k=-\infty}^{\infty} \int_0^a f(x + ka) dx$$
$$= \cdots + \int_0^a f(x-a) dx + \int_0^a f(x) dx + \int_0^a f(x+a) dx + \cdots$$
$$= \cdots + \int_{-a}^0 f(x) dx + \int_0^a f(x) dx + \int_a^{2a} f(x) dx + \cdots$$
$$= \int_{-\infty}^{\infty} f(x) dx .$$

We use marginalization of a Gaussian distribution at (c). Finally, we apply the formulas for circular moments of WN distributions as given in [7], [4]. $\qquad\square$

*B. Proof of Lemma 3*

*Proof.* The equations for mean and uncertainty are obtained by solving the system of equations

$$\begin{bmatrix} m_{1,1} \\ m_{1,2} \end{bmatrix} = \begin{bmatrix} \exp(in\mu_1 - n\sigma_1^2/2) \\ \exp(in\mu_2 - n\sigma_2^2/2) \end{bmatrix}$$

for $\underline{\mu}$ and $\sigma_1, \sigma_2$. To get the correlation parameter $\rho$, we solve

$$\rho_c = \frac{\sinh(\sigma_1\sigma_2\rho)}{\sqrt{\sinh(\sigma_1^2)\sinh(\sigma_2^2)}}$$

for $\rho$. If $\rho \in (-1, 1)$, this yields a valid solution and the resulting TWN distribution has indeed the desired first toroidal moment and circular correlation coefficient. If, conversely, $\rho \notin (-1, 1)$, there exists no TWN distribution with the desired properties, because it is a necessary condition that $\rho$ fulfills this equation. The intuition why such a TWN distribution does not necessarily exist is depicted in Fig. 2b. $\qquad\square$

## REFERENCES

[1] R. E. Kalman, "A New Approach to Linear Filtering and Prediction Problems," *Transactions of the ASME Journal of Basic Engineering*, vol. 82, pp. 35–45, 1960.

[2] S. J. Julier and J. K. Uhlmann, "Unscented Filtering and Nonlinear Estimation," *Proceedings of the IEEE*, vol. 92, no. 3, pp. 401–422, Mar. 2004.

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

[4] S. R. Jammalamadaka and A. Sengupta, *Topics in Circular Statistics*. World Scientific, 2001.

[5] M. Azmani, S. Reboul, J.-B. Choquel, and M. Benjelloun, "A Recursive Fusion Filter for Angular Data," in *IEEE International Conference on Robotics and Biomimetics (ROBIO 2009)*, 2009, pp. 882–887.

[6] G. Stienne, S. Reboul, M. Azmani, J. Choquel, and M. Benjelloun, "A Multi-temporal Multi-sensor Circular Fusion Filter," *Information Fusion*, vol. 18, pp. 86–100, Jul. 2013.

[7] 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.

[8] G. Kurz, I. Gilitschenski, S. J. Julier, and U. D. Hanebeck, "Recursive Estimation of Orientation Based on the Bingham Distribution," in *Proceedings of the 16th International Conference on Information Fusion (Fusion 2013)*, Istanbul, Turkey, Jul. 2013.

[9] I. Gilitschenski, G. Kurz, S. J. Julier, and U. D. Hanebeck, "Unscented Orientation Estimation Based on the Bingham Distribution," *arXiv preprint: Systems and Control (cs.SY)*, 2013. [Online]. Available: http://arxiv.org/abs/1311.5796

[10] J. Glover and L. P. Kaelbling, "Tracking 3-D Rotations with the Quaternion Bingham Filter," MIT, Tech. Rep., Mar. 2013.

[11] W. Feiten, M. Lang, and S. Hirche, "Rigid Motion Estimation using Mixtures of Projected Gaussians," in *Proceedings of the 16th International Conference on Information Fusion (Fusion 2013)*, 2013.

[12] J. M. Selig, *Geometric Fundamentals of Robotics*, 2nd ed., ser. Monographs in Computer Science, F. B. S. David Gries, Ed. New York: Springer, 2005.

[13] 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.

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

[15] S. R. Jammalamadaka and Y. Sarma, "A Correlation Coefficient for Angular Variables," *Statistical theory and data analysis II*, pp. 349–364, 1988.

[16] K. V. Mardia, "Statistics of Directional Data," *Journal of the Royal Statistical Society. Series B (Methodological)*, vol. 37, no. 3, pp. 349–393, 1975.

[17] 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.

[18] P. E. Jupp and K. V. Mardia, "A General Correlation Coefficient for Directional Data and Related Regression Problems," *Biometrika*, vol. 67, no. 1, pp. 163–173, 1980.

[19] N. I. Fisher and A. J. Lee, "A Correlation Coefficient for Circular Data," *Biometrika*, vol. 70, no. 2, pp. 327–332, 1983.

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

[21] A. Genz, "Numerical Computation of Rectangular Bivariate and Trivariate Normal and t Probabilities," *Statistics and Computing*, vol. 14, no. 3, pp. 251–260, 2004.

[22] L. F. Shampine, "Matlab Program for Quadrature in 2D," *Applied Mathematics and Computation*, vol. 202, no. 1, pp. 266–274, 2008.

[23] U. D. Hanebeck, "PGF 42: Progressive Gaussian Filtering with a Twist," in *Proceedings of the 16th International Conference on Information Fusion (Fusion 2013)*, Istanbul, Turkey, Jul. 2013.

[24] G. Kurz, I. Gilitschenski, and U. D. Hanebeck, "The Partially Wrapped Normal Distribution for SE(2) Estimation (to appear)," in *Proceedings of the 2014 IEEE International Conference on Multisensor Fusion and Information Integration(MFI 2014)*, Beijing, China, Sep. 2014.