# Unscented von Mises–Fisher Filtering

Gerhard Kurz[1], Igor Gilitschenski[2], and Uwe D. Hanebeck[1]

**Abstract**

We introduce the Unscented von Mises–Fisher Filter (UvMFF), a nonlinear filtering algorithm for dynamic state estimation on the $n$-dimensional unit hypersphere. Estimation problems on the unit hypersphere occur in computer vision, for example when using omnidirectional cameras, as well as in signal processing. As approaches in literature are limited to very simple system and measurement models, we propose a deterministic sampling scheme on the unit hypersphere, which allows us to handle nonlinear system and measurement models. The proposed approach can be seen as a hyperspherical variant of the Unscented Kalman Filter (UKF). The advantages of the novel method are shown by means of simulations.

## I. Introduction

Estimation problems in which the state is represented by a point on the unit sphere or unit hypersphere have gained interest recently. A number of publications have considered the use of the von Mises–Fisher (VMF) distribution [1], [2] for hyperspherical estimation problems. Applications include visual tracking on the unit sphere [3], omnidirectional cameras [4], [5], estimation of crystal orientations in crystallography [6], high angular resolution diffusion MRI [7], clustering of beam directions for radiation therapy [8], and signal processing for microphone arrays [9].

Some authors such as [3] have proposed recursive filtering algorithms based on the VMF distribution. State-of-the-art methods for dealing with VMF distributions are, however, not capable of propagating the distribution though nonlinear functions. To solve this problem, we propose a novel filter based on a deterministic sampling scheme that is reminiscent of the Unscented Kalman Filter (UKF) [10]. Compared with stochastic sampling, deterministic sampling yields higher accuracy with a lower number



[1] Gerhard Kurz and Uwe D. Hanebeck 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). Igor Gilitschenski is with the Autonomous Systems Laboratory (ASL), Institute of Robotics and Intelligent Systems, Swiss Federal Institute of Technology (ETH) Zurich, 8092 Zurich, Switzerland (e-mail: igilitschenski@ethz.ch).

of samples [11], [12]. This work constitutes a generalization of the nonlinear circular filters based on deterministic sampling [13], [14] to a higher number of dimensions. Furthermore, it is related to a similar approach that we published for the Bingham distribution [15].

The contributions of this paper can be summarized as follows. We present a novel deterministic sampling scheme for the von Mises–Fisher distribution with an arbitrary number of dimensions. Then, we propose a nonlinear filtering algorithm based on the novel sampling scheme. Finally, we evaluate the proposed methods in comparison with the UKF.

## II. Hyperspherical Statistics

The von Mises–Fisher (VMF) distribution [1] is given by the unimodal probability density function (pdf)

$$\mathcal{VMF}(\underline{x}; \underline{\mu}, \kappa) = c_d(\kappa) \cdot \exp(\kappa \cdot \underline{\mu}^T \underline{x}) \ , \tag{1}$$

where $\underline{x} \in S^{d-1} = \{\underline{x} \in \mathbb{R}^d : \|\underline{x}\| = 1\}$ is located on the unit hypersphere in $\mathbb{R}^d$ where $d \geq 2$. The parameter $\underline{\mu} \in S^{d-1}$ specifies the location of the mode of the distribution and $\kappa \geq 0$ specifies its concentration. Moreover, the term $c_d(\kappa)$ refers to the normalization constant given by

$$c_d(\kappa) = \frac{\kappa^{d/2-1}}{(2\pi)^{d/2} I_{d/2-1}(\kappa)} \ , \tag{2}$$

where $I_v(x)$ is the modified Bessel function of first kind and order $v$ [16, Sec. 9.6].

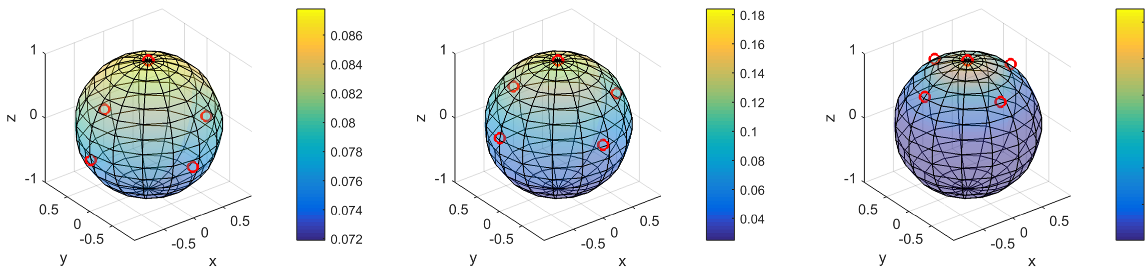
The VMF distribution is isotropic, i.e., it is rotationally symmetric around $\underline{\mu}$. This can be seen by $\underline{\mu}^T \underline{x} = \cos(\angle(\underline{\mu}, \underline{x}))$ because the value of the pdf only depends on the angle between $\underline{\mu}$ and $\underline{x}$. The von Mises distribution [17] arises as a special case of the VMF distribution for $d = 2$.

On the unit circle, it is common to consider trigonometric moments [13], and this concept can be generalized to the unit hypersphere. In this case, we consider the expectation

$$\underline{m} = \mathbb{E}(\underline{x}) = \int_{S^{d-1}} \underline{x} \cdot f(\underline{x}) \, d\underline{x} \ , \tag{3}$$

also called the mean resultant vector. For the VMF distribution, it can be shown that $\underline{m} = \underline{\mu} \cdot A_d(\kappa)$, where $A_d(\kappa) = \frac{I_{d/2}(\kappa)}{I_{d/2-1}(\kappa)}$. The direction of $\underline{m}$ can be seen as a hyperspherical mean, and it is identical to the direction of $\underline{\mu}$ in case of a VMF distribution. The length of $\underline{m}$ determines the uncertainty of the distribution and is related to the $\kappa$ parameter of the VMF distribution.

Let us now consider the problem of parameter estimation for the VMF distribution. Given a set of $n$ weighted samples $\underline{x}_1, \dots, \underline{x}_n$ with weights $w_1, \dots, w_n > 0$, where $\sum_{k=1}^n w_k = 1$, we can calculate the mean resultant vector according to $\underline{m} = \sum_{k=1}^n w_k \cdot \underline{x}_k$. It has been shown that matching the mean resultant vector of a VMF distribution to the mean resultant vector of the samples yields the maximum




$\kappa = 0.1$ $\kappa = 1$ $\kappa = 4$

Fig. 1. Examples of deterministic sampling of a VMF distribution with $\underline{\mu} = [0, 0, 1]^T$ and different values of $\kappa$.

likelihood estimate of the given samples, i.e., moment matching is equivalent to maximum likelihood estimation [18, Sec. A.1]. Solving for the parameters of the VMF distribution yields

$$\underline{\mu} = \frac{\underline{m}}{\|\underline{m}\|}, \quad \kappa = A_d^{-1}(\|\underline{m}\|) \ , \tag{4}$$

where $\underline{m} \neq \underline{0}$ is the mean resultant vector of the samples[1]. Parameter estimation and an efficient algorithm for computation of $A_d^{-1}(\cdot)$ are discussed in [19]. We later refer to this procedure as ParameterEstimation$(\cdot)$.

## III. Deterministic Sampling on the Hypersphere

In this section, we derive a deterministic sampling scheme for the VMF distribution. First, we only consider the case where $\underline{\mu} = [1, 0, \dots, 0]^T$. We will later show how to generalize the proposed approach to arbitrary $\underline{\mu} \in S^{d-1}$.

The basic idea of the proposed approach consists in matching the mean resultant vector $\underline{m}$ of a set of carefully chosen samples to that of the VMF distribution. We consider $2d-1$ equally weighted[2] samples, because closed-form solutions are possible for this number of samples and because it corresponds to the number of samples used by the UKF [10] in dimension $d-1$. These samples are given by

$$\begin{aligned}
\underline{s}_1^T &= [1, 0, \dots, 0] \ , \\
\underline{s}_2^T &= [\cos(\alpha), \sin(\alpha), 0, \dots, 0] \ , \\
\underline{s}_3^T &= [\cos(\alpha), -\sin(\alpha), 0, \dots, 0] \ , \\
\underline{s}_4^T &= [\cos(\alpha), 0, \sin(\alpha), 0, \dots, 0] \ , \\
\underline{s}_5^T &= [\cos(\alpha), 0, -\sin(\alpha), 0, \dots, 0] \ ,
\end{aligned}$$

[1] If $\underline{m} = \underline{0}$, $\underline{\mu}$ is undefined and $\kappa = 0$, i.e., the VMF distribution is uniform.

[2] It would be possible to choose a different weight for the sample at $\underline{\mu}$ by introducing a scaling parameter similar to the UKF [10] or the unscented Bingham filter [15]. For simplicity, we omit this parameter as a solution with equally weighted samples always exists.

$$\vdots$$

$$\underline{s}_{2d-2}^T = [\cos(\alpha), 0, \ldots, 0, \sin(\alpha)] \ ,$$

$$\underline{s}_{2d-1}^T = [\cos(\alpha), 0, \ldots, 0, -\sin(\alpha)] \ ,$$

where $\alpha \in (0, \pi)$ is an angle that remains to be determined. We calculate the mean resultant vector $\underline{m} = [m_1, \ldots, m_d]^T$

$$\underline{m} = \frac{1}{2d-1} \sum_{k=1}^{2d-1} \underline{s}_k = \frac{1}{2d-1} [1 + (2d-2)\cos(\alpha), 0, \ldots, 0]^T \ .$$

Using $m_1 \geq 0$ for $\underline{\mu} = [1, 0, \ldots, 0]^T$, we get $\|\underline{m}\| = m_1$ and

$$\|\underline{m}\| = \frac{|1 + (2d-2)\cos(\alpha)|}{2d-1} \Leftrightarrow \cos(\alpha) = \frac{\|\underline{m}\|(2d-1) - 1}{2d-2}$$

Observe that $\|\underline{m}\| \in [0, 1]$. Because we have

$$\frac{\|\underline{m}\|(2d-1) - 1}{2d-2} \geq \frac{0 - 1}{2d-2} \geq -1$$

and

$$\frac{\|\underline{m}\|(2d-1) - 1}{2d-2} \leq \frac{(2d-1) - 1}{2d-2} = 1 \ ,$$

the term $\frac{\|\underline{m}\|(2d-1)-1}{2d-2}$ is always in $[-1, 1]$. Thus, a solution always exists and we can obtain $\alpha$ from

$$\alpha = \arccos\left(\frac{\|\underline{m}\|(2d-1) - 1}{2d-2}\right) \ . \tag{5}$$

To generalize this solution to a VMF distribution with arbitrary $\underline{\mu} \in S^{d-1}$, we rotate all samples using a rotation matrix $\mathbf{Q}$ that ensures $\underline{\mu} = \mathbf{Q}\underline{s}_1$. This rotation matrix is not uniquely determined, so we use the QR-decomposition of a suitable initial matrix $\mathbf{M}$ whose first column is $\underline{\mu}$ to obtain $\mathbf{Q}$ (see also [20]).

Pseudocode of the proposed sampling method is given in Algorithm 1, where we collect all samples in a matrix $\mathbf{S} = [\underline{s}_1, \ldots, \underline{s}_{2d-1}]$. Some examples are shown in Fig. 1. It is worth mentioning that for $d = 2$ (i.e., on the unit circle), this sampling scheme coincides with the scheme proposed in [13].

## IV. Nonlinear Recursive Filtering

A novel nonlinear recursive filtering algorithm is now developed based on the VMF distribution. To deal with the nonlinearity, we use the deterministic sampling scheme introduced in the previous section. In the following, we do not consider system and measurement functions $S^{d-1} \times S^{d-1} \to S^{d-1}$ based on a group operation on the unit hypersphere as it is sometimes done for $S^1$ and $S^3$, e.g., in [21], [22]. Instead, we consider the transition density $f(\underline{x}_{k+1}|\underline{x}_k)$ and the measurement likelihood $f(\underline{z}_k|\underline{x}_k)$.

### A. Prediction Step

In this section, we distinguish between several system models, i.e., different types of transition densities.

**Algorithm 1:** Deterministic sampling.

**Input**: VMF parameters $\underline{\mu}$, $\kappa$

**Output**: samples $\mathbf{S}$

$d \leftarrow \dim(\underline{\mu})$;

/* obtain samples for $\underline{\mu} = [1, 0, \ldots, 0]^T$ */

$\mathbf{S} \leftarrow \mathbf{0}_{d \times (2d-1)}$;

$\mathbf{S}(1,1) \leftarrow 1$;

$m_1 \leftarrow I_{d/2}(\kappa) / I_{d/2-1}(\kappa)$ ;

$\alpha \leftarrow \arccos(((2d-1)m_1 - 1)/(2d-2))$;

**for** $i \leftarrow 1$ **to** $d-1$ **do**

  $\mathbf{S}(1, 2i) \leftarrow \cos(\alpha)$; $\mathbf{S}(1, 2i+1) \leftarrow \cos(\alpha)$;

  $\mathbf{S}(i+1, 2i) \leftarrow \sin(\alpha)$; $\mathbf{S}(i+1, 2i+1) \leftarrow -\sin(\alpha)$;

/* rotate samples around $\underline{\mu}$ */

$\mathbf{M} \leftarrow [\underline{\mu}, \underline{0}, \ldots \underline{0}]$;

$[\mathbf{Q}, \mathbf{R}] \leftarrow \text{QrDecomposition}(\mathbf{M})$;

**if** $\mathbf{R}(1,1) < 0$ **then**

  /* ensure that first column of $\mathbf{Q}$ is $\underline{\mu}$ rather than $-\underline{\mu}$ */

  $\mathbf{Q} \leftarrow -\mathbf{Q}$;

$\mathbf{S} \leftarrow \mathbf{Q} \cdot \mathbf{S}$;

**return** $\mathbf{S}$;

*1) Identity System Model:* First, we consider a simple identity system model with an optional fixed rotation defined by a rotation matrix $\mathbf{Q}_k \in SO(d) \subset \mathbb{R}^{d \times d}$. We assume the transition density to be given by a VMF distribution according to

$$f(\underline{x}_{k+1} | \underline{x}_k) = \mathcal{VMF}(\underline{x}_{k+1}; \underline{\mu} = \mathbf{Q}_k \cdot \underline{x}_k, \kappa_k^w) \ . \tag{6}$$

This system can be seen as a hyperspherical analogue to an identity system model with additive noise, even though we do not have a true addition operation on the hypersphere. According to the Chapman–Kolmogorov equation, the predicted density $f^p(\underline{x}_{k+1})$ is obtained from the previous estimated density $f^e(\underline{x}_k)$ according to

$$f^p(\underline{x}_{k+1}) = \int_{S^{d-1}} f(\underline{x}_{k+1} | \underline{x}_k) f^e(\underline{x}_k) d\underline{x}_k \ . \tag{7}$$

It has been shown that the resulting density is not, in general, the density of a VMF distribution [23].

The true density can be approximated using a VMF density with parameters

$$\underline{\mu}_{k+1}^{p} = \mathbf{Q}_k \underline{\mu}_k^e, \quad \kappa_{k+1}^{p} = A_d^{-1}(A_d(\kappa_k^e) A_d(\kappa_k^w)))$$

as given in [9]. Further discussion can be found in [23].

*2) Nonlinear System Model with VMF noise:* More generally, we can assume that the transition density is given by

$$f(\underline{x}_{k+1} | \underline{x}_k) = \mathcal{VMF}(\underline{x}_{k+1}; \underline{\mu} = a_k(\underline{x}_k), \kappa_k^w) \ , \tag{8}$$

where $a_k : S^{d-1} \to S^{d-1}$ is an arbitrary function that describes the system dynamics. This can be seen as an analogue to a nonlinear system model with additive noise. Analogous to the circular prediction algorithm in [13], we can derive the method given in Algorithm 2.

**Algorithm 2:** Prediction with VMF noise

**Input**: parameters of estimate $\underline{\mu}_k^e, \kappa_k^e$, concentration parameter of noise $\kappa_k^w$, function $a_k(\cdot)$

**Output**: parameters of prediction $\underline{\mu}_{k+1}^p, \kappa_{k+1}^p$

$\underline{s}_1^e, \dots, \underline{s}_{2d-1}^e \leftarrow$ DeterministicSampling$(\underline{\mu}_k^e, \kappa_k^e)$;

$\underline{s}_1^p, \dots, \underline{s}_{2d-1}^p \leftarrow a_k(\underline{s}_1^e), \dots, a_k(\underline{s}_{2d-1}^e)$;

$\underline{\mu}, \kappa \leftarrow$ ParameterEstimation$(\underline{s}_1^p, \dots, \underline{s}_{2d-1}^p)$;

$\underline{\mu}_{k+1}^p \leftarrow \underline{\mu}$;

$\kappa_{k+1}^p \leftarrow A_d^{-1}(A_d(\kappa) A_d(\kappa_k^w)))$;

**return** $\underline{\mu}_{k+1}^p, \kappa_{k+1}^p$;

*3) Nonlinear System Model with Arbitrary Noise:* We can derive a VMF-assumed filter for a generalization of the previous system model where we allow the use of an arbitrary noise distribution. In this case, the transition density is given by

$$\begin{aligned} f(\underline{x}_{k+1} | \underline{x}_k) &= \int_W f(\underline{x}_{k+1} | \underline{x}_k, \underline{w}_k) f^w(\underline{w}_k) dw_k \\ &= \int_W \delta(x_{k+1} - a_k(x_k, w_k)) f^w(\underline{w}_k) dw_k \ , \end{aligned} \tag{9}$$

where $a_k : S^{n-1} \times W \to S^{n-1}$ is an arbitrary function, and $W$ is the space on which the noise distribution is defined. We assume that a set of weighted samples from the noise distribution is given. Then, we calculate the Cartesian product of all state and all noise samples and propagate each combination through the system function. The resulting samples are then approximated by a VMF density (see Algorithm 3). This method is similar to the non-additive system noise scenario in [14].

**Algorithm 3:** Prediction with arbitrary noise

**Input**: parameters of estimate $\underline{\mu}_k^e, \kappa_k^e$, noise samples $\underline{s}_1^w, \dots, \underline{s}_L^w$ with weights $\gamma_1, \dots, \gamma_L$, function $a_k(\cdot, \cdot)$

**Output**: parameters of prediction $\underline{\mu}_{k+1}^p, \kappa_{k+1}^p$

$\underline{s}_1, \dots, \underline{s}_{2d-1} \leftarrow$ DeterministicSampling$(\underline{\mu}_k^e, \kappa_k^e)$;

**for** $j \leftarrow 1$ **to** $2d-1$ **do**

  **for** $k \leftarrow 1$ **to** $L$ **do**

    $\underline{s}_{(j-1)\cdot L+k}^p \leftarrow a_k(\underline{s}_j, \underline{s}_k^w)$;

    $w_{(j-1)\cdot L+k} \leftarrow \frac{\gamma_k}{2d-1}$;

$\underline{\mu}_{k+1}^p, \kappa_{k+1}^p \leftarrow$ ParameterEstimation$(\underline{s}_1^p, \dots, \underline{s}_{L\cdot(2d-1)}^p,$ $w_1, \dots, w_{L\cdot(2d-1)})$;

**return** $\underline{\mu}_{k+1}^p, \kappa_{k+1}^p$;

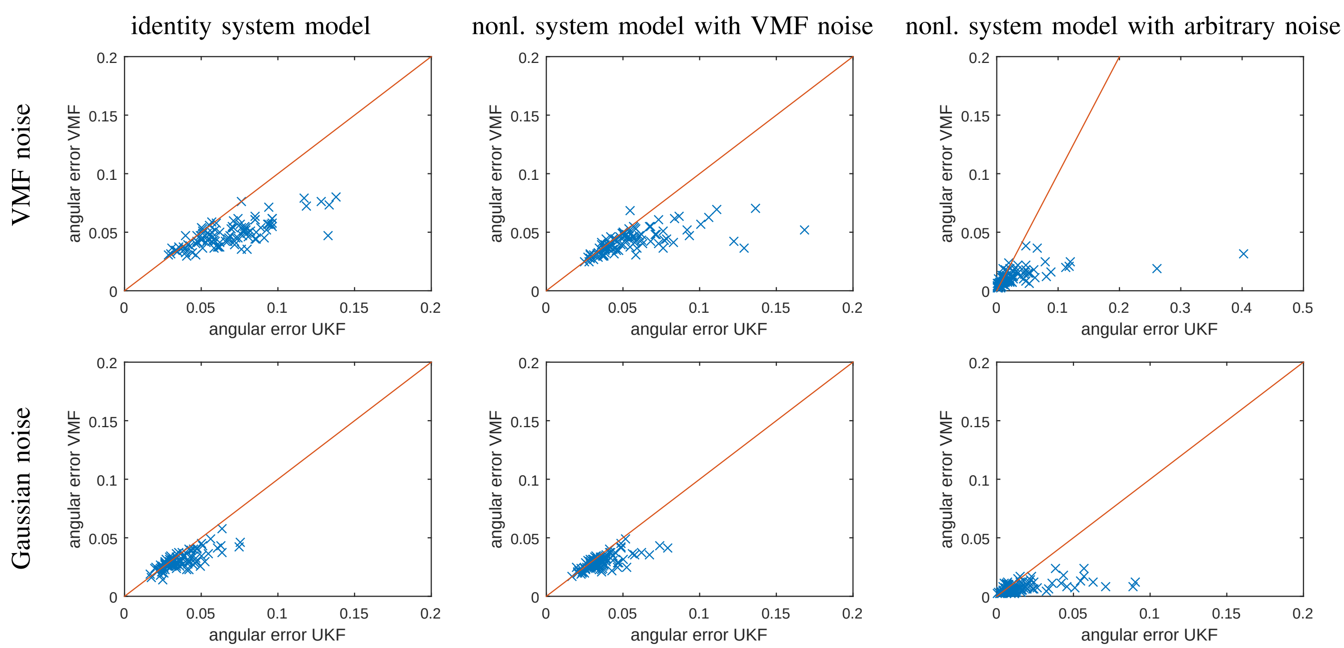



Fig. 2. Evaluation results. Each cross represents one Monte Carlo run. Points below the diagonal (red) indicate that the VMF filter performed better whereas points above the diagonal indicate that the UKF performed better.

### *B. Measurement Update Step*

Similar to the system models discussed above, it is also possible to consider different measurement models.

*1) Identity Measurement Model:* A common measurement model is given by the identity with an offset given by a rotation matrix $\mathbf{Q}_k$. We assume that the likelihood is given by an unnormalized VMF distribution

$$f(\underline{z}_k|\underline{x}_k) \propto \mathcal{VMF}(\underline{z}_k; \underline{\mu}_k^v = \mathbf{Q}_k \cdot \underline{x}_k, \kappa_k^v) \ . \tag{10}$$

In this case, we can apply Bayes' theorem to obtain

$$f(\underline{x}_k|\underline{z}_k) \propto f(\underline{z}_k|\underline{x}_k) f(\underline{x}_k) \ . \tag{11}$$

Because both the prior $f(\underline{x}_k) = \mathcal{VMF}(\underline{x}, \underline{\mu}_k^p, \kappa_k^p)$ and the likelihood $f(\underline{z}_k|\underline{x}_k)$ are VMF-distributed, we can use the exact multiplication formula for VMF distributions (see [3, eq. (2.23)] ). The parameters of the renormalized product of two VMF distributions with parameters $\underline{\mu}^A, \kappa^A$ and $\underline{\mu}^B, \kappa^B$ are given by

$$\underline{\mu} = \frac{\kappa^A \underline{\mu}^A + \kappa^B \underline{\mu}^B}{\|\kappa^A \underline{\mu}^A + \kappa^B \underline{\mu}^B\|}, \quad \kappa = ||\kappa^A \underline{\mu}^A + \kappa^B \underline{\mu}^B|| \ . \tag{12}$$

*2) Nonlinear Measurement Model:* If a nonlinear measurement model is given, more sophisticated approaches are necessary. If the measurement model is given by

$$f(\underline{z}_k|\underline{x}_k) = \mathcal{VMF}(\underline{z}_k; \underline{\mu}_k^v = h_k(\underline{x}_k), \kappa_k^v) \tag{13}$$

with a nonlinear measurement function $h_k : S^{d-1} \to S^{d-1}$, the techniques based on approximation of the measurement function described in [24] can be used. More generally, if an arbitrary likelihood $f(\underline{z}_k|\underline{x}_k)$ is given, we can apply the progressive measurement update proposed in [25].

## V. Evaluation

The proposed filter is evaluated in simulations on the sphere $S^2 \subset \mathbb{R}^3$. We provide a comparison with an Unscented Kalman Filter (UKF) [10] where we constrain the mean vector of the Gaussian distribution to unit length after every prediction and every measurement update. An evaluation of the propagation accuracy can be found in the supplementary material for this paper.

We consider three different system functions. First, we investigate the identity system model (6). Then we consider a nonlinear model with VMF noise (8), where $a_k(\cdot)$ is the nlerp-function (normalized linear interpolation) known from computer graphics [26, eq. (4.62)]

$$a_k(\underline{x}_k, \alpha) = \frac{\alpha \cdot \underline{x} + (1-\alpha)\underline{u}_k}{\|\alpha \cdot \underline{x} + (1-\alpha)\underline{u}_k\|} \tag{14}$$

(assuming $\underline{x}_k \neq -\underline{u}$ ), which interpolates between $\underline{x}_k$ and $\underline{u}_k$. We choose time-variant $\underline{u}_k = [\cos(c \cdot k), \sin(c \cdot k), 0]^T$, where $c = 0.01$ and $\alpha = 0.05$. Finally, we use a system model with arbitrary noise (9), where the system function is given by the nlerp-function, and $\alpha$ is distributed according to a discrete distribution with probabilities $\mathbb{P}(\alpha = 0.01) = 0.6$, $\mathbb{P}(\alpha = 0.05) = 0.2$, and $\mathbb{P}(\alpha = 0.1) = 0.2$. In all cases, we restrict ourselves to the identity measurement model (10).

The noise parameters are given by $\kappa_k^w = 100$, $\kappa_k^v = 5$ and the initial estimate is set to $\underline{\mu}_0^e = [1, 0, 0]^T$, $\kappa_0^e = 0.1$, whereas the true initial state is uniformly sampled from $S^2$. For the UKF, the noise parameters were converted to Gaussians using Monte Carlo integration. In order to avoid an unfair disadvantage for the UKF as a result of assuming the wrong noise, we performed all simulations twice, once with VMF noise and once with comparable (normalized) Gaussian noise obtained by Monte Carlo integration.

The system was simulated for $K = 100$ time steps. We consider the mean cosine error measure (see [27, eq. (1.3.7)])

$$\frac{1}{K} \sum_{k=1}^{K} 1 - \cos(\angle(\underline{x}_k, \underline{x}_k^{\text{true}})) \ . \tag{15}$$

The results of 100 Monte Carlo runs are depicted in Fig. 2. It can be seen that the performance of the proposed filter is better than the performance of the UKF in most of the runs. The advantage seems to be smaller in the case of Gaussian noise compared with the case with VMF noise, but the proposed method is still clearly superior.

## VI. Conclusion

We have presented a novel deterministic sampling scheme for the von Mises–Fisher distribution on the unit hypersphere. Furthermore, we have shown how to derive nonlinear filtering algorithms based on this sampling scheme. MATLAB implementations of these algorithms will be made available as part of `libDirectional` [28].

The proposed algorithm was evaluated with three different system models and Gaussian as well as VMF-distributed noise. It was shown to outperform a spherical version of the UKF in all considered scenarios.

## References


[1] R. Fisher, “Dispersion on a Sphere,” *Proceedings of the Royal Society of London. Series A, Mathematical and Physical Sciences*, vol. 217, no. 1130, pp. 295–305, 1953.

[2] N. I. Fisher, T. Lewis, and B. J. Embleton, *Statistical Analysis of Spherical Data*. Cambridge University Press, 1987.

[3] A. Chiuso and G. Picci, “Visual Tracking of Points as Estimation on the Unit Sphere,” in *The Confluence of Vision and Control*. Springer, 1998, vol. 237, pp. 90–105.

[4] I. Markovic, F. Chaumette, and I. Petrovic, “Moving Object Detection, Tracking and Following Using an Omnidirectional Camera on a Mobile Robot,” in *Proceedings of the 2014 IEEE International Conference on Robotics and Automation (ICRA 2014)*, Hong-Kong, Jun. 2014.

[5] I. Markovic, M. Bukal, J. Cesic, and I. Petrovic, “Direction-only Tracking of Moving Objects on the Unit Sphere via Probabilistic Data Association,” in *Proceedings of the 17th International Conference on Information Fusion (Fusion 2014)*, Salamanca, Spain, Jul. 2014.

[6] Y.-H. Chen, D. Wei, G. Newstadt, M. DeGraef, J. Simmons, and A. Hero, “Parameter Estimation in Spherical Symmetry Groups,” *IEEE Signal Processing Letters*, vol. 22, no. 8, pp. 1152–1155, 2015.

[7] A. Bhalerao and C.-F. Westin, “Hyperspherical von Mises–Fisher Mixture (HvMF) Modelling of High Angular Resolution Diffusion MRI,” in *Medical Image Computing and Computer-Assisted Intervention (MICCAI 2007)*. Springer, 2007, pp. 236–243.

[8] M. Bangert, P. Hennig, and U. Oelfke, “Using an Infinite von Mises–Fisher Mixture Model to Cluster Treatment Beam Directions in External Radiation Therapy,” in *Ninth International Conference on Machine Learning and Applications (ICMLA)*. IEEE, 2010, pp. 746–751.

[9] J. Traa and P. Smaragdis, “Multiple Speaker Tracking With the Factorial von Mises–Fisher Filter,” in *IEEE International Workshop on Machine Learning for Signal Processing (MLSP)*, Sep. 2014.

[10] 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.

[11] G. Kurz, I. Gilitschenski, and U. D. Hanebeck, “Deterministic Approximation of Circular Densities with Symmetric Dirac Mixtures Based on Two Circular Moments,” in *Proceedings of the 17th International Conference on Information Fusion (Fusion 2014)*, Salamanca, Spain, Jul. 2014.

[12] U. D. Hanebeck, M. F. Huber, and V. Klumpp, “Dirac Mixture Approximation of Multivariate Gaussian Densities,” in *Proceedings of the 2009 IEEE Conference on Decision and Control (CDC 2009)*, Shanghai, China, Dec. 2009.

[13] 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.

[14] ——, “Recursive Bayesian Filtering in Circular State Spaces (to appear),” *IEEE Aerospace and Electronic Systems Magazine*, 2016.

[15] I. Gilitschenski, G. Kurz, S. J. Julier, and U. D. Hanebeck, “Unscented Orientation Estimation Based on the Bingham Distribution,” *IEEE Transactions on Automatic Control*, vol. 61, no. 1, pp. 172–177, Jan. 2016.

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

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

[18] A. Banerjee, I. S. Dhillon, J. Ghosh, and S. Sra, “Clustering on the Unit Hypersphere Using von Mises–Fisher Distributions,” *Journal of Machine Learning Research*, vol. 6, pp. 1345–1382, 2005.

[19] S. Sra, “A Short Note on Parameter Approximation for von Mises–Fisher Distributions: And a Fast Implementation of Is(x),” *Computational Statistics*, vol. 27, no. 1, pp. 177–190, 2012.

[20] G. Kurz and U. D. Hanebeck, “Stochastic Sampling of the Hyperspherical von Mises–Fisher Distribution Without Rejection Methods,” in *Proceedings of the IEEE ISIF Workshop on Sensor Data Fusion: Trends, Solutions, Applications (SDF 2015)*, Bonn, Germany, Oct. 2015.

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

[22] G. Kurz, I. Gilitschenski, S. Julier, and U. D. Hanebeck, “Recursive Bingham Filter for Directional Estimation Involving 180 Degree Symmetry,” *Journal of Advances in Information Fusion*, vol. 9, no. 2, pp. 90–105, Dec. 2014.

[23] F. Chatelain and N. Le Bihan, “Von Mises–Fisher Approximation of Multiple Scattering Process on the Hypersphere,” in *2013 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2013, pp. 6461–6465.

[24] I. Gilitschenski, G. Kurz, and U. D. Hanebeck, “Non-Identity Measurement Models for Orientation Estimation Based

on Directional Statistics," in *Proceedings of the 18th International Conference on Information Fusion (Fusion 2015)*, Washington D. C., USA, Jul. 2015.

[25] G. Kurz, I. Gilitschenski, and U. D. Hanebeck, "Nonlinear Measurement Update for Estimation of Angular Systems Based on Circular Distributions," in *Proceedings of the 2014 American Control Conference (ACC 2014)*, Portland, Oregon, USA, Jun. 2014.

[26] E. Lengyel, *Mathematics for 3D Game Programming and Computer Graphics*, 3rd ed. Cengage Learning, 2011.

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

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