

Nonlinear von Mises–Fisher Filtering Based on Isotropic Deterministic Sampling

Kailai Li, Florian Pfaff, and Uwe D. Hanebeck

Abstract—We present a novel deterministic sampling approach for von Mises–Fisher distributions of arbitrary dimensions. Following the idea of the unscented transform, samples of configurable size are drawn isotropically on the hypersphere while preserving the mean resultant vector of the underlying distribution. Based on these samples, a von Mises–Fisher filter is proposed for nonlinear estimation of hyperspherical states. Compared with existing von Mises–Fisher-based filtering schemes, the proposed filter exhibits superior hyperspherical tracking performance.

I. INTRODUCTION

Modeling uncertainties on directional domains requires consideration of the underlying manifold structure [1]. Being inherently defined on the unit hyperspheres¹, the von Mises–Fisher (vMF) distribution has been a popular statistical tool for hyperspherical inference in various application scenarios, such as omnidirectional vision [2], [3], scene understanding [4], protein analysis [5], and remote sensing [6], etc.

For recursive hyperspherical estimation with nonlinear dynamics, vMF-based filters typically rely on sample approximations. For instance, a random sampling approach was introduced for the vMF distribution on \mathbb{S}^2 in [7]. For vMF distributions of arbitrary dimensions, approaches with generally unbounded run time were introduced using rejection sampling in [8], [9]. These approaches were further improved for deterministic run time in [10], with analytic expressions for an odd number of dimensions.

In contrast to random sampling-based approaches, filtering schemes with deterministic samples yield repeatable results and show better accuracy for nonlinear recursive estimation [11], [12]. Deterministic sampling schemes on directional manifolds have been established for various distributions from directional statistics [1]. These methods are all based on moment matching for fitting prior and posterior densities to given propagated or reweighted samples during the prediction and update steps, respectively. Samples are drawn from corresponding continuous densities that preserve moments up to a certain order [13], [14].

Following the idea of the unscented transform (UT) [11], a deterministic sampling approach was proposed for the Bingham distribution in [15] with hyperspherical samples maintaining the moments of the first two orders. Compared

with random sample-based filters, the resulting unscented Bingham filter has shown superior tracking accuracy and run-time efficiency for quaternion-based nonlinear orientation estimation. Similar ideas were also realized on the circular domain. In [16], the wrapped normal and von Mises distributions were approximated by wrapped Dirac mixtures with three components up to the first trigonometric moment for nonlinear angular estimation. Furthermore, a UT-based sampling scheme was proposed for the von Mises–Fisher distribution on $\mathbb{S}^{n-1} \subset \mathbb{R}^n$ in [17]: $2(n-1)+1$ deterministic samples are drawn to maintain the mean resultant vector. These samples are used in the unscented vMF filter (UvMFF) for nonlinear orientation estimation.

To improve the performance of directional estimation with strong nonlinearities induced by complex system and measurement models, much effort has been devoted to generating larger numbers of deterministic samples. For circular densities such as the wrapped normal and von Mises distributions, a method was proposed in [18] to draw five deterministic samples according to a scaling factor to preserve the first and second trigonometric moments. Superposition is used to merge sample sets with different scaling factors into a larger set while maintaining the first two moments. In [19], an arbitrary number of deterministic samples is generated by minimizing a von Mises kernel-based statistical distance between the underlying continuous circular density and its approximating circular Dirac mixture while preserving higher trigonometric moments. Another optimization-based approach was proposed in [20], where circular distributions are approximated by discrete samples via Voronoi quantization. To improve the computational time for large sample sizes, a binary tree-based sampling approach was proposed in [18] via recursive division of the circular domain. The numbers of Dirac components in each interval are set to be proportional to the local probability mass with a post-processing for preserving the first trigonometric moment.

Within the scope of deterministic sampling for hyperspherical continuous distributions, approaches generating configurable numbers of samples were only proposed for the Bingham distribution. In [13], [14], spherical geometry was exploited to establish a tangent space around the mode. There, samples are drawn via a sampling scheme originally proposed for multivariate Gaussian distributions in [12]. An extra moment correction is needed after mapping the samples back to the hypersphere. The sampling scheme itself relies on minimizing a distance measure built upon the localized cumulative distribution (LCD) [21]. By utilizing larger numbers

This work is supported by the German Research Foundation (DFG) under grant HA 3789/16-1. The authors are with the Intelligent Sensor-Actuator-Systems Laboratory (ISAS), Institute for Anthropomatics and Robotics, Karlsruhe Institute of Technology (KIT), Germany. E-Mails: {kailai.li, florian.pfaff, uwe.hanebeck}@kit.edu

¹We use the term “hypersphere” to denote both spherical and hyperspherical domains throughout this paper.

of samples, improved performance was shown in recursive quaternion estimation.

So far, no method for generating deterministic samples of arbitrary sizes for the von Mises–Fisher distribution is available. Suitable optimization-based schemes require sophisticated statistical distances between the underlying continuous densities and their discrete sample approximations. For distributions from directional statistics, such distances typically require costly numerical integration on hyperspheres. As the sample set is restricted to the surface of a hypersphere, special optimization algorithms are required, such as Riemannian optimization [22]. Furthermore, optimizing large sample sets can lead to undesired local minima with poor approximation quality. In summary, these factors prohibit effective deployments of optimization-based schemes to online vMF filtering.

In this paper, we propose an efficient deterministic sampling approach for von Mises–Fisher distributions on hyperspheres. Unlike existing sampling schemes for hyperspherical distributions, our method does not rely on optimization. User-configurable numbers of samples are drawn with an isotropic layout to preserve the mean resultant vector of a vMF distribution. Based on these samples, a novel von Mises–Fisher filter is proposed for nonlinear hyperspherical estimation.

The remainder of this paper is formulated as follows. In Sec. II, fundamentals about the von Mises–Fisher distribution will be introduced. The proposed deterministic sampling approach for the vMF distribution will be explained in Sec. III. After that, we will develop a novel von Mises–Fisher filter in Sec. IV followed by a quantitative evaluation on nonlinear spherical estimation in Sec. V. The work will be concluded in Sec. VI.

II. PRELIMINARIES

A. The von Mises–Fisher Distribution

The von Mises–Fisher distribution $\mathcal{VMF}(\underline{\mathbf{x}}; \underline{\nu}, \kappa)$ is defined on the unit hypersphere $\mathbb{S}^{n-1} \subset \mathbb{R}^n$ and has the following probability density function (PDF)

$$f_{\mathcal{VMF}}(\underline{\mathbf{x}}) = N_n \exp(\kappa \underline{\nu}^\top \underline{\mathbf{x}}), \quad \underline{\mathbf{x}} \in \mathbb{S}^{n-1}, \quad (1)$$

with $\underline{\nu} \in \mathbb{S}^{n-1}$ being the mode of the distribution and $\kappa \geq 0$ controlling the concentration. N_n is the normalization constant and is computed as

$$N_n = \left(\int_{\mathbb{S}^{n-1}} f_{\mathcal{VMF}}(\underline{\mathbf{x}}) d\underline{\mathbf{x}} \right)^{-1} = \frac{\kappa^{n/2-1}}{(2\pi)^{n/2} \mathcal{I}_{n/2-1}(\kappa)}.$$

Here, $\mathcal{I}_{n/2-1}$ denotes the modified Bessel function of the first kind and order $n/2 - 1$. Thus, the vMF distribution is unimodal and has an isotropic dispersion on the hypersphere. Its mean resultant vector is defined as the expectation

$$\underline{\alpha} = \mathbb{E}(\underline{\mathbf{x}}) = \int_{\mathbb{S}^{n-1}} \underline{\mathbf{x}} f_{\mathcal{VMF}}(\underline{\mathbf{x}}) d\underline{\mathbf{x}}, \quad (2)$$

which can be interpreted as the generalization of the trigonometric moment of circular densities on the hypersphere.

Furthermore, we have $\underline{\alpha} = \underline{\nu} \mathcal{A}_n(\kappa)$, with

$$\mathcal{A}_n(\kappa) = \frac{\mathcal{I}_{n/2}(\kappa)}{\mathcal{I}_{n/2-1}(\kappa)}, \quad (3)$$

namely, the direction of the hyperspherical mean is identical to the mode of the vMF distribution. Fitting a vMF distribution via moment matching to the mean resultant vector is equivalent to a maximum likelihood estimation (MLE) [23, Sec. A.1]. Therefore, the parameters of the vMF distribution can be obtained with $\underline{\nu} = \underline{\alpha}/\|\underline{\alpha}\|$ and $\kappa = \mathcal{A}_n^{-1}(\|\underline{\alpha}\|)$. An efficient algorithm for computing the inverse of the Bessel function ratio \mathcal{A}_n^{-1} is given in [24].

B. Hyperspherical Geometry

As shown in (1), the vMF distribution measures uncertainty according to the length of the geodesic curve on the hypersphere, i.e., the arc length between $\underline{\mathbf{x}}$ and $\underline{\nu}$. Any point $\underline{\mathbf{x}} \in \mathbb{S}^{n-1}$ can be mapped to the tangent space $\mathbb{T}_{\underline{\nu}}\mathbb{S}^{n-1}$ at the mode $\underline{\nu}$ via the logarithm map

$$\tilde{\underline{\mathbf{x}}} = \text{Log}_{\underline{\nu}}(\underline{\mathbf{x}}) = (\underline{\mathbf{x}} - \cos(\alpha)\underline{\nu}) \frac{\alpha}{\sin(\alpha)} \in \mathbb{T}_{\underline{\nu}}\mathbb{S}^{n-1}. \quad (4)$$

Here, $\alpha = \arccos(\underline{\nu}^\top \underline{\mathbf{x}})$ is essentially the arc length between the mode $\underline{\nu}$ and the point $\underline{\mathbf{x}}$. It is trivial to prove that $\|\underline{\alpha}\| = \|\text{Log}_{\underline{\nu}}(\underline{\mathbf{x}})\|$. Therefore, the geodesic curve length is preserved in the tangent space under the logarithm map. Conversely, any point $\tilde{\underline{\mathbf{x}}}_t \in \mathbb{T}_{\underline{\nu}}\mathbb{S}^{n-1}$ can be retracted back to the hypersphere via the exponential map

$$\underline{\mathbf{x}} = \text{Exp}_{\underline{\nu}}(\tilde{\underline{\mathbf{x}}}) = \cos(\|\tilde{\underline{\mathbf{x}}}\|)\underline{\nu} + \frac{\sin(\|\tilde{\underline{\mathbf{x}}}\|)}{\|\tilde{\underline{\mathbf{x}}}\|}\tilde{\underline{\mathbf{x}}} \in \mathbb{S}^{n-1}. \quad (5)$$

W.r.t. a local orthonormal basis $\mathcal{B}_{\underline{\nu}}$ at $\underline{\nu}$, the tangent space $\mathbb{T}_{\underline{\nu}}\mathbb{S}^{n-1}$ can be viewed as a $(n-1)$ -ball of radius π , i.e., \mathbb{B}_{π}^{n-1} .

III. ISOTROPIC DETERMINISTIC SAMPLING ON HYPERSPHERES

Given a vMF distribution in the form of (1), we aim to approximate the continuous density with a set of m deterministic samples $\mathbb{X} = \{\underline{\mathbf{x}}_i\}_{i=1}^m \subset \mathbb{S}^{n-1} \subset \mathbb{R}^n$. The samples are equally weighted with weights set to $1/m$ so that each sample contributes equally to the estimation result. Moreover, the samples are to preserve the mean resultant vector of the vMF distribution given in (2). This guarantees the best approximation in the sense of maximum likelihood estimation. When approximating an arbitrary hyperspherical density with a vMF distribution, moment matching to the mean resultant vector also minimizes the information loss in the sense of Kullback–Leibler divergence [25]. We avoid optimization-based sampling schemes for a better trade-off between the approximation accuracy and computational cost.

As the vMF distribution has an isotropic dispersion on the hypersphere, we design the layout of the deterministic samples to be isotropic as well. The whole sample set is composed of one *sun sample*, the sample located at the mode, and the *planet samples*, samples surrounding the sun sample in a circle-wise manner. We configure l hyperspherical circles with uniform intervals of δ w.r.t. the longitude, where each

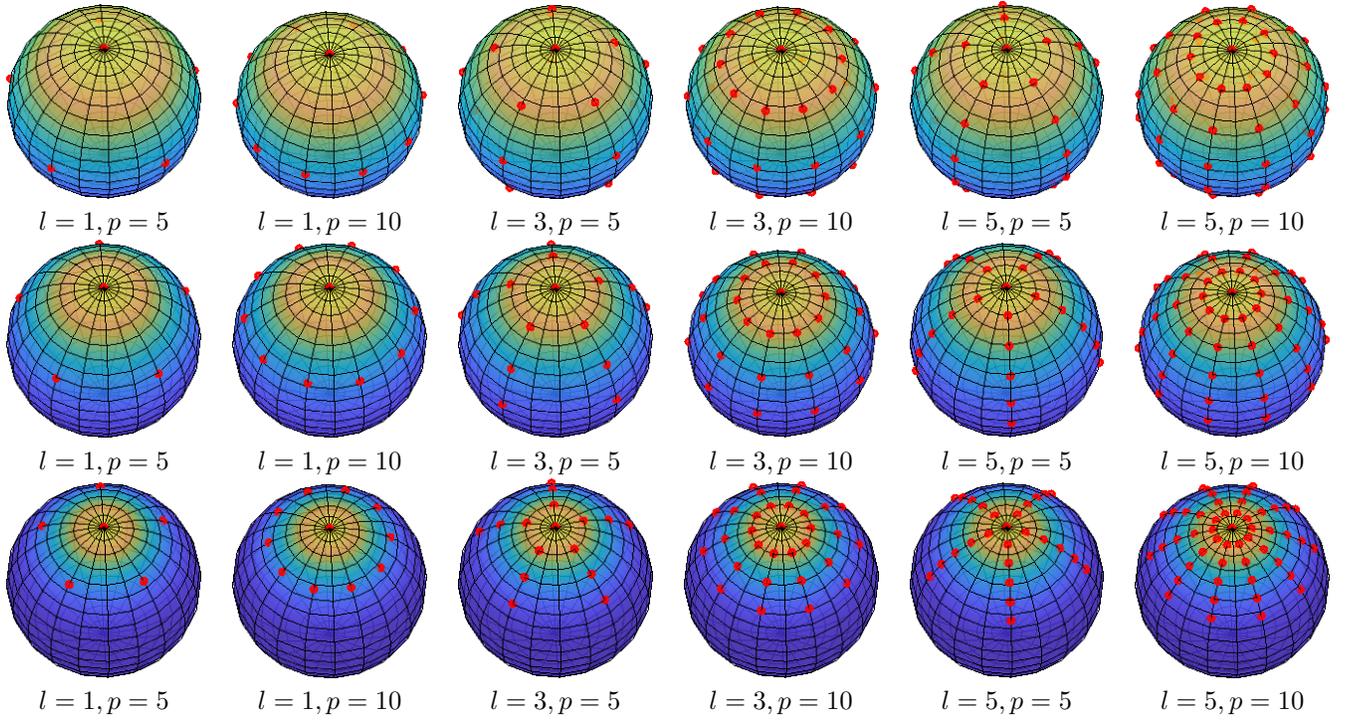


Fig. 1: Isotropic deterministic samples (red dots) of equal weights given by the proposed sampling scheme on \mathbb{S}^2 . The vMF distributions on the same line have the same concentration parameter ($\kappa = \{0.6, 2, 6\}$ from top to bottom).

Algorithm 1: Isotropic Deterministic Sampling

Input: $\mathcal{VMF}(\underline{\mathbf{x}}; \underline{\nu}, \kappa)$, layer number l , per-layer resolution p

Output: deterministic sample set \mathbb{X}

- 1 $\mathbb{X} \leftarrow \underline{\nu}$;
 - 2 $\mathcal{B}_{\underline{\nu}} \leftarrow \text{getBasis}(\underline{\nu})$;
 - 3 $\{\tilde{\sigma}_s\}_{s=1}^p \leftarrow \text{equalPartition}(\mathbb{S}^{n-2}, p)$;
 - 4 $\delta \leftarrow \text{computeInterval}(p, l)$;
 - 5 **for** $r \leftarrow 1$ **to** l **do**
 - 6 **for** $s \leftarrow 1$ **to** p **do**
 - 7 $\underline{\sigma}_s \leftarrow \text{Exp}_{\underline{\nu}}(r\delta \mathcal{B}_{\underline{\nu}} \tilde{\sigma}_s)$;
 - 8 $\mathbb{X} \leftarrow \mathbb{X} \cup \underline{\sigma}_s$;
 - 9 **return** \mathbb{X} ;
-

circle contains p uniformly distributed samples. In total, $m = pl + 1$ deterministic samples are obtained. When undergoing the logarithm map in (4) at mode $\underline{\nu}$, points on the r th circle are mapped to the r th layer of the hypersphere $\mathbb{S}_{r\delta}^{n-2}$ of radius $r\delta$ within the hyperball \mathbb{B}_{π}^{n-1} of $\mathbb{T}_{\underline{\nu}}\mathbb{S}^{n-1}$.

The sampling procedure is detailed in Alg. 1. The sun sample is added first (Alg. 1, line 1). Subsequently, we compute the null space of the mode $\underline{\nu}$ to obtain $\mathcal{B}_{\underline{\nu}} \in \mathbb{R}_{n \times (n-1)}$ as the local basis of the tangent space $\mathbb{T}_{\underline{\nu}}\mathbb{S}^{n-1}$ at $\underline{\nu}$ (Alg. 1, line 2). To guarantee a uniform placement of the samples on the same hyperspherical circle, we apply the equal area partitioning algorithm of [26] on the unit hypersphere \mathbb{S}^{n-2} with the per-layer resolution p (Alg. 1,

line 3). The obtained points $\{\tilde{\sigma}_s\}_{s=1}^p \subset \mathbb{R}^{n-1}$ are of unit length and evenly distributed on \mathbb{S}^{n-2} . When multiplied by the scaling factor $r\delta$ w.r.t. the local basis $\mathcal{B}_{\underline{\nu}}$ and mapped back to \mathbb{S}^{n-1} via the exponential map in (5), they preserve the even placement on the hyperspherical circle surrounding the mode $\underline{\nu}$ (Alg. 1, line 5-8).

The equal area partitioning algorithm is recursive and does not rely on any optimization. It runs efficiently for hyperspheres of arbitrary dimensions and has been applied to multiple scenarios of nonlinear state estimation [27]–[30]. As indicated in [26], points given by the equal partitioning are numerically zero-centered, i.e., $\sum_{s=1}^p \tilde{\sigma}_s = \mathbf{0}$, when p is an even number or an odd number larger than $2(n-2) - 1$. Per-layer resolution values that are odd and smaller than $2(n-2) - 1$ are not considered as the original UvMFF in [17] already exploits $2n - 1$ samples.

As explained in Sec. I, the isotropic samples need to maintain the mean resultant vector of the underlying vMF distribution to guarantee an approximation in the MLE sense. Intuitively, the moment constraint can be satisfied by a proper scaling interval δ due to the isotropic dispersion of the vMF distribution. Given the equally partitioned samples $\tilde{\sigma}_s \in \mathbb{S}^{n-2}$ w.r.t. the local basis $\mathcal{B}_{\underline{\nu}}$, we scale the samples to the r th layer by the factor $r\delta$ and perform the exponential map in (5), namely,

$$\begin{aligned} \underline{\sigma}_{r,s} &= \text{Exp}_{\underline{\nu}}(r\delta \mathcal{B}_{\underline{\nu}} \tilde{\sigma}_s) \\ &= \cos(r\delta) \underline{\nu} + \sin(r\delta) \mathcal{B}_{\underline{\nu}} \tilde{\sigma}_s \in \mathbb{S}^{n-1}. \end{aligned}$$

Thus, the samples' hyperspherical mean $\underline{\mu}$ can be derived as

$$\begin{aligned}\underline{\mu} &= \frac{1}{pl+1} \left(\underline{\nu} + \sum_{r=1}^l \sum_{s=1}^p \underline{\sigma}_{r,s} \right) \\ &= \frac{1}{pl+1} \left(\underline{\nu} + \sum_{r=1}^l \sum_{s=1}^p (\cos(r\delta) \underline{\nu} + \sin(r\delta) \mathcal{B}_{\underline{\nu}} \tilde{\sigma}_s) \right).\end{aligned}$$

Considering that points from the equal area partitioning are zero-centered for typical configurations, we obtain

$$\underline{\mu} = \frac{1}{pl+1} \left(1 + p \sum_{r=1}^l \cos(r\delta) \right) \underline{\nu},$$

which is then imposed to be identical to the mean resultant vector $\underline{\alpha}$ in (2), namely $\underline{\mu} \stackrel{!}{=} \underline{\alpha} = \underline{\nu} \mathcal{A}_n(\kappa)$, with $\mathcal{A}_n(\kappa)$ given in (3). Therefore, the following equation can be derived

$$\sum_{r=1}^l \cos(r\delta) = \frac{(pl+1) \mathcal{A}_n(\kappa) - 1}{p}, \delta \in (0, \pi/l].$$

The finite series on the left-hand side can be simplified with Lagrange's trigonometric identity [31, Sec. 2.4.1.6] and we obtain

$$\frac{\sin((l+1/2)\delta)}{2\sin(\delta/2)} = \frac{(pl+1) \mathcal{A}_n(\kappa) - 1}{p} + \frac{1}{2}$$

for $\delta \in (0, \pi/l]$. The left-hand side can be rewritten as $\pi \mathcal{D}_l(\delta)$, with $\mathcal{D}_l(\delta) = \frac{\sin((l+1/2)\delta)}{2\pi \sin(\delta/2)}$ being the Dirichlet kernel [32] and we finally obtain

$$\mathcal{D}_l(\delta) = \frac{(pl+1) \mathcal{A}_n(\kappa) - 1}{\pi p} + \frac{1}{2\pi}.$$

It is trivial to confirm that the equation above has one single root in the range of $(0, \pi/l]$. By applying typical numerical equation solvers, we can obtain the δ preserving the mean resultant vector of the underlying vMF distribution (Alg. 1, line 4). In Fig. 1, we show a few examples given by the proposed sampling algorithm for different vMF distributions and different configurations.

IV. THE VON MISES–FISHER FILTER

We consider the following nonlinear hyperspherical filtering problem. The system model is given as

$$\underline{\mathbf{x}}_{k+1} = \underline{a}(\underline{\mathbf{x}}_k, \underline{\mathbf{w}}_k), \quad (6)$$

with the hyperspherical state $\underline{\mathbf{x}}_k, \underline{\mathbf{x}}_{k+1} \in \mathbb{S}^{n-1}$, system noise $\underline{\mathbf{w}}_k \in \mathbb{W}$, and hyperspherical transition function $\underline{a}: \mathbb{S}^{n-1} \times \mathbb{W} \rightarrow \mathbb{S}^{n-1}$. The measurement model is given by

$$\underline{\mathbf{z}}_k = \underline{h}(\underline{\mathbf{x}}_k, \underline{\mathbf{v}}_k), \quad (7)$$

with measurement $\underline{\mathbf{z}}_k \in \mathbb{Z}$, measurement noise $\underline{\mathbf{v}}_k \in \mathbb{V}$, and measurement function $\underline{h}: \mathbb{S}^{n-1} \times \mathbb{V} \rightarrow \mathbb{Z}$.

Given an arbitrary nonlinear system model in (6), the following prior density can be derived

$$\begin{aligned}f^P(\underline{\mathbf{x}}_{k+1}) &= \int_{\mathbb{S}^{n-1}} \int_{\mathbb{W}} f(\underline{\mathbf{x}}_{k+1}, \underline{\mathbf{x}}_k, \underline{\mathbf{w}}_k) d\underline{\mathbf{w}}_k d\underline{\mathbf{x}}_k \\ &= \int_{\mathbb{S}^{n-1}} f^e(\underline{\mathbf{x}}_k) \int_{\mathbb{W}} f(\underline{\mathbf{x}}_{k+1}, \underline{\mathbf{w}}_k | \underline{\mathbf{x}}_k) d\underline{\mathbf{w}}_k d\underline{\mathbf{x}}_k \\ &= \int_{\mathbb{S}^{n-1}} f^e(\underline{\mathbf{x}}_k) \int_{\mathbb{W}} \delta(\underline{\mathbf{x}}_{k+1} - \underline{a}(\underline{\mathbf{x}}_k, \underline{\mathbf{w}}_k)) f^w(\underline{\mathbf{w}}_k) d\underline{\mathbf{w}}_k d\underline{\mathbf{x}}_k.\end{aligned}$$

Similar to the sampling-based nonlinear prediction step in [17, Alg. 3], we represent the previous posterior with a Dirac mixture supported by the isotropic deterministic samples from the proposed approach, namely $f^e(\underline{\mathbf{x}}_k) = \sum_{i=1}^m \omega_i \delta(\underline{\mathbf{x}}_k - \underline{x}_{k,i}^e)$. Furthermore, given an arbitrary noise distribution represented by the Dirac mixture $f^w(\underline{\mathbf{w}}_k) = \sum_{j=1}^{m_w} \omega_j^w \delta(\underline{\mathbf{w}}_k - \underline{w}_{k,j}^w)$, the predicted vMF is fitted to a Cartesian product of state and noise samples that are propagated through the transition function, i.e.,

$$f^P(\underline{\mathbf{x}}_{k+1}) \approx \sum_{i=1}^m \sum_{j=1}^{m_w} \omega_i \omega_j^w \delta(\underline{\mathbf{x}}_{k+1} - \underline{x}_{k+1,ij}^P).$$

Here, $\underline{x}_{k+1,ij}^P = \underline{a}(\underline{x}_{k,i}^e, \underline{w}_{k,j}^w)$ denotes the prior samples.

For the update step, we reweight each prior sample $\underline{x}_{k,i}^P$ by fusing its likelihood given the measurement \hat{z}_k via

$$\begin{aligned}f^e(\underline{\mathbf{x}}_k | \hat{z}_k) &\propto f^L(\hat{z}_k | \underline{\mathbf{x}}_k) f^P(\underline{\mathbf{x}}_k) \\ &= \sum_{i=1}^m \omega_i f^L(\hat{z}_k | \underline{x}_{k,i}^P) \delta(\underline{\mathbf{x}}_k - \underline{x}_{k,i}^P).\end{aligned} \quad (8)$$

Based on the reweighted samples, the posterior vMF distribution can be obtained again by means of moment matching.

V. EVALUATION

For highly nonlinear systems, a small sample set is usually not sufficient for describing the corresponding densities. In addition, it often leads to sample degeneration in the measurement update step, especially in the case of low-noise observations. In contrast to methods using only a few samples, such as one generating $2n - 3$ samples in [17], the proposed deterministic sampling approach gives a more accurate representation of the transition densities and likelihoods while satisfying the condition of unscented transform. Thus, an improved performance of nonlinear vMF filtering can be expected.

We consider the following simulation scenario in the evaluation. The system state $\underline{\mathbf{x}}_k = [\mathbf{x}_k^1, \mathbf{x}_k^2, \mathbf{x}_k^3]^\top \in \mathbb{S}^2$ is a spatial orientation vector that undergoes uncertain rotations given by the following unit quaternion

$$\underline{\mathbf{w}}_k = \left[\cos\left(\frac{\text{acos}(\mathbf{x}_k^3) + \theta_k}{2}\right), \sin\left(\frac{\text{acos}(\mathbf{x}_k^3) + \theta_k}{2}\right) \underline{u}_k^\top \right]^\top. \quad (9)$$

Given the uncertain input $\underline{\mathbf{w}}_k$ above, the system model is formulated as $\underline{\mathbf{x}}_{k+1} = \underline{\mathbf{w}}_k \otimes \underline{\mathbf{x}}_k \otimes \underline{\mathbf{w}}_k^*$, with \otimes being the Hamilton product and $\underline{\mathbf{w}}_k^*$ the conjugate of $\underline{\mathbf{w}}_k$ [14]. The rotation axis \underline{u}_k of quaternion (9) is fixed. The rotation angle, however, is synthesized by the polar angle of the current

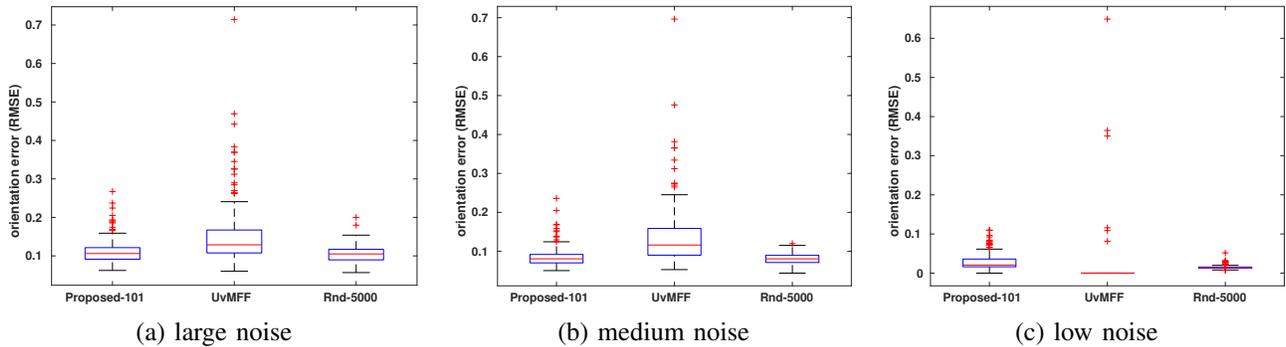


Fig. 2: Comparison of the proposed sampling scheme with UT-only and random samplings for nonlinear spherical tracking.

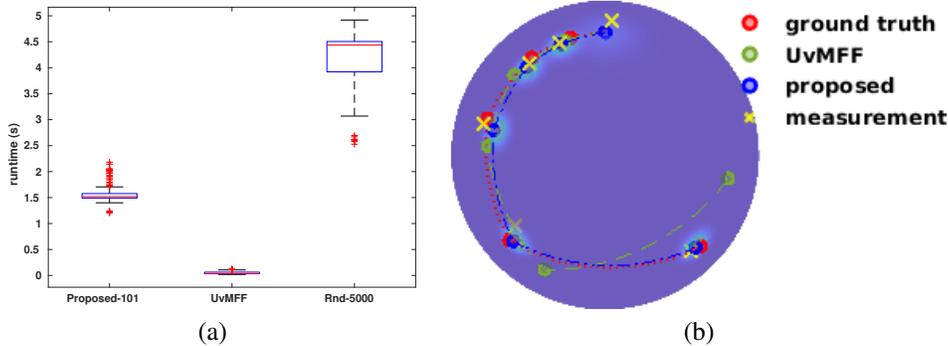


Fig. 3: (a) Run time with large measurement noise. (b) Typical run of spherical tracking with medium measurement noise.

state and an uncertain angular input θ_k . We further assume that the angular input is von Mises-distributed, namely, $\theta_k \sim \mathcal{VM}(\mu_\theta, \kappa_\theta)$, with μ_θ and κ_θ being the mean and concentration of the von Mises distribution. Note that such a system model is highly nonlinear and induces a state-dependent transition density. The measurement model is the observation of the state \mathbf{x}_k with a zero-mean Gaussian-distributed noise following

$$\mathbf{z}_k = \mathbf{x}_k + \mathbf{v}_k, \mathbf{v}_k \sim \mathcal{N}(\mathbf{0}, \Sigma_v). \quad (10)$$

Here, Σ_v denotes the covariance matrix of the Gaussian distribution. The additive noise term above then induces a likelihood function in the form of $f^L(\mathbf{z}_k | \mathbf{x}_k) = f_{\mathcal{N}}(\mathbf{z}_k - \mathbf{x}_k)$, which is then deployed for updating the prior samples via (8).

We compare the vMF filter using the proposed deterministic sampling scheme with the unscented vMF Filter (UvMFF) in [17] and a vMF filter using random sampling. The uncertain angular input θ_k in (9) follows a von Mises distribution of $\mu_\theta = \pi/6$ and $\kappa_\theta = 50$. The fixed rotation axis is $\underline{u}_k = [1/\sqrt{3}, 1/\sqrt{3}, 1/\sqrt{3}]^\top$. The initial state of all the filters are set to be a vMF distribution with $\underline{\nu} = [0, 0, 1]^\top$ and $\kappa = 50$. To further evaluate the filtering performances, three levels of measurement noise in (10) are simulated with $\Sigma_v = \lambda \mathbf{I}_{3 \times 3}$, $\lambda = \{10^{-2}, 5 \cdot 10^{-3}, 10^{-4}\}$. Here, a smaller λ indicates a more peaky likelihood function, therefore strengthening the effect of potential sample degeneration. The proposed vMF filter exploits $pl + 1 = 20 \cdot 5 + 1 = 101$ deterministic samples, whereas the stochastic modeling-based one relies on 5000 random samples. For spherical filtering ($n = 3$), the originally proposed UvMFF has a fixed

sample size of $2(n - 1) + 1 = 5$. In all three test cases, 200 Monte Carlo runs are performed with 10 time steps each.

The estimation error is quantified as the arc length between the estimate $\hat{\mathbf{x}}_k$ and the ground truth \mathbf{x}_k , i.e., $e = \text{acos}(\hat{\mathbf{x}}_k^\top \mathbf{x}_k)$. As shown in Fig. 2 with the boxplot of MATLAB, the proposed deterministic sampling scheme gives better estimation accuracy than the UvMFF under all three noise levels. In particular, the UvMFF almost totally fails under low measurement noise level (no box is plotted) due to the peaky likelihood function. As several outliers in Fig 2-(c) show, it can rarely track the orientation state and induces large estimation error. Moreover, 101 deterministic samples given by the proposed sampling scheme achieves similar tracking accuracy as 5000 random samples.

As further justified in Fig. 3-(a) w.r.t. computational cost for filtering under the large measurement noise level, the proposed sampling scheme provides a good trade-off between run-time efficiency and accuracy. Fig. 3-(b) further shows one example run of the evaluated spherical tracking under medium measurement noise level. In addition to the spherical location estimates, the posterior distribution from the proposed filter is also plotted at each step for justifying the estimated uncertainty. Compared with the original UT-based method [17], the proposed sampling scheme considerably improves the performance of nonlinear vMF filtering.

VI. CONCLUSION

In this work, we propose a novel deterministic sampling scheme with user-configurable sample sizes for von Mises–Fisher distributions of arbitrary dimensions. It provides

equally weighted samples located isotropically on the hypersphere while preserving the mean resultant vector of the underlying vMF distribution. When deployed for nonlinear vMF filtering, the proposed scheme provides a more accurate representation of the transition densities and likelihoods, improving estimation performance on hyperspheres.

The proposed method could be enhanced in two ways: (1) For non-identity measurement models with a known likelihood function, the progressive update method [33], [34] could be used in conjunction with the proposed sampling scheme for further improvement on filtering performance. (2) Instead of relying on single vMF distributions, the proposed sampling scheme could be extended to mixtures for capturing more complex real-world distributions [35]–[37].

REFERENCES

- [1] K. V. Mardia and P. E. Jupp, *Directional Statistics*. John Wiley & Sons, 2009, vol. 494.
- [2] I. Marković, F. Chaumette, and I. Petrović, “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)*. IEEE, 2014, pp. 5630–5635.
- [3] H. Guan and W. A. Smith, “Structure-from-Motion in Spherical Video Using the von Mises-Fisher Distribution,” *IEEE Transactions on Image Processing*, vol. 26, no. 2, pp. 711–723, 2016.
- [4] J. Straub, J. Chang, O. Freifeld, and J. Fisher III, “A Dirichlet Process Mixture Model for Spherical Data,” in *Artificial Intelligence and Statistics*, 2015, pp. 930–938.
- [5] P. D. Hoff, “Simulation of the Matrix Bingham–von Mises–Fisher Distribution, with Applications to Multivariate and Relational Data,” *Journal of Computational and Graphical Statistics*, vol. 18, no. 2, pp. 438–456, 2009.
- [6] D. Lungu and O. Ersoy, “Unsupervised Classification of Hyperspectral Images on Spherical Manifolds,” in *Industrial Conference on Data Mining*. Springer, 2011, pp. 134–146.
- [7] W. Jakob, “Numerically Stable Sampling of the von Mises-Fisher Distribution on S^2 (and Other Tricks),” Interactive Geometry Lab, ETH Zürich, Tech. Rep., 2012.
- [8] G. Ulrich, “Computer Generation of Distributions on the M-Sphere,” *Journal of the Royal Statistical Society: Series C (Applied Statistics)*, vol. 33, no. 2, pp. 158–163, 1984.
- [9] A. T. Wood, “Simulation of the von Mises Fisher Distribution,” *Communications in Statistics-Simulation and Computation*, vol. 23, no. 1, pp. 157–164, 1994.
- [10] 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.
- [11] S. J. Julier and J. K. Uhlmann, “Unscented Filtering and Nonlinear Estimation,” *Proceedings of the IEEE*, vol. 92, no. 3, pp. 401–422, 2004.
- [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] K. Li, D. Frisch, B. Noack, and U. D. Hanebeck, “Geometry-Driven Deterministic Sampling for Nonlinear Bingham Filtering,” in *Proceedings of the 2019 European Control Conference (ECC 2019)*, Naples, Italy, June 2019.
- [14] K. Li, F. Pfaff, and U. D. Hanebeck, “Hyperspherical Deterministic Sampling Based on Riemannian Geometry for Improved Nonlinear Bingham Filtering,” in *Proceedings of the 22nd International Conference on Information Fusion (Fusion 2019)*, Ottawa, Canada, July 2019.
- [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] 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, June 2013.
- [17] —, “Unscented von Mises–Fisher Filtering,” *IEEE Signal Processing Letters*, vol. 23, no. 4, pp. 463–467, Apr. 2016.
- [18] G. Kurz, I. Gilitschenski, R. Y. Siegwart, and U. D. Hanebeck, “Methods for Deterministic Approximation of Circular Densities,” *Journal of Advances in Information Fusion*, vol. 11, no. 2, pp. 138–156, Dec. 2016.
- [19] U. D. Hanebeck and A. Lindquist, “Moment-based Dirac Mixture Approximation of Circular Densities,” in *Proceedings of the 19th IFAC World Congress (IFAC 2014)*, Cape Town, South Africa, Aug. 2014.
- [20] I. Gilitschenski, G. Kurz, U. D. Hanebeck, and R. Siegwart, “Optimal Quantization of Circular Distributions,” in *Proceedings of the 19th International Conference on Information Fusion (Fusion 2016)*, Heidelberg, Germany, July 2016.
- [21] U. D. Hanebeck and V. Klumpp, “Localized Cumulative Distributions and a Multivariate Generalization of the Cramér–von Mises Distance,” in *Proceedings of the 2008 IEEE International Conference on Multisensor Fusion and Integration for Intelligent Systems (MFI 2008)*, Seoul, Republic of Korea, Aug. 2008, pp. 33–39.
- [22] N. Boumal, B. Mishra, P.-A. Absil, and R. Sepulchre, “Manopt, a Matlab Toolbox for Optimization on Manifolds,” *Journal of Machine Learning Research*, vol. 15, pp. 1455–1459, 2014.
- [23] 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, no. Sep, pp. 1345–1382, 2005.
- [24] 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.
- [25] G. Kurz, F. Pfaff, and U. D. Hanebeck, “Kullback-Leibler Divergence and Moment Matching for Hyperspherical Probability Distributions,” in *Proceedings of the 19th International Conference on Information Fusion (Fusion 2016)*, Heidelberg, Germany, July 2016.
- [26] P. Leopardi, “A Partition of the Unit Sphere Into Regions of Equal Area and Small Diameter,” *Electronic Transactions on Numerical Analysis*, vol. 25, no. 12, pp. 309–327, 2006.
- [27] K. Li, F. Pfaff, and U. D. Hanebeck, “Grid-Based Quaternion Filter for $SO(3)$ Estimation,” in *Proceedings of the 2020 European Control Conference (ECC 2020)*, Saint Petersburg, Russia, May 2020.
- [28] F. Pfaff, K. Li, and U. D. Hanebeck, “The Spherical Grid Filter for Nonlinear Estimation on the Unit Sphere,” in *Proceedings of the 1st Virtual IFAC World Congress (IFAC-V 2020)*, July 2020.
- [29] —, “A Hyperhemispherical Grid Filter for Orientation Estimation,” in *Proceedings of the 23rd International Conference on Information Fusion (Fusion 2020)*, Virtual, July 2020.
- [30] G. Kurz and U. D. Hanebeck, “Linear Regression Kalman Filtering Based on Hyperspherical Deterministic Sampling,” in *Proceedings of the 56th IEEE Conference on Decision and Control (CDC 2017)*, Melbourne, Australia, Dec. 2017.
- [31] A. Jeffrey and H. H. Dai, *Handbook of Mathematical Formulas and Integrals*. Elsevier, 2008.
- [32] A. M. Bruckner, J. B. Bruckner, and B. S. Thomson, *Real Analysis*. Prentice Hall, Upper Saddle River, NJ, 1997.
- [33] U. D. Hanebeck, K. Briechle, and A. Rauh, “Progressive Bayes: A New Framework for Nonlinear State Estimation,” in *Proceedings of SPIE, AeroSense Symposium*, vol. 5099, Orlando, Florida, USA, May 2003, pp. 256 – 267.
- [34] K. Li, D. Frisch, S. Radtke, B. Noack, and U. D. Hanebeck, “Wavefront Orientation Estimation Based on Progressive Bingham Filtering,” in *Proceedings of the IEEE ISIF Workshop on Sensor Data Fusion: Trends, Solutions, Applications (SDF 2018)*, Oct. 2018.
- [35] J. Straub, T. Campbell, J. P. How, and J. W. Fisher, “Small-Variance Nonparametric Clustering on the Hypersphere,” in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2015, pp. 334–342.
- [36] H. Möls, K. Li, and U. D. Hanebeck, “Highly Parallelizable Plane Extraction for Organized Point Clouds Using Spherical Convex Hulls,” in *Proceedings of the 2020 IEEE International Conference on Robotics and Automation (ICRA 2020)*, Paris, France, May 2020.
- [37] K. Li, F. Pfaff, and U. D. Hanebeck, “Hyperspherical Unscented Particle Filter for Nonlinear Orientation Estimation,” in *Proceedings of the 1st Virtual IFAC World Congress (IFAC-V 2020)*, July 2020.