

Geometry-driven Deterministic Sampling for Nonlinear Bingham Filtering

Kailai Li, Daniel Frisch, Benjamin Noack, and Uwe D. Hanebeck

Abstract— We propose a geometry-driven deterministic sampling method for Bingham distributions in arbitrary dimensions. With flexibly adjustable sampling sizes, the novel scheme can generate equally weighted samples that satisfy requirements of the unscented transform and approximate higher-order shape information of the Bingham distribution. By leveraging retraction techniques from Riemannian geometry, the sigma points are constrained to preserve the second-order moment. Meanwhile, samples in each principal direction are located in a way that minimizes a distance measure between the on-tangent-plane Dirac mixtures and the underlying on-manifold density. For that, the modified Cramér–von Mises distance based on the localized cumulative distribution (LCD) is employed. We further integrate the proposed approach into a quaternion-based orientation estimation framework. Compared to the existing unscented sampling approach drawing only fixed and limited numbers of sigma points, simulation results show that the proposed scheme enables better accuracy and robustness for nonlinear Bingham filtering.

I. INTRODUCTION

Reliable orientation filtering is crucial for state estimation and control problems that appear in various application scenarios, e.g., point registration [1], multilateration [2], pose estimation as well as robotic perception and manipulation [3]–[8]. Due to the underlying nonlinearity and periodicity, however, recursive estimators for robust and accurate orientation estimation cannot be trivially developed. Mathematically speaking, the planar and spatial orientations belong to the special orthogonal groups $SO(2)$ and $SO(3)$, respectively. Conventional stochastic filtering algorithms, e.g., the well-known extended Kalman filter (EKF) [9] or the unscented Kalman filter (UKF) [10], rely on local linearization of the nonlinearity with noise terms assumed to be Gaussian-distributed. However, such approximations are problematic under high nonlinearity and strong system noise as they lack the consideration of the underlying nonlinear group structure. For practical use, sensors providing higher-order motion information are usually needed (e.g., IMU), and sensor fusion techniques are required [11].

Moreover, orientations can be parameterized in different ways. Popular options, such as the Euler angles, suffer from gimbal lock. The singularity can be resolved via over-parameterization, e.g., by using the well-known rotation matrices. However, they can sacrifice numerical instabilities due to the large degree of redundancy (3×3 elements for 3-DoF orientations). In contrast, unit quaternions provide a natural way to parameterize orientations without singularity, and only bring one degree of redundancy. In this paper,

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

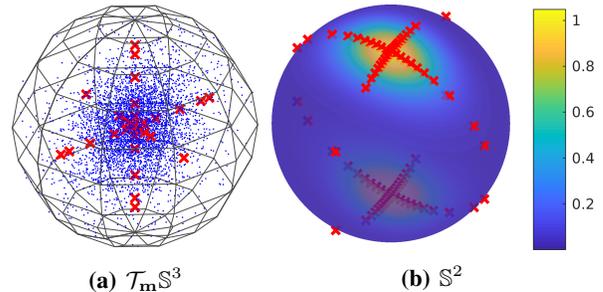


Fig. 1: Deterministic samples (red ‘x’) from the proposed approach for \mathbb{S}^3 (shown on $\mathcal{T}_m \mathbb{S}^3$ via logarithm map) and \mathbb{S}^2 . Small blue dots are randomly drawn to show the density.

quaternions on \mathbb{S}^1 (unit circle) and \mathbb{S}^3 (unit hypersphere) are employed to represent planar and spatial orientations.

In [4], [5], [12], the Bingham distribution [13] was used for stochastically modeling uncertain unit quaternions, and applied further for orientation estimation. As the distribution is inherently defined on \mathbb{S}^n , Bingham-based filtering algorithms no longer require local linearization of the manifold [14]. Corresponding nonlinear Bingham filters are typically sampling-approximation-based, in which the distribution is estimated via moment matching (for the first two moments). Compared to Monte Carlo-based filtering frameworks, deterministic schemes can guarantee reproducible results and are more efficient.

In [15], a deterministic sampling approach for quaternion Bingham distributions has been proposed following the idea of the unscented transform (UT), and enables efficient recursive orientation estimation. However, the sigma points are limited to be of fixed size (on \mathbb{S}^n only $2n + 1$ unduplicated samples are generated¹), and can only approximate moments up to the second order. This is sufficient for moment matching-based parameter fitting of the distribution, but lacks consideration of the higher-order shape information. Nonlinear Bingham filtering approaches with small sample sizes can easily suffer from degeneration issue in a sampling-reweighting-approximation filtering workflow [2].

There has been effort to adapt the UT-based sampling scheme to approximate higher-order moments for Gaussian distributions, where certain statistical distance metrics are minimized based on Dirac mixtures of deterministic samples. The samples are either located on the principal axes [16] or generally in the domain of definition [17]. In [18], a sample reduction technique was introduced based on the localized cumulative distribution (LCD), enabling Dirac mixture approximation for non-continuous densities. However,

¹Here, antipodally symmetric sigma points are considered as “duplicated” as they represent the same orientation.

the aforementioned approaches are all proposed for linear domains and cannot be trivially extended to distributions on nonlinear manifolds, such as a Bingham on the hypersphere. Though in [19], a sampling approach was introduced for approximating the von Mises distribution on the circular domain, it specifically relies on the trigonometric moment and cannot be generalized for hyperspheres. Besides, some of these approaches do not maintain the first two moments [20]. However, it is still our goal to have the sampling scheme follow the idea of the unscented transform to ensure that no information gets lost for parameter fitting.

In this paper, a novel deterministic sampling scheme for the Bingham distribution is proposed (examples shown in Fig. 1). Taking ideas from Riemannian geometry, the geometric structure of the Bingham distribution on \mathbb{S}^n can be depicted on the tangent space at the mode. Optimization-based Dirac mixture approximation is then performed in each principal direction under constraints deduced by the unscented transform via retractions. The on-tangent-plane optimization is performed to minimize an LCD-based Cramér-von Mises distance [18] between the Dirac mixtures and the underlying on-manifold density. Our contributions are:

- A generic deterministic sampling scheme is proposed for Bingham distributions in arbitrary dimensions. The samples are located in principal directions and preserve the first two moments with consideration of approximating higher-order moments of the Bingham density.
- The scheme allows flexibly given sizes of samples, which are equally weighted for uniform contributions. The accuracy and robustness of nonlinear Bingham filtering can thus be improved.

The remainder of the paper is as follows. Preliminaries about quaternion representation, Bingham distribution, and spherical geometry are introduced in Sec. II. The general geometry-driven sampling scheme is proposed in Sec. III. In Sec. IV, geometric interpretation of the unscented transform is derived, based on which the LCD-driven shape approximation technique is introduced. The novel sampling scheme is then evaluated based on simulations for nonlinear Bingham filtering in Sec. V. The work is finally concluded in Sec. VI.

II. PRELIMINARIES

A. Unit Quaternions and Orientation Parameterization

Unit quaternions [2] can be seen as a reparameterization of the axis-angle representation for spatial orientations. A rotation with angle θ around the axis of a unit vector $\underline{u} \in \mathbb{R}^3$ can be represented by the quaternion as

$$\mathbf{x} = [\cos(\theta/2), \underline{u}^\top \sin(\theta/2)]^\top \in \mathbb{R}^4, \quad (1)$$

according to which any $\underline{v} \in \mathbb{R}^3$ can be rotated to \underline{v}' via

$$\underline{v}' = \mathbf{x} \otimes \underline{v} \otimes \mathbf{x}^*. \quad (2)$$

Here, \otimes denotes the Hamilton product [21] and $\mathbf{x}^* = \text{diag}(1, -1, -1, -1) \mathbf{x}$ is the conjugate of \mathbf{x} . As the norm of a quaternion is defined as $\sqrt{\mathbf{x} \otimes \mathbf{x}^*}$, the vector in (1) is thus of unit length and belongs to the unit hypersphere $\mathbb{S}^3 = \{\mathbf{x} \in \mathbb{R}^4 \mid \|\mathbf{x}\| = 1\}$. Moreover, as (2) indicates, two antipodally symmetric unit quaternions on the hypersphere, e.g., \mathbf{x} and $-\mathbf{x}$, denote the same rotation.

B. Bingham Distribution

As introduced in [4], [5], [15], [22], the Bingham distribution defined on $\mathbb{S}^n \subset \mathbb{R}^{n+1}$ has the following form

$$f_B(\mathbf{x}; \mathbf{M}, \mathbf{Z}) = \frac{1}{N(\mathbf{Z})} \exp(\mathbf{x}^\top \mathbf{M} \mathbf{Z} \mathbf{M}^\top \mathbf{x}), \quad (3)$$

with the diagonal matrix $\mathbf{Z} = \text{diag}(z_1, \dots, z_n, z_0)$ determining the concentration as well as the normalization constant $N(\mathbf{Z})$, and the real orthogonal matrix $\mathbf{M} \in \mathbb{R}^{(n+1) \times (n+1)}$ the orientation on \mathbb{S}^n . The matrices \mathbf{Z} and \mathbf{M} can be generated via eigendecomposition of a negative semi-definite matrix \mathbf{C}_B . For convenience, the concentration elements are usually readjusted in the ascending order, namely $z_1 \leq \dots \leq z_n \leq z_0 \leq 0$, and column vectors of \mathbf{M} are reordered correspondingly. The mode of the Bingham is given as the column vector of \mathbf{M} associated with the largest value of matrix \mathbf{Z} . Intuitively, a Bingham distribution is derived by conditioning a zero-mean Gaussian distribution in \mathbb{R}^{n+1} on the sphere \mathbb{S}^n followed by the re-normalization. Thus, its first-order moment is zero and the second-order moment, namely the covariance, is

$$\text{cov}_B = \mathbf{M} \cdot \text{diag}(d_1, \dots, d_n, d_0) \cdot \mathbf{M}^\top, \quad (4)$$

with $d_i = \frac{\partial N(\mathbf{Z})}{\partial z_i} / N(\mathbf{Z})$, $i = 0, \dots, n$ denoting each element of the diagonal matrix ($\sum_{i=0}^n d_i = 1$). The distribution is thus able to model the antipodally symmetric uncertainty of unit quaternions as $f_B(\mathbf{x}) = f_B(-\mathbf{x})$ denote the same density value.

C. Spherical Geometry

As the sphere \mathbb{S}^n is a compact Riemannian manifold with constant unit curvature, the local geometric structure can be depicted on the tangent plane [23]. Given $\mathbf{x}, \mathbf{y} \in \mathbb{S}^n \subset \mathbb{R}^{n+1}$, \mathbf{y} can be mapped onto the tangent plane at \mathbf{x} , denoted as $\mathcal{T}_{\mathbf{x}} \mathbb{S}^n$, via the logarithm map

$$\mathbf{y}_t = \text{Log}_{\mathbf{x}}(\mathbf{y}) = (\mathbf{y} - \cos(\alpha) \mathbf{x}) \frac{\alpha}{\sin(\alpha)}, \quad (5)$$

with $\alpha = \arccos(\mathbf{x} \bullet \mathbf{y})$ being the arc length between \mathbf{x} and \mathbf{y} . The inverse operation, namely the exponential map, retracts $\mathbf{y}_t \in \mathcal{T}_{\mathbf{x}} \mathbb{S}^n$ back to \mathbb{S}^n via

$$\mathbf{y} = \text{Exp}_{\mathbf{x}}(\mathbf{y}_t) = \cos(\|\mathbf{y}_t\|) \mathbf{x} + \frac{\sin(\|\mathbf{y}_t\|)}{\|\mathbf{y}_t\|} \mathbf{y}_t. \quad (6)$$

The distance metric denoted as geodesic between $\mathbf{x}, \mathbf{y} \in \mathbb{S}^n$ is essentially the arc length α , which is preserved by the logarithm map, i.e., $d(\mathbf{x}, \mathbf{y}) = \alpha = \|\text{Log}_{\mathbf{x}}(\mathbf{y})\|$. Based on spherical geometry, we illustrate the geometric structure of the unit quaternion manifold as follows.

Remark II.1 *Geometric structure of unit quaternion manifold representing spatial rotations. When tangent space is shifted between the identity $\mathbf{1} = [1, 0, 0, 0]^\top$ and the Bingham mode \mathbf{m} , the following rules hold*

$$\begin{aligned} \mathbf{x}_t &= \text{Log}_{\mathbf{m}}(\mathbf{x}) = \mathbf{m} \otimes \text{Log}_{\mathbf{1}}(\mathbf{m}^{-1} \otimes \mathbf{x}) \in \mathcal{T}_{\mathbf{m}} \mathbb{S}^3, \\ \mathbf{x} &= \text{Exp}_{\mathbf{m}}(\mathbf{x}_t) = \mathbf{m} \otimes \text{Exp}_{\mathbf{1}}(\mathbf{m}^{-1} \otimes \mathbf{x}_t) \in \mathbb{S}^3. \end{aligned}$$

The logarithm map at identity $\mathbf{1}$ can be derived according to (5) as $\text{Log}_{\mathbf{1}}(\mathbf{x}) = [0, (\theta/2) \underline{u}^\top]^\top$. As \underline{u} is of unit length

and $\theta \in [0, \pi]$, points mapped to the tangent space of \mathbb{S}^3 are restricted in a ball of radius $\pi/2$. For arbitrary unit quaternions \mathbf{x}, \mathbf{y} , their Hamilton product can be represented as ordinary matrix multiplication [6], [7], namely for $\mathbf{x} = [x_1, x_2, x_3, x_4]^\top \in \mathbb{S}^3$, we have $\mathbf{x} \otimes \mathbf{y} = \mathbf{Q} \mathbf{y}$, with

$$\mathbf{Q} = \begin{bmatrix} x_1 & -x_2 & -x_3 & -x_4 \\ x_2 & x_1 & -x_4 & x_3 \\ x_3 & x_4 & x_1 & -x_2 \\ x_4 & -x_3 & x_2 & x_1 \end{bmatrix} := [\mathbf{e}_0, \mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3].$$

It can be confirmed that $\mathbf{Q} \in \text{SO}(4)$, i.e., the 4-dimensional rotation group, as $\mathbf{Q} \mathbf{Q}^\top = \mathbf{Q}^\top \mathbf{Q} = \mathbf{I} \in \mathbb{R}^{4 \times 4}$ and $\det \mathbf{Q} = 1$. Furthermore, \mathbf{Q} essentially depicts the local structure of \mathbb{S}^3 at \mathbf{x} . Here, the first column vector \mathbf{e}_0 locates the tangent space $\mathcal{T}_{\mathbf{x}} \mathbb{S}^3$, and $\{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$ is one orthonormal basis that can be used as the local coordinate system [23] of $\mathcal{T}_{\mathbf{x}} \mathbb{S}^3$. In other words, we have $\mathcal{T}_{\mathbf{x}} \mathbb{S}^3 = \mathcal{T}_{\mathbf{e}_0} \mathbb{S}^3 = \text{span}(\{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\})$.

III. GEOMETRY-DRIVEN SAMPLING SCHEME FOR BINGHAM DISTRIBUTIONS

A. Spherical Sampling Scheme

Unlike the deterministic sampling approach given in [15], we propose a geometry-driven scheme for drawing samples that satisfy requirements of the unscented transform on Bingham distributions in arbitrary dimensions. The idea of the unscented transform is to use the so-called sigma points to represent the distribution while preserving its first two moments. As the Bingham is antipodally symmetric, the sigma points can be drawn by first sampling on the hemisphere associated with one mode to preserve the second moment, then doubling with the antipodes. By doing so, the sigma points can have zero mean.

For simplifying derivations below, we denote the matrix \mathbf{M} in (3) columnwise as $\mathbf{M}_B = [\mathbf{m}_1, \dots, \mathbf{m}_n, \mathbf{m}]$, with the last column \mathbf{m} being the mode as introduced in Sec. II-B. Since \mathbf{M}_B is real orthogonal, the set $\{\mathbf{m}_i\}_{i=1:n}$ inherently provides an orthonormal basis of the tangent plane $\mathcal{T}_{\mathbf{m}} \mathbb{S}^n$, such that $\mathcal{T}_{\mathbf{m}} \mathbb{S}^n = \text{span}(\{\mathbf{m}_i\}_{i=1:n})$. The proposed sampling scheme is shown in Alg. 1 and explained as follows:

- 1) locate center sigma point to mode, i.e., $\sigma_0 = \mathbf{m}$,
- 2) generate scaling factors \underline{r}_i on each axis i to localize sigma points in a way that preserves the first two moments and adapts to higher-order shape information (line 3, Alg. 1),
- 3) locate two mirrored sigma points on each principal direction \mathbf{m}_i according to each scaling factor $\underline{r}_{i,j}$, such that $\sigma_{t,ij}^\pm = \pm \underline{r}_{i,j} \mathbf{m}_i \in \mathcal{T}_{\mathbf{m}} \mathbb{S}^n$, and
- 4) perform retractions from the tangent plane back to the manifold, i.e., $\sigma_{ij}^\pm = \mathcal{R}_{\mathbf{m}}(\sigma_{t,ij}^\pm) \in \mathbb{S}^n$.

The scaling factors are given as a set of vectors $\{\underline{r}_i\}_{i=1:n}$ for each principal direction \mathbf{m}_i , and elements for each vector \underline{r}_i belong to a range Ω_i determined by the retraction, namely

$$\underline{r}_{i,j} \in \Omega_i, \quad j = 1, \dots, l. \quad (7)$$

Here, each vector \underline{r}_i indicates $2 \times l$ locations on the principal direction \mathbf{m}_i (including the mirrored ones). As columns of \mathbf{M}_B are of unit length, we have $\|\sigma_{t,ij}^\pm\| = \underline{r}_{i,j}$. Derivation of each scaling vector \underline{r}_i given size l for equally weighted samples will be introduced in Sec. IV.

Algorithm 1 Spherical Sampling Scheme

```

procedure sphericalSampling( $f_B, l$ )
1:  $\{\mathbf{m}_1, \dots, \mathbf{m}_n, \mathbf{m}\} \leftarrow \text{extractColumns}(\mathbf{M}_B)$ ;
2:  $\sigma_0 \leftarrow \mathbf{m}$ ;
3:  $\{\underline{r}_{i,j} \in \Omega_i\}_{i=1:n, j=1:l} \leftarrow \text{getScalingFactors}(f_B, l)$ ;
4: for  $i = 1, \dots, n$  do
5:   for  $j = 1, \dots, l$  do
6:      $\sigma_{t,ij}^\pm \leftarrow \pm \underline{r}_{i,j} \mathbf{m}_i$ ;
7:      $\sigma_{ij}^\pm \leftarrow \mathcal{R}_{\mathbf{m}}(\sigma_{t,ij}^\pm)$ ;
8:   end for
9: end for
10: return  $\{\sigma_0, \{\sigma_{ij}^\pm\}_{i=1:n, j=1:l}\}$ 
end procedure

```

B. Spherical Retraction

A retraction is employed to preserve the principal direction when mapping sigma points from the tangent plane back to the manifold. We define the operator $\mathcal{R}_{\mathbf{m}} : \mathcal{T}_{\mathbf{m}} \mathbb{S}^n \mapsto \mathbb{S}^n$ for retractions at the mode of the Bingham distribution. The general introductions about retraction in Riemannian geometry can be found in [23]–[25]. We hereby give three options based on the spherical geometry (shown in Fig. 2): exponential, orthographic, and gnomonic retraction.

1) *Exponential Retraction*: The exponential map in the context of Riemannian geometry induces a natural retraction for the hypersphere [23]. Given one sigma pair $\sigma_{t,ij}^\pm = \pm \underline{r}_{i,j} \mathbf{m}_i \in \mathcal{T}_{\mathbf{m}} \mathbb{S}^n$, the mapped point via the exponential retraction can be directly derived from (6) as

$$\begin{aligned} \mathcal{R}_{\mathbf{m}}^{\text{exp}}(\sigma_{t,ij}^\pm) &= \cos(\|\sigma_{t,ij}^\pm\|) \mathbf{m} + \frac{\sin(\|\sigma_{t,ij}^\pm\|)}{\|\sigma_{t,ij}^\pm\|} \sigma_{t,ij}^\pm \\ &= \mathbf{m} \cos \underline{r}_{i,j} \pm \mathbf{m}_i \sin \underline{r}_{i,j} \in \mathbb{S}^n. \end{aligned}$$

The logarithm and exponential map preserve the length of geodesic connecting the mode \mathbf{m} to the other points on \mathbb{S}^n , and the dispersion of Bingham is antipodally symmetric. Thus, the corresponding scaling factors in (7) for exponential retraction are therefore bounded, namely $\underline{r}_{i,j} \in [0, \pi/2]$.

For hyperspheres, different types of cartographic projections, which are originally used to visualize the spherical surface on a plane, also induce possible retractions. More specifically, inspired by the orthographic and gnomonic projection, we introduce the following two projection-like retractions [24], respectively.

2) *Orthographic Retraction*: The orthographic projection projects a point on the hypersphere orthogonally to the tangent plane. Therefore, the corresponding retraction is given as its inverse operation, i.e.,

$$\begin{aligned} \mathcal{R}_{\mathbf{m}}^{\text{ort}}(\sigma_{t,ij}^\pm) &= \sigma_{t,ij}^\pm + \sqrt{1 - \|\sigma_{t,ij}^\pm\|^2} \mathbf{m} \\ &= \pm \underline{r}_{i,j} \mathbf{m}_i + \sqrt{1 - \underline{r}_{i,j}^2} \mathbf{m} \in \mathbb{S}^n. \end{aligned}$$

The scaling factors are thus limited by the largest possible norm of vectors on the tangent plane induced by orthographic projections, i.e., $\underline{r}_{i,j} \in [0, 1]$.

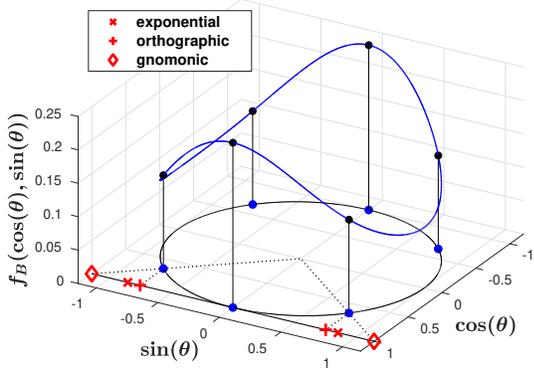


Fig. 2: Different retractions generating the same sigma points (blue dots on the circle) for a Bingham distribution on \mathbb{S}^1 . Here, the tangent plane (a 1-D line) is located at $\theta = 0$.

3) *Gnomonic Retraction:* The gnomonic retraction is defined as the inverse operation of gnomonic projection, i.e.,

$$\mathcal{R}_{\mathbf{m}}^{\text{gno}}(\sigma_{t,ij}^{\pm}) = \frac{\sigma_{t,ij}^{\pm} + \mathbf{m}}{\sqrt{1 + \|\sigma_{t,ij}^{\pm}\|^2}} = \frac{\pm r_{i,j} \mathbf{m}_i + \mathbf{m}}{\sqrt{1 + r_{i,j}^2}} \in \mathbb{S}^n,$$

which essentially induces the nearest on-sphere point of $\sigma_{t,ij}^{\pm} \in \mathcal{T}_{\mathbf{m}} \mathbb{S}^n$. As points on the equator of \mathbb{S}^n are mapped to infinity via gnomonic projection, the range for the corresponding scaling factors is unbounded, namely $r_{i,j} \in \mathbb{R}_+$.

IV. ON-TANGENT-PLANE DIRAC MIXTURE

APPROXIMATION UNDER ON-MANIFOLD CONSTRAINTS

The aforementioned spherical sampling scheme relies on a set of properly given scaling vectors $\{r_i\}_{i=1:n}$ satisfying requirements of the unscented transform and considering higher-order moments of the Bingham distribution. We thus formulate it as an optimization problem, where the statistical distance between the sigma Dirac mixtures and the underlying Bingham density is minimized on $\mathcal{T}_{\mathbf{m}} \mathbb{S}^n$ under the constraint of preserving the second moment on \mathbb{S}^n .

A. Geometric Interpretation of the Unscented Transform

For setting up the constraints for optimizing the scaling vectors $\{r_i\}_{i=1:n}$, we impose second-order moments in (4), i.e., $\text{cov}_{\mathbf{B}} = d_0 \mathbf{m} \mathbf{m}^{\top} + \sum_{i=1}^n d_i \mathbf{m}_i \mathbf{m}_i^{\top}$, to be identical to the covariance of the on-manifold sigma points

$$\mathbf{P} = w_0 \sigma_0 \sigma_0^{\top} + \sum_{i=1}^n \sum_{j=1}^l w_{ij} (\sigma_{ij}^+ (\sigma_{ij}^+)^{\top} + \sigma_{ij}^- (\sigma_{ij}^-)^{\top}).$$

Here, w_{ij} denotes the weight of each sigma pair σ_{ij}^{\pm} retracted from the tangent plane such that $w_0 + 2 \sum_{i=1}^n \sum_{j=1}^l w_{ij} = 1$. Appendix A shows the sample covariance for each retraction. Given the weights, the scaling factors thus essentially provide a geometric interpretation of the unscented transform, namely how far should the sigma points be located from the mode along the principal curves on \mathbb{S}^n [26], such that the second moment is preserved. Given $\sum_{i=0}^n d_i = 1$, constraints for each spherical retraction can be derived as follows.

- exponential retraction:
 $d_i = 2 \sum_{j=1}^l w_{ij} (\sin r_{i,j})^2, \quad r_{i,j} \in [0, \pi/2];$

- orthographic retraction:

$$d_i = 2 \sum_{j=1}^l w_{ij} r_{i,j}^2, \quad r_{i,j} \in [0, 1];$$

- gnomonic retraction:

$$d_i = 2 \sum_{j=1}^l w_{ij} r_{i,j}^2 / (1 + r_{i,j}^2), \quad r_{i,j} \in [0, +\infty).$$

When $l = 1$, for instance, it gives the classical unscented transform as shown in the following example.

Example IV.1 *Classical unscented transform for Bingham distribution on \mathbb{S}^n . The weighting factors [10] are given as*

$$w_0 = \lambda / (n + \lambda), \quad w_i = 1 / (2(n + \lambda)),$$

with $i = 1, \dots, n$ and λ the controlling factors as usual, such that $w_0 + 2 \sum_{i=1}^n w_i = 1$. By placing one pair of mirrored sigma points ($l = 1$) in each principal direction, one can derive, for instance, the scaling factors under exponential retraction as $r_i = \arcsin \sqrt{d_i / (2w_i)}$. Since the resulting scaling factors are bounded ($r_i \in [0, \pi/2]$), λ is also bounded. Scaling factors from the other two retractions can be derived similarly. Fig. 2 shows the case of $n = 1$ and $\lambda = 0.5$ (equally weighted on \mathbb{S}^1 circle) for a Bingham distribution with $\mathbf{C}_{\mathbf{B}} = \text{diag}(-1, -5)$.

B. On-tangent-plane Dirac Mixture Approximation

Under the constraint induced by the unscented transform in each principal direction, we adapt the sigma points according to the higher-order shape information of the Bingham distribution by means of Dirac mixture approximation. Here, the samples are equally weighted, such that uniform contributions among the sigma points are guaranteed. Therefore, the weights are $w = 1 / (2 \times n \times l + 1)$, and the constraints in Sec. IV-A can be simplified to $\mathcal{C}(r_i) = 0, r_{i,j} \in \Omega_i$ for each individual retraction as:

- exponential retraction:

$$\begin{aligned} \mathcal{C}^{\text{exp}}(r_i) &= 2 \sum_{j=1}^l (\sin r_{i,j})^2 - (2nl + 1) d_i, \\ \Omega_i^{\text{exp}} &= [0, \pi/2]; \end{aligned}$$

- orthographic retraction:

$$\begin{aligned} \mathcal{C}^{\text{ort}}(r_i) &= 2 \sum_{j=1}^l r_{i,j}^2 - (2nl + 1) d_i, \\ \Omega_i^{\text{ort}} &= [0, 1]; \end{aligned}$$

- gnomonic retraction:

$$\begin{aligned} \mathcal{C}^{\text{gno}}(r_i) &= 2 \sum_{j=1}^l r_{i,j}^2 / (1 + r_{i,j}^2) - (2nl + 1) d_i, \\ \Omega_i^{\text{gno}} &= [0, +\infty). \end{aligned}$$

We further manipulate the sigma points in each principal direction to minimize their statistical distance to the underlying Bingham density. First, we uniformly discretize the range $\Omega_i^{\pm} := \{\pm \omega \mid \omega \in \Omega_i\}$ based on (7) with a certain resolution K to get support points in each principal direction i , namely

$$\{\underline{v}_{i,k} \mathbf{m}_i\}_{k=1:K} \subset \mathcal{T}_{\mathbf{m}} \mathbb{S}^n, \quad \underline{v}_{i,k} \in \Omega_i^{\pm}. \quad (8)$$

The optimization is then to minimize the distance between Dirac mixtures given by the sigma points $\tilde{f}_i^{\sigma}(x, r_i) = \frac{1}{2l+1} \sum_{j=1}^l \{\delta(x + r_{i,j}) + \delta(x - r_{i,j})\}$, and the one given by the support points $\tilde{f}_i^s(x) = \sum_{k=1}^K p_{ik} \delta(x - \underline{v}_{i,k})$, with $x \in \Omega_i^{\pm}$. Here, $\delta(\cdot)$ denotes the Dirac delta function and p_{ik} is the Bingham density $f_{\mathbf{B}}(\mathcal{R}_{\mathbf{m}}(\underline{v}_{i,j} \mathbf{m}_i))$ at the retracted support points after normalization. For this approximation, the sample reduction technique proposed in [18] can be

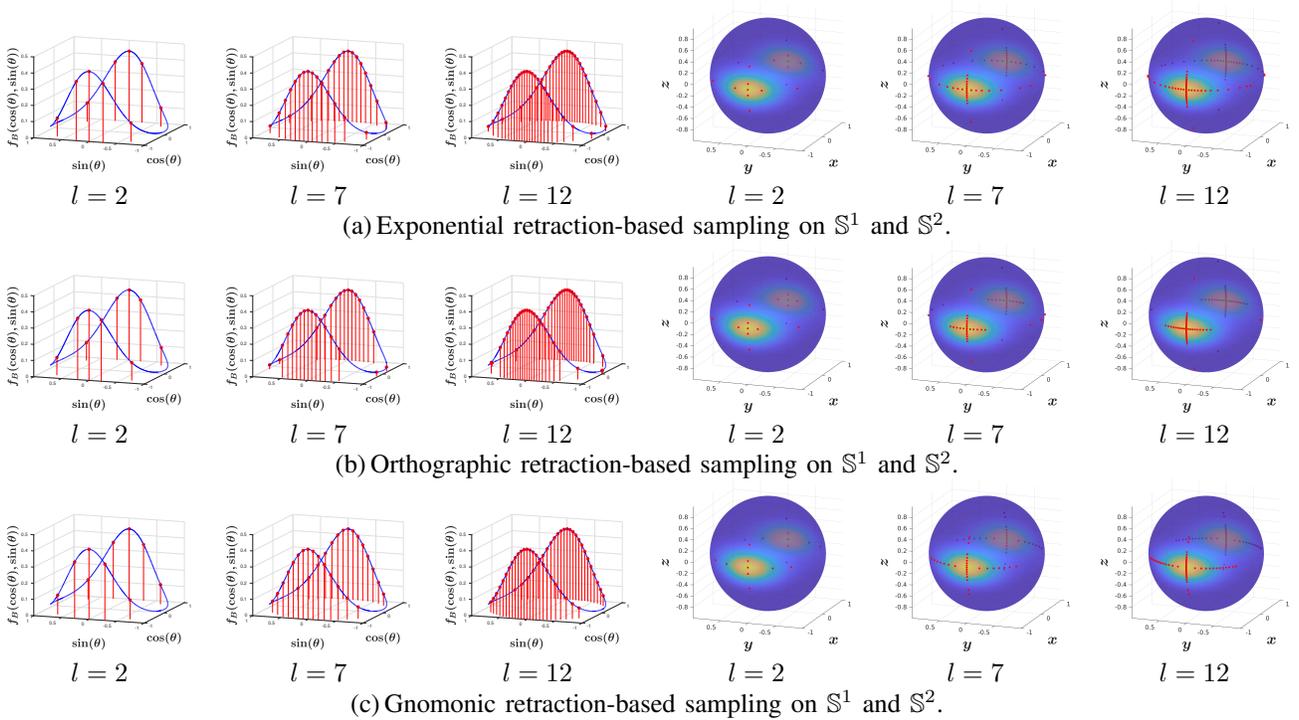


Fig. 3: Results given by the proposed deterministic sampling scheme generating flexibly given numbers of sigma points. For each dimensionality of \mathbb{S}^n ($n = 1, 2$), same Bingham distribution is approximated based on the exponential, orthographic, and gnomonic retraction, respectively.

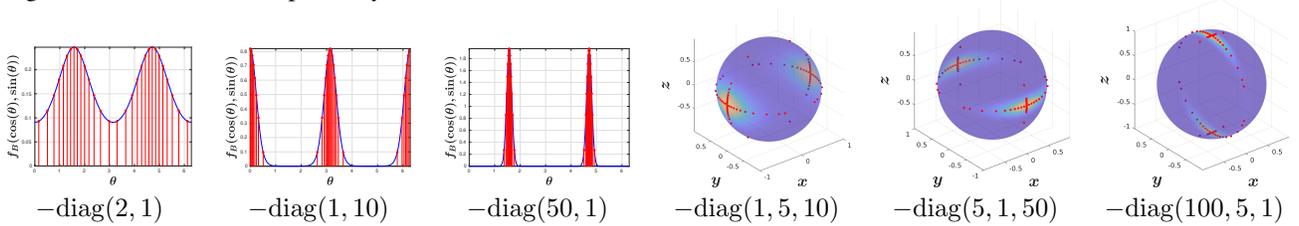


Fig. 4: Results of the proposed deterministic sampling scheme based on exponential retraction for Bingham distributions with different parameterizations \mathbf{C}_B ($l = 7$).

employed, where the densities are converted to the LCD form

$$\begin{aligned} \tilde{F}_i^\sigma(\underline{r}_i, m, b) &= \frac{1}{2nl+1} \left\{ \sum_{j=1}^l \exp\left(-\frac{(\underline{r}_{i,j} - m)^2}{2b^2}\right) + \right. \\ &\quad \left. \exp\left(-\frac{m^2}{2b^2}\right) + \sum_{j=1}^l \exp\left(-\frac{(\underline{r}_{i,j} + m)^2}{2b^2}\right) \right\}, \\ \tilde{F}_i^s(m, b) &= \sum_{k=1}^K \exp\left(-\frac{(\underline{v}_{i,k} - m)^2}{2b^2}\right). \end{aligned}$$

Here, m denotes continuous test centers, around which the local mass of the distribution in range b is measured by the kernel $\mathcal{K}(x - m, b) = \exp\left(-\frac{1}{2}\left(\frac{x-m}{b}\right)^2\right)$ (see Appendix B). The distance between Dirac mixtures LCDs can be measured by the modified Cramér-von Mises distance [18]

$$\begin{aligned} \mathcal{D}_i^2(\underline{r}_i) &= \int_{\mathbb{R}_+} h(b) \int_{\mathbb{R}} \left(\tilde{F}_i^\sigma(\underline{r}_i, m, b) - \tilde{F}_i^s(m, b) \right)^2 dm db, \\ \text{with } h(b) &= \begin{cases} 1, & b \in [0, b_{\max}] \\ 0, & \text{elsewhere} \end{cases} \end{aligned}$$

controlling the kernel size of the weighting function. The aforementioned distance metric has a closed-form solution, and can be practically used as the target function in the following optimization problem

$$\underline{r}_i^* = \arg \min_{\underline{r}_i} \mathcal{D}_i(\underline{r}_i), \text{ s.t. } \mathcal{C}_i(\underline{r}_i) = 0, \underline{r}_{i,j} \in \Omega_i,$$

with constraint functions \mathcal{C}_i and boundaries Ω_i as introduced in Sec. IV-B. To solve this constrained optimization problem, we employ the Matlab function *fmincon* of default setup. Fig. 3 shows results using the three retractions introduced in Sec. III-B. Numbers of samples are $2 \times n \times l + 1$, with $n = 1, 2$ and $l = 2, 7, 12$. The Bingham parameters are $\mathbf{C}_B^{\text{circle}} = -\text{diag}(2, 5)$, and $\mathbf{C}_B^{\text{sphere}} = -\text{diag}(1, 5, 15)$. Fig. 4 shows the results of varying parameters \mathbf{C}_B with $n = 1, 2$ and fixed sample size.

V. EVALUATION

We evaluate the aforementioned sampling scheme in the context of nonlinear Bingham filtering for orientation estimation. We hereby set up a system model $\mathbf{x}_k = \mathbf{x}_{k-1} \otimes \mathbf{u}_k \otimes \mathbf{w}_k$, with \mathbf{x}_k denoting the quaternion state, \mathbf{u}_k the

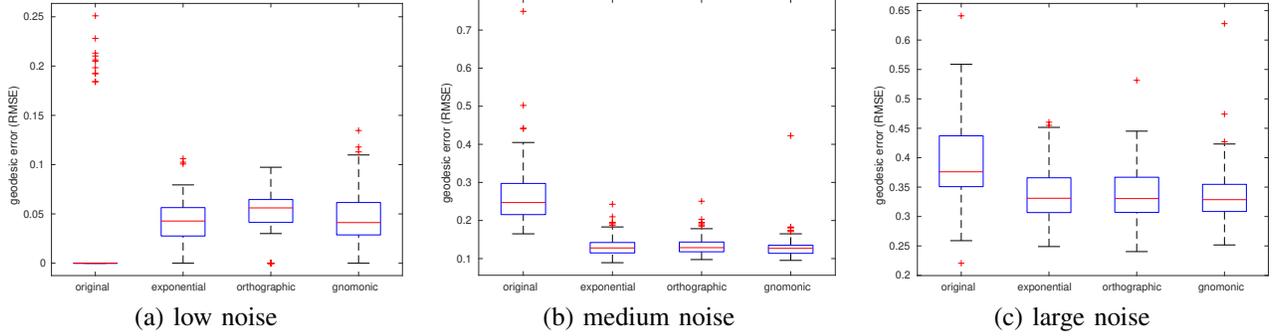


Fig. 5: Tracking accuracy of nonlinear Bingham filtering for planar orientation estimation under different noise levels. We compare the proposed scheme using all three retractions with the original deterministic sampling approach in [15].

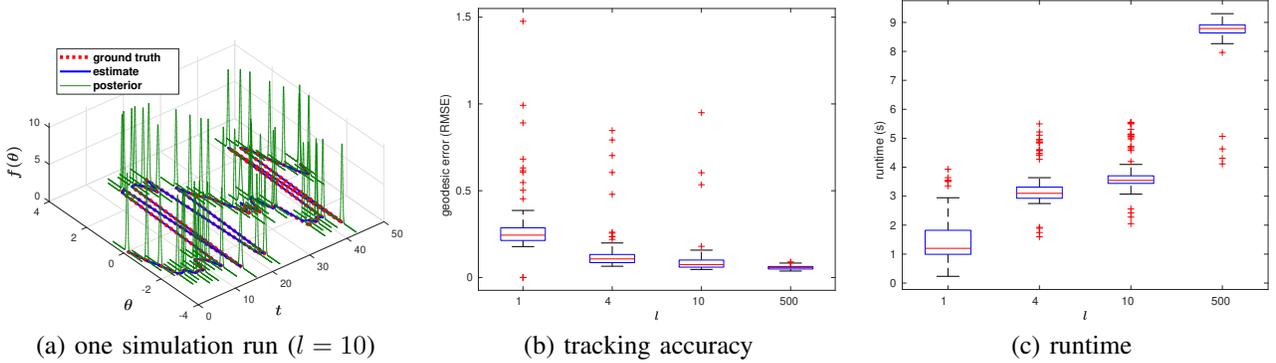


Fig. 6: Tracking accuracy and runtime results based on gnomonic retraction with different sample sizes of $2 \times l + 1$. Here, $l = 1$ indicates the original UT-based sampling approach in [15] and $l = 500$ indicates a naively implemented Bingham particle filter. A representative tracking result is shown in (a).

quaternion system input, and \mathbf{w}_k the noise term assumed to be Bingham-distributed. A non-identity measurement model is given as $\mathbf{z}_k = \mathbf{x}_k \otimes \mathbf{n}_0 \otimes \mathbf{x}_k^* + \mathbf{v}_k$, with \mathbf{z}_k being the orientation measurement by rotating the initial orientation \mathbf{n}_0 with the quaternion state \mathbf{x}_k according to (2). An additive noise term \mathbf{v}_k is used which is assumed to be zero-mean Gaussian-distributed, namely $\mathbf{v}_k \sim \mathcal{N}(\mathbf{0}, \Sigma^{\mathbf{v}})$. We employ the Bingham filtering framework originally proposed in [15], however, with a modified update step for the non-identity measurement model. Here, the prior can be computed by composing Bingham distributions. In the update step, deterministic samples from the prior are first drawn, and then reweighted according to the following likelihood

$$f(\mathbf{z}_k | \mathbf{x}_k) = f_{\mathbf{v}_k}(\mathbf{z}_k - \mathbf{x}_k \otimes \mathbf{n}_0 \otimes \mathbf{x}_k^*). \quad (9)$$

The posterior is estimated by approximating the reweighted samples via moment matching as introduced in [15]. Our simulation is done for planar orientation estimation based on 100 Monte Carlo runs of 50 time steps. The Bingham noise is parameterized by $\mathbf{C}_B^{\mathbf{w}} = -\text{diag}(1, 10)$, indicating a large zero-centered uncertainty as shown in Fig. 4.

We first evaluate the estimation accuracy by using the proposed sampling scheme in comparison with the original one in [15] under different measurement noise levels. Here, we have $\Sigma_{\text{high}}^{\mathbf{v}} = 10^{-1} \cdot \text{diag}(1, 5)$, $\Sigma_{\text{medium}}^{\mathbf{v}} = 10^{-2} \cdot \text{diag}(1, 5)$, and $\Sigma_{\text{low}}^{\mathbf{v}} = 10^{-4} \cdot \text{diag}(1, 5)$. The error metric is defined based on the geodesic in arc length as introduced in Sec. II-C. Fig. 5 shows the RMSE error for the original unscented

transform given in [15], where only 3 samples are drawn, and for the proposed sampling scheme which gives $2 \times 10 + 1 = 21$ samples adaptive to the shape of the Bingham density. For different noise levels, the proposed sampling scheme gives more accurate results by using the three retractions. Particularly, the original approach fails to track under low measurement noise level due to the sample degeneration issue as shown in Fig. 5(a). The three retractions show similar tracking accuracy.

Furthermore, we evaluate the performance of the proposed sampling scheme with different sample sizes regarding accuracy and efficiency. The measurement noise is $\Sigma^{\mathbf{v}} = 10^{-3} \cdot \text{diag}(1, 10)$. The proposed sampling scheme can give accurate tracking result as shown in Fig. 6(a). Moreover, it shows better tracking accuracy with more deterministic samples than the UT-based Bingham filter ($l = 1$) originally proposed in [15]. The Bingham filter using the proposed deterministic sampling approach with $l = 10$ shows similar tracking accuracy as the one using random sampling [27] with $l = 500$, however, with much less runtime.

VI. CONCLUSIONS

In this paper, we proposed a novel deterministic sampling scheme for the Bingham distribution based on spherical geometry. The approach works for arbitrary dimensions and can generate flexibly adjustable numbers of equally weighted sigma points that satisfy requirements of the unscented transform, meanwhile adapt to higher-order moments. The

basic spherical sampling scheme proposed in this work could potentially be applied to other distributions defined on the sphere or other nonlinear manifolds as well as to distributed estimation schemes for sensor networks [28]. The simulations show that our adaptive sampling approach improves the accuracy and robustness of nonlinear Bingham filtering for quaternion-based orientation estimation compared to the basic unscented filter in [15]. Still, the proposed sampling scheme has much potential to exploit. Its efficiency could be considerably improved by including closed-form gradients into optimizations as in [18]. The current sampling scheme can be extended by approximating the density between principal directions, such that a better filtering performance can be expected.

APPENDIX

A. Covariance Given by Different Retractions

Sigma points covariance derived by different retractions are summarized as follows:

- exponential retraction (Sec. III-B.1):

$$\mathbf{P}^{\text{exp}} = \left(w_0 + 2 \sum_{i=1}^n \sum_{j=1}^l w_{ij} (\cos r_{i,j})^2 \right) \mathbf{m} \mathbf{m}^\top + 2 \sum_{i=1}^n \sum_{j=1}^l w_{ij} (\sin r_{i,j})^2 \mathbf{m}_i \mathbf{m}_i^\top;$$
- orthographic retraction (Sec. III-B.2):

$$\mathbf{P}^{\text{ort}} = \left(w_0 + 2 \sum_{i=1}^n \sum_{j=1}^l w_{ij} (1 - r_{i,j}^2) \right) \mathbf{m} \mathbf{m}^\top + 2 \sum_{i=1}^n \sum_{j=1}^l w_{ij} r_{i,j}^2 \mathbf{m}_i \mathbf{m}_i^\top;$$
- gnomonic retraction (Sec. III-B.3):

$$\mathbf{P}^{\text{gno}} = \left(w_0 + 2 \sum_{i=1}^n \sum_{j=1}^l \frac{w_{ij}}{1+r_{i,j}^2} \right) \mathbf{m} \mathbf{m}^\top + 2 \sum_{i=1}^n \sum_{j=1}^l \frac{w_{ij} r_{i,j}^2}{1+r_{i,j}^2} \mathbf{m}_i \mathbf{m}_i^\top.$$

B. Localized Cumulative Distribution (LCD)

For a probability density function $f : \mathbb{R}^n \rightarrow \mathbb{R}_+$ modeling the uncertainty of variable $\underline{x} \in \mathbb{R}^n$, the corresponding localized cumulative distribution (LCD) is in the following form

$$F(\underline{m}, \underline{b}) = \int_{\mathbb{R}^n} f(\underline{x}) \mathcal{K}(\underline{x} - \underline{m}, \underline{b}) d\underline{x},$$

with $\underline{m} \in \mathbb{R}^n$ and $\underline{b} \in \mathbb{R}_+^n$ denoting the location and the size of kernel function $\mathcal{K}(\underline{x} - \underline{m}, \underline{b})$.

REFERENCES

- [1] R. A. Srivatsan, M. Xu, N. Zevallios, and H. Choset, "Probabilistic Pose Estimation Using a Bingham Distribution-based Linear Filter," *The International Journal of Robotics Research*, June 2018.
- [2] K. Li, D. Frisch, S. Radtke, B. Noack, and U. D. Hanebeck, "Wave-front 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.
- [3] M. Bloesch, M. Burri, S. Omari, M. Hutter, and R. Siegwart, "Iterated Extended Kalman Filter Based Visual-inertial Odometry Using Direct Photometric Feedback," *The International Journal of Robotics Research*, vol. 36, no. 10, pp. 1053–1072, 2017.
- [4] J. Glover and L. P. Kaelbling, "Tracking the Spin on a Ping Pong Ball with the Quaternion Bingham Filter," in *2014 IEEE International Conference on Robotics and Automation (ICRA 2014)*, May 2014, pp. 4133–4140.
- [5] J. Glover, G. Bradski, and R. Rusu, "Monte Carlo Pose Estimation with Quaternion Kernels and the Bingham Distribution," in *Proceedings of Robotics: Science and Systems*, 2011.
- [6] K. Li, G. Kurz, L. Bernreiter, and U. D. Hanebeck, "Nonlinear Progressive Filtering for SE(2) Estimation," in *Proceedings of the 21st International Conference on Information Fusion (Fusion 2018)*, Cambridge, United Kingdom, July 2018.
- [7] —, "Simultaneous Localization and Mapping Using a Novel Dual Quaternion Particle Filter," in *Proceedings of the 21st International Conference on Information Fusion (Fusion 2018)*, Cambridge, United Kingdom, July 2018.
- [8] K. Li, F. Pfaff, and U. D. Hanebeck, "Geometry-Driven Stochastic Modeling of SE(3) States Based on Dual Quaternion Representation," in *Proceedings of the 2019 IEEE International Conference on Multisensor Fusion and Integration for Intelligent Systems (MFI 2019)*, May 2019.
- [9] J. Civera, O. G. Grasa, A. J. Davison, and J. Montiel, "1-Point RANSAC for Extended Kalman Filtering: Application to Real-time Structure from Motion and Visual Odometry," *Journal of Field Robotics*, vol. 27, no. 5, pp. 609–631, 2010.
- [10] S. J. Julier and J. K. Uhlmann, "Unscented Filtering and Nonlinear Estimation," *Proceedings of the IEEE*, vol. 92, no. 3, pp. 401–422, 2004.
- [11] A. I. Mourikis and S. I. Roumeliotis, "A Multi-State Constraint Kalman Filter for Vision-aided Inertial Navigation," in *Proceedings 2007 IEEE International Conference on Robotics and Automation*, April 2007, pp. 3565–3572.
- [12] J. Glover and S. Popovic, "Bingham Procrustean Alignment for Object Detection in Clutter," in *Proceedings of the 2013 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS 2013)*, 2013, pp. 2158–2165.
- [13] C. Bingham, "An Antipodally Symmetric Distribution on the Sphere," *Ann. Statist.*, vol. 2, no. 6, pp. 1201–1225, 11 1974.
- [14] C. Forster, L. Carlone, F. Dellaert, and D. Scaramuzza, "On-Manifold Preintegration for Real-Time Visual-Inertial Odometry," *IEEE Transactions on Robotics*, vol. 33, no. 1, pp. 1–21, 2017.
- [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. F. Huber and U. D. Hanebeck, "Gaussian Filter based on Deterministic Sampling for High Quality Nonlinear Estimation," in *Proceedings of the 17th IFAC World Congress (IFAC 2008)*, vol. 17, no. 2, Seoul, Republic of Korea, July 2008.
- [17] 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.
- [18] U. D. Hanebeck, "Optimal Reduction of Multivariate Dirac Mixture Densities," *at - Automatisierungstechnik, Oldenbourg Verlag*, vol. 63, no. 4, pp. 265–278, Apr. 2015.
- [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] O. C. Schrempf, D. Brunn, and U. D. Hanebeck, "Density Approximation Based on Dirac Mixtures with Regard to Nonlinear Estimation and Filtering," in *Proceedings of the 2006 IEEE Conference on Decision and Control (CDC 2006)*, San Diego, California, USA, Dec. 2006.
- [21] W. R. Hamilton, "On Quaternions; or on a New System of Imaginaries in Algebra," *Philosophical Magazine Series 3*, vol. 25, no. 163, pp. 10–13, 1844.
- [22] K. V. Mardia and P. E. Jupp, *Directional Statistics*. John Wiley & Sons, 2009, vol. 494.
- [23] P.-A. Absil, R. Mahony, and R. Sepulchre, *Optimization Algorithms on Matrix Manifolds*. Princeton University Press, 2009.
- [24] P.-A. Absil and J. Malick, "Projection-like Retractions on Matrix Manifolds," *SIAM Journal on Optimization*, vol. 22, no. 1, pp. 135–158, 2012.
- [25] S. Sra and R. Hosseini, "Conic Geometric Optimization on the Manifold of Positive Definite Matrices," *SIAM Journal on Optimization*, vol. 25, no. 1, pp. 713–739, 2015.
- [26] S. Hauberg, "Principal Curves on Riemannian Manifolds," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 38, no. 9, pp. 1915–1921, 2016.
- [27] J. T. Kent, A. M. Ganeiber, and K. V. Mardia, "A New Method to Simulate the Bingham and Related Distributions in Directional Data Analysis with Applications," *arXiv preprint arXiv:1310.8110*, 2013.
- [28] S. Radtke, K. Li, B. Noack, and U. D. Hanebeck, "Comparative Study of Track-to-Track Fusion Methods for Cooperative Tracking with Bearing-Only Measurements," in *Proceedings of the 2019 IEEE International Conference on Multisensor Fusion and Integration for Intelligent Systems (MFI 2019)*, May 2019.