Efficient Bingham Filtering Based on Saddlepoint Approximations

Igor Gilitschenski[*], **Gerhard Kurz**[*], **Simon J. Julier**[†], and **Uwe D. Hanebeck**[*]

[*]Intelligent Sensor-Actuator-Systems Laboratory (ISAS)
Institute for Anthropomatics and Robotics
Karlsruhe Institute of Technology (KIT), Germany
gilitschenski@kit.edu, gerhard.kurz@kit.edu, uwe.hanebeck@ieee.org

[†]Virtual Environments and Computer Graphics Group
Department of Computer Science
University College London (UCL), United Kingdom
s.julier@ucl.ac.uk

***Abstract*—In this paper, we address the problem of developing computationally efficient recursive estimators on the periodic domain of orientations using the Bingham distribution. The Bingham distribution is defined directly on the unit hypersphere. As such, it is able to describe both large and small uncertainties in a unified framework. In order to tackle the challenging computation of the normalization constant, we propose a method using its saddlepoint approximations and an approximate MLE based on the Gauss-Newton method. In a set of simulation experiments, we demonstrate that the Bingham filter not only outperforms both Kalman and particle filters, but can also be implemented efficiently.**

***Keywords*—Bingham distribution, directional statistics, moment matching, maximum likelihood estimation.**

I. Introduction

Inference on periodic domains is important for many applications particularly for orientation estimation, e.g., in smartphones, UAVs, and augmented reality. This is often achieved using low cost, noisy sensing systems such as MEMS IMUs or GPS. High levels of uncertainty need to be considered in scenarios involving low-quality sensors, weak prior information, or strong environmental disturbances. These scenarios require stochastic filtering techniques to be capable of dealing with both periodicity of the underlying domain of orientations and strong noise.

State of the art algorithms usually require some modification in order to account for the periodicity of the underlying domain. They are based on nonlinear modifications of the Kalman filter such as the extended Kalman filter (EKF) or the unscented Kalman filter (UKF) [1]. It is possible to introduce constraints and use nonlinear projection to achieve feasible estimation results [2]. There exists an extensive body of research based on these approaches. Extended Kalman filters for quaternion-based orientation estimation are presented in [3], [4], [5]. A comparison of UKF and EKF for orientation estimation is presented in [6]. Further algorithms based on the Kalman filter are discussed in [7], [8], [9]. These approaches assume all uncertainties to be Gaussian and linearity of the underlying state-space, which is inherently wrong when system state and measurements are defined on a periodic domain. Due to local linearity of the considered manifolds, filters based on these assumptions yield good results for low system and measurement noise. They are prone to problems when the considered estimation scenario involves highly uncertain quantities.

A. Motivation

A new approach has emerged which uses distribution assumptions inherently considering the geometry of the underlying manifold in order to handle low-quality sensors and high measurement noise. This is done by using probability distributions from directional statistics [10], such as the von Mises distribution on the circle or the Bingham distribution on the hypersphere. Filtering techniques based on these distribution assumptions are usually more complicated than algorithms based on the Gaussian distribution. However, they are promising because they result in robust estimators even for scenarios involving strong system and measurement noise [11], [12], [13], [14], [15].

Stochastic filtering based on the Bingham distribution motivates the contribution of this work. The Bingham distribution is an antipodally symmetric distribution defined on arbitrary dimensional hyperspheres [16]. It is of interest because estimating angles or orientations can be seen as a special case of state estimation on hyperspheres. Estimation of angles is performed by using a two-dimensional Bingham distribution (which is basically a distribution defined on the circle) and rescaling it if the 180° symmetry is not desired (however, this rescaling procedure results in a von Mises distribution). Orientations in 3D can be estimated using unit quaternions for representing an orientation and considering the four-dimensional Bingham distribution (it is a natural representation of uncertainty over the unit sphere of quaternions). In this situation, antipodal symmetry of the Bingham distribution is a desirable property reflecting the fact that unit quaternions q and $-q$ represent the same orientation.

The first filters based on the Bingham distribution made use of the fact that it is closed under Bayesian inference and proposed an approach based on moment matching to account for an orientational and angular equivalent of adding two random variables [17], [14]. Furthermore, a UKF like sampling scheme was proposed in [18] for handling more complex system models. One of the remaining challenges when applying filters based on the Bingham distribution is the efficient computation of its normalization constant, its derivatives, and efficient computation of Bingham distribution parameters from given moments or given samples. The normalization constant and its derivatives are of particular importance because their computation is involved in several operations such as computation of the covariance of a Bingham distributed random vector and parameter estimation. This computation is currently avoided by using precomputed lookup tables [19]. Unfortunately, this approach might be not feasible for systems involving a limited amount of memory.

B. Contribution

In this work, we propose a way to avoid the use of precomputed lookup tables or Markov Chain Monte Carlo techniques for computationally efficient stochastic filtering based on the Bingham distribution. This is challenging, because the prediction step requires a numerical moment matching procedure that involves multiple evaluations of the Bingham normalization constant and its derivatives. Therefore, recently proposed saddlepoint approximations of the Bingham normalization constant are used. We show that the optimization problem involved in obtaining these saddlepoint approximations always converges. Furthermore, a Gauss-Newton based computation scheme for moment matching and maximum likelihood estimation (MLE) is derived. In this process, we compute a special case of the relationship between the Bingham normalization constant and its derivatives. The evaluation compares the proposed computation scheme to other recent approaches and we show that the resulting filter yields a superior estimation quality over a particle filter and the UKF at a comparable computation time.

The remainder of the paper is structured as follows. In the next section, we provide the necessary definitions and formally state the considered problems. The problem of computing the Bingham normalization constant and a moment matching procedure is discussed in Section III. The proposed algorithm is evaluated in Section IV by comparing the computation time, the relative computation error and by simulating a filter run in an orientation estimation scenario. Finally, the work is concluded in Section V.

II. Considered Problem

In this work, we consider the Bingham distribution, given by

$$f(\underline{x}) = N(\mathbf{C})^{-1} \cdot \exp(\underline{x}^\top \mathbf{C}\, \underline{x}) \ ,$$

where $\underline{x} \in S_{d-1} \subset \mathbb{R}^d$ and $\mathbf{C} \in \mathbb{R}^{d \times d}$ is symmetric negative definite (S_{d-1} is used to denote the unit sphere in $\mathbb{R}^d$). Thus, it naturally arises when restricting a zero-mean Gaussian random variable to unit length. Typically, we use a representation based on an eigendecomposition of $\mathbf{C}$ given by $\mathbf{M}\,\mathbf{Z}\,\mathbf{M}^\top$. Thus, the density is usually written as

$$f(\underline{x}) = N(\mathbf{C})^{-1} \exp(\underline{x}^\top \mathbf{M}\,\mathbf{Z}\,\mathbf{M}^\top \underline{x})$$

and a Bingham distributed random vector is usually denoted by $\underline{x} \sim \mathrm{Bingham}(\mathbf{M}, \mathbf{Z})$. This representation has the benefit of separating the location of the modes (encoded in $\mathbf{M}$) from the magnitude of the dispersion (encoded in $\mathbf{Z}$).

For estimation of parameters from empirical moments of some random samples and for moment matching procedures, we are interested in the computation of the moments of a Bingham distributed random vector $\underline{x} \sim \mathrm{Bingham}(\mathbf{M}, \mathbf{Z})$. Due to antipodal symmetry it is easily seen that $\mathbb{E}(\underline{x}) = \underline{0}$. Thus, the second moment corresponds to the covariance of a Bingham distributed random vector. For the second moment, in [16], Bingham obtained

$$\mathbb{E}\left(\underline{x} \cdot \underline{x}^\top\right) = \mathbf{M} \cdot \mathrm{diag}\left(\omega_1, \ldots, \omega_n\right) \cdot \mathbf{M}^\top \ , \tag{1}$$

where

$$\omega_i = \frac{\frac{\partial}{\partial z_i} N(\mathbf{Z})}{N(\mathbf{Z})} \tag{2}$$

and $\mathbf{Z} = \mathrm{diag}(z_1, \ldots, z_n)$. Thus, the second moment can be used to compute the Bingham distribution parameters. Furthermore, $\mathbf{M}$ and ω_i result from the eigendecomposition of the second moment $\mathbb{E}\left(\underline{x} \cdot \underline{x}^\top\right)$. However, for the computation of $\mathbf{Z}$, a numerical procedure is required, which requires the evaluation of the normalization constant and its derivatives in each iteration.

This procedure is of particular importance in filters based on the Bingham distribution [17], [14]. Thus, the problem considered in the following is threefold. First, we are interested in the efficient computation of the normalization constant. Second, we are interested in the computation of its derivatives. Finally, we need an efficient estimation procedure for estimating the values of the parameter matrix $\mathbf{Z}$ from a given second moment.

III. Efficient Computations Involving the Bingham Normalization Constant

The prediction step is the most expensive part of filters based on the Bingham distribution, because it involves a spherical equivalent to convolution. Unfortunately, this operation does not result in a Bingham distribution. Thus, approximation procedures are required. The most expensive part is the computation of second moments for moment matching. Thus, it is of particular interest to speed up the whole optimization procedure, which motivates taking a closer look at the Bingham normalization constant and its derivatives.

The Bingham normalization constant can be expressed as a hypergeometric function of matrix argument [20]

$$N(\mathbf{Z}) = \left|S^{n-1}\right| \cdot {}_1F_1\left(\frac{1}{2}; \frac{n}{2}; \mathbf{Z}\right) \ ,$$

where $\left|S^{n-1}\right|$ denotes the area of the surface of a unit ball in $\mathbb{R}^n$. A naïve approach would be to perform numerical integration, which is computationally burdensome. The hypergeometric function is also computed by the algorithm proposed in [21] through a series expansion, which becomes burdensome for many interesting cases (e.g., when the expected angular uncertainty is sufficiently small). For the particular case of the Bingham normalization constant, performance issues can be handled by using precomputed lookup tables as in [22], which might not be desirable when only a very limited amount of memory is available, e.g., for an n-dimensional Bingham

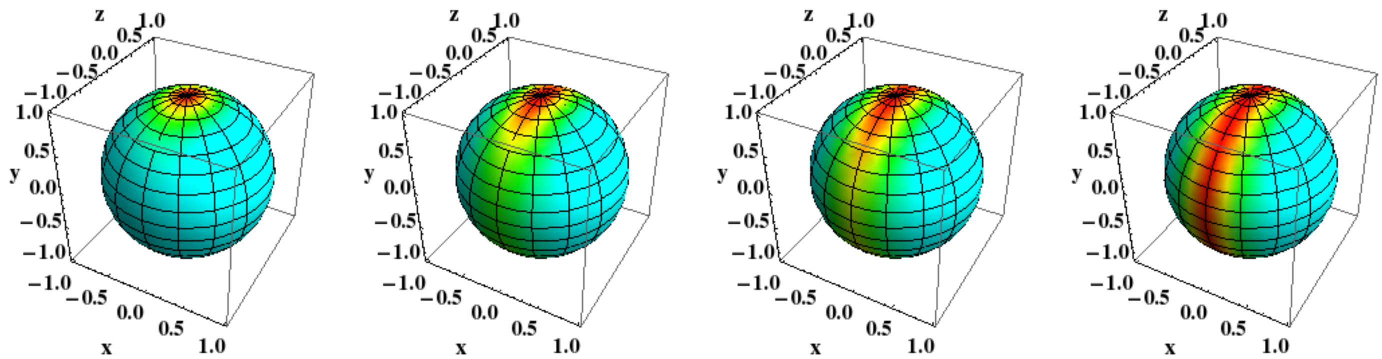

Figure 1: The unnormalized probability density function of the Bingham distribution in three dimensions is represented as a heat map on the unit ball. Due to the antipodal symmetry this plots remain the same for arbitrary $180°$ rotations.

distribution, the lookup table used in [19] uses 73^{n-1} entries and performs interpolation to approximate the true value of the normalization constant. Holonomic gradient descent was also applied for the exact computation of the Bingham normalization constant [23]. However, it does not offer sufficient performance for real-time applications when computing certain parameter combinations.

A. Approximation of Normalization Constant and its Derivatives

In order to avoid precomputation, we consider an approach based on saddlepoint approximations, which were proposed in [24] for approximation of probability density functions. This approach can be used for computing an approximation of the normalization constant of the Bingham distribution when the entries of the matrix $\mathbf{Z} = \mathrm{diag}(z_1, \ldots, z_n)$ are negative. Kume and Wood have shown in [25] that the Bingham normalization constant can be approximated using a third-order saddlepoint approximation by

$$\begin{aligned}\tilde{N}(\mathbf{Z}) :=& 2^{1/2}\pi^{(n-1)/2}\left(K^{(2)}(\hat{t}, \mathbf{Z})\right)^{-1/2} \\ & \left(\prod_{i=0}^{p}\left(\sqrt{-z_i} - \hat{t}\right)^{-1/2}\right)\exp(-\hat{t} + T) \ ,\end{aligned}$$

where

$$T := \rho_4(\hat{t}, \mathbf{Z})/8 - 5\rho_3(\hat{t}, \mathbf{Z})/24 \ ,$$

with $\rho_j(t, \mathbf{Z}) := K^{(j)}(\hat{t})/K^{(2)}(\hat{t})$, $z_i \leq 0$, and

$$K^{(j)}(t, \mathbf{Z}) := \sum_{i=1}^{n}\left(\frac{(j-1)!}{2}\frac{1}{(\sqrt{-z_i} - t)^j}\right) \ .$$

Furthermore $\hat{t}$ is a unique solution of $K^{(1)}(t, \mathbf{Z}) = 1$ on $(-\infty, z^*)$ with $z^* = \min_i\left(\sqrt{-z_i}\right)$.

In order to find an efficient way to compute $\hat{t}$, we first restate a result from [25], which yields

$$z^* - \frac{n}{2} < \hat{t} < z^* - \frac{1}{2} \ .$$

Now, we take a closer look at the behaviour of $K^{(1)}(t, \mathbf{Z})$ on the considered interval (it is not necessary to consider the more general case of $K^{(j)}(t, \mathbf{Z})$ here, because only $K^{(1)}(t, \mathbf{Z})$ is involved in the optimization).

Proposition 1. *For a fixed $\mathbf{Z}$, the function $K^{(1)}(t, \mathbf{Z})$ is convex on $\hat{t} \in (-\infty, z^*)$, where z^* is defined as above.*

Proof: We show convexity by considering the second derivative of $K^{(1)}(t, \mathbf{Z})$, which is given by

$$\frac{\delta^2 K^{(1)}(t, \mathbf{Z})}{\delta t^2} = \sum_{i=1}^{n}\frac{1}{\left(\sqrt{-z_i} - t\right)^3} \ .$$

The convexity follows immediately from $\delta^2 K^{(1)}(t, \mathbf{Z})/\delta t^2 > 0$ for $t < z^*$. □

Using this result, we obtain guaranteed convergence and error bounds when applying Newton's method for solving $K^{(1)}(t, \mathbf{Z}) = 1$ on our considered interval [26].

This result can be generalized to the computation of $N(\mathbf{Z})$ with possibly non-negative entries in $\mathbf{Z}$. From the definition of the normalization constant, one can easily see

$$N(\mathbf{Z} + c\mathbf{I}) = N(\mathbf{Z}) \cdot \exp(c) \ , \qquad (3)$$

where $\mathbf{I}$ denotes the $n \times n$ identity matrix. For computing the normalization constant when $\mathbf{Z}$ has nonnegative entries, we can simply choose $-c$ to be larger then the maximum entry of $\mathbf{Z}$. Then, computing $N(\mathbf{Z} + c\mathbf{I})$ and applying (3) yields the desired result.

Another main challenge in handling the Bingham distribution is the computation of its derivatives, because they are used for computing the covariance and thus required for parameter estimation based on moment matching. In [27], Kume and Wood demonstrated that a relationship exists between the derivatives of a Bingham normalizing constant and a normalizing constant of Bingham distributions of higher dimensions. For our purposes, we derive a special case of Proposition 1 from [27].

Corollary 2. *We consider once again the Bingham normalization constant $N(\mathbf{Z})$ with $\mathbf{Z} = \mathrm{diag}(z_1, \ldots, z_d)$. Then*

$$\frac{\partial N(\mathbf{Z})}{\partial z_i} = \frac{1}{2\pi} N(\mathrm{diag}(z_1, \ldots, z_{i-1}, z_i, z_i, z_i, z_{i+1}, \ldots, z_d)) \ .$$

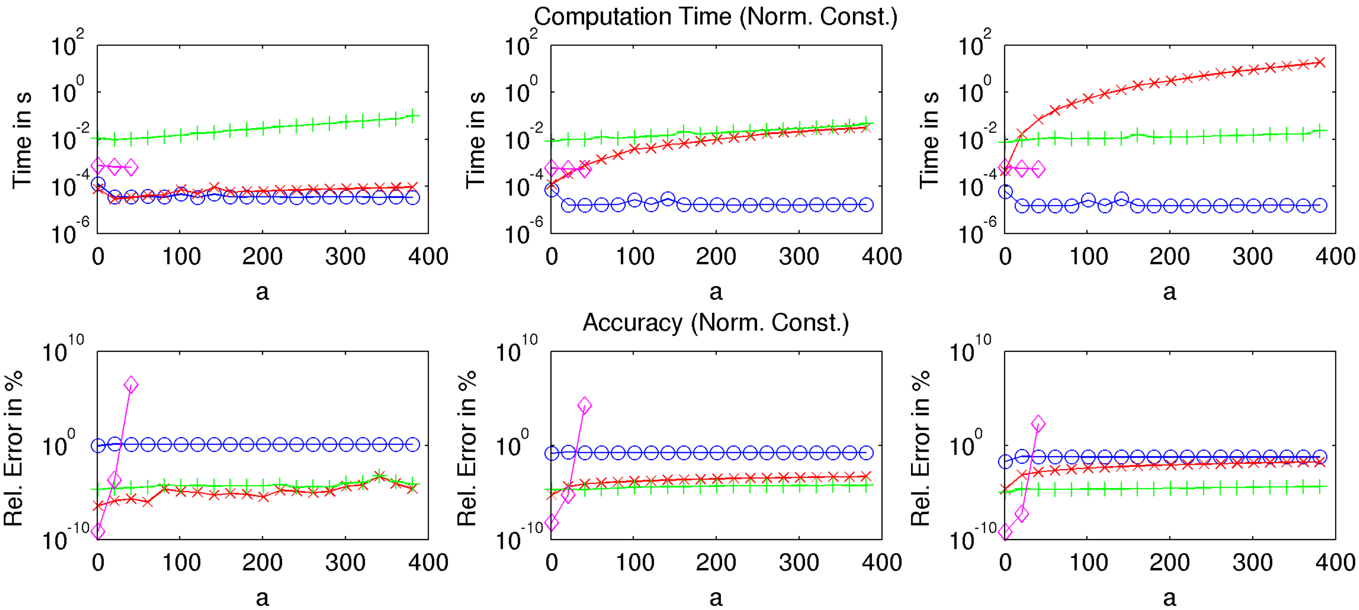

Figure 2: Comparison of the computation of the normalizing constants. For the cases $\mathbf{Z} = -\text{diag}(a, a, a, 0.1)$ (first column), $\mathbf{Z} = -\text{diag}(a, a, 0.1, 0.1)$ (second column) and $\mathbf{Z} = -\text{diag}(a, 0.1, 0.1, 0.1)$ (third column). These evaluation involves computations based on a series expansion (red), saddlepoint approximations (blue), holonomic gradient descent (green), and the approach proposed by Koev et al. (magenta).

Proof: The original result states that using a multi-index $\underline{k} = (k_1, \dots, k_d) \in \mathbb{N}_0^d$, the partial derivatives of a Bingham normalization constant are given by

$$\frac{\partial^{|i|} N(\text{diag}(z_1, \dots, z_d))}{\partial z_1^{k_1} \dots \partial z_d^{k_d}} = a(\underline{k}) N\left(\tilde{\mathbf{Z}}^{\underline{k}}\right) ,$$

where

$$a(\underline{k}) = \pi^{-|\underline{k}|} \prod_{i=1}^{d} \frac{\Gamma((1+2k_i)/2)}{\Gamma(1/2)} , \tag{4}$$

$|\underline{k}|$ is the entry sum of $\underline{k}$, and

$$\tilde{\mathbf{Z}}^{\underline{k}} = \text{diag}(\underbrace{z_1, \dots, z_1}_{(2k_1+1) \text{ times}}, \underbrace{z_2, \dots, z_2}_{(2k_2+1) \text{ times}}, \dots) .$$

Thus, the partial derivatives of a Bingham normalization constant can be stated as a rescaled Bingham normalization constant of another, higher dimensional, Bingham density. We consider the special case where for a fixed i, we have $k_i = 1$ and $k_j = 0$ for $i \neq j$. This simplifies (4) according to

$$a(\underline{k}) = \frac{\Gamma((1+2k_i)/2)}{\pi\,\Gamma(1/2)} = \frac{\Gamma(3/2)}{\pi\,\Gamma(1/2)} = \frac{1}{2\pi} .$$

□

An approximation of the derivative is obtained by replacing $N(\cdot)$ on the right hand side by the saddlepoint approximation $\tilde{N}(\cdot)$.

B. *Efficient Parameter Estimation*

The parameter estimation procedure for the Bingham distribution is also based on moment matching. That is, the second moment (1) is matched to the covariance matrix $\mathbf{C}$ of a set of samples. This is performed in two steps. First, the matrix $\mathbf{M}$ is obtained by eigendecomposition of $\mathbf{C}$. Second, $\mathbf{Z}$ is found by solving

$$\underbrace{\frac{\frac{\delta}{\delta z_i} N(\mathbf{Z})}{N(\mathbf{Z})}}_{=:f(\underline{z})} - e_i = 0 \tag{5}$$

for each $i = 1, \dots, n$, where e_i is the i-th eigenvalue of $\mathbf{C}$. The matrix $\mathbf{Z}$ can be interpreted as a vector of its diagonal entries (denoted by $\underline{z}$). The root finding problem (5) might not have a solution, when approximations of $N(\mathbf{Z})$ are used. However, it can be reformulated as

$$\min_{\underline{z}} ||f(\underline{z})||_2 . \tag{6}$$

Solving (6) can be done using the Gauss-Newton method. This approach requires the computation of the Jacobian $J_{f,k}$ of $f(\underline{x}_k)$ and therefore the computation of the gradient and the Hessian of $N(\mathbf{Z})$. Now, each iteration is computed by

$$\underline{x}^{(k+1)} = \underline{x}^{(k)} - \left(J_{f,k}^\top J_{f,k}\right)^{-1} J_{f,k}^\top f(\underline{x}^{(k)}) .$$

In this scheme, the pseudo-inverse is used. Furthermore, as an immediate consequence of (3) it follows $\text{Bingham}(\mathbf{M}, \mathbf{Z}) = \text{Bingham}(\mathbf{M}, \mathbf{Z} + c\mathbf{I})$ for all $c \in \mathbb{R}$. This relationship is used to help avoiding numerical cancelation errors (by setting $x_i^{(k)} := x_i^{(k)} - \max_j x_j^{(k)}$ in each iteration).

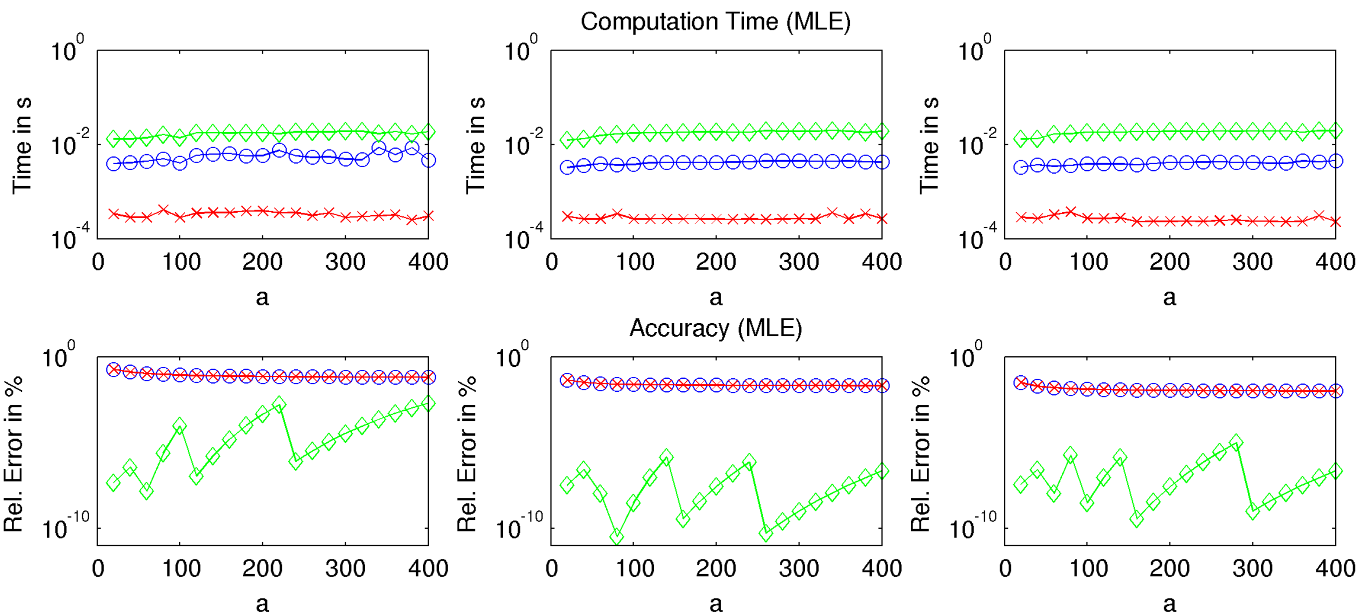

Figure 3: Comparison of the maximum likelihood estimation. For the cases $\mathbf{Z} = -\mathrm{diag}(a, a, a, 0.1)$ (first column), $\mathbf{Z} = -\mathrm{diag}(a, a, 0.1, 0.1)$ (second column) and $\mathbf{Z} = -\mathrm{diag}(a, 0.1, 0.1, 0.1)$ (third column). For the evaluation of the MLE approaches a Matlab implementation of the Gauss-Newton approach (blue), a C implementation of the Gauss-Newton approach (red), and a fsolve based approach (green) are shown.

IV. EVALUATION

The proposed methodology will be evaulated in two ways. First, the performance of the proposed method for computing the normalization constant and performing parameter estimation will be investigated. Second, the proposed filter will be compared to filtering methods involving a Gaussian assumption in an attitude estimation example.

A. *Normalization Constant and Maximum Likelihood Estimation*

The computation of the normalizing constant was evaluated by comparing the computation time and accuracy of the proposed method to some recently published computation methods. Ground truth was obtained by numerical direct integration in Matlab 2014a. A C implementation of the series expansion from [22] was used as a comparison method. Furthermore, we also used the method from [21] (where we summed over all terms with $|\kappa| \leq 100$, i.e., the algorithm provided by the authors was called using `mhg(100,2,0.5,2,Z)`) and an approach based on holonomic gradient descent [23]. The optimization procedure based on Newton's method for obtaining saddlepoint approximations was also implemented in C. All simulations were performed on a system with an Intel Core i7-2620M CPU and 8GB RAM. The performance of the series based computation differs depending on the structure of the Matrix $\mathbf{Z}$. Thus, we have considered three different scenarios in the evaluations. These scenarios differ in the number of diagonal entries in $\mathbf{Z}$ which are significantly different from 0. This consideration is sufficient because of the property (3).

The results of this comparison are shown in Fig. 2. The main benefit of the proposed approach is that its computation time is not dependent on the entries of $\mathbf{Z}$. At some point it outperforms all of the involved comparison methods. On the other hand, the accuracy seems worse at first. However, it is not dependent on the entries of $\mathbf{Z}$ either, and the relative error is below 1%. Thus, the proposed method offers a sufficiently good accuracy at low computational cost.

The resulting MLE algorithm was evaluated by generating a deterministic sample set as described in [18] and then using moment matching procedures for estimating the original parameters from these samples. The results are shown in Fig. 3. Similarly to the computation of the normalization constant, this was done for a different number of entries significantly different from 0 in the matrix $\mathbf{Z}$ and for different values of these entries. The proposed approach based on the Gauss-Newton method is less exact compared to `fsolve` (using a Levenberg–Marquardt algorithm). However, the errors are below 0.1% and thus once again sufficiently small for our purposes.

B. *Filtering*

In this subsection, our goal is to compare the Bingham filter when using the proposed computation methods to approaches making a Gaussian assumption. In order to asses the overall quality of the proposed approach we compare both computation time and estimation quality. Thus, a typical simulation run is performed in order to evaluate the resulting filter. Here, the Bingham filter is compared against a UKF and a Gaussian particle filter with 100 particles. The UKF and the particle filter are both implemented as in [18]. The only difference in the Bingham filter implementation is that all computations involving the normalization constant (e.g., moment computations and moment matching) were implemented using the proposed methods.

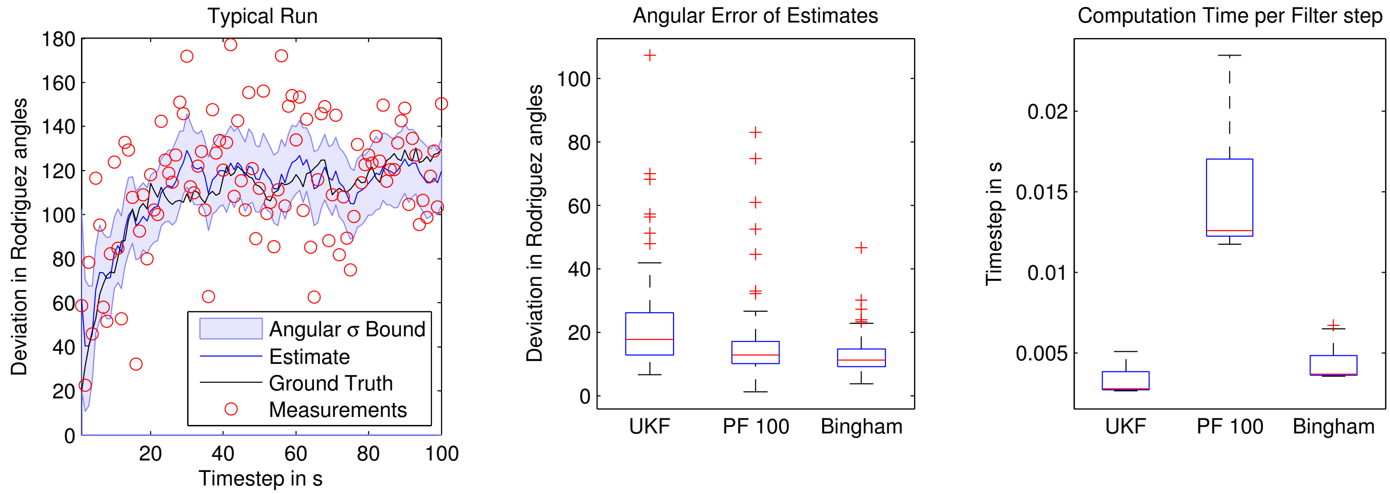

Figure 4: A typical run of the Bingham filter, the UKF, and a Gaussian particle filter with 100 particles and the corresponding errors of the estimates and the computation time of the filters (including both, prediction and measurement update step).

In our simulation setup, the system function stabilizes a noisy system state. Thus, the resulting system model is given by

$$\underline{x}_{t+1} = \underline{x}_t \oplus (\underline{x}_t^{-1} \oplus \underline{q})^c \oplus \underline{w}_t \ ,$$

which stabilizes the system state $\underline{x}_t$ towards a goal state $\underline{q}$ (here chosen as $(0.5, 0.5, 0.5, 0.5)^\top$) at the rate c (here 0.1). Here $\oplus$ denotes the quaternion multiplication. Identity matrices were chosen for $\mathbf{M}_w$ and $\mathbf{M}_v$ to obtain a zero mean equivalent to noise in linear systems. The dispersion parameters were chosen to be $\mathbf{Z}_w = \operatorname{diag}(-500, -500, -500, 0)$ and $\mathbf{Z}_v = \operatorname{diag}(-10, -10, -10, 0)$. This corresponds to expected Rodrigues angles uncertainty of $5°$ and $46°$ respectively. The expected angular deviation of a four dimensional Bingham distributed random vector $\underline{x}$ is given (in terms of Rodrigues angles) by

$$\mathbb{E}\left(2 \cdot \min(\operatorname{acos}(\underline{m}^\top \cdot \underline{x}), \pi - \operatorname{acos}(\underline{m}^\top \cdot \underline{x}))\right) \ ,$$

where $\underline{m}$ is one of the two modes of the Bingham distribution. The choice of the mode does not matter because the minimization procedure accounts for antipodal symmetry. In order to use correct parameters for the UKF and the Gaussian particle filter, parameters for corresponding Gaussian distributions were obtained by random sampling from the Bingham distributions involved and computing the respective means and covariances.

The simulation run is shown in Figure 4. For ease of understanding, angular deviation (based on Rodrigues angles [28]) from $(0, 0, 0, 1)$ is used to draw the trajectory. The filter itself was implemented in Matlab. When the Gauss-Newton based MLE (implemented in C) was used, one filter step (including time and measurement update) took 4ms. The same run was carried out using a `fsolve` based MLE where only the saddlepoint approximations were implemented in C. In that case, the overall filter step took $18ms$. The result of this comparison shows that the proposed filter offers better estimation results than approaches assuming the Gaussian distribution at a comparable computational cost.

V. Conclusion

In this work, we proposed a way to avoid the use of precomputed lookup tables when applying the Bingham filter for orientation estimation. This was achieved by showing that the optimization problem involved in computing the normalization constant always converges, and deriving a special case of the relationship between the Bingham normalization constant and its derivatives. Furthermore, we have proposed a numerical parameter estimation scheme for a moment matching procedure. The presented results make orientation estimation based on the Bingham filter possible in scenarios where use of large lookup tables is not feasible.

Acknowledgment

This work was partially supported by grants from the German Research Foundation (DFG) within the Research Training Groups RTG 1194 "Self-organizing Sensor-Actuator-Networks", and RTG 1126 "Soft-tissue Surgery: New Computer-based Methods for the Future Workplace"

References

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

[2] S. J. Julier and J. J. LaViola, Jr., "On Kalman Filtering With Nonlinear Equality Constraints," *IEEE Transactions on Signal Processing*, vol. 55, no. 6, pp. 2774–2784, Jun. 2007.

[3] J. Marins, E. Bachmann, R. McGhee, and M. Zyda, "An Extended Kalman Filter for Quaternion-based Orientation Estimation Using MARG Sensors," in *Proceedings 2001 IEEE/RSJ International Conference on Intelligent Robots and Systems*, vol. 4. IEEE, 2001, pp. 2003–2011.

[4] A. M. Sabatini, "Quaternion-based Extended Kalman Filter for Determining Orientation by Inertial and Magnetic Sensing," *IEEE transactions on bio-medical engineering*, vol. 53, no. 7, pp. 1346–56, Jul. 2006.

[5] I. Bar-Itzhack and Y. Oshman, "Attitude Determination from Vector Observations: Quaternion Estimation," *IEEE Transactions on Aerospace and Electronic Systems*, vol. AES-21, no. 1, pp. 128–136, Jan. 1985.

[6] J. LaViola, "A Comparison of Unscented and Extended Kalman Filtering for Estimating Quaternion Motion," in *Proceedings of the 2003 American Control Conference*, vol. 3, 2003, pp. 2435–2440.

[7] X. Yun, M. Lizarraga, E. Bachmann, and R. McGhee, "An Improved Quaternion-based Kalman Filter for Real-Time Tracking of Rigid Body Orientation," in *Proceedings 2003 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS 2003)*, vol. 2. IEEE, 2003, pp. 1074–1079.

[8] Y. S. Suh, "Orientation Estimation Using a Quaternion-Based Indirect Kalman Filter With Adaptive Estimation of External Acceleration," *IEEE Transactions on Instrumentation and Measurement*, vol. 59, no. 12, pp. 3296–3305, Dec. 2010.

[9] D. Choukroun, I. Bar-itzhack, and Y. Oshman, "Novel Quaternion Kalman filter," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 42, no. 1, pp. 174–190, Jan. 2006.

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

[11] 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.

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

[13] I. Gilitschenski, G. Kurz, and U. D. Hanebeck, "Circular Statistics in Bearings-only Sensor Scheduling," in *Proceedings of the 16th International Conference on Information Fusion (Fusion 2013)*, Istanbul, Turkey, Jul. 2013.

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

[15] M. Antone and S. Teller, "Scalable, absolute position recovery for omni-directional image networks," in *Proceedings of the 2001 IEEE Computer Society Conference on Computer Vision and Pattern Recognition. CVPR 2001*, vol. 1. IEEE Comput. Soc, 2001, pp. I–398–I–405.

[16] C. Bingham, "An Antipodally Symmetric Distribution on the Sphere," *The Annals of Statistics*, vol. 2, no. 6, pp. 1201–1225, Nov. 1974.

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

[18] I. Gilitschenski, G. Kurz, S. J. Julier, and U. D. Hanebeck, "Unscented Orientation Estimation Based on the Bingham Distribution," *arXiv preprint: Systems and Control (cs.SY)*, 2013.

[19] J. Glover, "libbingham bingham statistics library," 2013. [Online]. Available: http://code.google.com/p/bingham/

[20] C. S. Herz, "Bessel Functions of Matrix Argument," *Annals of Mathematics*, vol. 61, no. 3, pp. 474–523, May 1955.

[21] P. Koev and A. Edelman, "The Efficient Evaluation of the Hypergeometric Function of a Matrix Argument," *Math. Comp*, vol. 75, pp. 833–846, 2005.

[22] J. Glover, R. Rusu, and G. Bradski, "Monte Carlo Pose Estimation with Quaternion Kernels and the Bingham Distribution," in *Proceedings of Robotics: Science and Systems*, Los Angeles, CA, USA, Jun. 2011.

[23] T. Sei and A. Kume, "Calculating the Normalising Constant of the Bingham Distribution on the Sphere Using the Holonomic Gradient Method," *arXiv:1201.3239*, Apr. 2013.

[24] H. E. Daniels, "Saddlepoint Approximations in Statistics," *The Annals of Mathematical Statistics*, vol. 25, no. 4, pp. 631–650, Dec. 1954.

[25] A. Kume and A. T. A. Wood, "Saddlepoint Approximations for the Bingham and Fisher-Bingham Normalising Constants," *Biometrika*, vol. 92, no. 2, pp. 465–476, Jun. 2005.

[26] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge University Press, 2004.

[27] A. Kume and A. T. Wood, "On the Derivatives of The Normalising Constant of the Bingham Distribution," *Statistics & Probability Letters*, vol. 77, no. 8, pp. 832–837, Apr. 2007.

[28] B. K. P. Horn, "Closed-form solution of absolute orientation using unit quaternions," *Journal of the Optical Society of America A*, vol. 4, no. 4, p. 629, Apr. 1987.