

Fokker-Planck Prediction on the Cylindric Manifold using Tensor Decomposition of a Regular Grid

Daniel Frisch and Uwe D. Hanebeck

Intelligent Sensor-Actuator-Systems Laboratory (ISAS)

Institute for Anthropomatics and Robotics

Karlsruhe Institute of Technology (KIT), Germany

daniel.frisch@kit.edu, uwe.hanebeck@kit.edu

Abstract—The Fokker-Planck propagator is derived for prediction on cylindric manifolds. We exploit the low-rank tensor decomposition technique that is already being used in the Euclidean domain. With only a small change to the finite difference matrix, we can readily apply it to certain manifolds such as the cylinder. Our application example is estimating the angular position and velocity of a rotating shaft. This state estimation problem may seem linear at first glance, but since the underlying state space is nonlinear due to the periodicity of the angular coordinate, it is an inherently nonlinear estimation problem.

I. INTRODUCTION

High-quality nonlinear state estimation requires a density representation that is expressive and facilitates highly accurate prediction and filtering steps with reasonable computational complexity and storage requirements.

A. Assumed Density Particle Filtering

Densities may be represented using a parametric form, such as the Gaussian. Samples thereof allow for approximate nonlinear filtering and prediction [1], [2], where high-quality deterministic samples [3], [4], [5] maximize the posterior accuracy. However, the Gaussian assumption may be too restrictive, such that the results are not satisfying even for a very high number of samples [6].

On nonlinear manifolds, like cylinders, tori, and spheres, one has the additional effect that, due to periodicity, probability mass can “come back from the other side”, see Fig. 1g, thus rendering even linear-looking systems like constant velocity into inherently nonlinear problems. Assumed density particle filtering techniques for manifolds have so far been developed for spheres $\mathbb{S}^2$ [7], [8], [9], [10], the torus $\mathbb{T}^2$ [11], [12], [13], and the special Euclidean groups SE(2) [14], [15] and SE(3) [16], [17].

B. Particle Filter

Another possibility is a purely sample-based representation using purely a point cloud of samples, which is very flexible and can represent arbitrary densities. In general, the prediction step is easy to perform with particles, as they are just propagated by the system equation (and perturbed with some noise). The Bayesian filter step is however difficult as it introduces sample weights, leading to sample degeneracy [18], [19], [20]. And in higher dimensions one requires a large number of samples

to achieve high accuracy. Some of these problems can be mitigated via progressive resampling techniques [21], [22].

C. Regular Grid Representation

The most straightforward density representation is a simple regular grid that discretizes the relevant part of the respective state space. In that case, the filter step is very easy to compute, while the prediction step is more difficult. Prediction can be performed by solving the Fokker-Planck (FP) equation, which has been attempted in various ways [23], [24], [25], [26], [27], [28], [29], [30]. On certain periodic manifolds, prediction can also be done via trigonometric polynomials and the Fourier transform [31], [32]. This has been done on the torus [33], the sphere [34], [35], and the SO(2) [36], [37] and SE(3) [38] domains. Of course, all these methods have a problem with high dimensions, as the number of grid points or voxels (and thereby probability values to store) increases exponentially with the dimension, famously coined the “curse of dimensionality”.

D. Low-Rank Grid Representation

It is however often possible to capture most of the variability using low-rank tensor representations [39], [40]. Like for example in the two-dimensional (2D) case, which we focus on in this work, the probability grid (or tensor) can be associated with a matrix $\mathbf{P}$. We compute its singular value decomposition $\mathbf{P} = \mathbf{U}\mathbf{\Sigma}\mathbf{V}^\top = \sum_{i=1}^N \rho_{u,i} \rho_{v,i}^\top$, where $\rho_{u,i}$ and $\rho_{v,i}$ are the columns of $\mathbf{U}$ and $\mathbf{V}$, multiplied with $\sqrt{\sigma_i}$, respectively, where σ_i are the entries of diagonal $\mathbf{\Sigma}$. If everything is sorted such that the σ_i are appearing in decreasing order, we can obtain a truncated representation $\mathbf{P} \approx \sum_{i=1}^L \rho_{u,i} \rho_{v,i}^\top$ of lower rank $L < N$. The ρ are then called loading vectors, as they are the factors of the low-rank tensor decomposition. The question now is whether we can not only efficiently store the truncated representation, but also perform prediction and filtering steps on it, without the necessity to ever iterate through the full grid. The answer is yes, as has been proposed in [41], [42], [43].

E. Contribution

In this work, we extend the application to nonlinear periodic manifolds, in particular the cylinder. This may be used, for example, to represent a rotating shaft, representing its angular position φ on the periodic dimension and the rotation velocity ω on the real dimension, enabling us to estimate the rotation

speed using angular measurements from an encoder. Nonlinear periodic manifolds are in principle particularly suitable for grid-based techniques, as the periodicity inherently limits the extent of the state space: along periodic dimensions, we only need to decide for the grid resolution but not for the start and end point.

II. PREDICTION STEP

A. System Dynamics

We define the state on a cylindrical manifold as

$$\underline{x} = \begin{bmatrix} \varphi \\ \omega \end{bmatrix} \in \mathbb{S} \times \mathbb{R}, \quad (1)$$

where φ is the angular position defined on unit sphere $\mathbb{S} = \mathbb{R}/(2\pi\mathbb{Z})$ and ω is the angular velocity. Furthermore, analogous to the example in [41, Sec. V], we describe the system dynamics as nearly constant angular velocity

$$d\underline{x} = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \underline{x} dt + q \begin{bmatrix} 0 \\ 1 \end{bmatrix} dw, \quad (2)$$

with standard Wiener process w and its infinitesimal variance q . The FP equation of the time evolution of the probability density $p(\underline{x}, t)$ for this system dynamics is

$$\frac{\partial p}{\partial t} = -\omega \frac{\partial p}{\partial \varphi} + \frac{q}{2} \frac{\partial^2 p}{\partial \omega^2}. \quad (3)$$

The first term describes the “translation” of density mass along the φ direction based on the ω value, and the second part is a diffusion term that causes an overall “broadening” and “Gaussianization” of the density, caused by the additive noise. Note that even though the system dynamics look linear, the problem is inherently nonlinear due to the periodicity of φ .

B. State Space Discretization

We discretize the state space using a regular grid representation, where the angular position φ is represented on a periodic grid and the angular velocity ω on a real-valued grid. For the periodic dimension, we use a grid with N_φ points, which are uniformly spaced on $\mathbb{S}$, and for the real-valued dimension, we use a grid with N_ω points that are uniformly spaced in the interval $[\omega_{\min}, \omega_{\max}]$

$$\varphi_i = \frac{2\pi(i-1)}{N_\varphi}, \quad \omega_j = \omega_{\min} + \frac{(j-1) \cdot (\omega_{\max} - \omega_{\min})}{N_\omega - 1}, \quad (4)$$

for $i \in \{1, \dots, N_\varphi\}$ and $j \in \{1, \dots, N_\omega\}$. These values are stored in vectors $\underline{\varphi}$ and $\underline{\omega}$, respectively. The coordinates of the full grid are then given by the Cartesian product of these two vectors, resulting in a 2D grid of size $N_\varphi \times N_\omega$. The grid spacings are given by

$$\Delta\varphi = \frac{2\pi}{N_\varphi}, \quad \Delta\omega = \frac{\omega_{\max} - \omega_{\min}}{N_\omega - 1}. \quad (5)$$

The probability density $p(\underline{x})$ is represented as a matrix of probabilities $\mathbf{P} \in \mathbb{R}^{N_\varphi \times N_\omega}$, where the entry $\mathbf{P}_{ij}$ corresponds to grid point (φ_i, ω_j) . By introducing a linear index, we can also write the probability density as a vector $\underline{p} \in \mathbb{R}^{N_\varphi \cdot N_\omega}$,

where the entry p_k corresponds to the grid point (φ_i, ω_j) with $k = i + jN_\varphi$. That is, we use column-major order to store the matrix in a vector.

C. Fokker-Planck Equation Discretization

Now we can state the discretized FP equation in the form of a matrix equation. The time evolution of the probability density is given by

$$\frac{\partial \underline{p}}{\partial t} = \mathbb{L} \underline{p}, \quad (6)$$

where $\mathbb{L}$ is the FP-operator. Again following [41, Eq. 29], it can be written as

$$\mathbb{L} = \mathbb{L}_1 + \mathbb{L}_2 \quad (7)$$

$$= -\text{diag}(\underline{\omega}) \otimes \mathbf{D}_\varphi + \mathbf{D}_{\omega\omega} \otimes \left(\frac{q}{2} \mathbf{I}(N_\varphi)\right) \quad (8)$$

$$= \mathbf{A}_{\omega,1} \otimes \mathbf{A}_{\varphi,1} + \mathbf{A}_{\omega,2} \otimes \mathbf{A}_{\varphi,2}, \quad (9)$$

with Kronecker product $\otimes$ and N -dimensional identity matrix $\mathbf{I}(N)$. Thereby $\mathbf{D}_\varphi$ is the first-order finite difference matrix for the first derivative along the periodic dimension

$$\mathbf{D}_\varphi = \frac{1}{2\Delta\varphi} \begin{bmatrix} 0 & 1 & 0 & 0 & 0 & -1 \\ -1 & 0 & 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 1 & 0 & 0 \\ 0 & 0 & \ddots & \ddots & \ddots & 0 \\ 0 & 0 & 0 & -1 & 0 & 1 \\ 1 & 0 & 0 & 0 & -1 & 0 \end{bmatrix}. \quad (10)$$

Note that the entries in the upper-right and lower-left corners are non-zero to account for the periodicity of the angular position. The matrix $\mathbf{D}_{\omega\omega}$ is the second-order finite difference matrix for the second derivative along the real-valued dimension and is given by

$$\mathbf{D}_{\omega\omega} = \frac{1}{\Delta\omega^2} \begin{bmatrix} -2 & 1 & 0 & 0 & 0 & 0 \\ 1 & -2 & 1 & 0 & 0 & 0 \\ 0 & 1 & -2 & 1 & 0 & 0 \\ 0 & 0 & \ddots & \ddots & \ddots & 0 \\ 0 & 0 & 0 & 1 & -2 & 1 \\ 0 & 0 & 0 & 0 & 1 & -2 \end{bmatrix}. \quad (11)$$

Therefore, $\mathbb{L}_1$ is a block diagonal matrix

$$\mathbb{L}_1 = \begin{bmatrix} -\omega_0 \cdot \mathbf{D}_\varphi & 0 & \cdots & 0 \\ 0 & -\omega_1 \cdot \mathbf{D}_\varphi & \cdots & 0 \\ \vdots & 0 & \ddots & 0 \\ 0 & 0 & \cdots & -\omega_{N_\omega-1} \cdot \mathbf{D}_\varphi \end{bmatrix}, \quad (12)$$

and the contribution of $\mathbb{L}_1$ in (6) is taking the derivative of each column of $\mathbf{P}$ and multiplying with $-\omega_j$, corresponding to the first term in (3).

D. Diagonalizations

We are going to need the eigendecompositions $\mathbf{A} = \mathbf{V}\mathbf{D}\mathbf{V}^{-1}$ of the $\mathbf{A}$ matrices in (9), so here we describe how to obtain them.

$\mathbf{A}_{\varphi,1} = \mathbf{D}_{\varphi}$ (10) can be computed in closed form as it is a circulant matrix [44]. With $\gamma_{\varphi} = \exp\{2\pi i/N_{\varphi}\}$, the eigenvector matrix is, for $i, j \in \{1, \dots, N_{\varphi}\}$,

$$[\mathbf{V}_{\varphi,1}]_{i,j} = \frac{1}{\sqrt{N_{\varphi}}} \cdot \gamma_{\varphi}^{(i-1) \cdot (j-1)} . \quad (13)$$

The eigenvalues are, for $i \in \{1, \dots, N_{\varphi}\}$,

$$\lambda_i = \frac{1}{2\Delta_{\varphi}} \left(-\gamma_{\varphi}^{-(i-1)} + \gamma_{\varphi}^{-(N_{\varphi}-1) \cdot (i-1)} \right) \quad (14)$$

$$= \frac{1}{\Delta_{\varphi}} \sinh \left((i-1) \cdot \frac{2\pi i}{N_{\varphi}} \right) . \quad (15)$$

$\mathbf{A}_{\omega,1} = -\text{diag}(\underline{\omega})$ is trivial to diagonalize as it is already a diagonal matrix,

$$\mathbf{V}_{\omega,1} = \mathbf{I}(N_{\omega}) , \quad \mathbf{D}_{\omega,1} = -\text{diag}(\underline{\omega}) . \quad (16)$$

$\mathbf{A}_{\varphi,2} = \frac{q^2}{2} \cdot \mathbf{I}(N_{\varphi})$ is also trivial to diagonalize as it is a scalar multiple of the identity matrix,

$$\mathbf{V}_{\varphi,2} = \mathbf{I}(N_{\varphi}) , \quad \mathbf{D}_{\varphi,2} = \frac{q^2}{2} \cdot \mathbf{I}(N_{\varphi}) . \quad (17)$$

$\mathbf{A}_{\omega,2} = \mathbf{D}_{\omega\omega}$ is a symmetric Toeplitz matrix that we diagonalized by numerically computing the eigendecomposition of the non-periodic difference matrix for second derivatives.

III. FOKKER-PLANCK PREDICTORS

In this section, we describe various approaches to predict the probability density function on the grid using the FP equation.

A. Exact Prediction via Matrix Exponential

We can solve (6) up to numerical accuracy via the exponential of $\mathbb{L}$. The solution is given by

$$\underline{p}(t + \tau) = \exp(\tau \mathbb{L}) \underline{p}(t) , \quad (18)$$

where $\underline{p}(t)$ is the initial probability density, e.g., the last estimate of the filter. The matrix exponential can be computed by the eigendecomposition of $\mathbb{L}$

$$\mathbb{L} = \mathbf{V}\mathbf{D}\mathbf{V}^{-1} , \quad (19)$$

$$\exp(\tau \mathbb{L}) = \mathbf{V} \text{diag}(e^{\lambda_1 \tau}, e^{\lambda_2 \tau}, \dots, e^{\lambda_{N_{\varphi} \cdot N_{\omega}} \tau}) \mathbf{V}^{-1} , \quad (20)$$

where λ_i are the eigenvalues and the entries in diagonal $\mathbf{D}$. Of course, computing the eigendecomposition of a $(N_{\varphi} \cdot N_{\omega} \times N_{\varphi} \cdot N_{\omega})$ matrix can be very expensive.

B. Prediction via Standard ODE Solver

We may also solve the FP equation (6) numerically using an off-the-shelf ordinary differential equation (ODE) solver that uses techniques such as the Runge-Kutta method. However, this involves representing the full grid of probabilities in $p \in \mathbb{R}^{N_{\varphi} \cdot N_{\omega}}$, and matrix multiplications with $\mathbb{L} \in \mathbb{R}^{N_{\varphi} \cdot N_{\omega} \times N_{\varphi} \cdot N_{\omega}}$, which is even square the size of p , albeit sparse. This can lead to significant memory and computational overhead, especially for large grids. Still, it is much faster than the matrix exponential approach, so we will use the ODE method as ground truth.

C. Prediction via Tensorized Predictor

Now we come to the tensorized predictor. This involves reformulating the FP equation in a way that allows us to exploit low-rank structure of the operator and the densities [41], [43]. From

$$\mathbb{L} = \mathbf{A}_{\omega,1} \otimes \mathbf{A}_{\varphi,1} + \mathbf{A}_{\omega,2} \otimes \mathbf{A}_{\varphi,2} , \quad (21)$$

with the Lie-Suzuki-Trotter approximation of first order [45, Eq. 4], [41, Sec. 3] we obtain

$$\exp\{\tau \mathbb{L}\} \quad (22)$$

$$\approx \prod_{l=1}^2 \exp\{\tau \mathbf{A}_{\omega,l} \otimes \mathbf{A}_{\varphi,l}\} \quad (23)$$

$$= \prod_{l=1}^2 \exp\left\{ \tau (\mathbf{V}_{\omega,l} \otimes \mathbf{V}_{\varphi,l}) (\mathbf{D}_{\omega,l} \otimes \mathbf{D}_{\varphi,l}) (\mathbf{V}_{\omega,l}^{-1} \otimes \mathbf{V}_{\varphi,l}^{-1}) \right\} \\ = \prod_{l=1}^2 (\mathbf{V}_{\omega,l} \otimes \mathbf{V}_{\varphi,l}) \exp\{\tau \mathbf{D}_{\omega,l} \otimes \mathbf{D}_{\varphi,l}\} (\mathbf{V}_{\omega,l}^{-1} \otimes \mathbf{V}_{\varphi,l}^{-1})$$

where we used eigendecompositions $\mathbf{A} = \mathbf{V}\mathbf{D}\mathbf{V}^{-1}$ of the matrices $\mathbf{A}_{\omega,l}$ and $\mathbf{A}_{\varphi,l}$, for $l = 1, 2$, respectively, and a Kronecker product property [46, Sec. 10.2.1], [47]. The exponential term is then approximated via a Taylor series expansion [48]

$$\exp\{\tau \mathbf{D}\} \approx \sum_{p=0}^{N_T} \frac{\tau^p}{p!} \cdot \mathbf{D}^p . \quad (24)$$

This leads to

$$\exp\{\tau \mathbb{L}\} \quad (25)$$

$$\approx \sum_{p_1, p_2} \prod_{l=1}^2 (\mathbf{V}_{\omega,l} \otimes \mathbf{V}_{\varphi,l}) (\tilde{\mathbf{D}}_{\omega,l}^{p_l} \otimes \tilde{\mathbf{D}}_{\varphi,l}^{p_l}) (\mathbf{V}_{\omega,l}^{-1} \otimes \mathbf{V}_{\varphi,l}^{-1}) \\ \approx \sum_{p_1, p_2} \left(\prod_{l=1}^2 \mathbf{V}_{\omega,l} \tilde{\mathbf{D}}_{\omega,l}^{p_l} \mathbf{V}_{\omega,l}^{-1} \right) \otimes \left(\prod_{l=1}^2 \mathbf{V}_{\varphi,l} \tilde{\mathbf{D}}_{\varphi,l}^{p_l} \mathbf{V}_{\varphi,l}^{-1} \right) ,$$

where

$$\tilde{\mathbf{D}}^{p_l} = \left(\frac{\tau^{p_l}}{p_l!} \right)^{1/2} \mathbf{D}^{p_l} , \quad (26)$$

for indices (ω, l) and (φ, l) , respectively. The square root comes from dividing the factor in (24) between both matrices.

Now the initial density $\underline{p}(t)$ is introduced. We assume it is given in a low-rank tensor decomposition form

$$\underline{p}(t) = \sum_{l'=1}^{L'} \underline{\rho}_{\omega, l'}(t) \otimes \underline{\rho}_{\varphi, l'}(t) , \quad (27)$$

where $\underline{\rho}_{\varphi, l'}(t) \in \mathbb{R}^{N_{\varphi}}$ and $\underline{\rho}_{\omega, l'}(t) \in \mathbb{R}^{N_{\omega}}$ are the loading vectors of the decomposition and L' is the rank. Then we have

$$\underline{p}(t + \tau) = \exp\{\tau \mathbb{L}\} \underline{p}(t) \quad (28)$$

$$\approx \sum_{l', p_1, p_2} \left(\prod_{l=1}^2 \mathbf{V}_{\omega,l} \tilde{\mathbf{D}}_{\omega,l}^{p_l} \mathbf{V}_{\omega,l}^{-1} \right) \underline{\rho}_{\omega, l'}(t) \quad (29)$$

$$\otimes \left(\prod_{l=1}^2 \mathbf{V}_{\varphi,l} \tilde{\mathbf{D}}_{\varphi,l}^{p_l} \mathbf{V}_{\varphi,l}^{-1} \right) \underline{\rho}_{\varphi,l'}(t) \quad (30)$$

$$\approx \sum_{l',p_1,p_2} \underline{\rho}_{\omega,l'}(t+\tau) \otimes \underline{\rho}_{\varphi,l'}(t+\tau) , \quad (31)$$

where

$$\underline{\rho}_{\varphi,l',p_1,p_2}(t+\tau) = \left(\prod_{l=1}^2 \mathbf{V}_{\varphi,l} \tilde{\mathbf{D}}_{\varphi,l}^{p_l} \mathbf{V}_{\varphi,l}^{-1} \right) \underline{\rho}_{\varphi,l'}(t) , \quad (32)$$

$$\underline{\rho}_{\omega,l',p_1,p_2}(t+\tau) = \left(\prod_{l=1}^2 \mathbf{V}_{\omega,l} \tilde{\mathbf{D}}_{\omega,l}^{p_l} \mathbf{V}_{\omega,l}^{-1} \right) \underline{\rho}_{\omega,l'}(t) . \quad (33)$$

Summarizing, for every dimension (φ, ω) , we can directly transform the L' loading vectors $\underline{\rho}(t)$ of the initial density $\underline{p}(t)$ into $L' \cdot N_T^2$ loading vectors $\underline{\rho}(\tau+t)$ of the evolved density $\underline{p}(\tau+t)$ using the stated operations. To complete this, a rank reduction step may be required, which can be done using a singular value decomposition or similar methods.

D. Step Size and Taylor Order Tradeoff

A practical consideration is the choice of time step τ and the order N_T of the Taylor expansion. Using a small τ allows for a low-order Taylor expansion (small N_T), but requires more prediction sub-steps to reach a given time horizon. Conversely, using a larger τ requires a higher-order Taylor expansion to maintain accuracy, but fewer sub-steps. The optimal choice depends on the desired accuracy, computational resources, and the properties of the FP operator. In practice, a balance must be struck between the number of sub-steps and the computational cost of higher-order expansions. Adaptive step size or order selection strategies may further improve efficiency [49], [50], [51].

IV. FILTER STEP

The filter step via the Bayes formula can also be performed in a tensorized manner [52]. Assume the likelihood $p(\underline{z} | \underline{x})$, for measurement $\underline{z}$, can be stated in decomposed form with “loading functions” $\lambda_{\varphi,l'}(\varphi)$ and $\lambda_{\omega,l'}(\omega)$

$$p(\underline{z} | \varphi, \omega) = \sum_{l'=1}^{L'} \lambda_{\varphi,l'}(\varphi) \cdot \lambda_{\omega,l'}(\omega) . \quad (34)$$

Assume the prior density p^p is also given in continuous form with loading functions ρ

$$p^p(\varphi, \omega) = \sum_{l=1}^L \rho_{\varphi,l}(\varphi) \cdot \rho_{\omega,l}(\omega) . \quad (35)$$

Then the posterior density p^e can be computed using Bayes’ theorem

$$p^e(\varphi, \omega) \quad (36)$$

$$= \frac{1}{c} p(\underline{z} | \varphi, \omega) \cdot p^p(\varphi, \omega) \quad (37)$$

$$= \frac{1}{c} \left(\sum_{l'=1}^{L'} \lambda_{\varphi,l'}(\varphi) \cdot \lambda_{\omega,l'}(\omega) \right) \cdot \left(\sum_{l=1}^L \rho_{\varphi,l}^p(\varphi) \cdot \rho_{\omega,l}^p(\omega) \right)$$

$$= \frac{1}{c} \sum_{l'=1}^{L'} \sum_{l=1}^L \lambda_{\varphi,l'}(\varphi) \cdot \rho_{\varphi,l}^p(\varphi) \cdot \lambda_{\omega,l'}(\omega) \cdot \rho_{\omega,l}^p(\omega) \quad (38)$$

$$= \frac{1}{c} \sum_{l'=1}^{L'} \sum_{l=1}^L \rho_{\varphi,l,l'}^p(\varphi) \cdot \rho_{\omega,l,l'}^p(\omega) . \quad (39)$$

Thus, the number of loading functions in the posterior density has increased to $(L' \cdot L)$. The same is true for the discretized case, where the loading functions of the prior are replaced by their discrete counterparts. The posterior loading vectors are then the Hadamard products of all combinations of the likelihood and prior loading vectors

$$\underline{\rho}_{\varphi,l',l}^e = \lambda_{\varphi,l'}(\varphi) \odot \underline{\rho}_{\varphi,l}^p , \quad (40)$$

$$\underline{\rho}_{\omega,l',l}^e = \lambda_{\omega,l'}(\omega) \odot \underline{\rho}_{\omega,l}^p , \quad (41)$$

with the full posterior density vector

$$\underline{p}^e = \frac{1}{c} \sum_{l'=1}^{L'} \sum_{l=1}^L \underline{\rho}_{\omega,l',l}^e \otimes \underline{\rho}_{\varphi,l',l}^e . \quad (42)$$

The normalization constant c can also be easily computed, as described in Section V-A.

V. MOMENTS

A. Normalization Constant

The FP equation is linear, therefore the prediction does not require a correctly normalized $\underline{p}(t)$. However, for computing the moments of $\underline{p}$, we need to ensure that the density is normalized. The normalization constant of a low-rank tensor can be cheaply computed [43, Appendix], [52]. Assuming the density $p(\varphi, \omega)$ is given in low-rank decomposition form

$$p(\varphi, \omega) = \sum_{l=1}^L \rho_{\omega,l}(\omega) \rho_{\varphi,l}(\varphi) \quad (43)$$

using continuous “loading functions” $\rho_{\varphi}(\varphi), \rho_{\omega}(\omega)$, then its integral is

$$c = \iint \sum_{l=1}^L p(\varphi, \omega) d\varphi d\omega \quad (44)$$

$$= \sum_{l=1}^L \left(\int \rho_{\varphi,l}(\varphi) d\varphi \right) \cdot \left(\int \rho_{\omega,l}(\omega) d\omega \right) \quad (45)$$

$$= \sum_{l=1}^L \left(\sum_{i=1}^{N_{\varphi}} [\underline{\rho}_{\varphi,l}]_i \Delta_{\varphi} \right) \cdot \left(\sum_{j=1}^{N_{\omega}} [\underline{\rho}_{\omega,l}]_j \Delta_{\omega} \right) . \quad (46)$$

In pseudo-code notation, what we actually do with the loading vectors is simply

$$c = \Delta_{\varphi} \cdot \Delta_{\omega} \cdot \sum_{l=1}^L \left[\text{sum}(\underline{\rho}_{\varphi,l}) \cdot \text{sum}(\underline{\rho}_{\omega,l}) \right] . \quad (47)$$

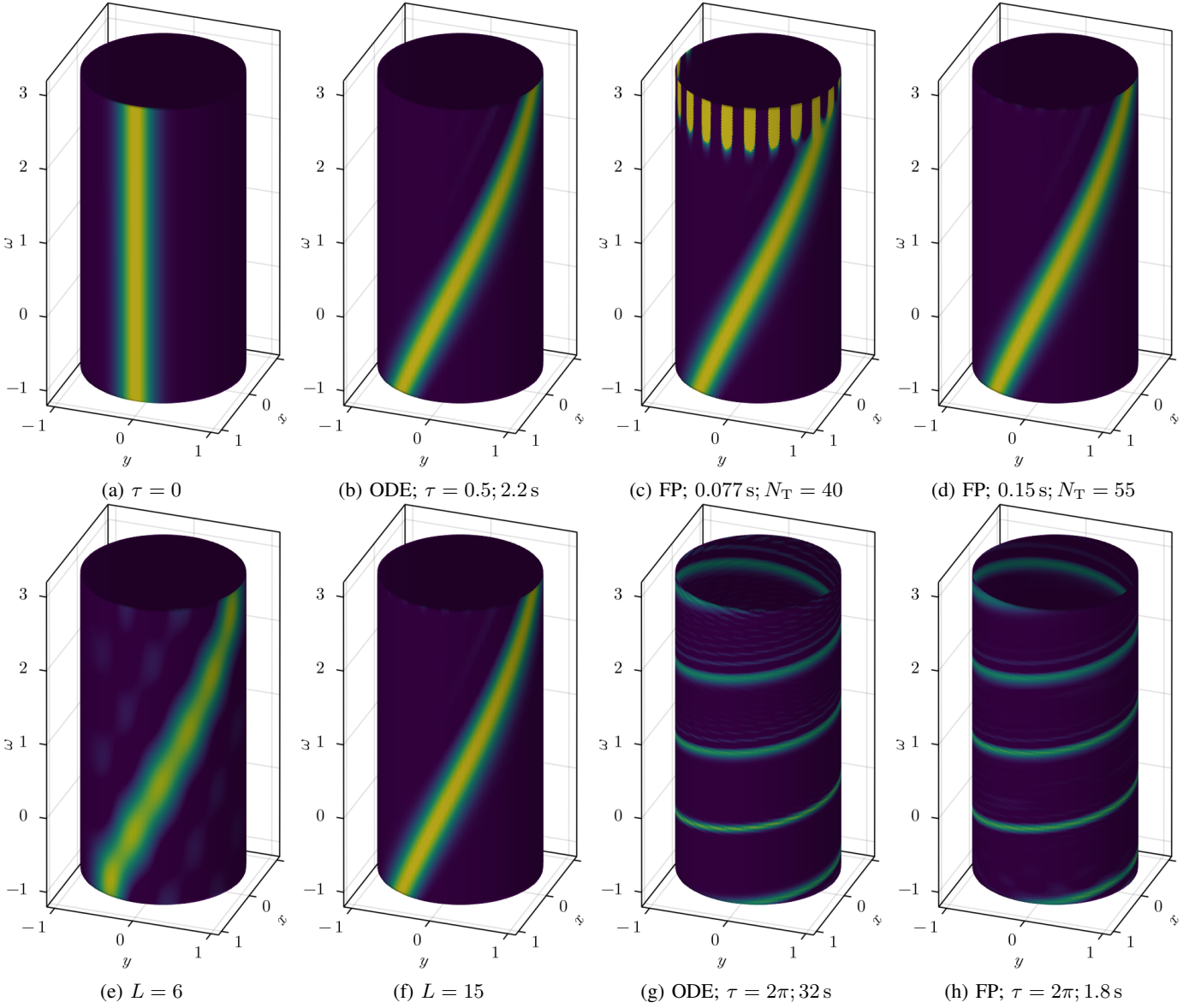


Fig. 1: Cylindric manifold with predicted densities. Areas with high density in yellow. Noted are propagated time τ , computation time (in seconds), order N_T of Taylor expansion, and rank L of representation. (a) initial density, (b) after propagating $\tau = 0.5$ with an off-the-shelf ODE solver, note that areas with positive ω are rotating in positive direction (and vice versa), (c) using the proposed, tensorized solver with $N_T = 40$ and a single step, (d) same but with $N_T = 55$, (e) SVD-reduction to 6 pairs of loading vectors, (f) SVD-reduction to 15 pairs of loading vectors, (g) now propagating to $\tau = 2\pi$ (such that for every unit of ω , one rotation is done) with an off-the-shelf ODE solver, (h) same with the proposed, tensorized solver with 102 sub-steps, $N_T = 4$, and rank 20, computing 17 times faster.

B. Mean of ω

With a similar procedure, the mean of ω can be computed. We have

$$\mathbb{E}\{\omega\} = \iint \sum_{l=1}^L \omega \cdot p(\varphi, \omega) \, d\varphi \, d\omega \quad (48)$$

$$= \sum_{l=1}^L \left(\int \rho_{\varphi,l}(\varphi) \, d\varphi \right) \cdot \left(\int \omega \cdot \rho_{\omega,l}(\omega) \, d\omega \right) \quad (49)$$

$$= \sum_{l=1}^L \left(\sum_{i=1}^{N_\varphi} [\underline{\rho}_{\varphi,l}]_i \Delta_\varphi \right) \cdot \left(\sum_{j=1}^{N_\omega} \omega_j \cdot [\underline{\rho}_{\omega,l}]_j \Delta_\omega \right).$$

Again, in pseudo-code notation, what we have to implement is simply

$$\mathbb{E}\{\omega\} = \Delta_\varphi \cdot \Delta_\omega \cdot \sum_{l=1}^L \left[\text{sum}(\underline{\rho}_{\varphi,l}) \cdot \text{sum}(\underline{\omega} \odot \underline{\rho}_{\omega,l}) \right]. \quad (50)$$

C. Variance of ω

The variance of ω can conveniently be computed from the noncentral second moment $E\{\omega^2\}$ and the mean $E\{\omega\}$ via

$$\text{Var}\{\omega\} = E\{\omega^2\} - (E\{\omega\})^2 . \quad (51)$$

And $E\{\omega^2\}$ can be obtained, very similarly to $E\{\omega\}$, via

$$E\{\omega^2\} \quad (52)$$

$$= \sum_{l=1}^L \left(\sum_{i=1}^{N_\varphi} [\underline{\rho}_{\varphi,l}]_i \Delta\varphi \right) \cdot \left(\sum_{j=1}^{N_\omega} \omega_j^2 \cdot [\underline{\rho}_{\omega,l}]_j \Delta\omega \right) \quad (53)$$

$$= \Delta_\varphi \cdot \Delta_\omega \cdot \sum_{l=1}^L \left[\text{sum}(\underline{\rho}_{\varphi,l}) \cdot \text{sum}(\underline{\omega} \odot \underline{\omega} \odot \underline{\rho}_{\omega,l}) \right] . \quad (54)$$

D. Circular Mean of φ

To compute the circular mean of φ , we first have to compute the first trigonometric moment [53, Sec. 2.4]

$$m_1 = \int_{\varphi=0}^{2\pi} p(\varphi) e^{i\varphi} d\varphi \quad (55)$$

$$= \int_{\varphi=0}^{2\pi} p(\varphi) \cos(\varphi) d\varphi + i \int_{\varphi=0}^{2\pi} p(\varphi) \sin(\varphi) d\varphi , \quad (56)$$

then the mean direction is given by [53, Sec. 2.2], [54, Sec. II.A]

$$E\{\varphi\} = \arg(m_1) . \quad (57)$$

In our particular case, we have the integral

$$m_1 = \iint e^{i\varphi} \cdot p(\varphi, \omega) d\varphi d\omega \quad (58)$$

$$= \sum_{l=1}^L \left(\int e^{i\varphi} \cdot \rho_{\varphi,l}(\varphi) d\varphi \right) \cdot \left(\int \rho_{\omega,l}(\omega) d\omega \right) \quad (59)$$

$$= \sum_{l=1}^L \left(\sum_{i=1}^{N_\varphi} e^{i\varphi_i} \cdot [\underline{\rho}_{\varphi,l}]_i \Delta\varphi \right) \cdot \left(\sum_{j=1}^{N_\omega} [\underline{\rho}_{\omega,l}]_j \Delta\omega \right) \quad (60)$$

$$= \Delta_\varphi \cdot \Delta_\omega \cdot \sum_{l=1}^L \left[\text{sum}(e^{i\varphi} \odot \underline{\rho}_{\varphi,l}) \cdot \text{sum}(\underline{\rho}_{\omega,l}) \right] , \quad (61)$$

and obtain the mean direction

$$E\{\varphi\} = \text{atan2}(\Re(m_1), \Im(m_1)) . \quad (62)$$

E. Circular Variance of φ

The circular variance can, like the circular mean, also be computed from the first trigonometric moment. It is given by [53, Sec. 2.3]

$$\text{Var}\{\varphi\} = 1 - |m_1| . \quad (63)$$

VI. EVALUATION

To demonstrate the method, we compute and visualize a number of densities on the cylinder. We define a 100×100 grid from 0 to 2π in φ and from -1 to 3 in ω . Therefore,

$$\underline{\rho} \in \mathbb{R}^{100} , \quad \mathbf{A} \in \mathbb{R}^{100 \times 100} , \quad \mathbb{L} \in \mathbb{R}^{10000 \times 10000} . \quad (64)$$

A. Initial Density

As an initial density, for φ , we assume the von Mises distribution [55], [56], $p(\varphi) = e^{\kappa \cos(\varphi)} / (2\pi I_0(\kappa))$, with $\kappa = 40$, where $I_0(\kappa)$ is the modified Bessel function of the first kind of order zero, and for ω a noninformative uniform distribution, $p(\omega) = \text{const}$. See Fig. 1a for a visualization. This density can be exactly represented with just one loading vector for φ and ω , i.e., $L' = 1$, where $p(\varphi, \omega) = p(\varphi) \cdot p(\omega)$.

B. Propagation

Then we propagate this density to $\tau = 0.5$ using an off-the-shelf ODE solver [57] and the proposed FP solver with $N_T = 40$ and $N_T = 55$, respectively. The results are shown in Figs. 1b to 1d. We see that in this particular case, one needs a Taylor expansion of order $N_T = 55$ to achieve a similar accuracy as the ODE solver and avoid the artifacts seen in Fig. 1c. In each case the tensorized FP propagator is much faster to compute, see computation times in subcaptions.

C. Post-Processing

With $N_T = 55$ Taylor terms, we obtain $L = 40^2 = 1600$ and $L = 55^2 = 3025$ loading vectors, so the resulting tensor actually has full rank, 100. After approximation to rank $L = 6$ (Fig. 1e) one can see some artifacts, but with $L = 15$ (Fig. 1f), it looks indistinguishable to the ground truth (Fig. 1b). To reduce computational load, this operation should be done via truncated SVD [58], [59].

D. Sub-Steps

Finally, we propagate to $\tau = 2\pi$, where one full rotation is done for every unit of ω . There the nonlinearity of the problem due to the periodicity of the domain becomes fully apparent, as no Gaussian estimator could reproduce the occurring densities. Again we compare the ODE solver (Fig. 1g) to the proposed FP propagator (Fig. 1h). The proposed propagator is much faster to compute, see computation times in the captions. We use 102 sub-steps, each followed by a SVD rank reduction to $L = 20$ loading vectors. Due to the smaller step size, N_T could be reduced from 55 to 4. In general, there is some tradeoff between N_T , number of sub-steps, and representation rank that has to be investigated further.

Julia source code of our implementation will be made available on CodeOcean and linked on the IEEE Xplore page.

VII. CONCLUSION

We derived the low-rank representation of the FP propagator for the constant velocity dynamics on the two-dimensional cylindric manifold, fully respecting the nonlinearity due to the periodicity of the domain. The tensorized solver was at least ten times faster than the an off-the-shelf ODE solver, not to mention an eigenvalue decomposition of the full FP operator matrix $\mathbb{L}$. We demonstrated that the FP propagator is able to efficiently and accurately predict densities on the cylinder. We saw that correlations between the coordinates are well captured and can be represented with only few loading vectors. This method can easily be extended to other system dynamics and similar

Riemannian manifolds in higher dimensions, in particular, the torus and hypertori, and Cartesian products between tori and Euclidean spaces other than the cylinder.

REFERENCES

- [1] Monica F. Bugallo and Petar M. Djurić. “Gaussian Particle Filtering in High-Dimensional Systems”. In: *2014 IEEE Workshop on Statistical Signal Processing (SSP)*. 2014, pp. 129–132. DOI: 10.1109/SSP.2014.6884592.
- [2] K. Ito and K. Xiong. “Gaussian Filters for Nonlinear Filtering Problems”. In: *IEEE Transactions on Automatic Control* 45.5 (2000), pp. 910–927. DOI: 10.1109/9.855552.
- [3] Daniel Frisch and Uwe D. Hanebeck. “The Generalized Fibonacci Grid as Low-Discrepancy Point Set for Optimal Deterministic Gaussian Sampling”. In: *Journal of Advances in Information Fusion* 18.1 (June 2023), pp. 16–34. URL: <https://isif.org/media/generalized-fibonacci-grid-low-discrepancy-point-set-optimal-deterministic-gaussian-sampling>.
- [4] Daniel Frisch and Uwe D. Hanebeck. “Deterministic Gaussian Sampling With Generalized Fibonacci Grids”. In: *Proceedings of the 24th International Conference on Information Fusion (Fusion 2021)*. Sun City, South Africa, Nov. 2021. DOI: 10.23919/FUSION49465.2021.9626975.
- [5] Jannik Steinbring, Martin Pander, and Uwe D. Hanebeck. “The Smart Sampling Kalman Filter with Symmetric Samples”. In: *Journal of Advances in Information Fusion* 11.1 (June 2016), pp. 71–90. URL: <https://isif.org/media/smart-sampling-kalman-filter-symmetric-samples>.
- [6] Leon Winheim and Uwe D. Hanebeck. “BNN Training as State Estimation: A Progressive Filtering Approach”. In: *Proceedings of the 2025 IEEE International Conference on Multisensor Fusion and Integration for Intelligent Systems (MFI 2025)*. College Station, Texas, Sept. 2025.
- [7] Daniel Frisch and Uwe D. Hanebeck. “Deterministic Sampling with Separation of Variables in Spherical Coordinates”. In: *Proceedings of the 28th International Conference on Information Fusion (FUSION 2025)*. Rio de Janeiro, Brazil, July 2025. DOI: 10.23919/FUSION65864.2025.11124009.
- [8] Daniel Frisch and Uwe D. Hanebeck. “Deterministic Von Mises-Fisher Sampling on the Sphere Using Fibonacci Lattices”. In: *Proceedings of the combined IEEE 2023 Symposium Sensor Data Fusion and International Conference on Multisensor Fusion and Integration (SDF-MFI 2023)*. Bonn, Germany, Nov. 2023. DOI: 10.1109/SDF-MFI59545.2023.10361396.
- [9] Kailai Li, Florian Pfaff, and Uwe D. Hanebeck. “Hyperspherical Dirac Mixture Reapproximation”. In: *arXiv preprint arXiv:2110.10411* (Oct. 2021). URL: <https://arxiv.org/abs/2110.10411>.
- [10] Daniel Frisch, Kailai Li, and Uwe D. Hanebeck. “Optimal Reduction of Dirac Mixture Densities on the 2-Sphere”. In: *Proceedings of the 1st Virtual IFAC World Congress (IFAC-V 2020)*. July 2020. DOI: 10.1016/j.ifacol.2020.12.1856.
- [11] Gerhard Kurz and Uwe D. Hanebeck. “Deterministic Sampling on the Torus for Bivariate Circular Estimation”. In: *IEEE Transactions on Aerospace and Electronic Systems* 53.1 (Feb. 2017), pp. 530–534. DOI: 10.1109/TAES.2017.2650079.
- [12] Gerhard Kurz, Florian Pfaff, and Uwe D. Hanebeck. “Nonlinear Toroidal Filtering Based on Bivariate Wrapped Normal Distributions”. In: *Proceedings of the 20th International Conference on Information Fusion (Fusion 2017)*. Xi’an, China, July 2017. DOI: 10.23919/ICIF.2017.8009831.
- [13] Gerhard Kurz and Uwe D. Hanebeck. “Toroidal Information Fusion Based on the Bivariate von Mises Distribution”. In: *Proceedings of the 2015 IEEE International Conference on Multisensor Fusion and Integration for Intelligent Systems (MFI 2015)*. San Diego, California, USA, Sept. 2015. DOI: 10.1109/MFI.2015.7295826.
- [14] Gerhard Kurz, Igor Gilitschenski, and Uwe D. Hanebeck. “The Partially Wrapped Normal Distribution for SE(2) Estimation”. In: *Proceedings of the 2014 IEEE International Conference on Multisensor Fusion and Information Integration (MFI 2014)*. Beijing, China, Sept. 2014. DOI: 10.1109/MFI.2014.6997733.
- [15] Kailai Li, Gerhard Kurz, Lukas Bernreiter, and Uwe 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. DOI: 10.23919/ICIF.2018.8455231.
- [16] Kailai Li, Florian Pfaff, and Uwe D. Hanebeck. “Unscented Dual Quaternion Particle Filter for SE(3) Estimation”. In: *IEEE Control Systems Letters* 5.2 (June 2020), pp. 647–652. DOI: 10.1109/LCSYS.2020.3005066.
- [17] Kailai Li, Florian Pfaff, and Uwe 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), part of 2019 IEEE International Conference on Industrial Cyber Physical Systems (ICPS 2019)*. Taipei, Republic of China, May 2019. DOI: 10.1109/ICPHYS.2019.8780254.
- [18] Christian Musso, Nadia Oudjane, and Francois Le Gland. “Improving Regularised Particle Filters”. In: *Sequential Monte Carlo Methods in Practice*. Ed. by Arnaud Doucet, Nando de Freitas, and Neil Gordon. New York, NY: Springer New York, 2001, pp. 247–271. DOI: 10.1007/978-1-4757-3437-9_12.
- [19] M. S. Arulampalam, S. Maskell, N. Gordon, and T. Clapp. “A Tutorial on Particle Filters for Online Nonlinear/Non-Gaussian Bayesian Tracking”. In: *IEEE Transactions on Signal Processing* 50.2 (2002), pp. 174–188. DOI: 10.1109/78.978374.
- [20] P.M. Djurić, J.H. Kotecha, Jianqui Zhang, Yufei Huang, T. Ghirmai, M.F. Bugallo, and J. Miguez. “Particle Filtering”. In: *IEEE Signal Processing Magazine* 20.5 (2003), pp. 19–38. DOI: 10.1109/MSP.2003.1236770.
- [21] Uwe D. Hanebeck and Martin Pander. “Progressive Bayesian Estimation with Deterministic Particles”. In: *Proceedings of the 19th International Conference on Information Fusion (Fusion 2016)*. Heidelberg, Germany, July 2016. URL: <https://ieeexplore.ieee.org/document/7528132>.
- [22] Uwe D. Hanebeck. “Progressive Bayesian Particle Flows Based on Optimal Transport Map Sequences”. In: *Proceedings of the 26th International Conference on Information Fusion (Fusion 2023)*. Charleston, USA, June 2023. DOI: 10.23919/FUSION52260.2023.10224077.
- [23] S. Challa and Y. Bar-Shalom. “Nonlinear Filter Design Using Fokker-Planck-Kolmogorov Probability Density Evolutions”. In: *IEEE Transactions on Aerospace and Electronic Systems* 36.1 (2000), pp. 309–315. DOI: 10.1109/7.826335.
- [24] A Masud and LA Bergman. “Solution of the Four Dimensional Fokker-Planck Equation: Still a Challenge”. In: *Proceedings of ICOSAR*. 2005, pp. 1911–1916.
- [25] H. Zhang and D. Laneuville. “Grid Based Solution of Zakai Equation with Adaptive Local Refinement for Bearings-only Tracking”. In: *2008 IEEE Aerospace Conference*. 2008, pp. 1–8. DOI: 10.1109/AERO.2008.4526436.
- [26] Mrinal Kumar and Suman Chakravorty. “A Nonlinear Filter Based on Fokker Planck Equation and MCMC Measurement Updates”. In: *49th IEEE Conference on Decision and Control (CDC)*. 2010, pp. 7357–7362. DOI: 10.1109/CDC.2010.5717524.
- [27] Mrinal Kumar and Suman Chakravorty. “Nonlinear Filter Based on the Fokker-Planck Equation”. In: *Journal of Guidance, Control, and Dynamics* 35.1 (2012), pp. 68–79. DOI: 10.2514/1.54070.
- [28] L. Pichler, A. Masud, and L. A. Bergman. “Numerical Solution of the Fokker-Planck Equation by Finite Difference and Finite Element Methods—A Comparative Study”. In: *Computational*

Methods in Stochastic Dynamics: Volume 2. Ed. by Manolis Papadrakakis, George Stefanou, and Vissarion Papadopoulos. Dordrecht: Springer Netherlands, 2013, pp. 69–85. DOI: 10.1007/978-94-007-5134-7_5.

- [29] Gang Qian, Khurram Shafique, and Ping Wang. “Fusion of Nonlinear Motion Dynamics Using Fokker-Planck Equation and Projection Filter”. In: *17th International Conference on Information Fusion (FUSION)*. 2014, pp. 1–7.
- [30] Xiaogang Wang, Wutao Qin, Yu Wang, and Naigang Cui. “Bayesian Filtering Based on Solving the Fokker-Planck Equation”. In: *Aircraft Engineering and Aerospace Technology* 90.2 (Mar. 2018), pp. 312–319. DOI: 10.1108/AEAT-09-2016-0161.
- [31] Florian Pfaff, Kailai Li, and Uwe D. Hanebeck. “Fourier Filters, Grid Filters, and the Fourier-Interpreted Grid Filter”. In: *Proceedings of the 22nd International Conference on Information Fusion (Fusion 2019)*. Ottawa, Canada, July 2019. DOI: 10.23919/FUSION43075.2019.9011318.
- [32] Florian Pfaff, Kailai Li, and Uwe D. Hanebeck. “Conditional Densities and Likelihoods for Hypertoroidal Densities Based on Trigonometric Polynomials”. In: *Proceedings of the 2021 IEEE International Conference on Multisensor Fusion and Integration for Intelligent Systems (MFI 2021)*. Karlsruhe, Germany, Sept. 2021. DOI: 10.1109/MFI52462.2021.9591160.
- [33] Florian Pfaff, Kailai Li, and Uwe D. Hanebeck. “Estimating Correlated Angles Using the Hypertoroidal Grid Filter”. In: *Proceedings of the 2020 IEEE International Conference on Multisensor Fusion and Integration for Intelligent Systems (MFI 2020)*. Virtual, Sept. 2020. DOI: 10.1109/MFI49285.2020.9235220.
- [34] Florian Pfaff, Kailai Li, and Uwe 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. DOI: 10.1016/j.ifacol.2020.12.031.
- [35] Florian Pfaff, Kailai Li, and Uwe D. Hanebeck. “A Hyperhemispherical Grid Filter for Orientation Estimation”. In: *Proceedings of the 23rd International Conference on Information Fusion (Fusion 2020)*. Virtual, July 2020. DOI: 10.23919/FUSION45008.2020.9190611.
- [36] Kailai Li, Florian Pfaff, and Uwe D. Hanebeck. “Grid-Based Quaternion Filter for SO(3) Estimation”. In: *Proceedings of the 2020 European Control Conference (ECC 2020)*. Virtual, May 2020. DOI: 10.23919/ECC51009.2020.9143723.
- [37] Gerhard Kurz, Florian Pfaff, and Uwe D. Hanebeck. “Discretization of SO(3) Using Recursive Tesseract Subdivision”. In: *Proceedings of the 2017 IEEE International Conference on Multisensor Fusion and Integration for Intelligent Systems (MFI 2017)*. Daegu, Republic of Korea, Nov. 2017. DOI: 10.1109/MFI.2017.8170406.
- [38] Florian Pfaff, Kailai Li, and Uwe D. Hanebeck. “The State Space Subdivision Filter for SE(3)”. In: *Proceedings of the 25th International Conference on Information Fusion (Fusion 2022)*. Linköping, Sweden, July 2022. DOI: 10.23919/FUSION49751.2022.9841384.
- [39] Tamara G. Kolda and Brett W. Bader. “Tensor Decompositions and Applications”. In: *SIAM Review* 51.3 (2009), pp. 455–500. DOI: 10.1137/07070111X.
- [40] Frank L. Hitchcock. “The Expression of a Tensor or a Polyadic as a Sum of Products”. In: *Journal of Mathematics and Physics* 6.1-4 (1927), pp. 164–189. DOI: 10.1002/sapm192761164.
- [41] Bruno Demissie, Muhammad Altamash Khan, and Felix Govaers. “Nonlinear Filter Design Using Fokker-Planck Propagator in Kronecker Tensor Format”. In: *2016 19th International Conference on Information Fusion (FUSION)*. 2016, pp. 1–8.
- [42] Muhammad Altamash Khan, Martin Ulmke, Bruno Demissie, Felix Govaers, and Wolfgang Koch. “Combining Log-Homotopy Flow With Tensor Decomposition Based Solution for Fokker-Planck Equation”. In: *2016 19th International Conference on Information Fusion (FUSION)*. 2016, pp. 2229–2236.
- [43] Felix Govaers, Bruno Demissie, Altamash Khan, Martin Ulmke, and Wolfgang Koch. “Tensor Decomposition-Based Multitarget Tracking in Cluttered Environments”. In: *Journal of Advances in Information Fusion* 14.1 (2019), p. 86. URL: <https://isif.org/media/tensor-decomposition-based-multitarget-tracking-cluttered-environments>.
- [44] Robert M. Gray. “Toeplitz and Circulant Matrices: A Review”. In: *Foundations and Trends® in Communications and Information Theory* 2.3 (2006), pp. 155–239. DOI: 10.1561/0100000006.
- [45] Masuo Suzuki. “Fractal decomposition of exponential operators with applications to many-body theories and Monte Carlo simulations”. In: *Physics Letters A* 146.6 (1990), pp. 319–323. DOI: [https://doi.org/10.1016/0375-9601\(90\)90962-N](https://doi.org/10.1016/0375-9601(90)90962-N).
- [46] K. B. Petersen and M. S. Pedersen. *The Matrix Cookbook*. Nov. 2012. URL: <http://www2.compute.dtu.dk/pubdb/pubs/3274-full.html>.
- [47] Charles F. Van Loan. “The Ubiquitous Kronecker Product”. In: *Journal of Computational and Applied Mathematics* 123.1 (2000), pp. 85–100. DOI: 10.1016/S0377-0427(00)00393-9.
- [48] Cleve Moler and Charles Van Loan. “Nineteen Dubious Ways to Compute the Exponential of a Matrix”. In: *SIAM Review* 20.4 (1978), pp. 801–836. DOI: 10.1137/1020098.
- [49] George Corliss and David Lowery. “Choosing a Stepsize for Taylor Series Methods for Solving ODE’S”. In: *Journal of Computational and Applied Mathematics* 3.4 (1977), pp. 251–256. DOI: 10.1016/S0377-0427(77)80016-2.
- [50] Gustaf Söderlind. “Automatic Control and Adaptive Time-Stepping”. In: *Numerical Algorithms* 31.1 (2002), pp. 281–310. DOI: 10.1023/A:1021160023092.
- [51] J.C. Butcher. “Numerical methods for ordinary differential equations in the 20th century”. In: *Journal of Computational and Applied Mathematics* 125.1 (2000), pp. 1–29. DOI: 10.1016/S0377-0427(00)00455-6.
- [52] Yifei Sun and Mrinal Kumar. “Nonlinear Bayesian Filtering Based on Fokker-Planck Equation and Tensor Decomposition”. In: *2015 18th International Conference on Information Fusion (Fusion)*. 2015, pp. 1483–1488.
- [53] Kanti V Mardia and Peter E Jupp. *Directional Statistics*. John Wiley & Sons, 2009.
- [54] Florian Pfaff, Gerhard Kurz, and Uwe D. Hanebeck. “Multi-modal Circular Filtering Using Fourier Series”. In: *Proceedings of the 18th International Conference on Information Fusion (Fusion 2015)*. Washington D.C., USA, July 2015. URL: <https://ieeexplore.ieee.org/document/7266630>.
- [55] K.V. Mardia. “Angular Data and Frequency Distributions”. In: *Statistics of Directional Data*. Ed. by K.V. MARDIA. Probability and Mathematical Statistics: A Series of Monographs and Textbooks. Academic Press, 1972, pp. 1–17. DOI: 10.1016/B978-0-12-471150-1.50007-8.
- [56] Ronald Fisher. “Dispersion on a Sphere”. In: *Proceedings of the Royal Society of London. Series A, Mathematical and Physical Sciences* 217.1130 (1953), pp. 295–305. URL: <http://www.jstor.org/stable/99186>.
- [57] Christopher Rackauckas and Qing Nie. “DifferentialEquations.jl—a Performant and Feature-Rich Ecosystem for Solving Differential Equations in Julia”. In: *Journal of Open Research Software* 5.1 (2017). DOI: 10.5334/jors.151.
- [58] Rasmus Munk Larsen. “Lanczos Bidiagonalization With Partial Reorthogonalization”. In: *DAIMI Report Series* 27.537 (Dec. 1998). DOI: 10.7146/dpb.v27i537.7070.
- [59] N. Halko, P. G. Martinsson, and J. A. Tropp. “Finding Structure with Randomness: Probabilistic Algorithms for Constructing Approximate Matrix Decompositions”. In: *SIAM Review* 53.2 (2011), pp. 217–288. DOI: 10.1137/090771806.