

Nonlinear Measurement Update for Estimation of Angular Systems Based on Circular Distributions

Gerhard Kurz¹, Igor Gilitschenski¹, and Uwe D. Hanebeck¹

Abstract—In this paper, we propose a novel progressive nonlinear measurement update for circular states. This generalizes our previously published circular filter that so far was limited to identity measurement equations. The new update method is based on circular distributions in order to capture the periodic properties of a circular system better than conventional approaches that rely on standard Gaussian distributions. Besides the progressive measurement update, we propose two additional measurement updates that are obtained by adapting traditional filters to the circular case. Simulations show the superiority of the proposed progressive approach.

I. INTRODUCTION

Many applications involve the estimation of angular quantities based on noisy measurements. Typical examples include the orientation of a vehicle, the wind direction, or the angle of a robotic joint. Sometimes these angles can be observed directly, but this is usually not the case. In certain applications, an angle has to be estimated but only a quantity that depends nonlinearly on the angle can be observed. This issue poses the problem of angular estimation with nonlinear measurement functions.

Traditional approaches to estimation of angular quantities are typically based on the use of classical filters such as the Kalman filter [1] or nonlinear extensions thereof such as the extended Kalman filter (EKF) or the unscented Kalman filter (UKF) [2]. However, classical approaches are usually based on the assumption of a Gaussian distribution, which is inaccurate in the angular case. There are two critical issues with these approaches. First, they neglect the periodic nature of the circle and can cause problems when the periodic boundary is crossed. Second, the true probability distribution has a different shape, which is not precisely Gaussian. Directional statistics [3], [4] is a subdiscipline of statistics that deals with directional rather than discrete or real-valued quantities. It can be used to remedy these problems by relying on probability distributions defined on the unit circle.

In 2009, Azmani et. al introduced a recursive angular filter based on the von Mises distribution [5], [6]. However, this filter is limited to identity system and measurement functions. In 2013, we published an angular filter based on the von Mises distribution, the wrapped normal distribution, and the wrapped Dirac mixture distribution. This filter can deal with arbitrary nonlinear system functions while still requiring the measurement function to be the identity [7]. We have applied

the filter to constrained object tracking [8] and the sampling scheme has been used for sensor scheduling [9].

In this paper, we introduce an extended angular filter based on the wrapped normal distribution, and the wrapped Dirac mixture distribution. It can handle both nonlinear system and measurement functions. Moreover, the von Mises distribution is no longer used, which makes the filter computationally more efficient and easier to implement, since numerical computations involving Bessel functions can be avoided. Consequently, all computations can be performed in closed form, which is an advantage compared to most circular filters.

II. RECURSIVE FILTER

We consider a dynamic system with the angular state $x_k \in [0, 2\pi)$ at time step k . The system equation is given by

$$x_{k+1} = a(x_k) + w_k \mod 2\pi$$

with nonlinear system function $a : [0, 2\pi) \rightarrow [0, 2\pi)$ and additive system noise $w_k \sim f^w(\cdot)$. The measurement equation is given by

$$\hat{z}_k = h(x_k) + v_k$$

with nonlinear measurement function $h : [0, 2\pi) \rightarrow H$ and additive noise $v_k \sim f^v(\cdot)$.

The measurement space H may be the unit circle S^1 or the real vector space \mathbb{R}^n , but the presented principles are more general and can be applied to other measurement spaces as well. Here, we restrict ourselves to additive noise, because this allows to calculate the likelihood $f(\hat{z}_k|x_k)$ according to

$$f(\hat{z}_k|x_k) = f^v(\hat{z}_k - h(x_k)) ,$$

where f^v is the probability density function (pdf) of the noise (see Appendix A). If the noise is not additive, calculation of the likelihood may be more complicated. However, the proposed algorithms can be applied to any type of noise as long as it is possible to calculate the likelihood explicitly.

In the following, we derive a recursive filtering algorithm to estimate the system state x_k . For this purpose, we need to perform two operations, prediction and update. The prediction step takes the previously estimated state x_{k-1}^e and obtains the predicted state x_k^p according to the system function. Subsequently, the update step calculates the estimated state x_k^e by including the measurement \hat{z}_k .

III. PREREQUISITES

Before we describe our contribution, we introduce the relevant concepts from circular statistics.

¹ G. Kurz, I. Gilitschenski, and U. D. Hanebeck are with the Intelligent Sensor-Actuator-Systems Laboratory (ISAS), Institute for Anthropomatics and Robotics, Karlsruhe Institute of Technology, Germany (e-mail: gerhard.kurz@kit.edu, gilitschenski@kit.edu, uwe.hanebeck@ieee.org).

A. Continuous and Discrete Probability Distributions

In this section, we define the probability distributions required for the proposed filtering algorithm and show some of their relevant properties.

Definition 1 (Wrapped normal distribution)

A wrapped normal (WN) distribution is given by its pdf

$$f(x; \mu, \sigma) = \frac{1}{\sqrt{2\pi}\sigma} \sum_{k=-\infty}^{\infty} \exp\left(-\frac{(x - \mu + 2k\pi)^2}{2\sigma^2}\right)$$

with parameters $\mu \in [0, 2\pi)$ and $\sigma > 0$.

The WN distribution is obtained by taking a Gaussian distribution and wrapping it around the unit circle while summing up the probabilities for equivalent angles. This distribution is of particular interest because it fulfills a central limit theorem in the circular case. For this reason, it is reasonable to assume that angular noise is WN distributed.

WN distributions are closed under convolution (i.e., addition of independent random variables). For two WN distributions with parameters (μ_1, σ_1) and (μ_2, σ_2) the convolution $f(x; \mu_1, \sigma_1) * f(x; \mu_2, \sigma_2)$ is given by a WN distribution (μ, σ) with parameters $\mu = \mu_1 + \mu_2 \bmod 2\pi$ and $\sigma = \sqrt{\sigma_1^2 + \sigma_2^2}$.

Definition 2 (Wrapped Dirac mixture distribution)

A wrapped Dirac mixture (WD) distribution with L components is given by its pdf

$$f(x; \beta_1, \dots, \beta_L, w_1, \dots, w_L) = \sum_{j=1}^L w_j \delta(x - \beta_j)$$

with Dirac positions $\beta_1, \dots, \beta_L \in [0, 2\pi)$ and weighting coefficients $w_1, \dots, w_L \geq 0$ with $\sum_{j=1}^L w_j = 1$.

Unlike the continuous WN distribution, the WD distribution is a discrete distribution. It can be imagined as a collection of weighted samples of the true distribution. This makes it easy to propagate a WD distribution through a nonlinear function. Consequently, we use the WD distribution for certain intermediate steps in the prediction as well as update algorithms.

The parameters of a WN distribution can be estimated by matching the first circular moment of a WD distribution as described in Appendix B.

Definition 3 (Circular moment)

For a random variable $x \in S^1$ distributed according to the pdf $f(\cdot)$, the n -th circular moment is given by

$$\mathbb{E}(\exp(ix)^n) = \int_0^{2\pi} \exp(inx) f(x) dx \in \mathbb{C},$$

where i is the imaginary unit.

Lemma 1 The n -th circular moment of a WN distribution with parameters (μ, σ) is given by

$$\exp\left(in\mu - \frac{n^2\sigma^2}{2}\right).$$

The n -th circular moment of a WD distribution with parameters $(\beta_1, \dots, \beta_L, w_1, \dots, w_L)$ is given by

$$\sum_{j=1}^L w_j \exp(in\beta_j).$$

B. Nondeterministic Sampling

In order to obtain a WD distribution from a WN distribution, we draw samples from the WN distribution. We distinguish between two possible approaches, nondeterministic and deterministic sampling.

Nondeterministic sampling refers to methods that draw samples randomly from a probability distribution. For example, random numbers could be generated according to a Gaussian distribution [10].

To sample circular distributions, usually the Metropolis-Hastings algorithm can be applied [11]. However, in the case of a WN distribution, nondeterministic sampling can be reduced to sampling of a Gaussian distribution with identical parameters (μ, σ) and considering the samples modulo 2π .

C. Deterministic Sampling

Unlike nondeterministic sampling where samples are drawn randomly, deterministic sampling tries to choose samples that approximate the original distribution according to some similarity measure. The advantage is that a much smaller number of samples is usually sufficient to get a good approximation of the original density. Furthermore, deterministic sampling allows the derivation of algorithms that yield reproducible results.

Deterministic sampling can, for example, be performed by moment matching as in the case of the UKF [2]. Rather than conventional moment matching, we use circular moment matching, which we previously developed in [7]. For $L = 3$ Dirac components, the deterministic samples of a WN distribution can be calculated in closed form (see Appendix B).

IV. NONLINEAR MEASUREMENT UPDATE

Since the prediction step has previously been published in [7], we will focus on the measurement update. For the sake of completeness, we give the prediction step in Appendix C.

In order to perform the nonlinear measurement update, we use an approach that we call WN assumed density filtering. These are similar to Gaussian assumed density filters such as [12], which rely on the assumption that the posterior density $f(x_k | \hat{z}_k)$ is Gaussian. In a similar fashion, we assume the posterior density $f(x_k | \hat{z}_k)$ to be a WN density.

According to Bayes' rule, we get

$$f(x_k | \hat{z}_k) = \frac{f(\hat{z}_k | x_k) \cdot f(x_k)}{f(\hat{z}_k)} \propto \underbrace{f(\hat{z}_k | x_k)}_{\text{likelihood}} \cdot \underbrace{f(x_k)}_{\text{prior}}.$$

Since we assume that the likelihood can be calculated explicitly, we only have to multiply it with the prior density and fit a WN density to the result. As multiplying the likelihood and the prior WN density can not be evaluated analytically, we approximate the prior WN density with a WD distribution and perform the multiplication componentwise.

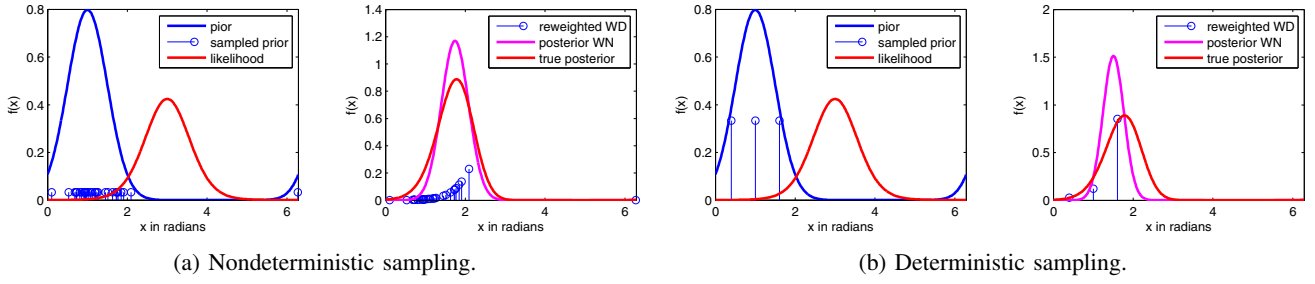


Fig. 1: Example of measurement update

We introduce three different methods that are all based on this idea and that have different advantages and disadvantages.

A. Nondeterministic Update

To perform the nondeterministic update, we draw a certain number of L random samples from the original distribution. The WD distribution is reweighted by multiplying each component with the likelihood $f(\hat{z}_k|x_k)$ and subsequently renormalized. Finally, a WN distribution is fitted to the reweighted WD distribution. An example with $L = 30$ is depicted in Fig. 1a and the algorithm is given in Algo. 1. If we assume that the likelihood can be evaluated in constant time, the overall runtime of the algorithm is $\mathcal{O}(L)$, i.e., it increases linearly with the number of samples.

This algorithm bears close resemblance to the Gaussian particle filter [13] except that we use WN distributions instead of Gaussians. For this reason, it also shares its weaknesses, namely the fact that a much larger number of samples is needed to match the quality of deterministic sampling. Furthermore, the algorithm still fails if no samples are located in an area in which the likelihood function has large values. This can happen when the likelihood is very narrow or far away from the prior state estimate. Even worse, the filter can fail under good circumstances with a certain (low) probability because the samples are drawn in an unfortunate way and are not representative of the true density.

Input: \hat{z}_k (measurement),
 (μ_k^p, σ_k^p) (predicted distribution of state),
 $f(\hat{z}_k|x_k)$ (likelihood function),
 L (number of samples)
Output: (μ_k^e, σ_k^e) (estimated distribution of state)
 $(\beta_1, \dots, \beta_L, w_1, \dots, w_L) \leftarrow \text{sampleNondeterm}(\mu_k^p, \sigma_k^p)$;
for $j \leftarrow 1$ **to** L **do**
 $w_j \leftarrow w_j \cdot f(\hat{z}_k|\beta_j)$;
end
 $(\mu_k^e, \sigma_k^e) \leftarrow \text{matchWn}(\beta_1, \dots, \beta_L, w_1, \dots, w_L)$;
Algorithm 1: Nondeterministic measurement update.

B. Naive Deterministic Update

A naive and computationally fast method is to use the previous approach in conjunction with deterministic sampling. First, the prior WN distribution is sampled deterministically to obtain a WD distribution with $L = 3$ components. The rest of the algorithm remains the same. An example is depicted

in Fig. 1b and the algorithm is given in Algo. 2. If the likelihood can be evaluated in constant time, the runtime of this algorithm is in $\mathcal{O}(1)$.

The disadvantage of this algorithm becomes obvious when looking at examples where at least two of the three samples are located in area in which the likelihood function has small values. In this case, the algorithm tends to become numerically unstable or the posterior distribution is estimated poorly. In the context of particle filters, this phenomenon is referred to as particle degeneracy [14].

Input: \hat{z}_k (measurement),
 (μ_k^p, σ_k^p) (predicted distribution of state),
 $f(\hat{z}_k|x_k)$ (likelihood function)
Output: (μ_k^e, σ_k^e) (estimated distribution of state)
 $(\beta_1, \dots, \beta_3, w_1, \dots, w_3) \leftarrow \text{sampleDeterm}(\mu_k^p, \sigma_k^p)$;
for $j \leftarrow 1$ **to** 3 **do**
 $w_j \leftarrow w_j \cdot f(\hat{z}_k|\beta_j)$;
end
 $(\mu_k^e, \sigma_k^e) \leftarrow \text{matchWn}(\beta_1, \dots, \beta_3, w_1, \dots, w_3)$;
Algorithm 2: Naive deterministic measurement update.

C. Progressive Deterministic Update

Because of the drawbacks of the previous algorithms, we introduce a third approach. This approach extends the deterministic method by a progressive filtering scheme that gradually includes the likelihood. The proposed algorithm constitutes a circular version of the so called progressive Gaussian filter (PGF42) [15]. A similar approach is also used by the progressive Dirac mixture filter [16].

The key idea is to take Bayes' rule and to express the likelihood as a product of likelihoods

$$\begin{aligned} f(x_k|\hat{z}_k) &\propto f(\hat{z}_k|x_k) \cdot f(x_k) \\ &= (f(\hat{z}_k|x_k)^{\lambda_1} \cdot \dots \cdot f(\hat{z}_k|x_k)^{\lambda_s}) \cdot f(x_k) , \end{aligned}$$

where $\Lambda := \sum_{j=1}^s \lambda_j = 1$. By choosing the values of $\lambda_1, \dots, \lambda_s$ appropriately, it can be ensured that reweighting does not cause particle degeneracy. More precisely, we can guarantee that the quotient between the smallest new weight l_{\min} and the largest new weight l_{\max} does not fall below some predetermined threshold τ , i.e.,

$$\frac{l_{\min}}{l_{\max}} \geq \tau .$$

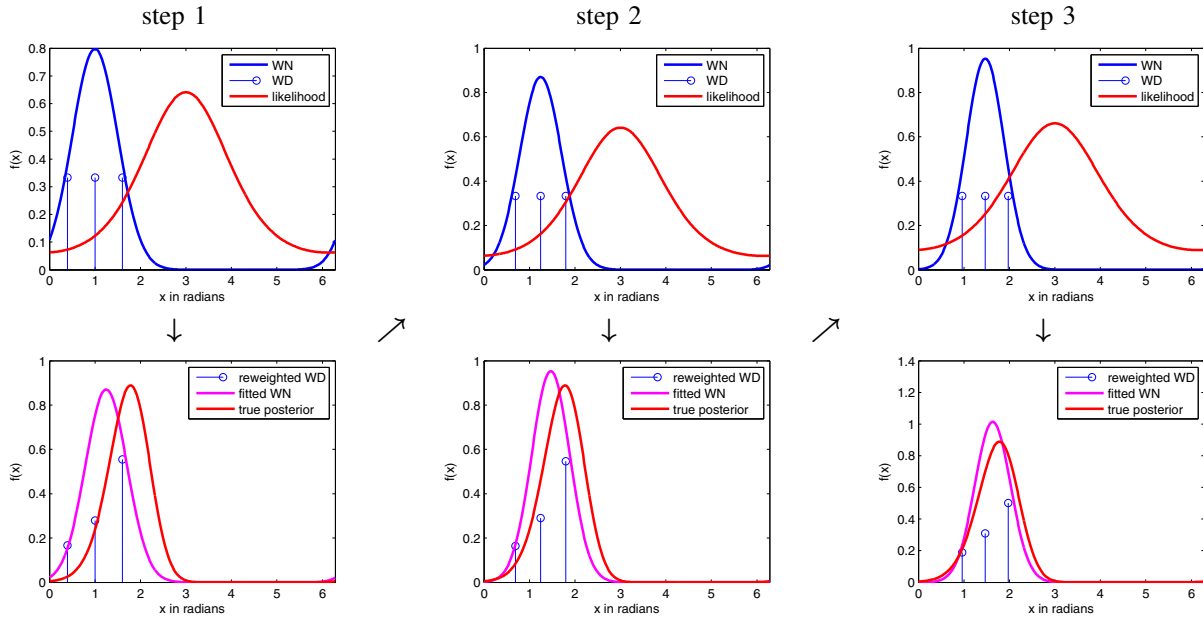


Fig. 2: Example of measurement update with progressive sampling. On the top left, the original prior is depicted and on the bottom right the final result can be seen. The threshold is chosen as $\tau = 0.3$ here.

Since we start with a uniformly weighted distribution, we have

$$\begin{aligned} \frac{l_{\min}}{l_{\max}} &= \frac{\min_{j=1,\dots,3} f(\hat{z}_k|\beta_j)^\lambda}{\max_{j=1,\dots,3} f(\hat{z}_k|\beta_j)^\lambda} \\ &= \left(\frac{\min_{j=1,\dots,3} f(\hat{z}_k|\beta_j)}{\max_{j=1,\dots,3} f(\hat{z}_k|\beta_j)} \right)^\lambda \geq \tau, \end{aligned}$$

which leads to

$$\lambda \leq \frac{\log(\tau)}{\log\left(\frac{\min_{j=1,\dots,3} f(\hat{z}_k|\beta_j)}{\max_{j=1,\dots,3} f(\hat{z}_k|\beta_j)}\right)}.$$

From this inequality, we can calculate the largest permissible step size λ . After obtaining the reweighted WD distribution this way, we fit a WN distribution and use deterministic sampling to obtain a new WD distribution, where all components have equal weight. Obviously, each of these steps is an approximation. The process is repeated until $\lambda_1, \dots, \lambda_s$ sum up to one. The algorithm is given in Algo. 3 and illustrated in Fig. 2.

V. EVALUATION

In order to evaluate the proposed methods, we performed Monte Carlo simulations in a scenario with two-dimensional measurements. We use WN distributed noise w_k and the system function

$$a_k(x_k) = x_k + c_1 \sin(x_k) + c_2.$$

This equation is motivated by a physical model of a robotic arm affected by gravity, which rotates around a joint [7]. The measurement equation is given by

$$\hat{z}_k = \begin{bmatrix} \cos(x_k) \\ \sin(x_k) \end{bmatrix} + \underline{v}_k$$

Input: \hat{z}_k (measurement),
 (μ_k^p, σ_k^p) (predicted distribution of state),
 $f(\hat{z}_k|x_k)$ (likelihood function), threshold parameter τ
Output: (μ_k^e, σ_k^e) (estimated distribution of state)
 $\Lambda \leftarrow 1$;
 $(\mu, \sigma) \leftarrow (\mu_k^p, \sigma_k^p)$;
while $\Lambda > 0$ **do**
 $(\beta_1, \dots, \beta_3, w_1, \dots, w_3) \leftarrow \text{sampleDeterm}(\mu, \sigma)$;
 $w_{\min} \leftarrow \min_{j=1,\dots,3} (f(\hat{z}_k|\beta_j))$;
 $w_{\max} \leftarrow \max_{j=1,\dots,3} (f(\hat{z}_k|\beta_j))$;
 $\lambda \leftarrow \min\left(\Lambda, \frac{\log(\tau)}{\log(w_{\min}/w_{\max})}\right)$;
 for $j \leftarrow 1$ **to** 3 **do**
 $w_j \leftarrow w_j \cdot f(\hat{z}_k|\beta_j)^\lambda$;
 end
 $(\mu, \sigma) \leftarrow \text{matchWn}(\beta_1, \dots, \beta_3, w_1, \dots, w_3)$;
 $\Lambda \leftarrow \Lambda - \lambda$;
end
 $(\mu_k^e, \sigma_k^e) \leftarrow (\mu, \sigma)$;
Algorithm 3: Progressive measurement update.

with Gaussian noise $\underline{v}_k \sim \mathcal{N}(\underline{0}, \mathbf{C}_k^v)$. This can be imagined as the projection of the end of the robot arm onto the x and y axes.

For the system function, we use the parameters $c_1 = 0.1, c_2 = 0.15$. The system noise is distributed according to a WN distribution with parameters $(\mu_k^w, \sigma_k^w) = (0, 0.2)$ and the initial estimate is a WN distribution with parameters $(\mu_0^e, \sigma_0^e) = (1, 1)$. We simulated 1000 Monte Carlo runs, each with a length of 100 time steps. For the nondeterministic filter, we used $L = 100$ samples and for the progressive filter, we used the threshold $\tau = 0.2$.

Since the main issue here is the evaluation of the nonlinear measurement update, we consider three different measurement noise levels, large, medium and small, which have covariances $\mathbf{C}_k^v = \text{diag}(1, 1)$, $\mathbf{C}_k^v = \text{diag}(0.1, 0.1)$, and $\mathbf{C}_k^v = \text{diag}(0.01, 0.01)$, respectively. An example run with large measurement noise is depicted in Fig. 3.

To provide a comparison between the proposed filtering algorithms, we compared the angular root mean square error (RMSE)

$$\sqrt{\frac{1}{100} \sum_{k=1}^{100} (\min(|\hat{x}_k - \mu_k^e|, 2\pi - |\hat{x}_k - \mu_k^e|))^2}$$

where \hat{x}_k is the true state at time k .

The results are depicted in Fig. 4. The nondeterministic filter and the progressive filter work well in all cases, but the naive deterministic filter fails for the case of small noise because of particle degeneration. Not surprisingly, the naive deterministic filter is the fastest. The progressive filter is faster than the nondeterministic filter because it uses significantly fewer samples. Thus, the progressive filter is a good choice regarding both performance and estimation quality.

VI. CONCLUSION

In this paper, we have presented a complete circular filter that can handle both nonlinear system and measurement functions. All required algorithms have been given in pseudo code to facilitate easy implementation. In particular, we would like to emphasize that all calculations are performed in closed form and there is no need for numerical approximations or the evaluation of complicated functions, such as the Bessel functions usually present in von Mises based approaches.

We have evaluated our results in simulations. The different proposed approaches have distinct advantages and disadvantages, but the progressive approach is shown to be superior in that it provides overall good results while being computationally efficient.

Future work may include deterministic sampling with more than three Dirac components and the extension to filtering multiple dependent angles.

ACKNOWLEDGMENT

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

APPENDIX

A. Derivation of Likelihood

The likelihood for additive noise is given according to

$$\begin{aligned} f(\hat{z}_k | x_k) &= \int_H f(\hat{z}_k, \underline{v}_k | x_k) d\underline{v}_k \\ &= \int_H f(\hat{z}_k | x_k, \underline{v}_k) f^v(\underline{v}_k) d\underline{v}_k \\ &= \int_H \delta(\hat{z}_k - h(x_k) - \underline{v}_k) f^v(\underline{v}_k) d\underline{v}_k \\ &= f^v(\hat{z}_k - h(x_k)). \end{aligned}$$

B. Moment Matching Between WN and WD

As shown in [7], the parameters (μ, σ) of a WN distribution can be estimated by matching the first circular moment of a WD distribution $(\beta_1, \dots, \beta_L, w_1, \dots, w_L)$ according to

$$\mu = \text{atan2} \left(\sum_{j=1}^L w_j \sin(\beta_j), \sum_{j=1}^L w_j \cos(\beta_j) \right)$$

and

$$\sigma = \sqrt{-2 \log \left(\sum_{j=1}^L w_j \cos(\beta_j - \mu_j) \right)}.$$

A WN distribution with parameters (μ, σ) can be approximated by the WD distribution with $L = 3$ components and parameters $w_1 = w_2 = w_3 = 1/3$ and

$$\beta_1 = \mu - \alpha, \quad \beta_2 = \mu, \quad \beta_3 = \mu + \alpha$$

where $\alpha = \arccos \left(\frac{3}{2} \exp \left(-\frac{\sigma^2}{2} \right) - \frac{1}{2} \right)$.

C. Nonlinear Prediction

Nonlinear prediction is performed in the same fashion as introduced in [7]. We assume w_k to be WN distributed with parameters (μ_k^w, σ_k^w) . The prediction can be performed as follows. First, the WN distribution representing the previous estimate is sampled deterministically to obtain a WD distribution with three components. Then, the samples are propagated through the nonlinear function $a_k(\cdot)$ and used to estimate the parameters of a new WN distribution. Finally, this distribution is convolved with the noise distribution (μ_k^w, σ_k^w) .

Input: (μ_k^e, σ_k^e) (estimated distribution of state),

$a_k(\cdot)$ (system function)

(μ_k^w, σ_k^w) (noise parameters)

Output: $(\mu_{k+1}^p, \sigma_{k+1}^p)$ (predicted distribution of state)

$(\beta_1, \dots, \beta_3, w_1, \dots, w_3) \leftarrow \text{sampleDeterm}(\mu_k^e, \sigma_k^e)$;

for $j \leftarrow 1$ **to** 3 **do**

$\beta_j \leftarrow a_k(\beta_j)$;

end

$(\mu, \sigma) \leftarrow \text{matchWn}(\beta_1, \dots, \beta_3, w_1, \dots, w_3)$;

$(\mu_k^p, \sigma_k^p) \leftarrow (\mu, \sigma) * (\mu_k^w, \sigma_k^w)$;

Algorithm 4: Prediction.

The complete algorithm for prediction is given in Algo. 4. If we assume that a_k can be evaluated in constant time, the runtime of the algorithm is in $\mathcal{O}(1)$, as we use a fixed number of Dirac components. All operations (except possibly the calculation of a_k) can be evaluated in closed form.

REFERENCES

- [1] R. E. Kalman, “A new approach to linear filtering and prediction problems,” *Transactions of the ASME Journal of Basic Engineering*, vol. 82, pp. 35–45, 1960.
- [2] S. J. Julier and J. K. Uhlmann, “Unscented filtering and nonlinear estimation,” *Proceedings of the IEEE*, vol. 92, no. 3, pp. 401–422, Mar. 2004.
- [3] K. V. Mardia and P. E. Jupp, *Directional Statistics*, 1st ed. Wiley, 1999.
- [4] S. R. Jammalamadaka and A. Sengupta, *Topics in Circular Statistics*. World Scientific Pub Co Inc, 2001.
- [5] M. Azmani, S. Reboul, J.-B. Choquel, and M. Benjelloun, “A recursive fusion filter for angular data,” in *IEEE International Conference on Robotics and Biomimetics (ROBIO 2009)*, 2009, pp. 882–887.

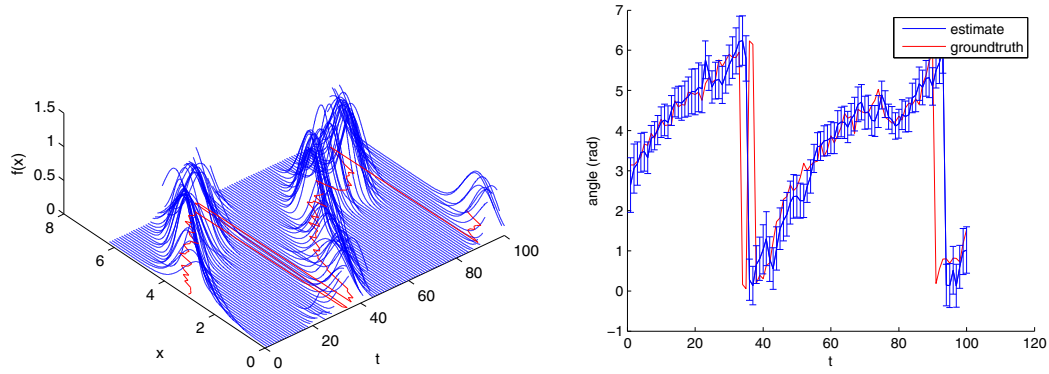


Fig. 3: An example run of the progressive filter with large measurement noise. On the left, we give the estimated WN distribution for each time step, and on the right, we give the estimate with σ -bounds.

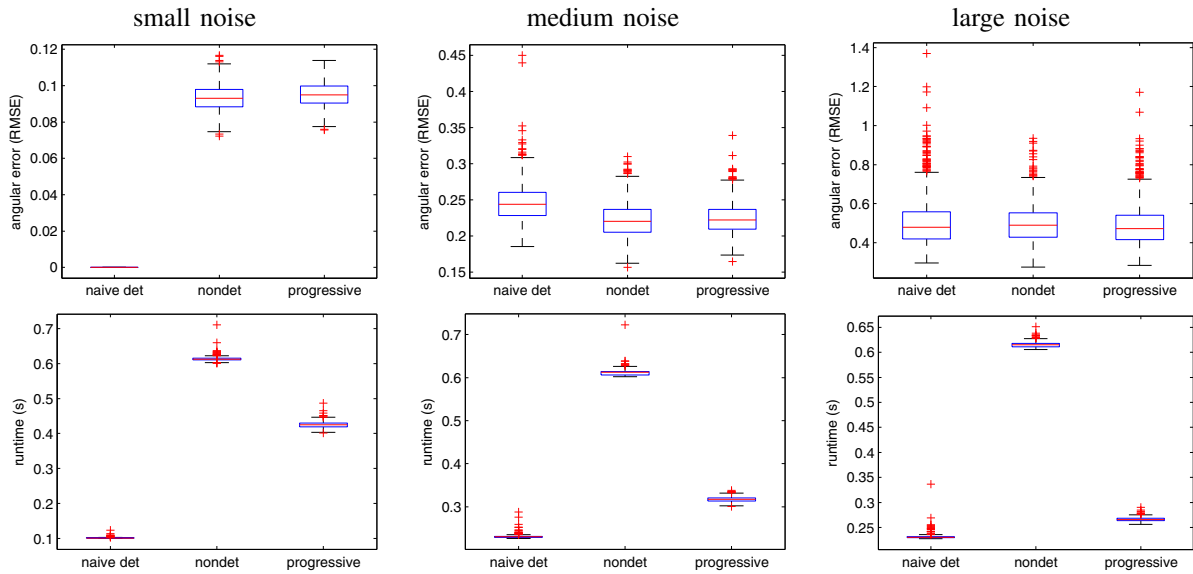


Fig. 4: Angular RMSE (in radians) and runtime (for 100 time steps) of the different filters depending on the measurement noise. The naive deterministic filter fails for small noise and does not yield valid results.

- [6] G. Stienne, S. Reboul, M. Azmani, J. Choquel, and M. Benjelloun, "A multi-temporal multi-sensor circular fusion filter," *Information Fusion*, vol. 18, pp. 86–100, Jul. 2013.
- [7] 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.
- [8] G. Kurz, F. Faion, and U. D. Hanebeck, "Constrained Object Tracking on Compact One-dimensional Manifolds Based on Directional Statistics," in *Proceedings of the Fourth IEEE GRSS International Conference on Indoor Positioning and Indoor Navigation (IPIN 2013)*, Montbeliard, France, Oct. 2013.
- [9] 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.
- [10] B. D. Ripley, "Computer generation of random variables: A tutorial," *International Statistical Review / Revue Internationale de Statistique*, vol. 51, no. 3, pp. pp. 301–319, 1983.
- [11] W. K. Hastings, "Monte carlo sampling methods using markov chains and their applications," *Biometrika*, vol. 57, no. 1, pp. 97–109, 1970.
- [12] M. F. Huber, F. Beutler, and U. D. Hanebeck, "Semi-Analytic Gaussian Assumed Density Filter," in *Proceedings of the 2011 American Control Conference (ACC 2011)*, San Francisco, California, USA, Jun. 2011.
- [13] J. H. Kotecha and P. Djuric, "Gaussian particle filtering," *IEEE Transactions on Signal Processing*, vol. 51, no. 10, pp. 2592–2601, 2003.
- [14] M. Arulampalam, S. Maskell, N. Gordon, and T. Clapp, "A tutorial on particle filters for online nonlinear/non-gaussian bayesian tracking," *IEEE Transactions on Signal Processing*, vol. 50, no. 2, pp. 174–188, 2002.
- [15] U. D. Hanebeck, "PGF 42: Progressive Gaussian Filtering with a Twist," in *Proceedings of the 16th International Conference on Information Fusion (Fusion 2013)*, Istanbul, Turkey, Jul. 2013.
- [16] P. Ruoff, P. Krauthausen, and U. D. Hanebeck, "Progressive Correction for Deterministic Dirac Mixture Approximations," in *Proceedings of the 14th International Conference on Information Fusion (Fusion 2011)*, Chicago, Illinois, USA, Jul. 2011.