Preprints of the 18th IFAC World Congress
Milano (Italy) August 28 - September 2, 2011

(Semi-)Analytic Gaussian Mixture Filter

Marco F. Huber [*], **Frederik Beutler** [**], **Uwe D. Hanebeck** [**]

[*] *Variable Image Acquisition and Processing Research Group, Fraunhofer Institute of Optronics, System Technologies and Image Exploitation (IOSB), Karlsruhe, Germany; (e-mail: marco.huber@ieee.org).*
[**] *Intelligent Sensor-Actuator-Systems Laboratory (ISAS), Institute for Anthropomatics, Karlsruhe Institute of Technology (KIT), Karlsruhe, Germany; (e-mail: beutler@ieee.org, uwe.hanebeck@ieee.org).*

Abstract: In nonlinear filtering, special types of Gaussian mixture filters are a straightforward extension of Gaussian filters, where linearizing the system model is performed individually for each Gaussian component. In this paper, two novel types of linearization are combined with Gaussian mixture filters. The first linearization is called analytic stochastic linearization, where the linearization is performed analytically and exactly, i.e., without Taylor-series expansion or approximate sample-based density representation. In cases where a full analytical linearization is not possible, the second approach decomposes the nonlinear system into a set of nonlinear subsystems that are conditionally integrable in closed form. These approaches are more accurate than fully applying classical linearization.

Keywords: State Estimation; Nonlinear Systems; Stochastic Systems; Decomposition Methods

1. INTRODUCTION

Determining the internal state of a dynamic system is essential in many applications, e.g., target tracking or simultaneous localization and mapping (SLAM). Due to disturbances in the measurement process and imperfect system models, the Bayesian estimation framework is often applied in order to deal with these uncertainties. For linear systems affected with Gaussian noise, the Kalman filter is the optimal estimator and Bayesian estimation can be performed in closed form. In case of nonlinearities and/or non-Gaussian noise, the Kalman filter is no longer optimal or even not applicable. In order to achieve an estimate for the internal state of the system, approximate state estimators have to be employed.

A common restriction of approximate state estimators is to assume a (jointly) Gaussian representation of the system state and the measurement. Estimators corresponding to these so-called Gaussian filters are for example the Extended Kalman Filter (EKF), the Unscented Kalman Filter (UKF, Julier et al. [2000]), analytic approaches (AGF, Beutler et al. [2010]), and semi-analytic approaches (SAGF, Huber et al. [2011]). Due to the Gaussian assumption, a linearization of the nonlinear system models is performed, which can be explicit (EKF) or implicit (UKF, AGF, SAGF) as discussed in Sec. 4.1. Obviously, in problems with strong nonlinearities that cause non-Gaussian densities, e.g., multi-modal or heavily skewed densities, the estimation accuracy of Gaussian filters is limited. But on the other hand, the computational complexity is only polynomial with the dimension of the state space.

Estimators especially designed for nonlinear non-Gaussian problems are for instance the grid filter or the particle filters (Arulampalam et al. [2002]), which make no specific assumptions about the shape of the density function. Hence, these approaches can be arbitrarily accurate, but their complexity is exponential.

A trade-off between both worlds are the so-called Gaussian mixture filters (GMFs, see Sec. 3), which are often applied to SLAM (Kwok et al. [2005], Lemaire et al. [2007]). Gaussian mixtures are very convenient for filtering purposes as with an increasing number of mixture components, a Gaussian mixture can approximate any density function with arbitrary accuracy (Maz'ya and Schmidt [1996]). GMFs utilize a weighted sum of Gaussian densities and typically rely on individual linearizations for each Gaussian component. By applying a bank of EKFs, individual linearization is achieved via component-wise first-order Taylor-series expansion (Alspach and Sorenson [1972]). A bank of UKFs corresponds to a sample-based stochastic linearization (Šimandl and Duník [2005]) for each component.

In this paper, two novel types of linearization are utilized for GMFs. For some classes of nonlinear functions, e.g., polynomial or trigonometric functions or combinations of them, stochastic linearization can be performed analytically and exactly (see Sec. 4.2). Instead of a sample point representation, the whole Gaussian density is propagated through the nonlinear function for linearization. The resulting filter is named the analytic Gaussian mixture filter (A-GMF). For arbitrary nonlinearities, the semi-analytic Gaussian mixture filter (SA-GMF) introduced in Sec. 4.3 utilizes a linearization approach that relies on a decomposition of the nonlinear system into integrable substructures such that one part of the problem can be solved

Copyright by the
International Federation of Automatic Control (IFAC)

analytically and for the remaining part a sample-based stochastic linearization is employed. This decomposition is much more general than the usual decomposition into conditional linear subsystems via Rao-Blackwellization. Since both GMFs at least partly rely on analytic linearization, much better local approximations and thus, more accurate estimation results are achieved compared to GMFs proposed by Alspach and Sorenson [1972] or Šimandl and Duník [2005]. This is demonstrated via simulations in Sec. 5, where a robot with tricycle kinematics is considered that measures distances to landmarks.

2. PROBLEM FORMULATION

A nonlinear dynamic system is described by its system and measurement equation according to

$$\underline{\boldsymbol{x}}_{k+1} = \underline{a}_k(\underline{\boldsymbol{x}}_k, \underline{u}_k, \underline{\boldsymbol{w}}_k) \ , \tag{1}$$

$$\underline{\boldsymbol{y}}_k = \underline{h}_k(\underline{\boldsymbol{x}}_k, \underline{\boldsymbol{v}}_k) \ , \tag{2}$$

where $\underline{\boldsymbol{x}}_k$ is the internal state, $\underline{\boldsymbol{w}}_k$ and $\underline{\boldsymbol{v}}_k$ are the white system and measurement noise processes, $\underline{\boldsymbol{y}}_k$ is the measurement process, and $\underline{u}_k$ a known control input. The known function $\underline{a}_k(\cdot,\cdot,\cdot)$ describes the evolution of the system over time and the known function $\underline{h}_k(\cdot,\cdot)$ the mapping between the internal state and the output of the system. Please note that actual measurement values $\hat{\underline{y}}_k$ are realizations of $\underline{\boldsymbol{y}}_k$.

The random variables, e.g., $\underline{\boldsymbol{x}}_k$, are denoted be bold font letters and are described by probability density functions $f(\cdot)$. For recursive estimation, the Bayesian estimation framework is applied, which consists of the prediction and filter step. In the following, a short overview of the Bayesian estimation framework is given.

2.1 Prediction Step

In the prediction step, the estimated density of the state $f^e(\underline{x}_k)$ is propagated to the next time step $k+1$ by means of the Chapman-Kolmogorov equation

$$f^p(\underline{x}_{k+1}) = \iint \underbrace{f(\underline{x}_{k+1}|\underline{x}_k, \underline{w}_k)}_{\delta(\underline{x}_{k+1} - \underline{a}_k(\underline{x}_k, \underline{u}_k, \underline{w}_k))} f^e(\underline{x}_k) f(\underline{w}_k) \, \mathrm{d}\underline{x}_k \, \mathrm{d}\underline{w}_k \ , \tag{3}$$

where $\delta(\cdot)$ is the Dirac delta distribution and $f^p(\underline{x}_{k+1})$ is the predicted density of the state.

2.2 Filter Step

In the filter step, the predicted density $f^p(\underline{x}_k)$ from the last prediction step is updated based on Bayes rule

$$f^e(\underline{x}_k) = \frac{1}{c_k} f(\hat{\underline{y}}_k|\underline{x}_k) f^p(\underline{x}_k) \ , \tag{4}$$

where $c_k = \int f(\hat{\underline{y}}_k|\underline{x}_k) f^p(\underline{x}_k) \, \mathrm{d}\underline{x}_k$ is a normalization constant and $f(\hat{\underline{y}}_k|\underline{x}_k)$ is the Likelihood function given by

$$f(\hat{\underline{y}}_k|\underline{x}_k) = \int \underbrace{f(\hat{\underline{y}}_k|\underline{x}_k, \underline{v}_k)}_{\delta(\hat{\underline{y}}_k - \underline{h}_k(\underline{x}_k, \underline{v}_k))} f(\underline{v}_k) \, \mathrm{d}\underline{v}_k \ .$$

The function $f(\hat{\underline{y}}_k|\underline{x}_k, \underline{v}_k)$ depends on the nonlinear function in (2) and the current measurement value $\hat{\underline{y}}_k$.

For simplicity and brevity, the time index k is omitted and for the prediction step, the variables $\underline{x}_k$ and $\underline{x}_{k+1}$ are replaced by $\underline{x}^e$ and $\underline{x}^p$, respectively.

3. GAUSSIAN MIXTURE FILTER

The basic idea behind a GMF is to apply linearization individually to each Gaussian component of a given Gaussian mixture for approximate prediction and filtering.

3.1 Prediction Step

In the prediction step, it is assumed that the result of the previous filter step is represented by means of the Gaussian mixture

$$f^e(\underline{x}^e) = \sum_{i=1}^{L} \omega_i^e \cdot \mathcal{N}(\underline{x}^e; \underline{\mu}_i^e, \mathbf{C}_i^e) \ , \tag{5}$$

where ω_i^e are non-negative weighting factors summing up to one and $\mathcal{N}(\underline{x}; \underline{\mu}, \mathbf{C})$ is a Gaussian density with mean vector $\underline{\mu}$ and covariance matrix $\mathbf{C}$. L is the number of components. The mixture in (5) is used in (3), which results in

$$f^p(\underline{x}^p) \approx \sum_{i=1}^{L} \omega_i^p \cdot f_i^p(\underline{x}^p) \ , \tag{6}$$

where the weighting factors [1] are $\omega_i^p = \omega_i^e$ and $f_i^p(\underline{x}^p)$ is the predicted density

$$f_i^p(\underline{x}^p) = \iint f(\underline{x}^p|\underline{x}^e, \underline{w}) \mathcal{N}(\underline{x}^e; \underline{\mu}_i^e, \mathbf{C}_i^e) f(\underline{w}) \, \mathrm{d}\underline{x}^e \, \mathrm{d}\underline{w} \tag{7}$$

of the ith component. In general, the integral in (7) cannot be solved in closed form. To simplify this problem, the predicted density of each component is individually approximated by a Gaussian density $f_i^p(\underline{x}^p) \approx \mathcal{N}(\underline{x}^p; \underline{\mu}_i^p, \mathbf{C}_i^p)$. Thus, it remains to calculate the moments $\underline{\mu}_i^p$ and $\mathbf{C}_i^p$ via moment matching. The mean $\underline{\mu}_i^p$ of the ith component for instance is

$$\underline{\mu}_i^p = \int \underline{x}^p \cdot f_i^p(\underline{x}^p) \, \mathrm{d}\underline{x}^p \ . \tag{8}$$

Using (7) in (8) and exploiting the sifting property of the Dirac delta distribution, the mean results in

$$\underline{\mu}_i^p = \iint \underline{a}(\underline{x}^e, \underline{u}, \underline{w}) \mathcal{N}(\underline{x}^e; \underline{\mu}_i^e, \mathbf{C}_i^e) f(\underline{w}) \, \mathrm{d}\underline{x}^e \, \mathrm{d}\underline{w} \ . \tag{9}$$

Similar to the mean, the covariance matrix is given by

$$\mathbf{C}_i^p = \iint (\underline{a}(\underline{x}^e, \underline{u}, \underline{w}) - \mu_i^p)(\underline{a}(\underline{x}^e, \underline{u}, \underline{w}) - \underline{\mu}_i^p)^{\mathrm{T}} \cdot \mathcal{N}(\underline{x}^e; \underline{\mu}_i^e, \mathbf{C}_i^e) f(\underline{w}) \, \mathrm{d}\underline{x}^e \, \mathrm{d}\underline{w} \ . \tag{10}$$

But still, the integrals in (9) and (10) cannot be solved analytically in general. A famous exception is the case, where $\underline{a}(\cdot,\cdot,\cdot)$ is linear and thus, the Kalman predictor can be applied individually for each component.

3.2 Filter Step

In the filter step, the Gaussian mixture (6) of the predicted state is used in (4), which results in

$$f^e(\underline{x}) = \frac{1}{c} \cdot f(\hat{\underline{y}}|\underline{x}) \cdot \sum_{i=1}^{L} \omega_i^p \cdot \mathcal{N}(\underline{x}; \underline{\mu}_i^p, \mathbf{C}_i^p) \ . \tag{11}$$

Due to the nonlinear measurement equation (2), the filter step cannot be solved analytically. For applying individual

[1] This weight update is exact only for linear system models, otherwise it is an approximation. For an improved weight update, see for example (Terejanu et al. [2008]).

approximations in (11), the equation is extended with $\mathcal{N}(\hat{\underline{y}}; \underline{\mu}_i^y, \mathbf{C}_i^y)$, which results in

$$f^e(\underline{x}) = \frac{1}{c}\sum_{i=1}^{L} w_i^p \cdot \mathcal{N}(\hat{\underline{y}}; \underline{\mu}_i^y, \mathbf{C}_i^y) \cdot \frac{f(\hat{\underline{y}}|\underline{x}^p)\mathcal{N}(\underline{x}^p; \underline{\mu}_i^p, \mathbf{C}_i^p)}{\mathcal{N}(\hat{\underline{y}}; \underline{\mu}_i^y, \mathbf{C}_i^y)} \approx \sum_{i=1}^{L} w_i^e \cdot \mathcal{N}(\underline{x}; \underline{\mu}_i^e, \mathbf{C}_i^e) \ ,$$

where the fraction is approximated with a Gaussian distribution $\mathcal{N}(\underline{x}; \underline{\mu}_i^e, \mathbf{C}_i^e)$ and the weights for the estimated density are given by

$$w_i^e = \frac{w_i^p \cdot \mathcal{N}(\hat{\underline{y}}; \underline{\mu}_i^y, \mathbf{C}_i^y)}{\sum_{i=1}^{L} w_i^p \cdot \mathcal{N}(\hat{\underline{y}}; \underline{\mu}_i^y, \mathbf{C}_i^y)} \ .$$

For calculating the estimated mean $\underline{\mu}_i^e$ and covariance $\mathbf{C}_i^e$ for each component i, it is assumed that the state and the measurement are jointly Gaussian. This assumption is typical for Gaussian filters and is only true for linear systems affected with Gaussian noise. Otherwise, it is an approximation. In doing so, the desired moments are

$$\begin{aligned} \underline{\mu}_i^e &= \underline{\mu}_i^p + \mathbf{C}_i^{x,y}(\mathbf{C}_i^y)^{-1}(\hat{\underline{y}} - \underline{\mu}_i^y) \ , \\ \mathbf{C}_i^e &= \mathbf{C}_i^p - \mathbf{C}_i^{x,y}(\mathbf{C}_i^y)^{-1}(\mathbf{C}_i^{x,y})^{\mathrm{T}} \ , \end{aligned} \tag{12}$$

which depend on the current measurement value $\hat{\underline{y}}$. To calculate the required parameters $\mathbf{C}_i^{x,y}$, $\mathbf{C}_i^y$, and $\underline{\mu}_i^y$ in (12), moment calculations similar to the prediction step are applied. The predicted measurement $\underline{\mu}_i^y$, the covariance of the measurement process $\mathbf{C}_i^y$, and the cross-covariance $\mathbf{C}_i^{x,y}$ between state and measurement can then be calculated via

$$\underline{\mu}_i^y = \iint \underline{h}(\underline{x}, \underline{v}) \cdot \mathcal{N}(\underline{x}; \underline{\mu}_i^p, \mathbf{C}_i^p) \cdot f(\underline{v}) \, \mathrm{d}\underline{x} \, \mathrm{d}\underline{v} \ , \tag{13}$$

$$\mathbf{C}_i^y = \iint (\underline{h}(\underline{x}, \underline{v}) - \underline{\mu}_i^y) \cdot (\underline{h}(\underline{x}, \underline{v}) - \underline{\mu}_i^y)^{\mathrm{T}} \cdot \mathcal{N}(\underline{x}; \underline{\mu}_i^p, \mathbf{C}_i^p) \cdot f(\underline{v}) \, \mathrm{d}\underline{x} \, \mathrm{d}\underline{v} \ , \tag{14}$$

$$\mathbf{C}_i^{x,y} = \iint (\underline{x} - \underline{\mu}_i^p) \cdot (\underline{h}(\underline{x}, \underline{v}) - \underline{\mu}_i^y)^{\mathrm{T}} \cdot \mathcal{N}(\underline{x}; \underline{\mu}_i^p, \mathbf{C}_i^p) \cdot f(\underline{v}) \, \mathrm{d}\underline{x} \, \mathrm{d}\underline{v} \ . \tag{15}$$

Unfortunately, these integrals cannot be solved in closed form in general. Thus, approximations in form of linearizations have to be applied, which are described in Sec. 4.

3.3 Gaussian Mixture Noise

So far, not much attention has been paid to the noise densities $f(\underline{v})$ and $f(\underline{w})$. If the noise is Gaussian, it can be processed jointly with the individual Gaussian components of the state density. One way to deal with arbitrary noise densities is to approximate them by means of a Gaussian mixture. In doing so, the noise can still be processed component-wise and the overall estimation procedure remains similar to the single Gaussian noise case. However, the calculation of the weighting factors differs, since the weighting factors of the noise mixture have to be considered as well. Furthermore, the number of components representing the Gaussian mixture of the state increases exponentially with the time. To bound this growth, Gaussian mixture reduction algorithms such as the one proposed in Huber and Hanebeck [2008a] have to be employed.

4. TYPES OF LINEARIZATION

For an approximate solution of the moment integrals in (9), (10) and (13)–(15), GMFs rely on individually linearizing the nonlinear functions $\underline{a}(\cdot,\cdot,\cdot)$ and $\underline{h}(\cdot,\cdot)$ for each component. Thus, the estimation performance of a GMF significantly depends on the quality of these linearizations. In the following, common types of linearization are briefly described. Then, two novel linearization approaches are proposed, namely the analytic stochastic linearization and the semi-analytic stochastic linearization.

For introducing the different linearization approaches, it is sufficient to restrict the focus on an abstract nonlinear transformation

$$\underline{\boldsymbol{y}} = \underline{g}(\underline{\boldsymbol{x}}) \ . \tag{16}$$

Here, the Gaussian random vector $\underline{\boldsymbol{x}} \sim \mathcal{N}(\underline{x}; \underline{\mu}^x, \mathbf{C}^x)$ is mapped to the random vector $\underline{\boldsymbol{y}}$. The nonlinear transformation $\underline{g}(\cdot)$ can be replaced by $\underline{a}(\cdot,\cdot,\cdot)$ in the prediction step and by $\underline{h}(\cdot,\cdot)$ in the filter step.

4.1 Classical Linearization

Generally, linearization approaches can be separated into explicit and implicit linearization. Classical linearizations that fall into these two classes are as follows.

Explicit Linearization For an explicit linearization, the nonlinear function $\underline{g}(\cdot)$ is linearized by applying a first-order Taylor-series expansion around the mean $\underline{\mu}^x$ as in the EKF. Thus, the linearized function is given by

$$\underline{g}(\underline{x}) \approx \underline{g}(\underline{\mu}^x) + \nabla_x \underline{g}(\underline{\mu}^x) \cdot (\underline{x} - \underline{\mu}^x) \ ,$$

where ∇_x is the gradient with respect to $\underline{x}$. Employing this type of linearization in a GMF corresponds to a bank of EKFs and the resulting GMF is called the Gaussian sum filter proposed in Alspach and Sorenson [1972].

Implicit Linearization In case of the implicit linearization, the nonlinear function in (16) remains untouched. Here, the word implicit indicates that if the mean $\underline{\mu}^y$ and covariance $\mathbf{C}^y$ of $\underline{\boldsymbol{y}}$ can be calculated, there exists an equivalent linear transformation

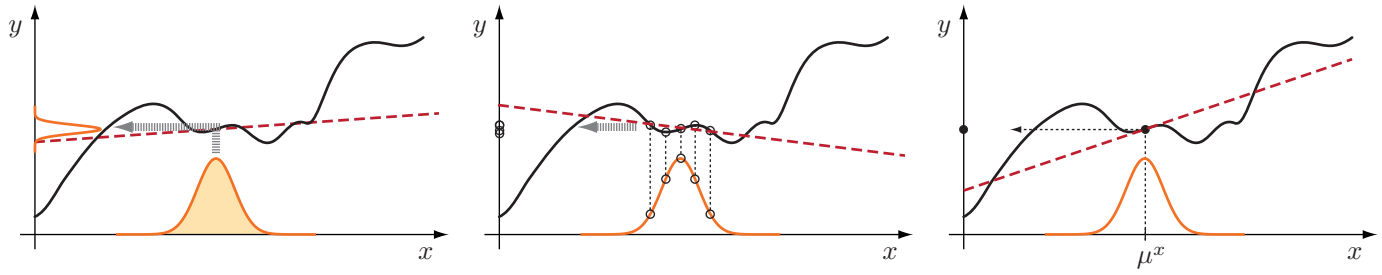
$$\underline{\boldsymbol{y}} = \underline{g}(\underline{\boldsymbol{x}}) \approx \mathbf{A} \cdot \underline{\boldsymbol{x}} + \underline{b} \tag{17}$$

that produces exactly the same mean and covariance of $\underline{\boldsymbol{y}}$ given a Gaussian random vector $\underline{\boldsymbol{x}}$.

A straightforward approach for implicit linearization is to approximate the Gaussian $\underline{\boldsymbol{x}}$ via a sample-based representation

$$\mathcal{N}(\underline{x}; \underline{\mu}^x, \mathbf{C}^x) \approx \sum_{j=1}^{N} w_j \cdot \delta(\underline{x} - \underline{\mu}_j) \ , \tag{18}$$

which exactly captures the mean $\underline{\mu}^x$ and covariance $\mathbf{C}^x$. Here, w_j is the weight of the jth sample point located at $\underline{\mu}_j$. In the following, the index j is used for indicating sample points, while the index i is used for components of a Gaussian mixture. The samples can be easily propagated through (16). Calculating the weighted sample mean and sample covariance for the propagated samples then allows approximately solving the moment integrals (9), (10) and (13)–(15). The linear transformation in (17)—if desired—can be obtained via least squares optimization as described in Lefebvre et al. [2005].



(a) Analytic stochastic linearization. (b) Sample-based linearization. (c) Linearization via Taylor-series expansion.

Fig. 1. Illustration of the different linearization approaches: the nonlinear function (black) and its linearized version (red dashed). (a) For the analytic stochastic linearization the entire Gaussian is propagated. (b) Sample-based linearization utilizes a sample representation of the Gaussian. (c) Taylor-series expansion linearizes the nonlinear function around a single point.

Several methods exist for calculating the sample representation in (18), e.g., the unscented transform (Julier et al. [2000]) or the deterministic sampling scheme (Huber and Hanebeck [2008b]). Combining the unscented transform with a GMF for instance, corresponds to a bank of UKFs and is described in detail in Šimandl and Duník [2005].

4.2 Analytic Stochastic Linearization

Special nonlinear functions for $\underline{g}(\cdot)$ facilitate to solve the moment integrals analytically and exactly, e.g., polynomials, trigonometric functions, and their combinations. In this case, the results of the moment integrals can be derived in an analytic form, as it is shown in the simulation example in Sec. 5. Thus, linearization is performed implicitly under the consideration of the entire Gaussian density of $\underline{\boldsymbol{x}}$. This is in contrast to the previously described sample-based linearization, which merely propagates a sample-based approximation of the Gaussian. Even more extreme is the explicit linearization used in the EKF, where linearization is only performed on the basis of the single point, that is the mean of the Gaussian. These differences are depicted in Fig. 1.

In the following, the combination of analytic stochastic linearization with a GMF is named analytic Gaussian mixture filter (A-GMF).

4.3 Semi-Analytic Stochastic Linearization

In order to extend the principle of analytic stochastic linearization to a wider class of nonlinear functions, the semi-analytic stochastic linearization approach is proposed. Here, sampled-based linearization is combined with the analytic stochastic linearization such that only some dimensions of the random vector $\underline{\boldsymbol{x}}$ are discretized by means of a sample representation. Thus, only some parts of the nonlinear transformation (16) have to be evaluated approximately for moment calculation.

For this purpose, the nonlinear equation (16) is rearranged according to

$$\underline{\boldsymbol{y}} = \underline{g}(\underline{\boldsymbol{x}}^a, \underline{\boldsymbol{x}}^b) \;, \tag{19}$$

where the Gaussian random vector $\underline{\boldsymbol{x}} = [(\underline{\boldsymbol{x}}^a)^{\mathrm{T}}, (\underline{\boldsymbol{x}}^b)^{\mathrm{T}}]^{\mathrm{T}}$ consists of the substates $\underline{\boldsymbol{x}}^a$, $\underline{\boldsymbol{x}}^b$ with mean and covariance

$$\underline{\mu}^x = \begin{bmatrix} \underline{\mu}^a \\ \underline{\mu}^b \end{bmatrix} \;,\; \mathbf{C}^x = \begin{bmatrix} \mathbf{C}^a & \mathbf{C}^{a,b} \\ \mathbf{C}^{b,a} & \mathbf{C}^b \end{bmatrix} \;.$$

As mentioned above, there exists no closed-form expression for the desired moments in general. However, the decomposition into $\underline{\boldsymbol{x}}^a$ and $\underline{\boldsymbol{x}}^b$ is chosen in such a way that the moment integrals can be calculated in closed form for any given fixed value of $\underline{\boldsymbol{x}}^b$. Hence, $\underline{g}(\cdot,\cdot)$ is denoted to be conditionally integrable if such a decomposition exists. For determining a sample-based representation of $\underline{\boldsymbol{x}}^b$, the sampling techniques mentioned in Sec. 4.1 are applied.

The analytic stochastic linearization and sampled-based linearization are extreme cases of the semi-analytic stochastic linearization: if $\underline{\boldsymbol{x}}^b$ has no entries, the semi-analytic stochastic linearization becomes an analytic stochastic linearization and if $\underline{\boldsymbol{x}}^a$ has no entries, the semi-analytic stochastic linearization degenerates into a sample-based linearization.

General Solution For the general transformation in (19), the desired moments of $\underline{\boldsymbol{y}}$ can be calculated as follows. At first, the joint density $f(\underline{x}, \underline{y})$ is separated by employing Bayes' rule

$$f(\underline{x}, \underline{y}) = \underbrace{\delta\big(\underline{y} - \underline{g}(\underline{x}^a, \underline{x}^b)\big)}_{= f(\underline{y}|\underline{x})} \cdot \underbrace{f(\underline{x}^a|\underline{x}^b) \cdot f(\underline{x}^b)}_{= f(\underline{x})} \;.$$

The conditional density $f(\underline{x}^a|\underline{x}^b) = \mathcal{N}\big(\underline{x}^a; \underline{\mu}^{a|b}, \mathbf{C}^{a|b}\big)$ is (conditionally) Gaussian with mean and covariance

$$\begin{aligned} \underline{\mu}^{a|b} &= \underline{\mu}^a + \mathbf{C}^{a,b} \cdot \big(\mathbf{C}^b\big)^{-1} \cdot \big(\underline{x}^b - \underline{\mu}^b\big) \;, \\ \mathbf{C}^{a|b} &= \mathbf{C}^a - \mathbf{C}^{a,b} \cdot \big(\mathbf{C}^b\big)^{-1} \cdot \mathbf{C}^{b,a} \;. \end{aligned} \tag{20}$$

To determine the mean $\underline{\mu}^y$, the Gaussian density $f(\underline{x}^b)$ of the substate $\underline{\boldsymbol{x}}^b$ is represented by means of a sample density as in (18), which allows a sample-based linearization. For integrating over $\underline{x}^b$, the sifting property of the Dirac delta distribution is exploited. This gives rise to

$$\underline{\mu}^y \approx \sum_{j=1}^{N} w_j \cdot \underline{\mu}_j^y \quad \text{with} \quad \underline{\mu}_j^y = \int \underline{g}(\underline{x}^a, \underline{\mu}_j^b) \cdot f(\underline{x}^a|\underline{\mu}_j^b)\, \mathrm{d}\underline{x}^a \tag{21}$$

for the mean of $\underline{\boldsymbol{y}}$ and analogously, the covariance of $\underline{\boldsymbol{y}}$ is approximated according to

$$\begin{aligned} \mathbf{C}^y &\approx \sum_{j=1}^{N} w_j \cdot \big(\mathbf{C}_j^y - \underline{\mu}_j^y (\underline{\mu}^y)^{\mathrm{T}} - \underline{\mu}^y (\underline{\mu}_j^y)^{\mathrm{T}} + \underline{\mu}^y (\underline{\mu}^y)^{\mathrm{T}}\big) \;, \\ \mathbf{C}_j^y &= \int \underline{g}\big(\underline{x}^a, \underline{\mu}_j^b\big) \cdot \underline{g}\big(\underline{x}^a, \underline{\mu}_j^b\big)^{\mathrm{T}} \cdot f\big(\underline{x}^a|\underline{\mu}_j^b\big)\, \mathrm{d}\underline{x}^a \;. \end{aligned} \tag{22}$$

It is worth mentioning that the integrals in (21) and (22) can be evaluated analytically as the function $\underline{g}(\cdot,\cdot)$ is chosen to be conditionally integrable. Solving these integrals is an off-line task and the solution is represented in parametric form for an efficient on-line evaluation.

Gaussian Estimation By putting it all together, prediction and filtering for the ith Gaussian component of the Gaussian mixture can now be derived.

In the prediction step, the predicted mean $\underline{\mu}_i^p$ (9) and covariance $\mathbf{C}_i^p$ (10) of $f_i^p(\underline{x}^p)$ have to be calculated. For this purpose, the system function (1) can be directly mapped to the nonlinear transformation (19) according to

$$\underline{\boldsymbol{x}}^p = \underline{a}\left(\underline{\boldsymbol{x}}^e, \underline{u}, \underline{\boldsymbol{w}}\right) = \underline{g}\big(\underline{\boldsymbol{x}}^a, \underline{\boldsymbol{x}}^b\big) \ .$$

Here, the system input $\underline{u}$ becomes a part of the function $\underline{g}(\cdot,\cdot)$ and the substates $\underline{\boldsymbol{x}}^a, \underline{\boldsymbol{x}}^b$ are augmented with the noise variables $\underline{\boldsymbol{w}}^a, \underline{\boldsymbol{w}}^b$, where $\underline{\boldsymbol{w}} = [(\underline{\boldsymbol{w}}^a)^{\mathrm{T}}, (\underline{\boldsymbol{w}}^b)^{\mathrm{T}}]^{\mathrm{T}}$, in order to consider additive and/or multiplicative noise. For calculating the mean $\underline{\mu}_i^p$ and covariance $\mathbf{C}_i^p$, (21) and (22) are used, respectively.

In the filter step, the measurement equation (2) is mapped to the nonlinear transformation (19) according to

$$\underline{\boldsymbol{y}} = \underline{h}(\underline{\boldsymbol{x}}, \underline{\boldsymbol{v}}) = \underline{g}\big(\underline{\boldsymbol{x}}^a, \underline{\boldsymbol{x}}^b\big) \ ,$$

where the measurement noise $\underline{\boldsymbol{v}}$ is spread across the substates $\underline{\boldsymbol{x}}^a, \underline{\boldsymbol{x}}^b$. It is worth mentioning that the decomposition of $\underline{\boldsymbol{x}}$ into the substates for the filter step is independent of the decomposition of the prediction step.

For determining the mean $\underline{\mu}_i^e$ and covariance $\mathbf{C}_i^e$ in (12) for the ith Gaussian component, the moments $\underline{\mu}_i^y$, $\mathbf{C}_i^y$ and $\mathbf{C}_i^{x,y}$ in (13)–(15) are required. According to (18), the density of the substate $\underline{\boldsymbol{x}}^b$ is approximated by the sample density

$$\mathcal{N}(\underline{x}^b; \underline{\mu}_i^b, \mathbf{C}_i^b) \approx \sum_{j=1}^{N} w_{ij} \cdot \delta(\underline{x}^b - \underline{\mu}_{ij}^b) \ .$$

By means of this sample representation, $\underline{\mu}_i^y$ and $\mathbf{C}_i^y$ can be calculated as in (21) and (22), respectively. The cross-covariance $\mathbf{C}_i^{x,y}$ needs further derivations. The cross-covariance $\mathbf{C}_i^{x,y} = [\mathbf{C}_i^{a,y}, \mathbf{C}_i^{b,y}]^{\mathrm{T}}$ consists of

$$\mathbf{C}_i^{a,y} = \sum_{j=1}^{N} w_{ij} \cdot \Big(\mathbf{C}_{ij}^{a,y} - \underline{\mu}_{ij}^{a|b}\big(\underline{\mu}_i^y\big)^{\mathrm{T}} + \underline{\mu}_i^a\big(\underline{\mu}_i^y - \underline{\mu}_{ij}^y\big)^{\mathrm{T}}\Big) \ ,$$

$$\mathbf{C}_i^{b,y} = \sum_{j=1}^{N} w_{ij} \cdot \big(\underline{\mu}_{ij}^b - \underline{\mu}_i^b\big) \cdot \big(\underline{\mu}_{ij}^y - \underline{\mu}_i^y\big)^{\mathrm{T}} \ ,$$

with

$$\mathbf{C}_{ij}^{a,y} = \int \underline{x}^a \cdot \underline{g}\big(\underline{x}^a, \underline{\mu}_{ij}^b\big)^{\mathrm{T}} \cdot f\big(\underline{x}^a | \underline{\mu}_{ij}^b\big) \, \mathrm{d}\underline{x}^a \ ,$$

where $\underline{\mu}_{ij}^{a|b}$ is calculated according to (20) with $\underline{x}^b$ replaced by $\underline{\mu}_{ij}^b$ and $\underline{\mu}_{ij}^y$ results from solving the integral in (21).

The combination of semi-analytic stochastic linearization with a GMF is named semi-analytic Gaussian mixture filter (SA-GMF) in the following.

5. SIMULATION RESULTS

In the simulations, a localization scenario is considered. A robot with tricycle kinematics measures the distance to one out of four landmarks per time step. The landmark considered for measurement is selected randomly with equal probability. The proposed A-GMF and SA-GMF are compared with the EKF-GMF (a.k.a. Gaussian sum filter, Alspach and Sorenson [1972]) and the UKF-GMF (Šimandl and Duník [2005]) for different numbers of components and measurement noise levels.

5.1 System and Measurement Model

The nonlinear kinematics model of the robot is given by

$$\begin{aligned} \boldsymbol{p}_{k+1}^x &= \boldsymbol{p}_k^x + (u_k^v + \boldsymbol{w}_k^v) \cdot \cos(\boldsymbol{\phi}_k + u_k^\alpha) \ , \\ \boldsymbol{p}_{k+1}^y &= \boldsymbol{p}_k^y + (u_k^v + \boldsymbol{w}_k^v) \cdot \sin(\boldsymbol{\phi}_k + u_k^\alpha) \ , \\ \boldsymbol{\phi}_{k+1} &= \boldsymbol{\phi}_k + (u_k^\alpha + \boldsymbol{w}_k^\alpha) \ , \end{aligned} \tag{23}$$

with state $\underline{\boldsymbol{x}}_k = [\boldsymbol{p}_k^x \ \boldsymbol{p}_k^y \ \boldsymbol{\phi}_k]^{\mathrm{T}}$, where $\boldsymbol{p}_k^x$ and $\boldsymbol{p}_k^y$ describe the Cartesian position of the robot and $\boldsymbol{\phi}_k$ its orientation. The known control inputs are the velocity u_k^v and the turning angle u_k^α. $\boldsymbol{w}_k^v$ and $\boldsymbol{w}_k^\alpha$ are noise processes affecting the corresponding control inputs. They are assumed to be zero-mean Gaussian with variances Q^v and Q^α, respectively.

The measured range $\boldsymbol{r}_k$ is given by the nonlinear measurement model

$$\boldsymbol{r}_k = \sqrt{(\boldsymbol{p}_k^x - L^x + \boldsymbol{v}_k^x)^2 + (\boldsymbol{p}_k^y - L^y + \boldsymbol{v}_k^y)^2} \ ,$$

where $\underline{L} = [L^x \ L^y]^{\mathrm{T}}$ is the position of the landmark, and $\underline{\boldsymbol{v}}_k = [\boldsymbol{v}_k^x \ \boldsymbol{v}_k^y]^{\mathrm{T}}$ is the measurement noise. The noise is assumed to be zero-mean Gaussian with covariance $\mathbf{C}^v$.

5.2 Estimator: A-GMF

The prediction step can be solved analytically as the model (23) consists of linear, bilinear, and polynomial functions. The predicted mean of the ith Gaussian component is

$$\begin{aligned} \underline{\mu}_i^p(1) &= \mu_i^e(1) + u^v \cdot e \cdot \cos(\beta), \\ \underline{\mu}_i^p(2) &= \underline{\mu}_i^e(2) + u^v \cdot e \cdot \sin(\beta), \\ \underline{\mu}_i^p(3) &= \underline{\mu}_i^e(3) + u^\alpha \ , \end{aligned}$$

where $\underline{\mu}_i^p(j)$ is the jth element of the vector $\underline{\mu}_i^p$ and $\mathbf{C}_i^e(m,n)$ is the element of the matrix $\mathbf{C}_i^e$ at row m and column n. The variables β and e are given by

$$\beta = u^\alpha + \underline{\mu}_i^e(3) \ \text{ and } \ e = \exp\left(-\tfrac{1}{2} \cdot \mathbf{C}_i^e(3,3)\right) \ ,$$

respectively. The analytic expressions of the predicted covariances can be found in Appendix A.

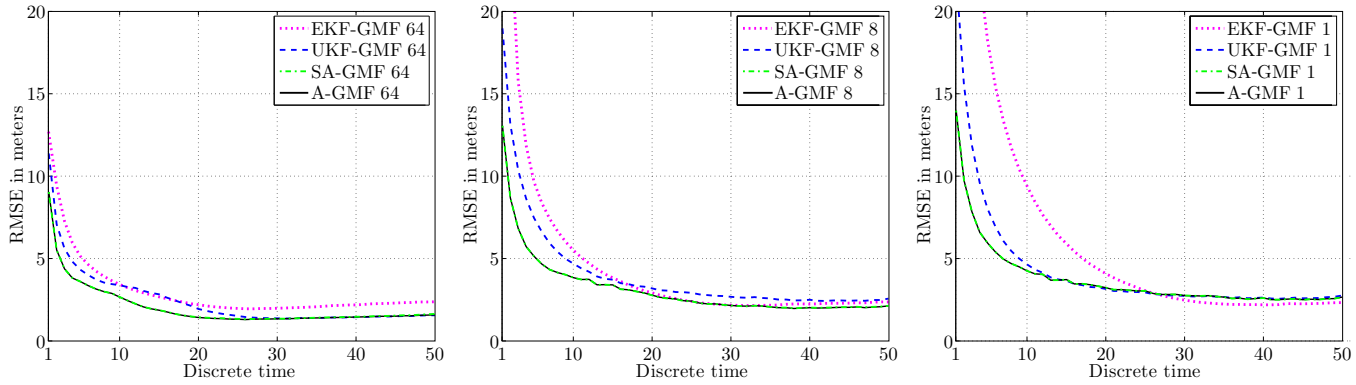
For an analytic moment calculation in the filter step, the measurement model needs to be squared, where the resulting new measurement is denoted by $\boldsymbol{y}_k := (\boldsymbol{r}_k)^2$. Based on the new squared measurement model, which now is polynominal, the required quantities in (13)-(15) are

$$\mu_i^y = \underline{A}^{\mathrm{T}} \cdot \underline{A} + \mathrm{trace}(\mathbf{T}) \ , \ \mathbf{C}_i^{x,y} = -2 \cdot \mathbf{C}_i^p \cdot \mathbf{P}^{\mathrm{T}} \cdot \underline{A} \ ,$$
$$C_i^y = \underline{1}_2^{\mathrm{T}} \cdot (4 \cdot (\underline{A} \cdot \underline{A}^{\mathrm{T}}) \circ \mathbf{T} + 2\mathbf{T} \circ \mathbf{T}) \cdot \underline{1}_2 \ ,$$

where $\circ$ is the element-wise product, $\underline{1}_2 = [1 \ 1]^{\mathrm{T}}$, and

$$\mathbf{P} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix} \ , \ \mathbf{T} = \mathbf{C}^v + \mathbf{P} \cdot \mathbf{C}_i^p \cdot \mathbf{P}^{\mathrm{T}} \ , \ \underline{A} = \underline{L} - \underline{\mu}_i^p \ .$$

The measurement value $\hat{y}$ is replaced by $(\hat{r})^2$.



(a) 64 components. (b) 8 components. (c) 1 component.

Fig. 2. The RMSE and its standard deviation over the 1000 trajectories at noise level 1 for different numbers of Gaussian components.

5.3 Estimator: SA-GMF

Even if the prediction step can be performed in closed form, the system model is written in conditionally linear form according to

$$\underline{\boldsymbol{x}}_{k+1} = \begin{bmatrix} u_k^v \cos(\boldsymbol{\phi}_k + u_k^\alpha) \\ u_k^v \sin(\boldsymbol{\phi}_k + u_k^\alpha) \\ \boldsymbol{\phi}_k + u_k^\alpha \end{bmatrix} + \begin{bmatrix} 1 & 0 & \cos(\boldsymbol{\phi}_k + u_k^\alpha) & 0 \\ 0 & 1 & \sin(\boldsymbol{\phi}_k + u_k^\alpha) & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \cdot \underline{\boldsymbol{x}}_k^a \ ,$$

in order to demonstrate to performace of the SA-GMF. Here, $\underline{\boldsymbol{x}}^a = [\boldsymbol{p}^x, \boldsymbol{p}^y, \boldsymbol{w}^v, \boldsymbol{w}^\alpha]^{\mathrm{T}}$ and $\underline{\boldsymbol{x}}^b = \boldsymbol{\phi}$. By approximating the orientation $\boldsymbol{\phi}$ by means of the sample-based representation, the nonlinear system model becomes a conditionally linear one. Thus, given a sample point $\underline{\mu}_j^b$, the prediction step with respect to $\underline{\boldsymbol{x}}^a$ for each Gaussian component can be solved via the Kalman predictor. The filter step is performed as in the A-GMF.

5.4 Setup

The initial position of the robot at time step $k = 0$ is $\underline{x}_0 = [5\ 3\ 0.2]^{\mathrm{T}}$. Furthermore, the known control inputs are constant and selected as $u^v = 0.1$ and $u^\alpha = 0.1$. The variances of the noise processes are $Q^v = 0.1$ and $Q^\alpha = 0.01$.

The measurement noise is isotropic, i.e., the covariance is $\mathbf{C}^v = (\sigma^v)^2 \cdot \mathbf{I}_2$ with $\mathbf{I}_2$ being the 2×2 identity matrix. For the variance $(\sigma^v)^2$, the three noise levels 0.5, 1, and 2 are considered. The positions $\underline{L}_i$, $i = 1, \ldots, 4$ of the four landmarks are

$$[\underline{L}_1\ \underline{L}_3\ \underline{L}_3\ \underline{L}_4] = \begin{bmatrix} 0 & 2 & 5 & 10 \\ 0 & 2 & 5 & 10 \end{bmatrix} \ .$$

The GMFs are initialized with different numbers of components, namely 1, 8, and 64 components. This initialization is performed via the splitting procedure proposed in Huber et al. [2008], where the Gaussian used for splitting has the covariance matrix $\mathbf{C} = \mathrm{diag}([1000\ 1000\ 30])$ and a mean sampled from the Gaussian $\mathcal{N}(\underline{x}; \underline{x}_0, \mathbf{C})$. For each combination of noise level and number of components, 1000 simulation runs are performed, where each run consists of 50 time steps.

5.5 Results

In Tab. 1, the average root mean square error (RMSE) with respect to the robot's position and the standard deviation of the average RMSE of all estimators for all noise levels and all numbers of components are listed. The A-GMF and the SA-GMF outperform the UKF-GMF and the EKF-GMF under all conditions. The RMSE for the noise level $(\sigma^v)^2 = 1$ is shown in Fig. 2. A-GMF and SA-GMF converge significantly faster compared to other two estimators and provide the lowest estimation errors.

In this simulation, A-GMF and SA-GMF perform almost identical as both use the same filter step and a sample-based representation in case of the SA-GMF is merely necessary for one dimension.

The EKF-GMF provides the worst estimation results in this scenario. This can be explained by the fact that the underlying linearization does not consider any uncertainty information and thus, the linearization error is neglected. Thanks to the sample-based linearization of the UKF-GMF, uncertainty information can be incorporated, which leads to a superior linearization and estimation compared to the EKF-GMF. However, only a finite number of sample points is used for moment calculation. In the A-GMF instead, the linearization is performed implicitly under the

Table 1. Average RMSE and standard deviation for the different estimators at different noise levels and numbers of components.

Noise 0.5	64	8	1
A-GMF	$\mathbf{2.02 \pm 1.34}$	$\mathbf{3.17 \pm 1.98}$	$\mathbf{3.76 \pm 2.01}$
SA-GMF	2.03 ± 1.34	$\mathbf{3.16 \pm 1.99}$	3.76 ± 2.09
UKF-GMF	2.41 ± 1.79	4.08 ± 3.02	4.31 ± 3.58
EKF-GMF	2.64 ± 1.86	4.08 ± 4.99	6.40 ± 9.64
Noise 1.0	64	8	1
A-GMF	$\mathbf{2.05 \pm 1.36}$	3.16 ± 1.97	$\mathbf{3.70 \pm 2.07}$
SA-GMF	$\mathbf{2.06 \pm 1.35}$	$\mathbf{3.15 \pm 1.97}$	$\mathbf{3.70 \pm 2.07}$
UKF-GMF	2.44 ± 1.80	4.04 ± 3.05	4.25 ± 3.61
EKF-GMF	2.95 ± 1.73	4.40 ± 4.43	6.72 ± 8.71
Noise 2.0	64	8	1
A-GMF	$\mathbf{2.16 \pm 1.35}$	$\mathbf{3.22 \pm 2.02}$	$\mathbf{3.78 \pm 2.16}$
SA-GMF	2.16 ± 1.36	$\mathbf{3.22 \pm 2.02}$	$\mathbf{3.78 \pm 2.16}$
UKF-GMF	2.61 ± 1.77	4.11 ± 3.08	4.35 ± 3.65
EKF-GMF	3.42 ± 1.55	4.95 ± 4.14	7.21 ± 7.60

consideration of the entire Gaussian density. This further improves estimation performance.

Furthermore, the computational complexity of the A-GMF and SA-GMF is lower compared to the UKF-GMF. This can be explained by the necessity of calculating matrix square roots for determining a sample-based representation of a Gaussian. This operation has cubic complexity and has to be applied in the UKF-GMF on the whole covariance matrix of each Gaussian component. It is not required for the A-GMF and in case of the SA-GMF, the covariance matrix is of reduced dimension as some dimensions of the state space are processed analytically. A further reason for the higher computational load of the UKF-GMF is the high number of on-line evaluations of the functions (1) and (2). In case of the A-GMF, all function evaluations are performed off-line, while for the SA-GMF the number of function evaluations is reduced due to the small dimension of the substate $\underline{\boldsymbol{x}}^b$.

6. CONCLUSIONS

This paper introduces two novel Gaussian mixture filters, which combine analytic stochastic linearization techniques with a Gaussian mixture density representation. The purely analytic stochastic linearization approach is designed for nonlinear filtering problems, where the integrals for calculating the mean and covariance can be solved in closed form. Semi-analytic linearization extends this approach to a wider class of filtering problems, where only some parts of the system model need to be analytically integrable. Thanks to the combination of these linearization approaches with a Gaussian mixture representation of the state density, the estimation performance is improved compared to Gaussian filters, especially in problems with severe nonlinearities. In the simulation it is shown that the proposed filter outperforms Gaussian mixture filters employing standard linearization techniques concerning estimation error and convergence.

Appendix A. ANALYTIC EXPRESSIONS FOR A-GMF

The analytic expressions of the elements of the predicted covariance matrix of the A-GMF are

$$\begin{aligned}
\mathbf{C}_i^p(1,1) =& \tfrac{1}{2}(u^v)^2 + \mathbf{C}_i^e(1,1) - (u^v e\cos(\beta))^2 \\
&+ \tfrac{1}{2}(u^v)^2 e^4\cos(2\beta) - 2u^v e\sin(\beta)\mathbf{C}_i^e(1,3) \\
&+ Q^v\tfrac{1}{2}(1+e^4\cos(2\beta)) \ , \\
\mathbf{C}_i^p(2,2) =& \tfrac{1}{2}(u^v)^2 + \mathbf{C}_i^e(2,2) - (u^v e\sin(\beta))^2 \\
&- \tfrac{1}{2}(u^v)^2 e^4\cos(2\beta) + 2u^v e\cos(\beta)\mathbf{C}_i^e(2,3) \\
&+ Q^v\tfrac{1}{2}(1-e^4\cos(2\beta)) \ , \\
\mathbf{C}_i^p(3,3) =& \mathbf{C}_i^e(3,3) + Q^\alpha \\
\mathbf{C}_i^p(1,2) =& \mathbf{C}_i^e(1,2) - u^v\cdot e\cdot\mathbf{C}_i^e(1,3)\cos(\beta) \\
&- u^v\cdot e\cdot\mathbf{C}_i^e(2,3)\sin(\beta) \\
&+ (u^v e)^2\cos(\beta)\sin(\beta)(e^2-1) + Q^v\tfrac{e^4}{2}\sin(2\beta) \ , \\
\mathbf{C}_i^p(1,3) =& \mathbf{C}_i^e(1,3) - u^v\cdot\mathbf{C}_i^e(3,3)\cdot e\cdot\sin(\beta) \ , \\
\mathbf{C}_i^p(2,3) =& \mathbf{C}_i^e(2,3) + u^v\cdot\mathbf{C}_i^e(3,3)\cdot e\cdot\cos(\beta) \ , \\
\mathbf{C}_i^p(2,1) =& \mathbf{C}_i^p(1,2),\ \mathbf{C}_i^p(3,1)=\mathbf{C}_i^p(1,3),\ \mathbf{C}_i^p(3,2)=\mathbf{C}_i^p(2,3) \ .
\end{aligned}$$

REFERENCES

D. Alspach and H. Sorenson. Nonlinear Bayesian Estimation using Gaussian Sum Approximations. *IEEE Transactions on Automatic Control*, 17(4):439 – 448, August 1972.

M. S. 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*, 50(2):174–188, February 2002.

F. Beutler, M. F. Huber, and U. D. Hanebeck. Optimal Stochastic Linearization for Range-based Localization. In *Proceedings of the 2010 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS 2010)*, Taipei, Taiwan, October 2010.

M. F. Huber and U. D. Hanebeck. Progressive Gaussian Mixture Reduction. In *Proceedings of the 11th International Conference on Information Fusion (Fusion 2008)*, pages 1–8, Cologne, Germany, July 2008a.

M. F. Huber and U. D. Hanebeck. Gaussian Filter based on Deterministic Sampling for High Quality Nonlinear Estimation. In *Proceedings of the 17th IFAC World Congress*, Seoul, Republic of Korea, July 2008b.

M. F. Huber, T. Bailey, H. Durrant-Whyte, and U. D. Hanebeck. On Entropy Approximation for Gaussian Mixture Random Vectors. In *Proceedings of the 2008 IEEE International Conference on Multisensor Fusion and Integration for Intelligent Systems (MFI 2008)*, pages 181–188, Seoul, Republic of Korea, August 2008.

M. F. Huber, F. Beutler, and U. D. Hanebeck. Semi-Analytic Gaussian Assumed Density Filter. In *Proceedings of the 2011 American Control Conference (ACC)*, San Francisco, California, June 2011.

S. Julier, J. Uhlmann, and H. F. Durrant-Whyte. A New Method for the Nonlinear Trannsformation of Means and Covariances in Filters and Estimators. *IEEE Transactions on Automatic Control*, 45(3):477–482, 2000.

N. M. Kwok, G. Dissanayake, and Q. P. Ha. Bearing-only SLAM Using a SPRT Based Gaussian Sum Filter. In *Proceedings of the 2005 IEEE International Conference on Robotics and Automation (ICRA 2005)*, pages 1109 – 1114, Barcelona, Spain, April 2005.

T. Lefebvre, H. Bruyninckx, and J. De Schutter. *Nonlinear Kalman Filtering for Force-Controlled Robot Tasks*. Springer Berlin, 2005.

T. Lemaire, C. Berger, I.-K. Jung, and S. Lacroix. Vision-Based SLAM: Stereo and Monocular Approaches. *International Journal of Computer Vision*, 74(3):343–364, 2007.

V. Maz'ya and G. Schmidt. On approximate approximations using Gaussian kernels. *IMA Journal of Numerical Analysis*, 16(1):13–29, 1996.

G. Terejanu, P. Singla, T. Singh, and P. D. Scott. A Novel Gaussian Sum Filter Method for Accurate Solution to the Nonlinear Filtering Problem. In *Proceedings of the 11th International Conference on Information Fusion (Fusion)*, Cologne, Germany, July 2008.

M. Šimandl and J. Duník. Sigma Point Gaussian Sum Filter Design Using Square Root Unscented Filters. In *Proceedings of the 16th IFAC World Congress*, pages 1–6, Prague, Czech Republic, 2005.