

The Sliced Gaussian Mixture Filter with Adaptive State Decomposition Depending on Linearization Error

Vesa Klumpp, Frederik Beutler,
and Uwe D. Hanebeck

Intelligent Sensor-Actuator-Systems Laboratory (ISAS),
Institute for Anthropomatics,
Karlsruhe Institute of Technology (KIT), Germany.
{klumpp,beutler}@kit.edu, uwe.hanebeck@ieee.org,

Dietrich Fränken

Data Fusion Algorithms and Software,
EADS Deutschland GmbH,
D-89077 Ulm, Germany.
dietrich.fraenken@eads.com

Abstract – In this paper, a novel nonlinear/non-linear model decomposition for the Sliced Gaussian Mixture Filter is presented. Based on the level of nonlinearity of the model, the overall estimation problem is decomposed into a “severely” nonlinear and a “slightly” nonlinear part, which are processed by different estimation techniques. To further improve the efficiency of the estimator, an adaptive state decomposition algorithm is introduced that allows decomposition according to the linearization error for nonlinear system and measurement models. Simulations show that this approach has orders of magnitude less complexity compared to other state of the art estimators, while maintaining comparable estimation errors.

Keywords: Nonlinear state estimation, state decomposition, Rao-Blackwellization.

1 Introduction

Recursive Bayesian state estimation plays an important role when dealing with measurements disturbed by noise. Examples of their field of application include robotics, localization, or tracking [1], among many others. The well-known Kalman filter [2] gives optimal results for these problems, when the models are linear. Unfortunately, often the considered problems can only be described by means of nonlinear models, which increases the need for appropriate nonlinear Bayesian estimators that are both computationally feasible and keep approximation errors low.

Different approaches to nonlinear Bayesian estimation exist. Sample-based estimators, such as particle filters [3], represent the estimated density by means of a finite set of samples, which allows to represent densities with arbitrary good precision. These filters can cope with “severely” nonlinear systems and complex multi-modal densities. Their drawback is that many samples may be needed in order to get reliable results, which is problematic especially for high-dimensional estimation problems [4].

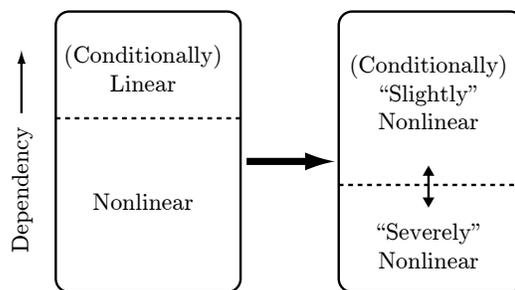


Figure 1: Modification of model decomposition. From nonlinear/linear decomposition towards an adaptive decomposition of overall nonlinear models.

For high-dimensional nonlinear estimation problems, often nonlinear filters with a Gaussian density assumption are utilized, such as the Extended Kalman filter [5], the Unscented Kalman filter [6], the linear regression Kalman filter [7], or the Gaussian Estimator [15]. This Gaussian assumption allows efficient processing, which is far less demanding than sample-based approaches, but it also limits the application of this class of filters. These filters are only capable of handling “slightly” nonlinear models, which makes them unsuitable for arbitrary estimation problems, where the estimated density can become multi-modal.

In order to reduce the computational complexity and maintain superior estimation performance, state space decomposition using Rao-Blackwellization [9] is a conceivable solution. This approach has been successfully pursued in special Rao-Blackwellized estimators, where the estimation problem is decomposed into a nonlinear and a conditionally linear part. There, the particular structure of the measurement and system models is employed in order to decompose the estimation problem, which allows to solve the overall problem in a (partially) decoupled way. For this, the marginal density of the nonlinear part is represented by means of a sample-based density, which results into linear models for the conditional part. Examples of this approach

are the Rao-Blackwellized – or marginalized – particle filter [10] or the Sliced Gaussian Mixture Filter, which has been presented in previous work [11, 12] by the authors. When the nonlinear part of the model is high-dimensional, these approaches suffer from the same growth of complexity as common sample-based methods. This only allows a limited dimensionality of the nonlinear part.

Contribution of this Paper: In this paper, the idea of using Rao-Blackwellization for nonlinear/linear model decomposition is extended to a more flexible nonlinear/nonlinear model decomposition, in order to allow a higher dimensionality of the nonlinear part. This is visualized in Fig. 1. For this, two classes of nonlinear models are defined, which depend on the level of nonlinearity of the model. For these model classes, different estimation techniques are applied. An algorithm for online state decomposition is introduced, which adaptively selects the suitable class of filters, while keeping estimation errors low and simultaneously reducing the computational complexity.

This approach is applied to the Sliced Gaussian Mixture Filter in order to extend its range of application and to increase its efficiency. When only a small substate is affected by the “severely” nonlinear model, a dramatic reduction in algorithm runtime can be observed, while maintaining superior estimation performance compared to other state of the art estimators.

2 Problem Formulation

The considered nonlinear measurement and system models for an n -dimensional state are given in the form

$$\hat{y}_k = h_k(\underline{x}_k) + \underline{v}_k \quad (1)$$

and

$$\underline{x}_{k+1} = a_k(\underline{x}_k) + \underline{w}_k, \quad (2)$$

which depend on nonlinear functions h_k and a_k . The additive measurement noise term \underline{v}_k , and the additive system noise term \underline{w}_k are zero-mean Gaussian distributed with covariance matrices \mathbf{C}_k^v and \mathbf{C}_k^w . The received measurement vector at time step k is denoted by \hat{y}_k .

The goal is to find a decomposition of the state $\underline{x}_k \in \mathbb{R}^n$ into two parts $\underline{x}_k = [(\underline{x}_k^m)^T, (\underline{x}_k^c)^T]^T$, the marginal substate $\underline{x}_k^m \in \mathbb{R}^d$ and the conditional substate $\underline{x}_k^c \in \mathbb{R}^{n-d}$, where the dependencies of the substates and the models are given by the factorization $f(\underline{x}_k) = f(\underline{x}_k^m, \underline{x}_k^c) = f(\underline{x}_k^m) \cdot f(\underline{x}_k^c | \underline{x}_k^m)$, i.e., the conditional part \underline{x}_k^c depends on \underline{x}_k^m .

For this state decomposition, Rao-Blackwellization is applied. The “hard” marginal part $f(\underline{x}_k^m)$ is represented by means of a sample-based density, i.e., a Dirac mixture density, for which the processing can be computationally demanding due to the large number of components needed. The conditional density $f(\underline{x}_k^c | \underline{x}_k^m)$ is

processed by means of “simple” nonlinear estimators, which assume a Gaussian density representation, but limit the characteristic of the nonlinear model due to implicit linearization.

In order to keep the computational complexity low and to limit the number of samples, the dimensionality d of \underline{x}_k^m should be as low as possible. For bounding estimation errors due to linearization of the conditional part, only certain parts of the overall nonlinear models may be processed by means of these “simple” nonlinear estimators. Thus, a tractable compromise between estimation accuracy and complexity has to be chosen, which makes the consideration of the induced linearization error necessary.

3 State Decomposition based on Linearization Error

3.1 Definition of Induced Linearization Error

A nonlinear mapping g

$$\underline{z} = g(\underline{x}) \approx \mathbf{G}\underline{x} + \underline{b}$$

can be approximated by means of a linear mapping \mathbf{G} with an offset \underline{b} [7]. The error term

$$\underline{e} = g(\underline{x}) - \mathbf{G}\underline{x} + \underline{b}$$

describes the difference between the nonlinear mapping g and its linearization. It has the covariance matrix

$$\mathbf{C}^e = \mathbf{C}^z - \mathbf{C}^{zx} (\mathbf{C}^x)^{-1} \mathbf{C}^{xz},$$

which quantifies the amount of induced linearization error.

The induced linearization error does not only depend on the nonlinear mapping g and its linearization, it also depends on the distribution of the random vector \underline{x} . For example, the induced linearization error is zero, if \underline{x} is distributed according to a single Dirac function. Then, the system state is exactly known and no linearization is necessary.

It can be stated that if the covariance matrix \mathbf{C}^e is zero, then the error term \underline{e} is distributed according to a single Dirac function [13]. In this case, \underline{e} is deterministic and the mapping g is an affine transformation.

When the approximation is determined by a statistical linearization, e.g., as performed approximatively in the linear regression Kalman filter [7], the Unscented Kalman filter [6], or the Gaussian Estimator [15], \underline{e} has zero mean and minimum covariance \mathbf{C}^e [7].

3.2 Induced Linearization Error for a given State Decomposition

For a particular \underline{x}_k^m , linearization is performed on the conditional measurement model

$$\hat{y}_k = h_k(\underline{x}_k^m, \underline{x}_k^c) + \underline{v}_k \quad (3)$$

and conditional system model

$$\underline{x}_{k+1} = \underline{a}_k(\underline{x}_k^m, \underline{x}_k^c) + \underline{w}_k. \quad (4)$$

The covariance of the induced linearization error can be stated for both measurement and system models and is denoted as $\mathbf{C}^{e|\underline{x}_k^m}$.

In order to assess the induced linearization error for the complete prior density, the *average accumulated linearization error* is given as

$$E = \text{trace} \left(\int f(\underline{x}_k^m) \cdot \mathbf{C}^{e|\underline{x}_k^m} d\underline{x}_k^m \right). \quad (5)$$

The trace was chosen to accumulate the induced linearization errors for every dimension when coping with multi-dimensional models.

It is now shown that E is zero, if and only if the conditional model is linear for every \underline{x}_k^m in the support of $f(\underline{x}_k^m)$. Because of the linearity of the trace operator, (5) is equal to

$$E = \int f(\underline{x}_k^m) \cdot \text{trace} \left(\mathbf{C}^{e|\underline{x}_k^m} \right) d\underline{x}_k^m.$$

The trace of a covariance matrix is non-negative, i.e., $\text{trace}(\mathbf{C}^{e|\underline{x}_k^m}) \geq 0$, and in the support of the density $f(\underline{x}_k^m) > 0$ holds. Furthermore, the trace of a covariance matrix is zero, if and only if the matrix is a zero matrix, i.e., $\text{trace}(\mathbf{C}^{e|\underline{x}_k^m}) = 0 \Leftrightarrow \mathbf{C}^{e|\underline{x}_k^m} = \mathbf{0}$. Thus, E is zero, if and only if the linearization error is zero for every \underline{x}_k^m in the support of $f(\underline{x}_k^m)$.

3.3 State Decomposition Algorithm

With the definition of the average accumulated linearization error for a given state decomposition, a suitable decomposition has to be found that keeps the linearization error low and the dimensionality d of the marginal part \underline{x}_k^m as low as possible.

A graph showing all possible state decompositions into two parts is visualized in Fig. 2. The state decompositions are given by a set of indices into the state vector $\underline{x}_k = [x_1, \dots, x_n]^T$, which describes the components of the marginal part \underline{x}_k^m in the vector \underline{x}_k . For example, the set $\{2, 3\}$ corresponds to the state decomposition $\underline{x}_k^m = [x_2, x_3]^T$, $\underline{x}_k^c = [x_1, x_4, x_5, \dots, x_n]^T$. All possible partitionings of the state vector \underline{x}_k into two parts are shown in an increasing order of d .

The optimal solution for decomposition can be found by gradually increasing the dimensionality of the marginal part d , until the linearization error is within acceptable bounds. This is performed online by a breadth-first search on the graph beginning with the highest node denoted by \emptyset going downwards.

Note that for nonlinear/linear decomposition, the optimal solution has minimum dimensionality d_{lin} and an average accumulated linearization error of zero. Now, in order to obtain a nonlinear/nonlinear decomposition, the average accumulated linearization error $E < \tau$, with

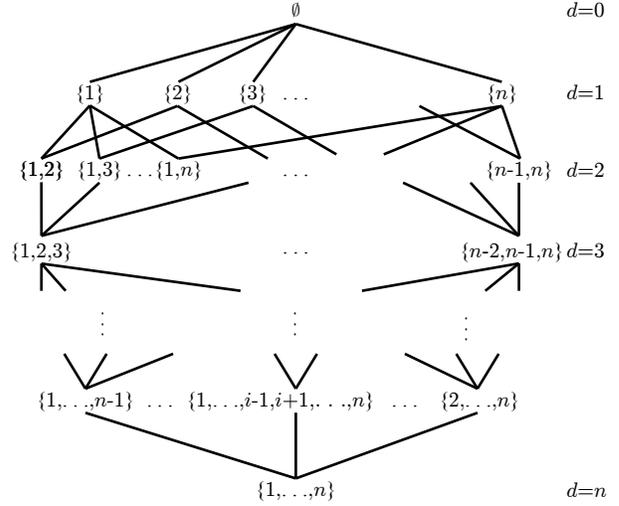


Figure 2: Graph structure visualizing all possible state space decompositions into two parts. The nodes are labelled with the sets of dimensions of one part.

a given threshold τ , is allowed. From all decompositions in Fig. 2, a decomposition with minimum dimensionality d_{nonlin} that satisfies $E < \tau$ is chosen.

The dimensionality d_{nonlin} resulting from nonlinear/nonlinear decomposition is not higher than the dimensionality d_{lin} from nonlinear/linear decomposition. By defining an threshold τ for the maximum acceptable linearization error, the dimensionality d_{nonlin} of the substate \underline{x}_k^m can be reduced, while keeping linearization errors small.

Simplification of the Algorithm: The breadth-first search in the graph can be computationally demanding, because a layer d in the graph consists of $\binom{n}{d}$ elements, which have to be evaluated. Furthermore, for growing d , the complexity for the calculation of (5) increases.

Two possible simplifications are introduced, which allow to efficiently find a suitable state decomposition:

1. Instead of performing a full breadth-first search on the graph, only a greedy graph search is performed. For a specific node, the child with the lowest linearization error is greedily selected until a solution is found. This reduces the number of state decompositions, for which the linearization error has to be evaluated, from $\binom{n}{k}$ per layer to a maximum of n .
2. The linearization error (5) can be approximated by employing certain Dirac mixture approximations of the marginal density $f(\underline{x}_k^m)$. Instead of representing $f(\underline{x}_k^m)$ by means of a high-quality Dirac mixture approximation for subsequent processing, only a crude approximation with few components is employed, for which the linearization error is evaluated. Examples are the unscented transfor-

mation [6] or the sample-based approximation in the Gaussian Estimator [15]. This results in a very fast approximation of the averaged, accumulative linearization error.

When applying these simplifications, finding the optimal solution is not guaranteed any longer. Another decomposition with a higher linearization error or higher dimensionality d could be the result of these simplifications. Although it is no longer guaranteed that the optimal state decomposition is found, these simplifications usually lead to results that are close or equal to the optimal solution. Furthermore, they still reduce the complexity of the state decomposition substantially, which is essential for efficient estimation.

4 SGMF for Nonlinear/Nonlinear Decomposition

In this section, the adaptive online state decomposition from Section 3 is applied to the Sliced Gaussian Mixture Filter, which is introduced in [11, 12].

4.1 Density Representation

Originating from a given Gaussian mixture density

$$\tilde{f}(\underline{x}_k) = \sum_{j=1}^M w_k^j \cdot \mathcal{N}(\underline{x}_k - \underline{\mu}_k^j, \mathbf{C}_k^j), \quad (6)$$

with weights w_k^j , means $\underline{\mu}_k^j$ and covariance matrices \mathbf{C}_k^j , the nonlinear measurement (1), and system model (2), a suitable state decomposition under consideration of the induced linearization error is determined according to Section 3.

For this state decomposition, the Gaussian mixture (6) can be approximated by means of a sliced Gaussian mixture density that employs a Dirac mixture representation of the marginal density $f(\underline{x}_k^m)$. The conditional density $f(\underline{x}_k^c | \underline{x}_k^m)$ is given in Gaussian mixture form, leading to the density representation

$$\begin{aligned} f(\underline{x}_k^m, \underline{x}_k^c) &= f(\underline{x}_k^m) \cdot f(\underline{x}_k^c | \underline{x}_k^m) \\ &= \sum_{i=1}^N \alpha_k^i \cdot \delta(\underline{x}_k^m - \underline{\xi}_k^i) \\ &\quad \cdot \left(\sum_{j=1}^{N^i} \beta_k^{ij} \cdot \mathcal{N}(\underline{x}_k^c - \underline{\mu}_k^{cij}, \mathbf{C}_k^{cij}) \right), \end{aligned} \quad (7)$$

with Dirac positions $\underline{\xi}_k^i$, means $\underline{\mu}_k^{cij}$, and covariance matrices \mathbf{C}_k^{cij} . A tractable solution for density approximation is discussed in [12], where a systematic Dirac mixture approximation algorithm is applied to the marginal $f(\underline{x}_k^m)$ and the approximation is extended over the complete state space by evaluating the given density (6).

The filter and prediction step employ this special density representation, which makes the density approximation step essential in the SGMF processing cycle, which is visualized in Fig. 3.

4.2 Density Processing

The density processing now slightly differs from [12]. The conditional substate \underline{x}_k^c depends on nonlinear parts of the models, which makes them intractable for the Kalman filter. Instead, due to the ‘‘slightly’’ nonlinear parts of the models, which have low linearization errors, special nonlinear estimators with a Gaussian density assumption can be employed for processing the Gaussian components in the slices.

Filter Step: For the substate \underline{x}_k^c , any nonlinear Bayesian estimator that fulfills certain requirements can be applied to the measurement model. The estimator has to provide the estimated density by means of a Gaussian component with mean and covariance matrix and it has to supply a weighting factor, i.e., the integral over the product of prior density and likelihood. The last condition is usually met by assuming a jointly Gaussian distribution about the prior density and the predicted measurement.

For a single slice, the component \underline{x}_k^m of the system state is given and the measurement model results in (3). In order to keep the linearization error low, every Gaussian component from a slice in (7) is processed separately. A nonlinear estimator with a jointly Gaussian distribution between a Gaussian over the conditional substate and the predicted measurement approximates

$$\begin{aligned} &f^{ij}(\underline{y}_k, \underline{x}_k^c | \underline{\xi}_k^i) \\ &= \mathcal{N}(\underline{y}_k - \underline{h}_k(\underline{\xi}_k^i, \underline{x}_k^c), \mathbf{C}_k^v) \\ &\quad \cdot \mathcal{N}(\underline{x}_k^c - \underline{\mu}_k^{cij}, \mathbf{C}_k^{cij}) \\ &\approx \mathcal{N}\left(\begin{bmatrix} \underline{y}_k \\ \underline{x}_k^c \end{bmatrix} - \begin{bmatrix} \tilde{\underline{y}}_k^{ij} \\ \underline{\mu}_k^{cij} \end{bmatrix}, \begin{bmatrix} \mathbf{C}_k^{yij} & \mathbf{C}_k^{ycij} \\ \mathbf{C}_k^{cyij} & \mathbf{C}_k^{cij} \end{bmatrix}\right) \end{aligned} \quad (8)$$

by means of a Gaussian density. The parameters $\tilde{\underline{y}}_k^{ij}$, \mathbf{C}_k^{yij} , and \mathbf{C}_k^{ycij} of the density (9) are usually calculated by linearizing the measurement model (3) by means of system linearization (EKF), stochastic linearization in case of sample-based methods (UKF, Gaussian Estimator), or with analytic moment calculation [16]. The estimated mean and covariance of the slice component are calculated by conditioning (8) to $\underline{y}_k = \hat{\underline{y}}_k$, which can be approximated by

$$f^{ij}(\underline{x}_k^c | \hat{\underline{y}}_k, \underline{\xi}_k^i) \approx \mathcal{N}(\underline{x}_k^c - \underline{\lambda}_k^{ij}, \mathbf{D}_k^{ij}),$$

with parameters $\underline{\lambda}_k^{ij}$ and \mathbf{D}_k^{ij} of the conditional Gaussian from Equation (9), i.e.,

$$\begin{aligned} \underline{\lambda}_k^{ij} &= \underline{\mu}_k^{cij} + \mathbf{C}_k^{cyij} \left(\mathbf{C}_k^{yij} \right)^{-1} (\hat{\underline{y}}_k - \tilde{\underline{y}}_k^{ij}), \\ \mathbf{D}_k^{ij} &= \mathbf{C}_k^{cij} - \mathbf{C}_k^{cyij} \left(\mathbf{C}_k^{yij} \right)^{-1} \mathbf{C}_k^{cyij}. \end{aligned}$$

When processing Gaussian mixtures, a weighting factor, i.e., the integral over the product of every Gaussian

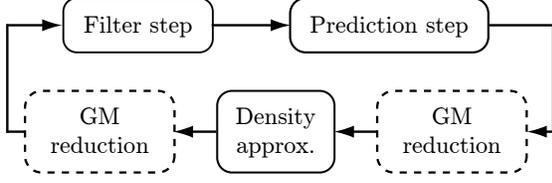


Figure 3: Essential steps in the SGMF processing sequence.

and the likelihood, has to be considered. In this case, because the linearization error is assumed to be small, it can be calculated according to

$$\gamma_k^{ij} = \int f(\hat{\underline{y}}_k, \underline{x}_k^c | \underline{\xi}_k^i) d\underline{x}_k^c \approx \mathcal{N}(\hat{\underline{y}} - \tilde{\underline{y}}_k^{ij}, \mathbf{C}_k^{yij}) .$$

Thus, the estimated sliced Gaussian mixture density is given as

$$f^e(\underline{x}_k^m, \underline{x}_k^c) = c_k \cdot \sum_{i=1}^N \alpha_k^i \cdot \delta(\underline{x}_k^m - \underline{\xi}_k^i) \cdot \left(\sum_{j=1}^{N^i} \beta_k^{ij} \cdot \gamma_k^{ij} \cdot \mathcal{N}(\underline{x}_k^c - \underline{\lambda}_k^{ij}, \mathbf{D}_k^{ij}) \right) , \quad (10)$$

with the normalizing factor

$$c_k = \left(\sum_{i=1}^N \sum_{j=1}^{N^i} \alpha_k^i \cdot \beta_k^{ij} \cdot \gamma_k^{ij} \right)^{-1} .$$

Prediction Step: For the prediction step, the only requirement for the nonlinear estimator applied to the slice components is that the predicted density is a Gaussian density.

The conditional system model is given by equation (4). Again, it is assumed that the induced linearization error is small in this case.

Just as in the filter step, every component of the slice is processed separately. The joint density of the estimated Gaussian component in (10) and its prediction is given as

$$f^{ij}(\underline{x}_{k+1}, \underline{x}_k^c | \underline{\xi}_k^i) = \mathcal{N}(\underline{x}_{k+1} - \underline{a}_k(\underline{\xi}_k^i, \underline{x}_k^c), \mathbf{C}_k^w) \cdot \mathcal{N}(\underline{x}_k^c - \underline{\lambda}_k^{ij}, \mathbf{D}_k^{ij}) , \quad (11)$$

which is Gaussian when the system model is linear. In order to receive a desired predicted Gaussian density, again different tasks can be performed, e.g., linearization of the system model as in the EKF, or by calculating mean and covariance of (11), for example by sampling $f(\underline{x}_k^c)$, as pursued in the UKF and the Gaussian Estimator. This can be described by an approximation of the marginal over the predicted state of the joint density (11) by a Gaussian

$$f^{ij}(\underline{x}_{k+1} | \underline{\xi}_k^i) = \int f^{ij}(\underline{x}_{k+1}, \underline{x}_k^c | \underline{\xi}_k^i) d\underline{x}_k^c \approx \mathcal{N}(\underline{x}_{k+1} - \underline{\mu}_{k+1}^{ij}, \mathbf{C}_{k+1}^{ij}) .$$

Finally, the predicted Gaussian mixture density is given by

$$f^p(\underline{x}_{k+1}) = \sum_{i=1}^N \sum_{j=1}^{N^i} \nu_{k+1}^{ij} \mathcal{N}(\underline{x}_{k+1} - \underline{\mu}_{k+1}^{ij}, \mathbf{C}_{k+1}^{ij}) ,$$

with the weighting factor

$$\nu_{k+1}^{ij} = \alpha_k^i \cdot \beta_k^{ij} \cdot \gamma_k^{ij} \cdot c_k .$$

Gaussian Mixture Reduction: For computational feasibility, a Gaussian mixture component reduction is necessary, which prevents exponential growth in the number of components, as discussed in [12]. Possible locations of the component reduction phase in the SGMF processing cycle are visualized in Fig. 3.

5 Simulations

5.1 Adaptive State Space Decomposition

This simulation shows the adaptive behaviour of the online state decomposition algorithm. The nonlinear system model for a ten-dimensional system is given by

$$\underline{x}_{k+1} = \begin{bmatrix} 0.2\mathbf{x}_k^{(1)} + 0.3\mathbf{x}_k^{(2)} + \left(\mathbf{x}_k^{(5)}\right)^2 \\ 0.2\mathbf{x}_k^{(2)} \mathbf{x}_k^{(8)} \\ \mathbf{x}_k^{(3)} + \cos\left(\mathbf{x}_k^{(4)}\right) \\ \sin^2\left(\mathbf{x}_k^{(4)}\right) \\ 0.2\left(\mathbf{x}_k^{(5)} + \mathbf{x}_k^{(6)} + \mathbf{x}_k^{(7)}\right) \\ \sqrt{|\mathbf{x}_k^{(6)}|} + \mathbf{x}_k^{(1)} \\ \mathbf{x}_k^{(6)} / \mathbf{x}_k^{(7)} + \sqrt{|\mathbf{x}_k^{(9)}|} \\ \mathbf{x}_k^{(3)} + \mathbf{x}_k^{(8)} \\ \left(0.2\mathbf{x}_k^{(2)} \mathbf{x}_k^{(9)}\right) / \left(\mathbf{x}_k^{(6)} + \mathbf{x}_k^{(7)}\right) \\ \sin\left(\mathbf{x}_k^{(10)}\right) + \mathbf{x}_k^{(3)} \end{bmatrix} + \underline{w}_k ,$$

with zero mean Gaussian distributed system noise \underline{w}_k , where $x^{(i)}$ denotes the i -th element of a vector \underline{x}_k . The measurement model is also nonlinear and is given by

$$\begin{bmatrix} \hat{\underline{y}}_k^{(1)} \\ \hat{\underline{y}}_k^{(2)} \\ \hat{\underline{y}}_k^{(3)} \end{bmatrix} = \begin{bmatrix} \mathbf{x}_k^{(1)} \mathbf{x}_k^{(4)} \mathbf{x}_k^{(7)} + \mathbf{x}_k^{(8)} \\ \left(\mathbf{x}_k^{(5)} + \mathbf{x}_k^{(10)}\right) \mathbf{x}_k^{(2)} + \mathbf{x}_k^{(4)} + \mathbf{x}_k^{(5)} + \mathbf{x}_k^{(6)} \\ \mathbf{x}_k^{(3)} \mathbf{x}_k^{(6)} + \sin\left(\mathbf{x}_k^{(4)}\right) \cos\left(\mathbf{x}_k^{(8)}\right) \end{bmatrix} + \underline{v}_k .$$

For adaptive state space decomposition, a threshold of $\tau = 0.015$ was chosen. The simulation was performed for ten time steps, where three-dimensional measurements are available at time steps 1,2,5,6,7,9, and 10.

For ten independent simulation runs, the average dimensionality d of \underline{x}_k^m is visualized in Fig. 4. For every time step, the system model adds uncertainty to the state estimate, and thus, a higher-dimensional substate

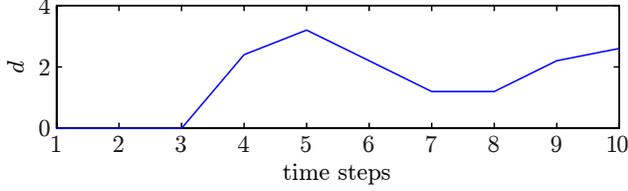


Figure 4: Average dimensionality d of \underline{x}_k^m for different time steps.

has to be processed by means of a Dirac mixture density due to the nonlinear system model. For the first time steps, $d = 0$ holds, and thus, the complete nonlinear model can be processed with a Gaussian assumed density filter. At the time steps 3 and 4, when no measurements are available, the uncertainty increases and the induced linearization error exceeds the threshold τ , which results in an increase of d . By incorporating measurements into the state estimate, this effect can be revoked and the dimensionality of d decreases. The same effect occurs for the missing measurement at time step 8.

For nonlinear/linear state space decomposition, d would be 7. The 7-dimensional nonlinear part would have to be processed by sample-based approaches and would lead to a much higher computational demand. Thus, in the proposed approach the dimensionality of the nonlinear marginal part is substantially reduced, which leads to a more efficient state estimate.

5.2 Decomposition for Reducing Computational Complexity

The simulation example pursued here is bearings-only tracking [1], or cooperative passive target tracking, where the task is to track a moving target by means of angular measurements from multiple sensors.

Simulation Setup: The simulation setup is adopted from [12], where the target is moving in a two-dimensional plane on an S-shaped curve passing two sensors located at the positions $[0 \text{ km}, 0 \text{ km}]^T$ and $[10 \text{ km}, 0 \text{ km}]^T$, as visualized in Fig. 5. The five-dimensional state space vector $\underline{x}_k = [p_k, q_k, \dot{p}_k, \dot{q}_k, \omega_k]^T$ consists of the target's position p_k, q_k , its velocity \dot{p}_k, \dot{q}_k , and turn rate ω_k . Nonlinear/linear decomposition results into $\underline{x}_k = [(\underline{x}_k^n)^T, (\underline{x}_k^l)^T]^T = [(\underline{x}_k^m)^T, (\underline{x}_k^c)^T]^T$, with $\underline{x}_k^n = [\omega_k, p_k, q_k]^T$, $\underline{x}_k^l = [\dot{p}_k, \dot{q}_k]^T$, and a nonlinear/nonlinear decomposition is given by $\underline{x}_k^m = \omega_k$, and $\underline{x}_k^c = [p_k, q_k, \dot{p}_k, \dot{q}_k]^T$.

The measurement equation relates the target position to the two measured angles originating from both sensors according to

$$\begin{bmatrix} \hat{\varphi}_k^{(1)} \\ \hat{\varphi}_k^{(2)} \end{bmatrix} = \begin{bmatrix} \text{atan2}\left(\mathbf{q}_k - \mathbf{q}_S^{(1)}, \mathbf{p}_k - \mathbf{p}_S^{(1)}\right) \\ \text{atan2}\left(\mathbf{q}_k - \mathbf{q}_S^{(2)}, \mathbf{p}_k - \mathbf{p}_S^{(2)}\right) \end{bmatrix} + \underline{\mathbf{v}}_k,$$

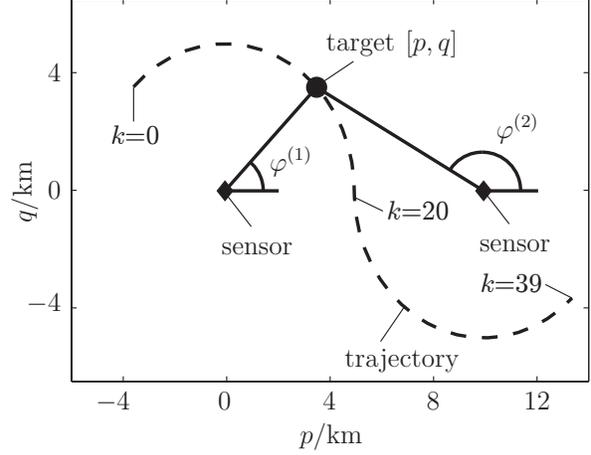


Figure 5: Visualization of cooperative passive target tracking with angular measurements.

where $\underline{\mathbf{v}}_k$ is zero-mean Gaussian distributed measurement noise with covariance matrix $\mathbf{C}_k^v = \text{diag}((1.5^\circ)^2, (1.5^\circ)^2)$. The motion of the target is modeled by

$$\begin{bmatrix} \mathbf{p}_{k+1} \\ \mathbf{q}_{k+1} \\ \dot{\mathbf{p}}_{k+1} \\ \dot{\mathbf{q}}_{k+1} \\ \omega_{k+1} \end{bmatrix} = \begin{bmatrix} T \text{si}(\omega_k T) & -T \text{co}(\omega_k T) \\ T \text{co}(\omega_k T) & T \text{si}(\omega_k T) \\ \cos(\omega_k T) & -\sin(\omega_k T) \\ \sin(\omega_k T) & \cos(\omega_k T) \\ 0 & 0 \end{bmatrix} \cdot \begin{bmatrix} \dot{\mathbf{p}}_k \\ \dot{\mathbf{q}}_k \end{bmatrix} + \begin{bmatrix} \mathbf{p}_k \\ \mathbf{q}_k \\ 0 \\ 0 \\ \omega_k \end{bmatrix} + \underline{\mathbf{w}}_k.$$

Note that the functions si and co are defined as

$$\text{si}(x) = \begin{cases} \frac{\sin x}{x}, & x \neq 0 \\ 1, & x = 0 \end{cases}, \quad \text{co}(x) = \begin{cases} \frac{1 - \cos x}{x}, & x \neq 0 \\ 0, & x = 0 \end{cases}.$$

The covariance \mathbf{C}_k^w of the system noise term is given as

$$\mathbf{C}_k^w = \begin{bmatrix} \rho_k^2 \mathbf{Q}_0 & \mathbf{0} \\ \mathbf{0}^T & T^2 a_{\omega_k}^2 \end{bmatrix}, \quad \mathbf{Q}_0 = \begin{bmatrix} T^3/3 \cdot \mathbf{I}_2 & T^2/2 \cdot \mathbf{I}_2 \\ T^2/2 \cdot \mathbf{I}_2 & T \cdot \mathbf{I}_2 \end{bmatrix},$$

with the identity matrix $\mathbf{I}_2 \in \mathbb{R}^{2 \times 2}$ and the parameters $T = 2 \text{ s}$, $\rho_k^2 = 9.6236 \text{ m}^2 \text{ s}^{-3}$, and $a_{\omega_k}^2 = 0.07 \text{ s}^{-4}$.

Simulation Results: In these simulations, the novel nonlinear/nonlinear state decomposition for the Sliced Gaussian Mixture Filter (N-SGMF) is compared to the nonlinear/linear state composition of the SGMF and the marginalized particle filter (MPF).

The numbers of samples and Gaussian components are chosen to yield about the same estimation performance for these three filters. The MPF uses 16384 particles with attached Gaussian components and for the SGMF 1024 slices were used, which were reduced to two components each, resulting in 2048 Gaussian components for the sliced Gaussian mixture density. The new N-SGMF approach takes 32 slices, where the conditional density for every slice was reduced to a

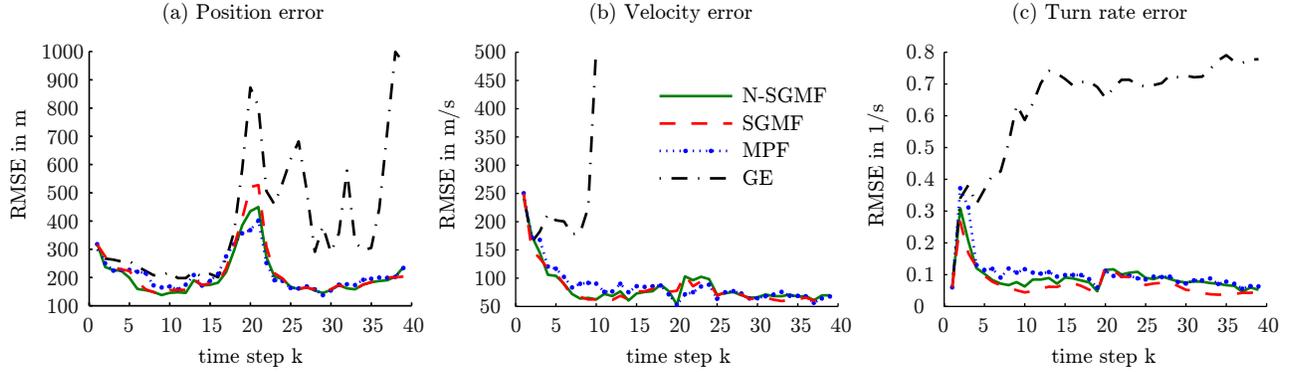


Figure 6: Comparison of the proposed algorithm (N-SGMF), the Sliced Gaussian Mixture Filter (SGMF), the marginalized particle filter (MPF), and the Gaussian Estimator (GE): The root-mean-square errors of the position, velocity, and turn rate averaged over 100 simulation runs for time steps 1 to 39.

single Gaussian, which was estimated with the Gaussian Estimator employing 17 sample points for 4 dimensions. Note that the conditional substate for nonlinear/linear state decomposition is three-dimensional, whereas for nonlinear/nonlinear decomposition it is only one-dimensional. Hence, fewer components are necessary for the N-SGMF. As reference, the Gaussian Estimator (GE) without any state decomposition is used with 21 sample points.

100 independent simulation runs were performed with 39 time steps for the S-shaped trajectory. Fig. 6 shows the root-mean-square error for the position, velocity, and turn rate over time for the compared algorithms. The figure shows similar estimation performance for the MPF, SGMF, and N-SGMF. The GE is not able to track the target satisfactorily over the complete time horizon.

Even though the estimation performance of the MPF, SGMF, and N-SGMF are similar, the number of used Gaussian components and the computational effort differ substantially. Table 1 lists the number of operations required. Here, only the numbers for one estimation cycle, i.e., approximation, filter step, and prediction step, were shown. The numbers only consider the processing of the conditional parts of the sliced Gaussian mixture density or the Gaussians associated to the particles in the MPF. Only matrix and vector operations were counted, omitting single scalar operations, the point-wise evaluation of a Gaussian mixture density, operations for random sampling, and Dirac mixture approximations. Furthermore, the costs of the Gaussian mixture reduction algorithm for the SGMF was left out, which contributes in a significant way to the SGMF run time. The table shows that the number of operations used in the proposed algorithm are about 50 to 100 times lower compared to the previous SGMF approach and still about 20 times lower than the MPF for most of the compared operation types.

This simulation shows the advantage of the proposed algorithm against state of the art estimators, such as

the previous SGMF approach and the MPF by applying the novel state decomposition. The number of operations and Gaussian components is dramatically reduced, while maintaining superior estimation quality. The new approach is able to handle complex nonlinear problems where simple estimators do not yield satisfactory results.

6 Conclusions

In this paper, a novel nonlinear/nonlinear state decomposition method for the SGMF is proposed. The state space is decomposed into two parts, where different estimation techniques are applied, depending on the level of nonlinearity of the model. The idea is to use nonlinear estimation techniques that make use of a Gaussian assumption for subspaces where the induced linearization error is low. By this means, the dimensionality of the remaining (highly) nonlinear subspace is decreased, where an expensive Dirac mixture approximation and processing has to be performed.

In order to obtain a suitable state decomposition of this type, linearization errors have to be considered. The average accumulated linearization error describes the error that occurs by nonlinear/nonlinear state decomposition and application of the estimators employing the Gaussian assumption. Because it depends on the prior density given, the state decomposition has to be performed online before density processing.

This state decomposition leads to more efficient processing of nonlinear models and reduces the number of slices and Gaussian components consumed in the SGMF dramatically. In comparison to other nonlinear state of the art estimators, the proposed approach is able to cope with complex nonlinear models, while consuming orders of magnitude less vector and matrix operations. Nevertheless, some challenges remain for the SGMF:

- Optimal density approximation: A sliced Gaussian mixture density approximation that minimizes approximation errors is desirable.

Table 1: Comparison of the number of vector and matrix operations for processing all conditional densities in the N-SGMF, SGMF, and MPF.

Number of operations	N-SGMF	SGMF	MPF
Eigen decomposition	64	0	0
Matrix inversion	1088	$2.63 \cdot 10^5$	16384
Matrix multiplication	5088	$2.71 \cdot 10^5$	147456
Matrix vector mult.	3104	$2.66 \cdot 10^5$	65536
Matrix addition	2112	$2.64 \cdot 10^5$	32768
Vector addition	4320	$2.64 \cdot 10^5$	16384

- Complete prevention of degeneration: Degeneration can occur in the filter step when the weights of the slices are updated according to the likelihood. This could be avoided by considering the likelihood directly in the density approximation process.
- Omitting the need for Gaussian mixture component reduction: Due to the increasing number of Gaussian mixture components during processing [12], a component reduction is needed. Other ways for approximation or processing the densities, which eliminate the need for a component reduction step, should be found.
- Consideration of linear sub-structures in the models: A hierarchical decomposition into three parts is possible. The conditional substate \underline{x}_k^c can further be split by a nonlinear/linear decomposition. This is not in the scope of this paper, but it can easily be incorporated into this framework, e.g., by employing [8].

References

- [1] B. Ristic and M. S. Arulampalam, “Tracking a Manoeuvring Rarget using Angle-only Measurements: Algorithms and Performance,” *Signal Processing*, vol. 83, no. 6, pp. 1223–1238, Jan. 2003.
- [2] 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.
- [3] M. S. Arulampalam, S. Maskell, N. Gordon, and T. Clapp, “A Tutorial on Particle Filters for On-line Nonlinear/Non-Gaussian Bayesian Tracking,” *IEEE Transactions on Signal Processing*, vol. 50, no. 2, pp. 174–188, Feb. 2002.
- [4] F. Daum and J. Huang, “Curse of Dimensionality and Particle Filters,” in *Proceedings of the IEEE Aerospace Conference*, 2003.
- [5] D. Simon, *Optimal State Estimation: Kalman, H-infinity, and Nonlinear Approaches*. John Wiley & Sons, 2006.
- [6] S. J. Julier and J. K. Uhlmann, “Unscented Filtering and Nonlinear Estimation,” *Proceedings of the IEEE*, vol. 92, no. 3, pp. 401–422, 2004.
- [7] T. Lefebvre, H. Bruyninckx, and J. D. Schutter, “Kalman Filters for Nonlinear Systems: a Comparison of Performance,” *International Journal of Control*, vol. 77, no. 7, pp. 639–653, May 2004.
- [8] F. Beutler, M. F. Huber, and U. D. Hanebeck, “Gaussian Filtering using State Decomposition Methods,” in *Proceedings of the 12th International Conference on Information Fusion (Fusion 2009)*, Seattle, Washington, July 2009.
- [9] A. Doucet, S. Godsill, and C. Andrieu, “On Sequential Monte Carlo Sampling Methods for Bayesian Filtering,” *Statistics and Computing*, vol. 10, no. 3, pp. 197–208, 2000.
- [10] T. Schön, F. Gustafsson, and P.-J. Nordlund, “Marginalized Particle Filters for Mixed Linear/Nonlinear State-Space Models,” *IEEE Transactions on Signal Processing*, vol. 53, no. 7, pp. 2279–2287, Jul. 2005.
- [11] V. Klumpp, F. Sawo, U. D. Hanebeck, and D. Fränken, “The Sliced Gaussian Mixture Filter for Efficient Nonlinear Estimation,” in *Proceedings of the 11th International Conference on Information Fusion (Fusion 2008)*, Cologne, Germany, Jul. 2008, pp. 1–8.
- [12] J. Hörst, F. Sawo, V. Klumpp, U. D. Hanebeck, and D. Fränken, “Extension of the Sliced Gaussian Mixture Filter with Application to Cooperative Passive Target Tracking,” in *Proceedings of the 12th International Conference on Information Fusion (Fusion 2009)*, Seattle, Washington, July 2009.
- [13] K. R. Parthasarathy, *Probability Measures on Metric Spaces*. Providence, RI: AMS Chelsea Publishing, 2005, vol. Reprinted.
- [14] U. D. Hanebeck, M. F. Huber, and V. Klumpp, “Dirac Mixture Approximation of Multivariate Gaussian Densities,” in *Proceedings of the 2009 IEEE Conference on Decision and Control (CDC 2009)*, Shanghai, China, Dec. 2009.
- [15] M. F. Huber and U. D. Hanebeck, “Gaussian Filter based on Deterministic Sampling for High Quality Nonlinear Estimation,” in *Proceedings of the 17th IFAC World Congress (IFAC 2008)*, vol. 17, no. 2, Seoul, Korea, Jul. 2008.
- [16] A. Papoulis and S. U. Pillai, *Probability, Random Variables and Stochastic Processes*, 4th ed. McGraw-Hill, 2002.