Tensor Decomposition-Based Multitarget Tracking in Cluttered Environments

FELIX GOVAERS
BRUNO DEMISSIE
ALTAMASH KHAN
MARTIN ULMKE
WOLFGANG KOCH

Many real-world applications of target tracking and state estimation are nonlinear filtering problems and can therefore not be solved by closed-form analytical solutions. In the recent past, tensor-based approaches have become increasingly popular due to very effective decomposition algorithms, which allow a compressed representation of discretized, high-dimensional data. It has been shown that by means of a Kronecker format of the Fokker–Planck equation, the Bayesian recursion for prediction and filtering can be solved for probability densities in a canonical polyadic decomposition (CPD). In this paper, the application of this approach on tracking multiple targets in a cluttered environment is presented. It is shown that intensity or probability hypothesis density-based filters can well be implemented using the CPD tensor format.

Manuscript received February 11, 2019; released for publication July 8, 2019.

The authors are with Fraunhofer FKIE, Wachtberg 53343, Germany. E-mail: {felix.govaers, bruno.demissie, altamash.khan, martin.ulmke, wolfgang.koch}@fkie.fraunhofer.de

1557-6418/19/$17.00 © 2019 JAIF

I. INTRODUCTION

As sensors become more and more ubiquitous, powerful and practical algorithms are required to extract relevant information from noisy, contradicting, ambiguous, or erroneous measurements. Often, the theory of Bayesian state estimation is a good choice to develop algorithms that cope with these challenges. Recent results of different research groups demonstrate that the common recursion of filtering and prediction can also be solved based on a decomposed tensor representation of the underlying multivariate density function. It is well known that by means of a CANDECOMP/PARAFAC decomposition (CPD) of a discretized density, a computationally effective representation can be achieved. Despite the fact that there are still open challenges, we might say that CPD representations are the most promising candidates to beat the curse of dimensionality in many research domains, including nonlinear filtering. In this paper, we apply the CPD-based state estimation to the problem of multitarget tracking in cluttered environments. We derive all update equations for a CPD tensor in the case of a single target and multiple targets when false measurements are present. An evaluation shows the performance of the approach.

The theory of target tracking has exposed a growing family of algorithms to compute the probability density function (pdf) of a system state based on noise-corrupted sensor observations. An estimate of the state is then obtained by taking the mean of the pdf.[1] The corresponding covariance matrix additionally provides a measure of accuracy for this estimate. Bayesian estimation is the framework of recursive filtering methodologies that allow us to process a current measurement by means of a *prior* or *initial* density and a measurement likelihood function that statistically describes the performance of the sensor. Thus, a tracking algorithm is an iterative updating scheme for calculating a conditional pdf $p(\mathbf{x}_k|\mathcal{Z}^k)$ that represents all available knowledge on the object state $\mathbf{x}_k$ at some time t_k, which typically is chosen as the present time. The densities are explicitly conditioned on the sensor data time series $\mathcal{Z}^k$. The iterative scheme consists of two processing steps per update cycle, referred to as *prediction* and *filtering*. The manipulation of the probability densities is given by the following basic equations (see [1], [2] for instance).

Prediction. Assuming the Markov property of the underlying object state, the prediction density $p(\mathbf{x}_k|\mathcal{Z}^{k-1})$ is obtained by combining the evolution model $p(\mathbf{x}_k|\mathbf{x}_{k-1})$ with the previous filtering density $p(\mathbf{x}_{k-1}|\mathcal{Z}^{k-1})$:

$$p(\mathbf{x}_{k-1}|\mathcal{Z}^{k-1}) \xrightarrow[\text{constraints}]{\text{evolution model}} p(\mathbf{x}_k|\mathcal{Z}^{k-1})$$

$$p(\mathbf{x}_k|\mathcal{Z}^{k-1}) = \int \mathrm{d}\mathbf{x}_{k-1}\; \underbrace{p(\mathbf{x}_k|\mathbf{x}_{k-1})}_{\text{evolution model}}\; \underbrace{p(\mathbf{x}_{k-1}|\mathcal{Z}^{k-1})}_{\text{previous filtering}}. \quad (1)$$

[1] Depending on the scenario, for instance, the expectation value, the maximum value, the median, or other statistics of the pdf can be used.

Filtering. The filtering density $p(\mathbf{x}_k|\mathcal{Z}^k)$ is obtained using the Bayes theorem by combining the sensor model $p(\mathbf{z}_k|\mathbf{x}_k)$, also called the "likelihood function," with the prediction density $p(\mathbf{x}_k|\mathcal{Z}^{k-1})$ according to

$$p(\mathbf{x}_k|\mathcal{Z}^{k-1}) \xrightarrow[\text{sensor model}]{\text{current sensor data}} p(\mathbf{x}_k|\mathcal{Z}^k)$$

$$p(\mathbf{x}_k|\mathcal{Z}^k) = \frac{p(\mathbf{z}_k|\mathbf{x}_k)\, p(\mathbf{x}_k|\mathcal{Z}^{k-1})}{\int d\mathbf{x}_k \underbrace{p(\mathbf{z}_k|\mathbf{x}_k)}_{\text{sensor model}} \underbrace{p(\mathbf{x}_k|\mathcal{Z}^{k-1})}_{\text{prediction}}}. \qquad (2)$$

According to this paradigm, an *object track* represents all relevant knowledge on a time-varying object state of interest, including its history and measures that describe the quality of this knowledge. As a technical term, "track" is therefore either a synonym for the collection of densities $p(\mathbf{x}_l|Z^l), l = 1, \ldots, k, \ldots$, or of suitably chosen parameters characterizing them, such as estimates and the corresponding estimation error covariance matrices.

An analytical solution to a recursive computation of these densities is given for instance by the Kalman filter in the case of linear Gaussian models [3]. For nonlinear scenarios, only approximate solutions are feasible. The first-order Taylor approximation is called the *extended Kalman filter* (EKF) that has low computation cost, due to its analytic solution of the prediction and filtering steps (see [1] for instance). The performance of the linearization can be improved by means of deterministic samples chosen at the local neighborhood of the current estimate. This algorithm is known as the *unscented Kalman filter* (UKF) [4]. The term *particle filter* (PF) subsumes all kinds of numerical solutions with nondeterministic samples. Here, knowledge about the state typically is represented by a set of state samples, which implies that the density is approximated by a Dirac mixture. Because the process noise terms are simulated by means of appropriately sampled random vectors, these methods are also known as *sequential Monte Carlo* (SMC) methods.

In the literature, a variety of particle filter algorithms can be found [5]. Still, the basic *sampling importance resampling* (SIR) particle filter [6] is often used due to its robustness. The main drawback of the SIR-PF is that it can suffer from impoverishment of the particle weights. For numerical reasons, resampling has to be used in order to avoid the particles to degenerate. More recently, new algorithms have been proposed based on a log-homotopy transition between the prior and the posterior. For instance, the Daum–Huang filters (see [7] and the references therein) model this transition phase in terms of a physical flow that is determined by a "force" induced by the measurement. This leads to a stochastic differential equation (SDE) that then can be solved numerically by introducing a discretized pseudotime evolving from the prior to the posterior pdf. However, the computation time for solving the SDE is often quite high for standard target tracking scenarios [8]. A different homotopy approach is provided by the *progressive filter* that was published by Hanebeck in [9]. In the progressive filter, an incremental inclusion of the likelihood function is achieved by a partition of the exponent of the likelihood going from zero to one. This prevents particle impoverishment by means of frequent resampling and an appropriately chosen step size. A Kullback–Leibler divergence-based approach to obtain the posterior particles is proposed in [10]. The resulting algorithm is similar to the *ensemble Kalman filter* (EnKF)-based filter proposed in [11], however, the additional noise term in the Kalman-based update is different. The EnKF adds some zero mean Gaussian-distributed noise to the measurement of each sample and applies Kalman filter update equations for each particle. As a consequence, a fast filter results that is consistent and performs well in nonlinear scenarios. The EnKF also has been extended to Gaussian mixtures in [12] and [13].

Instead of nondeterministic samples we can also use a grid of equidistant space points to represent the pdf in the field of view. The prior pdf is then obtained by solving the *Fokker–Planck equation* (FPE), which is equivalent to the integral formulation in (1). Challa and Bar-Shalom for instance use finite differences in [14] to obtain the solution of the FPE and show that a consistent result is obtained even for highly nonlinear problems with large noise variances. This static approach has not become as popular as the particle filters due to the higher computational load. There is a notable change in the way of thinking since it was discovered that separated representations of discretized multidimensional functions have surprisingly good approximation properties [15]. Nowadays it may even be seen as the only known way to overcome the curse of dimensionality [16], at least in cases, where the models can easily be approximated by factorized functions. In other words, approximations by separable functions are of particular interest when the dimensionality of the problem becomes large. To the authors' knowledge, a first attempt to integrate the tensor decomposition into a Bayesian estimation framework was given by Sun and Kumar [17]. Further development of the tensor-based approach and some tracking examples were given in [18] and [19].

In this paper, we demonstrate the performance of the tensor decomposition approach for realistic tracking problems with nonlinear measurement models. Further, we show that it can also be applied to multitarget tracking using the set theory-based approaches where an intensity function is computed instead of a pdf. Multiple numerical evaluations are shown to demonstrate that the tensor decomposition approach can achieve convincing results in terms of estimation accuracy in nonlinear filtering problems.

A. Structure

This paper is structured as follows. In Section II, we review the basic concept of nonlinear filtering using the

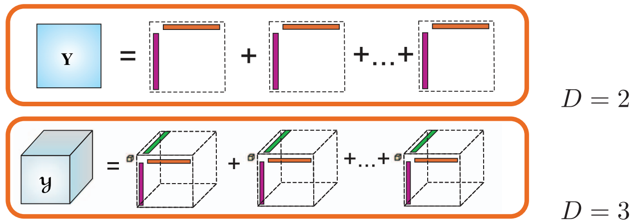

Fig. 1. Scheme of a tensor decomposition along its dimensions using the Kronecker product of vectors for two and three dimensions. Scheme taken from [21].

tensor decomposition approach. This is extended for the handling of nondetections and clutter measurements in Section III. The application on the set-based estimation theory for multitarget tracking is given in Section IV. Then, in Section V numerical examples are given to demonstrate the performance of the approach. The paper ends with conclusions in Section VI.

II. NONLINEAR STATE ESTIMATION BASED ON TENSOR DECOMPOSITION

It was already discovered in 1927 by Hitchcock that a D-way tensor can be represented as a sum of outer products [20]:

$$\mathcal{Y} = \sum_{l=1}^{L} \mathbf{y}_{1,l} \circ \cdots \circ \mathbf{y}_{D,l} \tag{3}$$

where L is the number of components. If for all $d \in \{1, \ldots, D\}$ $\mathbf{y}_{d,l}$ are of a given size $N_d \times 1$, in other words a vector, this representation of the tensor is called *canonical polyadic decomposition* or *CANDECOMP/PARAFAC decomposition*.[2] In this case, the $\mathbf{y}_{d,l}$ are called the *loading vectors*. This decomposition of a tensor is visualized for $D = 2$ and $D = 3$, respectively, in Fig. 1. It should be noted that this decomposition of a two-way tensor (matrix) can easily be obtained by the *singular value decomposition* (SVD). For higher dimensions the problem of finding a decomposed representation becomes NP-hard [22]. However, numerical solutions, such as the *alternating least squares* (ALS) algorithm, exist [23], which yield satisfying results for the problems addressed here in manageable time. For a fixed dimension d, it is assumed that the state space can be discretized into N_d grid points. These can be uniformly spaced with a fixed step size of Δ_{x_d} or chosen specifically for a numerical differentiation such as Chebyshev polynomials [18]. The probability density function restricted to the discretized state space points yields a D-way tensor, which approximates the original function:

$$p(\mathbf{x}_k|\mathcal{Z}^k) \approx [p([\mathbf{x}_d]_i|\mathcal{Z}^k)]_{n_1,\ldots,n_D}. \tag{4}$$

[2]An explanation of the abbreviations involved can be found in [21] and the references therein.

Throughout this paper, this tensor is represented in a decomposed form. Thus, the pdf at time t_k is approximated, again, by a CPD factorization:

$$[p([\mathbf{x}_d]_i|\mathcal{Z}^k)]_{n_1,\ldots,n_D} \approx \sum_{l=1}^{L} \rho_{1,l}^{(t_k)} \circ \cdots \circ \rho_{D,l}^{(t_k)} \tag{5}$$

where L is the number of components, which usually is a fixed user parameter and depends on the computational power of the fusion hardware and processing time constraints, "$\circ$" is the outer product and $\rho_{d,l}^{(t_k)}$ are the so-called *loading vectors* of dimension $N_d \times 1$ for each $d = 1, \ldots, D$ and $l = 1 \ldots, L$.

By means of an appropriate index function (see [19] for instance), an equivalent representation of a tensor can be achieved in its *vectorized* form:

$$p(\mathbf{x}_k|\mathcal{Z}^k) \approx \sum_{l=1}^{L} \bigotimes_{d=1}^{D} \rho_{d,l}^{(t_k)} \tag{6}$$

where "$\otimes$" is the Kronecker product. For the sake of notational simplicity, the latter form will be used throughout this paper.

It is assumed that the time evolution of the system is described by a continuous-time stochastic system given by

$$\mathrm{d}\mathbf{x} = \mathbf{f}(\mathbf{x}, t)\mathrm{d}t + \mathbf{G}(\mathbf{x}, t)\mathrm{d}\mathbf{w} \tag{7}$$

where $\mathbf{f}$ is the drift vector, $\mathbf{G}$ is the matrix of all diffusion coefficients, and $\mathrm{d}\mathbf{w}$ are the increments of a multivariate Brownian motion with covariance $\mathbf{Q}t$.

The measurement model is a general possibly nonlinear function $\mathbf{h}$ such that the observation at discrete instants of time t_k is given by

$$\mathbf{z}_k = \mathbf{h}(\mathbf{x}_k, t_k, \mathbf{v}_k) \tag{8}$$

where $\mathbf{v}_k$ is a random variable that represents the measurement noise of the sensor.

It is well known that the posterior pdf conditioned on all sensor data up to time t_k can be computed recursively by means of a prediction-filtering cycle. In recent publications, a tensor decomposition-based state estimation scheme has been proposed [17]–[19], which will be summarized in the remainder of this section.

A. Initialization

For the initialization, it is assumed that the initial pdf is given in a CPD form:

$$p(\mathbf{x}_0|\mathcal{Z}^0) = \sum_{l=1}^{L} \bigotimes_{d=1}^{D} \rho_{d,l}^{(t_0)}. \tag{9}$$

This can either be achieved by an analytical decomposition into sums of products of a given pdf evaluated at the discretization points or numerically by the ALS for instance.

B. Prediction

It is well known that the time evolution of the pdf is described by the FPE, for which the drift and diffusion parameters are given by the stochastic differential equation in (7):

$$\frac{\partial p}{\partial t} = -\sum_{i=1}^{D} \frac{\partial([\mathbf{f}]_i p)}{\partial x_i} + \frac{1}{2}\sum_{i,j=1}^{D} \frac{\partial^2([\mathbf{B}]_{i,j} p)}{\partial x_i \partial x_j} \tag{10}$$

where $\mathbf{B} = \mathbf{G}\mathbf{Q}\mathbf{G}^\top$ is the combined diffusion coefficient matrix. In the above equation, $[\mathbf{f}]_i$ denotes the ith entry of the drift vector and $[\mathbf{B}]_{i,j}$ is the entry in the ith row and jth column of the diffusion matrix. It is assumed that all components can be represented in a separable form such that

$$[\mathbf{f}]_i(\mathbf{x}) = \sum_{k=1}^{K_i} \prod_{d=1}^{D} f^i_{d,k}(x_d) \tag{11}$$

$$[\mathbf{B}]_{i,j}(\mathbf{x}) = \sum_{k=1}^{K_{i,j}} \prod_{d=1}^{D} B^{i,j}_{d,k}(x_d) \tag{12}$$

where K_i and $K_{i,j}$ are the number of components of the functions $[\mathbf{f}]_i$ and $[\mathbf{B}]_{i,j}$, respectively. By means of differentiation matrices and the FPE parameters in a separable form on the discretized grid we obtain an FPE operator $\mathbb{L}$ such that

$$\frac{\partial p}{\partial t} = \mathbb{L} p \tag{13}$$

where p is the pdf in a tensorized form. In the recent literature, two different approaches have been proposed to compute a numerical solution of the FPE for CPD tensors. The first solution is based on the ALS, where the FPE operator is augmented with a differentiation matrix $\mathbf{D}_t$ for the time dimension, which is accumulated with the pdf [17]:

$$(\mathbf{D}_t - \mathbb{L})p(\mathbf{x}, t) = 0. \tag{14}$$

Since the trivial solution $p = 0$ of (14) has to be avoided, constraints are added to the least square optimization such that the pdf is normalized and matches the previous pdf for the start time.

The second solution uses the tensor exponential of the FPE operator since it holds that

$$p(\mathbf{x}, t_k) = \exp\{\Delta_t \cdot \mathbb{L}\}\ p(\mathbf{x}, t_{k-1}) \tag{15}$$

where Δ_t is the time difference $t_k - t_{k-1}$. Here, the tensor exponential is approximated by means of a Taylor series [19].

C. Filtering

For the recursive prediction-filtering cycle, it is required to compute the posterior pdf, which incorporates the information of a given observation $\mathbf{z}_k$ at time t_k. This is achieved by means of the Bayes theorem

$$p(\mathbf{x}_k|\mathcal{Z}^k) = \frac{p(\mathbf{z}_k|\mathbf{x}_k)\cdot p(\mathbf{x}_k|\mathcal{Z}^{k-1})}{\int \mathrm{d}\mathbf{x}_k\ p(\mathbf{z}_k|\mathbf{x}_k)\cdot p(\mathbf{x}_k|\mathcal{Z}^{k-1})} \tag{16}$$

where $p(\mathbf{z}_k|\mathbf{x}_k)$ is the likelihood function. Since the likelihood evaluated on the discretized state space is a tensor, it is assumed that it is also given in a CPD form:

$$p(\mathbf{z}_k|\mathbf{x}_k) = \sum_{l'=1}^{L'} \bigotimes_{d=1}^{D} \lambda_{d,l'}. \tag{17}$$

As for the initial estimation pdf, this can be achieved by numerical methods if the likelihood function cannot be decomposed analytically. As a consequence, the posterior is given by

$$p(\mathbf{x}_k|\mathcal{Z}^k) = \frac{1}{c}\sum_{l=1}^{L}\sum_{l'=1}^{L'} \bigotimes_{d=1}^{D} \rho^{(k|k-1)}_{d,l} \odot \lambda_{d,l'} \tag{18}$$

$$=: \sum_{l=1}^{\tilde{L}} \bigotimes_{d=1}^{D} \rho^{(k|k)}_{d,l} \tag{19}$$

where "$\odot$" is the Hadamard (pointwise) product, $\tilde{L} = L \cdot L'$, and c is the normalization constant [17]. In order to keep the number of components L fixed, a mixture reduction technique such as tensor deflation has to be applied [21].

Note that the normalization constant of a CPD tensor p can easily be computed by the following summation[3]:

$$\int \mathrm{d}\mathbf{x}\ p(\mathbf{x}) = \sum_{l=1}^{L}\prod_{d=1}^{D} \Delta_{x_d} \sum_{i_d=1}^{N_d} [\rho_{d,l}]_{i_d} \tag{20}$$

where Δ_{x_d} is the discretization size of the state space in the dth dimension.

III. TRACKING IN CLUTTER

If the sensor produces multiple measurements at time t_k of a single target, the interpretation of the measurement origins is ambiguous. Measurements that come from unwanted objects or false detections are often referred to as *clutter*. If it is assumed that the target is detected with probability p_D, we have one additional interpretation assuming that the target is not detected at all. Therefore, let Z_k denote the set of m_k measurements $\mathbf{z}^1_k, \ldots, \mathbf{z}^{m_k}_k$ produced at time t_k.

[3] More details on implementation issues of the CPD tensor approach are given in the Appendix.

Let $j_k = 0$ denote the data interpretation hypothesis that the object has not been detected at all by the sensor at time t_k and so all measurements have to be considered as clutter. A hypothesis denoted by $j_k \in \{1, \ldots, m_k\}$ refers to the interpretation that the object has been detected, $\mathbf{z}_k^{j_k} \in Z_k$ being the corresponding measurement, and the remaining sensor data being clutter. Evidently, the set $\{0, \ldots, m_k\}$ describes mutually exclusive and exhaustive data interpretations. Due to the total probability theorem we have the following equation [24]:

$$p(Z_k, m_k|\mathbf{x}_k) = p_D\, p(Z_k, m_k|\mathbf{x}_k, O) + (1 - p_D)\, p(Z_k, m_k|\neg O) \quad (21)$$

where p_D is the probability of detection and O is the enumeration of a detection event. Often, it is assumed that the clutter measurements can be modeled statistically by a Poisson distribution in the number of measurements, which are uniformly distributed in the state space. Therefore, we have

$$\begin{aligned} p(Z_k, m_k|\neg O) &= p(Z_k|m_k, \neg O)p(m_k|\neg O) \\ &= p_F(m_k) \cdot |\mathrm{FoV}|^{-m_k} \end{aligned} \quad (22)$$

$$p_F(m_k) = \frac{\bar{m}^{m_k}}{m_k!} e^{-\bar{m}} \quad (23)$$

where $\bar{m}$ is the mean number of false measurements, p_F its Poisson distribution, and |FoV| is the size of the field of view. For the density conditioned on a detection, we can enumerate the data interpretation that the jth measurement is from the target:

$$p(Z_k|\mathbf{x}_k, O) = \sum_{j=1}^{m_k} p(Z_k|\mathbf{x}_k, O, j)p(j|O, \mathbf{x}_k). \quad (24)$$

Since hypothesis j implies that there are $m_k - 1$ false measurements, we have the following equation:

$$\begin{aligned} p(Z_k, m_k|\mathbf{x}_k, O, j) &= p(Z_k|m_k, \mathbf{x}_k, O, j)p(m_k|O) \\ &= p_F(m_k - 1)|\mathrm{FoV}|^{-(m_k-1)} \cdot p(\mathbf{z}_k^j|\mathbf{x}_k) \end{aligned} \quad (25)$$

where $p(\mathbf{z}_k^j|\mathbf{x}_k)$ is the density describing the sensor statistics for measuring a target with state $\mathbf{x}_k$. Furthermore, it is assumed that all hypotheses have the same prior probability, which yields

$$p(j|O, \mathbf{x}_k) = \frac{1}{m_k}. \quad (26)$$

Together we obtain

$$\begin{aligned} p(Z_k|\mathbf{x}_k) = p_D \frac{1}{m_k} p_F(m_k - 1)|\mathrm{FoV}|^{-(m_k-1)} \sum_{j=1}^{m_k} p(\mathbf{z}_k^j|\mathbf{x}_k) \\ + (1 - p_D)p_F(m_k) \cdot |\mathrm{FoV}|^{-m_k} \end{aligned} \quad (27)$$

$$= p_F(m_k)|\mathrm{FoV}|^{-m_k} \cdot \left(p_D \frac{|\mathrm{FoV}|}{\bar{m}} \sum_{j=1}^{m_k} p(\mathbf{z}_k^j|\mathbf{x}_k) + (1 - p_D) \right) \quad (28)$$

$$= \frac{p_F(m_k)|\mathrm{FoV}|^{-m_k}}{\rho_F} \cdot \left(p_D \sum_{j=1}^{m_k} p(\mathbf{z}_k^j|\mathbf{x}_k) + (1 - p_D)\rho_F \right) \quad (29)$$

where $\rho_F = \frac{\bar{m}}{|\mathrm{FoV}|}$ denotes the clutter density.

Since it is sufficient to model the likelihood function up to proportionality, we can neglect the constant factor and obtains

$$p(Z_k|\mathbf{x}_k) \propto (1 - p_D)\rho_F + p_D \sum_{j=1}^{m_k} p(\mathbf{z}_k^j|\mathbf{x}_k). \quad (30)$$

Again, it is assumed that a decomposed form of the sensor statistics $p(\mathbf{z}_k^j|\mathbf{x}_k)$ is available:

$$p(\mathbf{z}_k^j|\mathbf{x}_k) = \sum_{l'=1}^{L'} \bigotimes_{d=1}^{D} \lambda_{d,l'}^j. \quad (31)$$

This leads to a posterior pdf given by

$$\begin{aligned} p(\mathbf{x}_k|\mathcal{Z}^k) = \frac{1}{c}\Bigg((1 - p_D)\rho_F \sum_{l=1}^{L} \bigotimes_{d=1}^{D} \rho_{d,l}^{(k|k-1)} \\ + p_D \sum_{j=1}^{m_k} \sum_{l=1}^{L} \sum_{l'=1}^{L'} \bigotimes_{d=1}^{D} \rho_{d,l}^{(k|k-1)} \odot \lambda_{d,l'}^j \Bigg). \end{aligned} \quad (32)$$

Again, it is obvious that the posterior is already given in a decomposed form, however, the number of components has increased by a factor of $1 + L' \times m_k$ and therefore a deflation algorithm has to be applied for recursions with a constant number of loading vectors.

IV. MULTITARGET TRACKING

For a multitarget scenario, Bayesian filters can be derived by means of the theory of point set statistics. In the case that all targets are significantly separated, it is well possible to apply the single target filters from above on individual clusters. In all other cases, *probability hypothesis density* (PHD) [25] or intensity-based filters [26] have been proven particularly useful since association-free implementations are available that are easy to implement and of low computational complexity. Since there is no free lunch, this comes with a loss of target identities. However, methods have been proposed to overcome this problem.

The basic idea of the PHD or intensity filter is to model the conditional pdf of the set $\mathbf{x}_k^1, \ldots, \mathbf{x}_k^n$ of n target states as an inhomogeneous Poisson point process, that is, the number of targets is assumed to be Poisson distributed:

$$p(\mathbf{x}_k^1, \ldots, \mathbf{x}_k^n, n) = p(\mathbf{x}_k^1, \ldots, \mathbf{x}_k^n | n)\, p(n) \tag{33}$$

where

$$p(n) = \exp(-\mu)\frac{\mu^n}{n!} \tag{34}$$

$$\mu = \int \mathrm{d}\mathbf{x}\, f(x) \tag{35}$$

$$p(\mathbf{x}_k^1, \ldots, \mathbf{x}_k^n | n) = n! \cdot \prod_{j=1}^{n} p(\mathbf{x}_k^j). \tag{36}$$

The factor of $n!$ in (36) comes from the summation over all possible permutations of target identities, since the set is order free. The function f is the so-called intensity or PHD. Its integral μ is the mean number of targets and the spatial distribution of a target can be obtained by normalization of f:

$$p(\mathbf{x}_k^j) = \frac{f(\mathbf{x}_k^j)}{\int \mathrm{d}\mathbf{x}\, f(\mathbf{x})}. \tag{37}$$

Combining the above equations directly yields

$$p(\mathbf{x}_k^1, \ldots, \mathbf{x}_k^n, n) = \exp\left\{-\int \mathrm{d}\mathbf{x}\, f(x)\right\} \prod_{j=1}^{n} f(\mathbf{x}_k^j). \tag{38}$$

Therefore, f fully characterizes the point process.

In Bayesian set filters, it is sufficient to compute the prior and posterior intensity function f, if the Poisson assumption from above holds:

$$f_{k-1|k-1} \xrightarrow{\text{prediction}} f_{k|k-1} \tag{39}$$

$$f_{k|k-1} \xrightarrow{\text{filtering}} f_{k|k}. \tag{40}$$

In this section, it is shown that a recursive computation of the multitarget intensity function can be achieved by means of a decomposed tensor representation. For the initialization or previous filtering step, it is assumed that the intensity is given in a decomposed form:

$$f_{k-1|k-1} = \sum_{l=1}^{L} \bigotimes_{d=1}^{D} \rho_{d,l}^{(k-1|k-1)}. \tag{41}$$

Due to the relationship (37), the time evolution of the intensity function is described by the FPE:

$$\frac{\partial f}{\partial t} = -\sum_{i=1}^{D} \frac{\partial([\mathbf{f}]_i f)}{\partial x_i} + \frac{1}{2}\sum_{i,j=1}^{D} \frac{\partial^2([\mathbf{B}]_{i,j} f)}{\partial x_i \partial x_j}. \tag{42}$$

Another way to see this is the equivalence of the FPE to the Chapman–Kolmogorov equation [27], which is a more common way to compute the prior of a set theory-based filter [26]. As a consequence, it is possible to use one of the tensor propagators described in Section II-B.

Often, birth and death processes are modeled in addition to the system evolution to handle appearing or disappearing targets. Let $\boldsymbol{\nu}(\mathbf{x})$ be an intensity of a target birth point process given in a CPD representation

$$\boldsymbol{\nu}(\mathbf{x}) = \sum_{l_b=1}^{L_b} \bigotimes_{d=1}^{D} \nu_{d,l_b} \tag{43}$$

and $p_{\mathrm{TD}}(\mathbf{x})$ be the probability of target death such that the probability of survival is given by

$$p_{\mathrm{S}}(\mathbf{x}) = 1 - p_{\mathrm{TD}}(\mathbf{x}) = \sum_{l_d=1}^{L_d} \bigotimes_{d=1}^{D} p_{d,l_d}^{\mathrm{S}}. \tag{44}$$

If

$$f'_{k|k-1} = \sum_{l=1}^{L} \bigotimes_{d=1}^{D} \rho_{d,l}^{(k|k-1)\prime} \tag{45}$$

denotes the CPD tensor approximation of the intensity after the application of the time propagator, then the prior intensity is obtained by the superposition of the birth and death processes [26]. The result again is a CPD tensor, where deflation algorithms have to be applied to keep the number of loading vectors constant:

$$f_{k|k-1} = \sum_{l_b=1}^{L_b} \bigotimes_{d=1}^{D} \nu_{d,l_b} + \sum_{l_d=1}^{L_d} \sum_{l=1}^{L} \bigotimes_{d=1}^{D} p_{d,l_d}^{\mathrm{S}} \odot \rho_{d,l}^{(k|k-1)\prime}. \tag{46}$$

In many practical applications, it is sufficient to model the birth process as a constant birth rate $\boldsymbol{\nu}(\mathbf{x}) = \mathbf{b}$ and analogously set a constant probability of target death $p_{\mathrm{TD}}(\mathbf{x}) = \mathbf{d}$. In this case, the prior reduces to

$$f_{k|k-1} = \frac{\mathbf{b}}{\prod_{d=1}^{D} \Delta_{x_d} N_d} \bigotimes_{d=1}^{D} 1_d + (1-\mathbf{d}) \sum_{l=1}^{L} \bigotimes_{d=1}^{D} \rho_{d,l}^{(k|k-1)\prime} \tag{47}$$

where 1_d is the vector consisting of N_d ones.

A. Multitarget Filtering

Let $Z_k = \mathbf{z}_k^1, \ldots, \mathbf{z}_k^{m_k}$ be the set of observations produced at time t_k. As shown by Mahler [25], it is possible to approximate the posterior pdf of a multitarget Poisson point process by the following intensity function:

$$f_{k|k} = \left((1 - p_D) + \sum_{j=1}^{m_k} \frac{p(\mathbf{z}_k^j | \mathbf{x})\, p_D}{\lambda(\mathbf{z}_k^j)} \right) f_{k|k-1} \tag{48}$$

$$\lambda(\mathbf{z}_k^j) = \lambda_c(\mathbf{z}_k^j) + \int \mathrm{d}\mathbf{x}\, p(\mathbf{z}_k^j | \mathbf{x})\, p_D\, f_{k|k-1}(\mathbf{x}) \tag{49}$$

where $\lambda_c(\mathbf{z})$ is the Poisson intensity of the clutter point process. It is assumed that the prior intensity is given in a CPD form where $\rho_{d,l}^{(k|k-1)}$ are its loading vectors for $l = 1, \ldots, L$ and $d = 1, \ldots, D$. Using the sensor model from (31), we obtain the following update equation in a tensorized form:

$$f_{k|k} = \left((1 - p_D) \sum_{l=1}^{L} \bigotimes_{d=1}^{D} \rho_{d,l}^{(k|k-1)} \right.$$
$$\left. + \sum_{j=1}^{m_k} \frac{p_D}{\lambda(\mathbf{z}_k^j)} \sum_{l=1}^{L} \sum_{l'=1}^{L'} \bigotimes_{d=1}^{D} \rho_{d,l}^{(k|k-1)} \odot \lambda_{d,l'}^{j} \right) \tag{50}$$

where the expected measurement likelihood $\lambda(\mathbf{z}_k^j)$ is given by

$$\lambda(\mathbf{z}_k^j) = \lambda_c(\mathbf{z}_k^j) + \sum_{l=1}^{L} \sum_{l'=1}^{L'} p_D \prod_{d=1}^{D} \Delta_{x_d} \sum_{i=1}^{N_d} [\rho_{d,l}^{(k|k-1)}]_i \cdot [\lambda_{d,l'}^{j}]_i. \tag{51}$$

Here, the PHD filter update equations were used. The extension to higher order statistics filter for the number of targets (CPHD) [28], the iFilter [29], or (generalized) multi-Bernoulli filters [30] is straightforward.

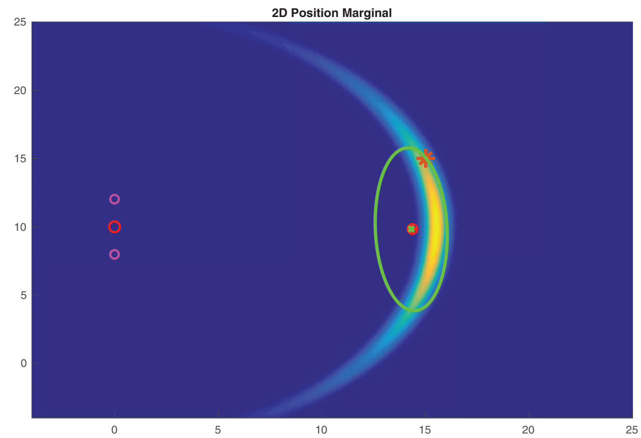

Fig. 2. Exemplary initial pdf based on two bistatic range measurements of a target (brown star) where the transmitter is located at the center (red circle) of two receivers (purple circles). It can be seen that the Gaussian approximation (green) does not reflect the measurement uncertainty in an appropriate way, in contrast to that the 2D marginal of the tensor representation (blue/yellow color range) gives a precise approximation to the true Bayes posterior.

V. NUMERICAL EVALUATION

In this section, we demonstrate the performance of the described algorithms by means of numerical simulations. It is divided into three parts describing the simulation setup and results of a nonlinear scenario (a), a scenario with ambiguous measurements (b), and a multitarget scenario (c), respectively. In all scenarios, a four-dimensional state space was used such that

$$\mathbf{x} = (x, y, \dot{x}, \dot{y})^\top. \tag{52}$$

The simulated target(s) move according to a discretized almost constant velocity model where the transition and process noise model is given by

$$\mathbf{x}_{k+1} = \mathbf{F}_{k|k-1}\mathbf{x}_k + \mathbf{w}_{k|k-1} \tag{53}$$

$$\mathbf{F}_{k|k-1} = \begin{pmatrix} 1 & T \\ 0 & 1 \end{pmatrix} \otimes \mathbf{I}_2 \tag{54}$$

$$\mathbf{w}_{k|k-1} \sim \mathcal{N}(\mathbf{O}, \mathbf{Q}_{k|k-1}) \tag{55}$$

$$\mathbf{Q}_{k|k-1} = q \cdot \begin{pmatrix} \mathbf{O} & \mathbf{O} \\ \mathbf{O} & T\mathbf{I}_2 \end{pmatrix} \tag{56}$$

where the power spectral density was set to $q = 0.1$.

A. Nonlinear Example

In the nonlinear scenario, a single target is observed by means of a bistatic radar with one transmitter antenna (Tx) and two receiving antennas (Rx_1 and Rx_2). Once a second, the bistatic ranges $|Tx - \mathbf{x}| + |Rx_i - \mathbf{x}|$ for $i = 1, 2$ are measured with additive Gaussian noise with a standard deviation of $\sigma_{br} = 1.0m$. The receivers were located at $(0, 8)^\top$, and $(0, 12)^\top$, respectively, and the transmitter was set in between at $(0, 10)^\top$.

The same measurements also were processed by an EKF and a bootstrap PF using importance resampling. The initial state and covariance for these filters was, respectively, estimated by the first and second moment of the likelihood function of the first measurement. The initial velocity was set to zero with a standard deviation of $2\frac{m}{s}$. The initial pdf of the position is shown exemplary in Fig. 2.

The results of the numerical evaluation are shown in Fig. 3. We have plotted the root mean squared error (RMSE) of 50 Monte Carlo simulations. Clearly, the tensor approach reaches the performance of the PF, which is close to the Cramer–Rao lower bound.

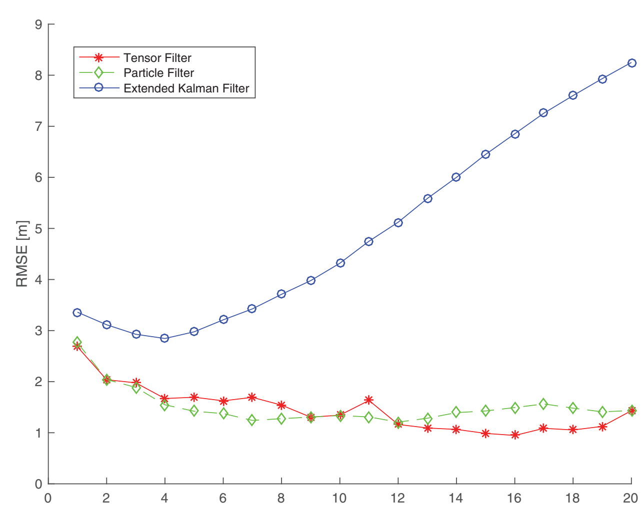

Fig. 3. Root mean squared error (RMSE) of 50 Monte Carlo simulations. It can be seen that the tensor decomposition-based approach has equal or better estimation performance than the extended Kalman filter or the particle filter.

B. Multitarget Example

The multitarget scenario demonstrates the ability of the tensor decomposition-based PHD filter as described in Section IV to estimate the number of targets and their respective states. Since the PHD intensity is a label-free statistic, an evaluation of the estimation errors would require some state extraction algorithm, which is out of the scope of this paper. As a consequence, the x–y marginals of the intensity function of an exemplary simulation are presented.

In the multitarget scenario, two targets are initialized with states $\mathbf{x}_0^1 = (0, 0, 1, 1)^\top$ and $(10, 10, -1, -1)^\top$, respectively.

Fig. 4 shows the intensity function in x–y coordinates after the initialization and after 50 time steps of 0.1 s.

C. Filtering in High Dimensions

The third example scenario can be considered a toy problem. It is specified to demonstrate the power of the tensor decomposition approach in high-dimensional filtering problems. These high dimensions can easily appear in practical scenarios such as SLAM-based [31] navigation data fusion.

The scenario demonstrated here is a one-step filtering of a given prior density using a linear measurement. For various dimensions D, the prior is given by a mean

$$\mathbf{x}_{0|0} = (50, \ldots, 50)^\top \tag{57}$$

and a covariance matrix $\mathbf{P}_{0|0} = [\mathbf{P}_{0|0}]_{i,j}$ for $i, j = 1, \ldots, D$, where

$$[\mathbf{P}_{0|0}]_{i,j} = 100, \text{ if } i == j \tag{58}$$

$$[\mathbf{P}_{0|0}]_{i,j} = 50, \quad \text{else.} \tag{59}$$

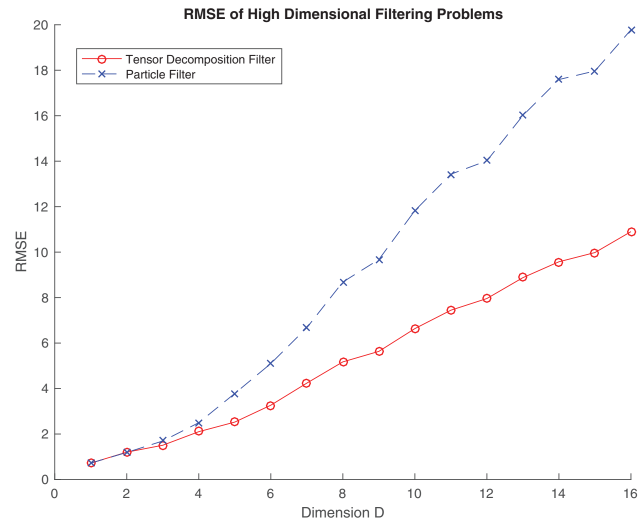

Fig. 5. Root mean squared error (RMSE) for 100 Monte Carlo simulations on a single measurement update for various dimensions.

The true state of the system is drawn randomly according to the prior distribution. The measurement vector is the true state corrupted by additive Gaussian noise, where the covariance matrix is given by the unity matrix $\mathbf{I}_D$. For the tensor approach, $L = 2{,}000$ components were used and for the particle filter $N_{\text{pf}} = 10{,}000$ particles were used.[4] For each dimension 100 Monte Carlo simulations were used.

It can be seen from Fig. 5 that the particle filter degrades due to the curse of dimensionality. The tensor decomposition approach clearly outperforms the particle filter in higher dimensions where it is obvious that the difference increases when the D grows. In Fig. 6, the mean processing times for both filters are summarized.

[4]The number of particles and components, respectively, was chosen such that the processing time is of equal magnitude.

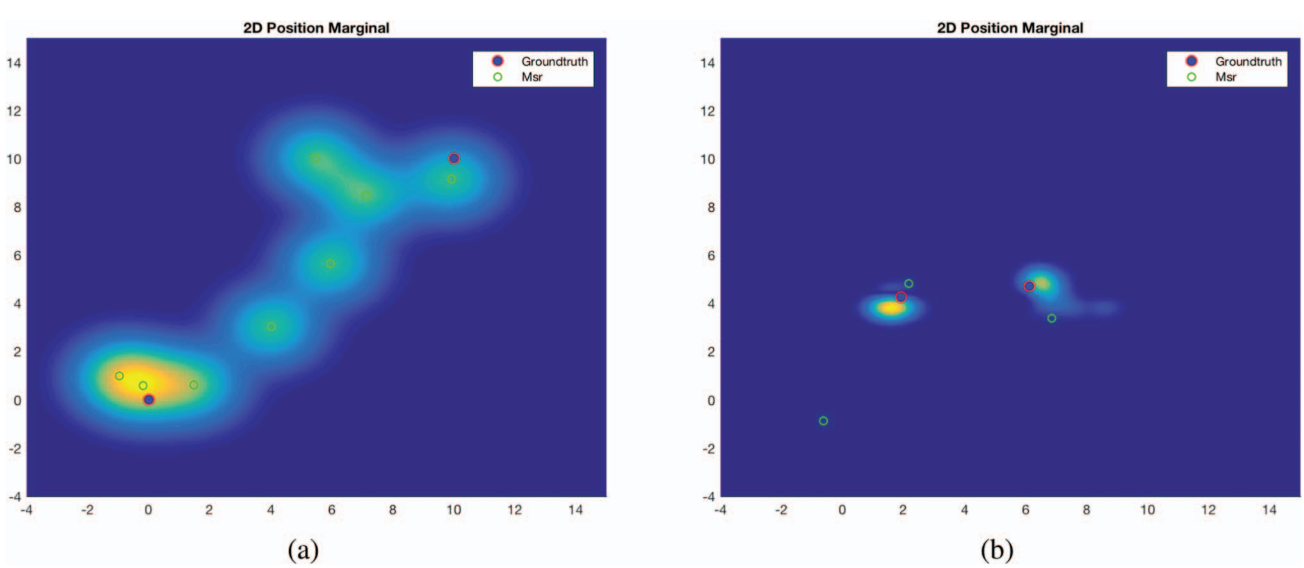

Fig. 4. Multitarget intensity function after the initialization (a) and after 50 time steps (b).

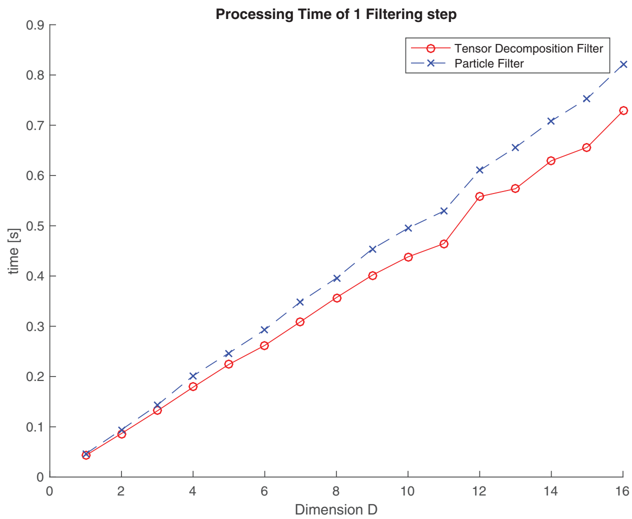

Fig. 6. Mean processing times for the tensor decomposition filter and the particle filter.

VII CONCLUSION

In this paper, we have applied the CPD-based approach of nonlinear Bayesian filtering on practical multitarget tracking problems with false measurements. By means of the *probabilistic data association* (PDA) likelihood, the update formula for single targets in cluttered environments for CPD tensors was derived. Then, it was shown that the CPD representation can also be used to apply set theory-based filters such as the PHD filter. In a numerical evaluation the performance was compared to the existing methods for nonlinear filtering. In a toy example it has been shown that the CPD is a powerful representation of a density function in particular if the problem is high dimensional.

APPENDIX

In the Appendix, we would like to provide some details and assistance for the interested reader who wants to implement some of the algorithms above. Since most engineers in data fusion and tracking work with Gaussian mixtures, particles, and optimization algorithms, the tensor decomposition-based data fusion can be considered quite new and unknown. This section should help to start from the scratch to implement the CPD tensors and the corresponding filters.

Implementation

A CPD tensor, which is given by

$$p(\mathbf{x}) = \sum_{l=1}^{L} \bigotimes_{d=1}^{D} \rho_{d,l}$$

is fully described by L *loading vectors* for each dimension $d = 1, \ldots, D$. It is useful to store these vectors as matrices, called *loading matrices*, $\mathbf{U}_1, \ldots, \mathbf{U}_D$. For a given dimension d, each loading vector $\rho_{d,l}$ has by definition N_{x_d} entries, therefore, the loading matrix $\mathbf{U}_d$ is of the size $N_{x_d} \times L$. In the following explanations, we will use the following short notation for the above CPD tensor:

$$p(\mathbf{x}) = [\mathbf{U}_1, \ldots, \mathbf{U}_D] = [\mathbf{U}_d]_{d=1}^{D}.$$

The multiplication with an exemplary decomposed likelihood function

$$\ell(\mathbf{z}, \mathbf{x}) = \bigotimes_{d=1}^{D} \lambda_d$$

then reduces to simple matrix multiplication:

$$p(\mathbf{x}) \cdot \ell(\mathbf{z}, \mathbf{x}) = [\mathrm{diag}\,[\lambda_d]\,\mathbf{U}_d]_{d=1}^{D}.$$

If the likelihood function has multiple components as in (17), the same operation is computed for each of the components and the resulting matrices are appended horizontally. Also it should be noted that likelihood functions of lower dimensions can easily be incorporated by setting $\lambda_d = \mathbb{1}_d$ for $d > d'$, where d' is the dimension of the likelihood.

Integration

The integral of a CPD tensor

$$\int \mathrm{d}\mathbf{x} \sum_{l=1}^{L} \bigotimes_{d=1}^{D} \rho_{d,l}$$

can easily be computed by means of cheap computational operations. This can be seen by the fact that

$$\int \mathrm{d}\mathbf{x} \sum_{l=1}^{L} \bigotimes_{d=1}^{D} \rho_{d,l} = \sum_{l=1}^{L} \int \mathrm{d}\mathbf{x} \bigotimes_{d=1}^{D} \rho_{d,l}$$
$$= \sum_{l=1}^{L} \sum_{i_1,\ldots,i_D} \prod_{d=1}^{D} [\rho_{d,l}]_{i_d} \cdot \Delta_{x_d}$$
$$= \sum_{l=1}^{L} \prod_{d=1}^{D} \Delta_{x_d} \sum_{i_d=1}^{N_d} [\rho_{d,l}]_{i_d}.$$

Computing the Mean Vector

Again, it is assumed that the pdf is given in a CPD-tensorized form:

$$p(\mathbf{x}) = \sum_{l=1}^{L} \bigotimes_{d=1}^{D} \rho_{d,l}.$$

The mean $\mathrm{E}\,[\mathbf{x}] = \hat{\mathbf{x}} = [\hat{\mathbf{x}}_d]_d$ is given by

$$\hat{\mathbf{x}}_d = \int \mathrm{d}\mathbf{x}\, \mathbf{x}_d \sum_{l=1}^{L} \bigotimes_{d=1}^{D} \rho_{d,l}.$$

Using the integration rule from above, we obtain

$$\hat{\mathbf{x}}_d = \sum_{l=1}^{L} \prod_{j \neq d} \left\{ \Delta_{x_j} \sum_{i_j=1}^{N_j} [\rho_{j,l}]_{i_j} \right\} \Delta_{x_d} \sum_{i_d=1}^{N_d} [\mathrm{diag}\,[[\mathbf{x}_d]]\, \rho_{d,l}]_{i_d}.$$

Computing the Covariance Matrix

Analogously, the covariance matrix $\operatorname{cov}[\mathbf{x}] = \mathbf{P} = [\mathbf{P}_{ij}]_{i,j}$ can be computed by $\frac{1}{2}D(D+1)$ integration operations. For given i and j, the covariance is given by

$$\mathbf{P}_{ij} = \int \mathrm{d}\mathbf{x}\, (\mathbf{x}_i - \hat{\mathbf{x}}_i)(\mathbf{x}_j - \hat{\mathbf{x}}_j) \sum_{l=1}^{L} \bigotimes_{d=1}^{D} \rho_{d,l}.$$

For $i \neq j$, we have

$$\mathbf{P}_{ij} = \sum_{l=1}^{L} \prod_{k \neq i \wedge k \neq j} \left\{ \Delta_{x_k} \sum_{i_k=1}^{N_k} [\rho_{k,l}]_{i_k} \right\} \cdot \Delta_{x_i} \sum_{i_i=1}^{N_i} [\operatorname{diag}[[\mathbf{T}_{\hat{\mathbf{x}}_i} \mathbf{x}_i]]\, \rho_{i,l}]_{i_i} \cdot \Delta_{x_j} \sum_{i_j=1}^{N_j} [\operatorname{diag}[[\mathbf{T}_{\hat{\mathbf{x}}_j} \mathbf{x}_j]]\, \rho_{j,l}]_{i_j}$$

where $\mathbf{T}_{\hat{\mathbf{x}}_j}$ is the affine translation such that $[(\mathbf{x}_j - \hat{\mathbf{x}}_j)]_i = \mathbf{T}_{\hat{\mathbf{x}}_j} \mathbf{x}_j$. The diagonal elements similarly are given by

$$\mathbf{P}_{ii} = \sum_{l=1}^{L} \prod_{k \neq i} \left\{ \Delta_{x_k} \sum_{i_k=1}^{N_k} [\rho_{k,l}]_{i_k} \right\} \cdot \Delta_{x_i} \sum_{i_i=1}^{N_i} [\operatorname{diag}[[([\mathbf{x}_i]_l - \hat{\mathbf{x}}_i)^2]_l]\, \rho_{i,l}]_{i_i}.$$

REFERENCES

[1] Y. Bar-Shalom, X. Li, and T. Kirubarajan
Estimation with Applications to Tracking and Navigation. Wiley-Interscience, New York, NY, 2001.

[2] W. Koch
Tracking and Sensor Data Fusion: Methodological Framework and Selected Applications. Springer, Heidelberg, Germany, 2014.

[3] R. E. Kalman
"A New Approach to Linear Filtering and Prediction Problems,"
Transactions of the ASME—Journal of Basic Engineering, vol. 82, no. Series D, pp. 35–45, 1960.

[4] S. J. Julier and J. K. Uhlmann
"New Extension of the Kalman Filter to Nonlinear Systems,"
in *Signal Processing, Sensor Fusion, and Target Recognition VI*, I. Kadar
Ed., vol. 3068, Jul. 1997, pp. 182–193.

[5] B. Ristic, S. Arulampalam, and N. Gordon
Beyond the Kalman Filter: Particle Filters for Tracking Applications. Artech House, Boston, MA, 2004.

[6] N. Gordon, D. Salmond, and A. Smith
"Novel Approach to Nonlinear/Non-Gaussian Bayesian State Estimation,"
Radar and Signal Processing, IEE Proceedings F, vol. 140, no. 2, pp. 107–113, Apr. 1993.

[7] F. Daum and J. Huang
"Particle Flow for Nonlinear Filters, Bayesian Decisions and Transport,"
in *2013 16th International Conference on Information Fusion (Fusion)*, Jul. 2013, pp. 1072–1079.

[8] M. Khan and M. Ulmke
"Non-Linear and Non-Gaussian State Estimation using log-Homotopy Based Particle Flow Filters,"
in *Sensor Data Fusion: Trends, Solutions, Applications (SDF), 2014*, Oct. 2014, pp. 1–6.

[9] U. Hanebeck
"Pgf 42: Progressive Gaussian Filtering with a Twist,"
in *2013 16th International Conference on Information Fusion (Fusion)* Jul. 2013, pp. 1103–1110.

[10] T. Yang, P. Mehta, and S. Meyn
"Feedback Particle Filter,"
IEEE Transactions on Automatic Control, vol. 58, no. 10, pp. 2465–2480, Oct. 2013.

[11] G. Evensen
"Sequential Data Assimilation with a Nonlinear Quasi-Geostrophic Model using Monte Carlo Methods to Forecast Error Statistics,"
Journal of Geophysical Research, vol. 99, pp. 10143–10162, 1994.

[12] S. Reich
"A Gaussian-Mixture Ensemble Transform Filter,"
Quarterly Journal of the Royal Meteorological Society, vol. 138, pp. 222–233, Jan. 2012.

[13] M. Glodek, M. Schels, and F. Schwenker
"Ensemble Gaussian Mixture Models for Probability Density Estimation (report),"
Computational Statistics, vol. 28, no. 1, p. 127(12), Feb. 2013.

[14] S. Challa and Y. Bar-Shalom
"Nonlinear Filter Design using Fokker Planck Kolmogorov Probability Density Evolutions,"
IEEE Transactions on AeroSpace and Electronic Systems, vol. 36, no. 1, pp. 309–315, Jan. 2000.

[15] G. Beylkin and M. J. Mohlenkamp
"Algorithms for Numerical Analysis in High Dimensions,"
SIAM Journal on Scientific Computing, vol. 26, pp. 2133–2159, 2005.

[16] B. N. Khoromskij
"Tensors-Structured Numerical Methods in Scientific Computing: Survey on Recent Advances,"
Chemometrics and Intelligent Laboratory Systems, vol. 110, no. 1, pp. 1–19, 2012.

[17] Y. Sun and M. Kumar
"Nonlinear Bayesian Filtering Based on Fokker Planck Equation and Tensor Decomposition,"
in *2015 18th International Conference on Information Fusion (Fusion)*, July 2015, pp. 1483–1488.

[18] M. A. Khan, M. Ulmke, B. Demissie, F. Govaers, and W. Koch
"Combining Log-Homotopy Flow with Tensor Decomposition Based Solution for Fokker Planck Equation,"
in *Proceedings of the 19th International Conference on Information Fusion*, Istanbul, Turkey, July 2016.

[19] B. Demissie, F. Govaers, and M. A. Khan
"Nonlinear Filter Design using Fokker-Planck Propagator in Kronecker Tensor Format,"
in *Proceedings of the 19th International Conference on Information Fusion*, 2016.

[20] F. L. Hitchcock
"The Expression of a Tensor or a Polyadic as a Sum of Products,"
Journal of Mathematics and Physics, vol. 6, no. 1–4, pp. 164–189, 1927 [Online]. Available: http://dx.doi.org/10.1002/sapm192761164.

[21] A.-H. Phan, P. Tichavsky, and A. Cichocki
"Tensor Deflation for CANDECOMP/PARAFAC. Part 1: Alternating SubSpace Update Algorithm,"
IEEE Transactions on Signal Processing, vol. 63, no. 22, pp. 5924–5938, Nov. 2015.

[22] C. J. Hillar and L.-H. Lim
"Most Tensor Problems are NP-Hard,"
Journal of the ACM, vol. 60, no. 6, pp. 45:1–45:39, Nov. 2013 [Online]. Available: http://doi.acm.org/10.1145/2512329.

[23] R. A. Harshman
"Foundations of the PARAFAC Procedure: Models and Conditions for an 'Explanatory' Multimodal Factor Analysis," UCLA Working Papers in Phonetics, 1970.

[24] W. Koch
in *Signal Processing Handbook*,
S. Stergiopoulos
Ed. CRC Press, Boca Raton, FL, 2000.

[25] R. Mahler
"Multitarget Bayes Filtering via First-Order Multitarget Moments,"
IEEE Transactions on Aeropace and Electronic Systems, vol. 39, no. 4, pp. 1152–1178, Nov. 2003.

[26] R. Streit
Poisson Point Processes: Imaging, Tracking, and Sensing. Springer, 2010.

[27] C. Gardiner
Stochastic Methods. Springer, 2009.

[28] R. P. S. Mahler
"Multitarget Bayes Filtering via First-Order Multitarget Moments,"
IEEE Transactions on Aeropace and Electronic Systems, vol. 39, no. 4, pp. 1152–1178, Oct. 2003.

[29] R. L. Streit and L. D. Stone
"Bayes Derivation of Multitarget Intensity Filters,"
in *2008 11th International Conference on Information Fusion*, June 2008, pp. 1–8.

[30] M. Beard, S. Reuter, K. Granstöm, B. T. Vo, B. N. Vo, and A. Scheel
"A Generalised Labelled Multi-Bernoulli Filter for Extended Multi-Target Tracking,"
in *2015 18th International Conference on Information Fusion (Fusion)*, July 2015, pp. 991–998.

[31] D. Törnqvist, T. B. Schön, R. Karlsson, and F. Gustafsson
"Particle Filter SLAM with High Dimensional Vehicle Model,"
Journal of Intelligent and Robotic Systems, vol. 55, no. 4, pp. 249–266, 2009 [Online]. Available: http://dx.doi.org/10.1007/s10846-008-9301-y.

Felix Govaers received the diploma in mathematics in 2007 and the Ph.D. degree in computer science in 2012, both from the University of Bonn, Bonn, Germany. Since 2009, he has been with the Department of Sensor Data Fusion and Information Fusion, Fraunhofer FKIE, where he led the research group distributed systems from 2014 to 2017. He is currently the Deputy Head with the Department of Sensor Data Fusion and Information Fusion, Fraunhofer FKIE. His research interests include data fusion for state estimation in nonlinear scenarios and in sensor networks, which includes track extraction, processing of delayed measurements, as well as the distributed Kalman filter and track-to-track fusion, and advances in state estimation, such as particle flow and homotopy filters and the random finite set theory approaches. His current research projects are tensor decomposition-based approaches to multitarget tracking. Since 2008, he has been an active member of the International Society of Information Fusion (ISIF) community. He organized the ISIF-cosponsored Sensor Data Fusion Workshop in Germany for many years as the Technical Program Chair. Since 2017, he has been an active member of the ISIF Board of Directors. At FUSION 2016, he was the Publication Chair of the organization team. Since 2014, he was an Associate Editor for the *IEEE Transactions on Aerospace and Electronic Systems*.

Bruno Demissie received the Diploma Degree in theoretical physics from the Ruhr-University-Bochum, Bochum, Germany and the Dr. rer. nat. degree from Carl-von-Ossietzky University, Oldenburg, Germany, in 1994 and 1997, respectively. He was a Postdoctoral Researcher with the Phillips-University, Marburg, Germany. From February 1998 until March 2000, he was a MINERVA fellow with the Weizmann Institute of Science, Rehovot, Israel. Since 2000, he has been with the Fraunhofer Institute for Communications, Information Processing and Ergonomics. He has been a member of the editorial board of the *International Journal of Electronics and Communications* (AEUE). Since 2015, he has been an Associate Editor and subsequently a Senior Area Editor for *IEEE Signal Processing Letters*.

Altamash Khan received the B.E. in electrical engineering from the National University of Sciences and Technology (NUST), Islamabad, Pakistan, in 2006 and the M.Sc. degree in wireless systems from KTH Royal Institute of Technology, Stockholm, Sweden in 2011. He also received the Ph.D. degree in computer science from the University of Bonn, Bonn, Germany in 2018. His doctoral research focused on the development of log-homotopy-based particle filters for providing efficient statistical estimation for complex systems. From April 2014 to March 2017, he was a Marie Curie Fellow at FKIE Fraunhofer, Wachtberg, Germany. Since July 2017, he has been with Zenuity, Gothenburg, Sweden, as an Algorithm Development Engineer for advanced driving assistance systems in the sensor fusion group. His research interests include nonlinear and non-Gaussian state estimation methods, multitarget tracking, signal processing for automotive applications, and wireless sensors networks.

Martin Ulmke received the Dipl.-Phy. (M.S.) and Dr. rer. nat. (Ph.D.) degrees from the Aachen Technical University (RWTH), Aachen, Germany, in 1991 and 1995, respectively. From 1995 to 1998, he was a Research Associate in condensed matter theory at the University of California in Davis, CA, USA and at the University of Augsburg, Augsburg, Germany. From 1998 to 2001, he was a System Engineer with MTU Aero Engines, Munich, Germany. Since 2001, he has been a part of the scientific staff with the Department of Sensor Data and Information Fusion, Fraunhofer FKIE, where he was the Head of Research Group "Distributed Sensor Systems".

Wolfgang Koch received the Dr. rer. nat. degree in theoretical physics from the Aachen Technical University, Aachen, Germany. For many years, he was the Head of the Department of Sensor Data and Information Fusion, Fraunhofer FKIE, an institute of the Fraunhofer Society, the largest institution for applied research in Europe. He has authored or coauthored a well-referenced textbook on target tracking and sensor data fusion, 16 handbook chapters, and above 200 journal and conference articles on various topics within the area of sensor data and information fusion. He was one of the coeditors of the two-volume handbook *Novel Radar Techniques and Applications*. He was a Fellow of the IEEE, where he was with the *IEEE Transactions on Aerospace and Electronic Systems* and was a member of the board of governors of the Aerospace and Electronics Systems Society. In 2015, he was the IEEE Distinguished Lecturer and Chairman of the German IEEE AESS chapter. In 2013, he was the President of the ISIF. His research interests include the NATO Science and Technology Organization. He received the Habilitation Degree in applied computer science from Bonn University, Bonn, Germany, where he gave lecture series on sensor data and information fusion. In 2016, he was the General Co-chair of the IEEE ISF International Conference on Information Fusion, Heidelberg, Germany.