Distributed Point-Mass Filter with Reduced Data Transfer Using Copula Theory

Jakub Matousek, Jindrich Dunik and Robin Forsling

Matousek, J., Dunik, J., Forsling, R., (2023), Distributed Point-Mass Filter with Reduced Data Transfer Using Copula Theory, *Proceedings of 2023 American Control Conference (ACC)*, , 1649-1654.
https://doi.org/10.23919/ACC55779.2023.10155942

Distributed Point-Mass Filter with Reduced Data Transfer Using Copula Theory

Jakub Matoušek, Jindřich Duník and Robin Forsling

***Abstract*— This paper deals with distributed Bayesian state estimation of generally nonlinear stochastic dynamic systems. In particular, distributed point-mass filter algorithm is developed. It is comprised of a basic part that is accurate but data intense and optional step employing advanced copula theory. The optional step significantly reduces data transfer for the price of a small accuracy decrease. In the end, the developed algorithm is numerically compared to the usually employed distributed extended Kalman filter.**

Keywords: Distributed estimation, point-mass filter, covariance intersection, data reduction.

I. Introduction

Distributed estimation of a dynamic state in a linear Gaussian setting is a well-studied topic. Common applications include target tracking where track estimates are fused into an improved estimate. In many cases, the Gaussian assumption does not hold and arbitrary *probability density functions* (PDFs) need to be accounted for. In such problems *Bayesian recursion relations* (BRRs, [1]) is an important methodology for filtering the dynamic state. BRRs can be realized using the *point-mass filter* (PMF)[1], see, e.g., [2]–[5]. An issue with implementing the BRRs scheme for distributed state estimation is the large amount of data that require to be transfered. This is particularly true in case of the PMF. For practical reasons it is therefore vital to investigate how to reduce data transfer for such distributed PMF. Another complicating factor in distributed estimation is cross-correlation between estimates, which is unknown in many cases [6].

Consider two Gaussian distributed estimates $p_1(\mathbf{x}) = \mathcal{N}(\hat{\mathbf{x}}_1, \mathbf{P}_1)$ and $p_2(\mathbf{x}) = \mathcal{N}(\hat{\mathbf{x}}_2, \mathbf{P}_2)$, where $\mathcal{N}(\hat{\mathbf{x}}, \mathbf{P})$ is a Gaussian distribution with mean $\hat{\mathbf{x}}$, and covariance $\mathbf{P}$. The $\mathbf{P}_1$ and $\mathbf{P}_2$ resp. their σ-ellipses are depicted in Fig. 1(a) for 2D case. To fuse these estimates exactly, the cross-correlations need to be known. If unknown, then *covariance intersection* (CI, [7]) is a popular linear fusion method. In CI the covariance matrix of the fused estimate is computed as $\mathbf{P}_f = \left(\omega\mathbf{P}_1^{-1} + (1-\omega)\mathbf{P}_2^{-1}\right)^{-1}$, where $\omega \in [0,1]$. One such $\mathbf{P}_f$ is plotted in Fig. 1(a).

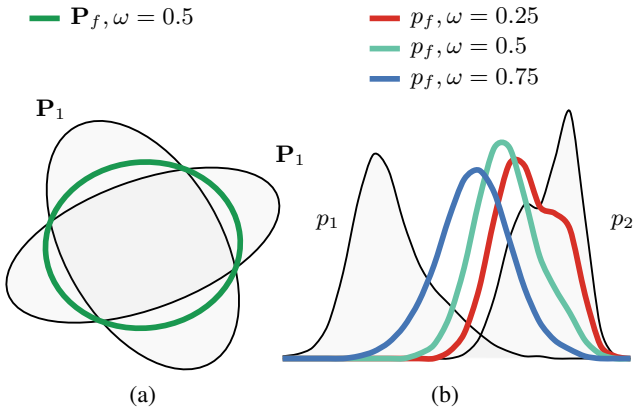

Fig. 1. Distributed estimate fusion. In (a) two Gaussian estimates are fused using CI despite cross-correlations being unknown. CI cannot be used in (b) where two non-Gaussian PDFs are fused.

For non-Gaussian PDFs, the original version of the CI is not applicable. A one-dimensional example of this case is illustrated in Fig. 1(b), where two PDFs $p_1(x)$ and $p_2(x)$ are to be fused. To handle unknown conditional probabilities between $p_1(x)$ and $p_2(x)$, one possibility is to use a generalization of the CI where the fused PDF is given by $p_f(x) \propto p_1(x)^{\omega} \cdot p_2(x)^{1-w}$ with $\omega \in [0,1]$ [8]. Several p_f corresponding to different values of ω are illustrated in Fig. 1(b). As representation of a general PDF requires many parameters, which is certainly true for the PDFs in the form of the point-mass density (PMD) used by the PMF, distributed fusion of general PDFs requires some kind of data transfer reduction technique. This problem has earlier been studied in the linear Gaussian case [9], [10]. However, in the general fusion context this problem remains untreated.

In this paper a *distributed PMF* is proposed. Data transfer reduction is enabled based on copula theory with a marginal degradation of the performance [11]. The usability of the proposed method is illustrated using simulations.

II. System Description, State Estimation

Following the discrete-time state-space model of a nonlinear stochastic dynamic system with additive noises

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

$$\mathbf{z}_k = \mathbf{h}_k(\mathbf{x}_k) + \mathbf{v}_k, \tag{1b}$$

J. Matoušek, and J. Duník are with Department of Cybernetics, Faculty of Applied Sciences, University of West Bohemia, Pilsen, Czech Republic. Robin Forsling is with Dept. of El. Eng., Linköping University, Sweden. E-mails: {dunikj,matoujak}@kky.zcu.cz (J. Duník, J. Matoušek), robin.forsling@liu.se (R. Forsling). This work has been supported by the project Improving the Quality of Internal Grant Schemes at the UWB, project registration number: CZ.02.2.69/0.0/0.0/19 073/0016931, and Student grant competition SGS-2022-022, and by the Industry Excellence Center LINK-SIC funded by VINNOVA and Saab AB

[1]Sometimes called histogram, or grid based filter.

is considered, where the vectors $\mathbf{x}_k \in \mathbb{R}^{n_x}$, $\mathbf{u}_k \in \mathbb{R}^{n_u}$, and $\mathbf{z}_k \in \mathbb{R}^{n_z}$ represent the *unknown* state of the system and the *known* input and measurement at time instant k, respectively. The state and measurement functions $\mathbf{f}_k : \mathbb{R}^{n_x \times n_u} \to \mathbb{R}^{n_x}$ and $\mathbf{h}_k : \mathbb{R}^{n_x} \to \mathbb{R}^{n_z}$ are supposed to be *known* vector transformations. Particular realizations of the state and measurement noises $\mathbf{w}_k$ and $\mathbf{v}_k$ are *unknown*, but their PDFs, i.e., the state noise PDF $p(\mathbf{w}_k)$ and the measurement noise PDF $p(\mathbf{v}_k)$, are supposed to be *known* and independent of the *known* initial state PDF $p(\mathbf{x}_0)$.

A. Bayesian State Estimation and Recursive Relations

The goal of the state estimation in the Bayesian framework is to find the filtering PDF of the state $\mathbf{x}_k$ conditioned on all measurements $\mathbf{z}^k = [\mathbf{z}_0, \mathbf{z}_1, \ldots, \mathbf{z}_k]$ up to the time instant k, i.e., the conditional PDF $p(\mathbf{x}_k|\mathbf{z}^k), \forall k$, is sought.

The general solution to the state estimation is given by the BRRs for the conditional PDFs[2] computation [12]

$$p(\mathbf{x}_k|\mathbf{z}^k) = \frac{p(\mathbf{x}_k|\mathbf{z}^{k-1})p(\mathbf{z}_k|\mathbf{x}_k)}{p(\mathbf{z}_k|\mathbf{z}^{k-1})}, \tag{2}$$

$$p(\mathbf{x}_k|\mathbf{z}^{k-1}) = \int p(\mathbf{x}_k|\mathbf{x}_{k-1})p(\mathbf{x}_{k-1}|\mathbf{z}^{k-1})d\mathbf{x}_{k-1}, \tag{3}$$

where $p(\mathbf{x}_k|\mathbf{z}^{k-1})$ is the one-step predictive PDF computed by the Chapman-Kolmogorov equation (3) and $p(\mathbf{x}_k|\mathbf{z}^k)$ is the filtering PDF computed by the Bayes' rule (2). The PDFs $p(\mathbf{x}_k|\mathbf{x}_{k-1})$ and $p(\mathbf{z}_k|\mathbf{x}_k)$ are the state transition PDF obtained from (1a) and the measurement PDF obtained from (1b), respectively. The PDF $p(\mathbf{z}_k|\mathbf{z}^{k-1}) = \int p(\mathbf{x}_k|\mathbf{z}^{k-1})p(\mathbf{z}_k|\mathbf{x}_k)\ d\mathbf{x}_k$ is the one-step predictive PDF of the measurement. The estimate of the state is given by the filtering and the predictive PDFs. The recursion (2), (3) starts from $p(\mathbf{x}_0|\mathbf{z}^{-1}) = p(\mathbf{x}_0)$.

Note that the fusion of the estimates can be carried out for arbitrary estimates (predictive, filtering, or initial estimates), therefore the conditioning and time step indices are, for the sake of clarity, omitted in the chapters dealing with fusion. In these chapters, bottom indices are used to index the fused estimates.

III. Information Fusion

Fusion is a subclass of estimation where the problem is to merge estimates from different sources [13]. Hence, estimation methods apply directly to the fusion problem. In distributed fusion or estimation, it is important to properly handle typically unknown cross-correlations between estimates. A common approach is then to use *conservative* estimation methods, which guarantee that the uncertainty associated with the computed fused estimate is not underestimated. In this section, two conservative linear fusion methods are reviewed first and then the problem of fusing PDFs is discussed.

[2]Considering the model (1a), (1b), the BRRs (2), (3) should be conditioned also on *available* sequence of the input $\mathbf{u}_k, \forall k$. However, for the sake of notational simplicity, the input signal is assumed to be implicitly part of the condition and it is not explicitly stated, i.e., $p(\mathbf{x}_{k+1}|\mathbf{x}_k) = p(\mathbf{x}_{k+1}|\mathbf{x}_k; \mathbf{u}_k)$, $p(\mathbf{x}_k|\mathbf{z}^k) = p(\mathbf{x}_k|\mathbf{z}^k; \mathbf{u}^{k-1})$, and $p(\mathbf{x}_{k+1}|\mathbf{z}^k) = p(\mathbf{x}_{k+1}|\mathbf{z}^k; \mathbf{u}^k)$.

A. Moment-Based Fusion

Linear fusion methods are often implemented with an (extended) Kalman filter (EKF) for measurement filtering, where the estimate is in the form of the first two conditional moments only, i.e., in the form of the mean and covariance matrix. This is practical as the standard computationally efficient moment-based fusion algorithms cannot take advantage of the knowledge of the whole conditional PDF. In the literature, a wide range of fusion methods can be found. Among them, the conservative linear fusion methods have attracted significant attention.

In a basic set-up, the goal of the moment-based fusion is to linearly fuse two stochastic variables (estimates) $\mathbf{x}_1$ and $\mathbf{x}_2$, which are related to the same (unknown) quantity and where each variable is described by the known mean and covariance matrix, i.e., $\hat{\mathbf{x}}_1 = \mathsf{E}(\mathbf{x}_1)$, $\mathbf{P}_1 = \mathsf{cov}(\mathbf{x}_1)$, $\hat{\mathbf{x}}_2 = \mathsf{E}(\mathbf{x}_2)$ and $\mathbf{P}_2 = \mathsf{cov}(\mathbf{x}_2)$. Let $\hat{\mathbf{x}}_f$ and $\mathbf{P}_f$ be the computed mean and covariance of the fused quantity $\mathbf{x}_f$. The linear fusion constraint means that

$$\mathbf{x}_f = \mathbf{K}\begin{pmatrix}\mathbf{x}_1 \\ \mathbf{x}_2\end{pmatrix}, \tag{4}$$

where $\mathbf{K}$ is a fusion gain matrix. The fused estimate is said to be *conservative* if

$$\mathbf{P}_f \succeq \mathsf{cov}\left(\mathbf{x}_f - \hat{\mathbf{x}}_f\right), \tag{5}$$

where $\mathbf{A} \succeq \mathbf{B}$ means $\mathbf{A} - \mathbf{B}$ is positive semi-definite [7].

Below two important special cases of a *conservative linear unbiased estimator* (CLUE, [14]) are provided.

1) Covariance Intersection (CI, [7]): The CI conservatively merges estimates under completely unknown cross-correlations. The formulas are given by

$$\hat{\mathbf{x}}_f = \mathbf{P}_f\left(w\mathbf{P}_1^{-1}\hat{\mathbf{x}}_1 + (1-w)\mathbf{P}_2^{-1}\hat{\mathbf{x}}_2\right), \tag{6}$$

$$\mathbf{P}_f = \left(w\mathbf{P}_1^{-1} + (1-w)\mathbf{P}_2^{-1}\right)^{-1}, \tag{7}$$

where $w \in [0,1]$ is derived by minimizing a loss function of $\mathbf{P}_f$. The CI is an optimal CLUE given that two estimates are fused and that the cross-correlations are completely unknown [15]. The CI framework can also be extended to a fusion of Gaussian mixtures used in the Gaussian sum filter [16].

2) Largest Ellipsoid (LE, [17]): The CI computes conservative estimates for all possible cross-correlations given that $\hat{\mathbf{x}}_1$ and $\hat{\mathbf{x}}_2$ are conservative. This also means that in many cases the CI is overly conservative, i.e., the computed covariance $\mathbf{P}_f$ is needlessly large. The LE method is a less conservative alternative to the CI[3]. It is not conservative with respect to completely unknown cross-correlations [20]. However, under restrictive assumptions on the cross-correlations,

[3]The LE method has also been called *ellipsoidal intersection* [18], or *safe fusion* [19].

the LE is an optimal CLUE and therefore it becomes an important alternative to the CI when the CI provides too conservative results [14]. The LE method is summarized in Algorithm 1.

Algorithm 1: Largest Ellipsoid Method

1) Factorize $\mathbf{P}_1 = \mathbf{U}_1\mathbf{D}_1\mathbf{U}_1^{\mathrm{T}}$ and let $\mathbf{T}_1 = \mathbf{D}_1^{-\frac{1}{2}}\mathbf{U}_1^{\mathrm{T}}$. Factorize $\mathbf{T}_1\mathbf{P}_2\mathbf{T}_1^{\mathrm{T}} = \mathbf{U}_2\mathbf{D}_2\mathbf{U}_2^{\mathrm{T}}$ and let $\mathbf{T}_2 = \mathbf{U}_2^{\mathrm{T}}$.
2) Transform using $\mathbf{T} = \mathbf{T}_2\mathbf{T}_1$ according to

$$\hat{\mathbf{z}}_1 = \mathbf{T}\hat{\mathbf{x}}_1, \qquad \mathbf{D}_1 = \mathbf{T}\mathbf{P}_1\mathbf{T}^{\mathrm{T}} = \mathbf{I},$$
$$\hat{\mathbf{z}}_2 = \mathbf{T}\hat{\mathbf{x}}_2, \qquad \mathbf{D}_2 = \mathbf{T}\mathbf{P}_2\mathbf{T}^{\mathrm{T}}.$$

3) For each $i = 1, \ldots, n$ of $\hat{\mathbf{z}}$ and diagonal $\mathbf{D}$ compute

$$\left([\hat{\mathbf{z}}]_i, [\mathbf{D}]_{ii}\right) = \begin{cases} \left([\hat{\mathbf{z}}_1]_i, 1\right), & \text{if } 1 \leq [\mathbf{D}_2]_{ii}, \\ \left([\hat{\mathbf{z}}_2]_i, [\mathbf{D}_2]_{ii}\right), & \text{if } 1 > [\mathbf{D}_2]_{ii}. \end{cases}$$

4) Fused estimate is given by $\hat{\mathbf{x}} = \mathbf{T}^{-1}\hat{\mathbf{z}}$ and $\mathbf{P} = \mathbf{T}^{-1}\mathbf{D}\mathbf{T}^{-\mathrm{T}}$

B. Probability Density Based Fusion

There is a considerable number of approaches that can be taken to fuse PDFs, the topic has been extensively studied by researchers last few decades. A very comprehensive survey considering the fusion of PDFs can be found in [21], where three main approaches have been identified

1) *Axiomatic approach*: The fusion rule is defined indirectly by a set of properties (axioms) that it is required to satisfy. In *this* paper, log-linear pooling is employed as it is computationally efficient and has the strongest link to the CI algorithm.
2) *Optimization approach*: The fusion rule is a result of optimization. This fusion approach is extremely computationally expensive, and therefore not considered in this paper.
3) *Supra-Bayesian approach*: This fusion center is considered a Bayesian observer that interprets the particular (nodes) PDFs as random observations. This approach requires some knowledge about the dependence of the fused PDFs, which is not available in the scenario considered in this paper.

IV. Motivation and Goals

The common methods for distributed and decentralized estimation (without fusion center), based on the local (or Gaussian) filters such as the EKF are moment-based. It means, these fusion methods cannot exploit the whole information encoded in the conditional PDF. Therefore, when the conditional PDFs estimated by the full-blown (or global) filters, such as the PMF of the particle filter, are available for the fusion, it might be highly beneficial to use PDF fusion methods instead. The global estimation methods are designed to deal with significant non-linearity/non-Gaussianity of the system.

The distributed version of the particle filter (dPF) has been researched and it is usually based on the Gaussian or Gaussian mixture approximation to the likelihood function. Then, the parameters of the Gaussian functions are sent between nodes. Therefore, either likelihood close to Gaussian is expected, or an algorithm has to be run to approximate the likelihood by a mixture [22], [23]. Alternatively, the dPF using adaptive encoding [24] can be employed, where several Lloyd-Max quantizers are trained, which makes the algorithm feasible only for lower dimensions. The distributed PMF has been presented, to the authors' knowledge, only in [25], where the same method of approximating the likelihood with a Gaussian mixture was employed, as in the dPF. This approach is, therefore, computationally prohibitive as well.

In the next section, we propose and discuss the concept of the distributed PMF (dPMF), which sends the estimated conditional PDFs between nodes, with the newly developed transfer reduction step based on the copula theory.

V. Distributed point-mass Filter

First a few preliminaries are presented that are crucial for understanding the proposed dPMF algorithm. Then the dPMF algorithm is described.

A. Point-Mass Density and Point-Mass Filter Outline

The PMF is based on an approximation of a conditional PDF $p(\mathbf{x}_k|\mathbf{z}^m)$, where $m = k$ for the filtering PDF and $m = k-1$ for the predictive PDF, by a *piece-wise constant* point-mass density $\hat{p}(\mathbf{x}_k|\mathbf{z}^m; \boldsymbol{\xi}_k)$ defined at the set of the discrete grid points $\boldsymbol{\xi}_k = \{\boldsymbol{\xi}_k^{(i)}\}_{i=1}^N, \boldsymbol{\xi}_k^{(i)} \in \mathbb{R}^{n_x}$, as follows

$$\hat{p}(\mathbf{x}_k|\mathbf{z}^m; \boldsymbol{\xi}_k) \triangleq \sum_{i=1}^{N} P_{k|m}(\boldsymbol{\xi}_k^{(i)}) S\{\mathbf{x}_k; \boldsymbol{\xi}_k^{(i)}, \boldsymbol{\Delta}_k\}, \quad (8)$$

with

- $P_{k|m}(\boldsymbol{\xi}_k^{(i)}) = c_k \tilde{P}_{k|m}(\boldsymbol{\xi}_k^{(i)})$, where $\tilde{P}_{k|m}(\boldsymbol{\xi}_k^{(i)}) = p(\boldsymbol{\xi}_k^{(i)}|\mathbf{z}^m)$ is the value of the conditional PDF $p(\mathbf{x}_k|\mathbf{z}^m)$ evaluated at the i-th grid point $\boldsymbol{\xi}_k^{(i)}$, $c_k = \delta_k \sum_{i=1}^N \tilde{P}_{k|m}(\boldsymbol{\xi}_k^{(i)})$ is a normalisation constant, and δ_k is the volume of the i-th point neighbourhood defined below,
- $\boldsymbol{\Delta}_k = [\boldsymbol{\Delta}_k(1), \boldsymbol{\Delta}_k(2), \ldots, \boldsymbol{\Delta}_k(n_x)]^{\mathrm{T}}$ defines a (hyper-) rectangular neighbourhood of a grid point $\boldsymbol{\xi}_k^{(i)}$, where the PDF $p(\mathbf{x}_k|\mathbf{z}^m)$ is assumed to be constant and has value $P_{k|m}(\boldsymbol{\xi}_k^{(i)})$, and
- $S\{\mathbf{x}_k; \boldsymbol{\xi}_k^{(i)}, \boldsymbol{\Delta}_k\}$ is the *selection* function defined as

$$S\{\mathbf{x}_k; \boldsymbol{\xi}_k^{(i)}, \boldsymbol{\Delta}_k\} = \begin{cases} 1, \text{if } |\mathbf{x}_k(j) - \boldsymbol{\xi}_k^{(i)}(j)| \leq \frac{\boldsymbol{\Delta}_k(j)}{2}, \\ 0, \text{otherwise.} \end{cases} \quad (9)$$

so that $\int S\{\mathbf{x}_k; \boldsymbol{\xi}_k^{(i)}, \boldsymbol{\Delta}_k\} d\mathbf{x}_k = \prod_{i=1}^{n_x} \boldsymbol{\Delta}_k(i) = \delta_k$.

The notation $\mathbf{x}(j)$ means the j-th element of the vector $\mathbf{x}$. Illustration of point-mass PDF approximation (8) with omitted time indices is shown in Fig. 2. The point-mass density (8) can, thus, be interpreted as a *sum* of weighted uniform distributions [26].

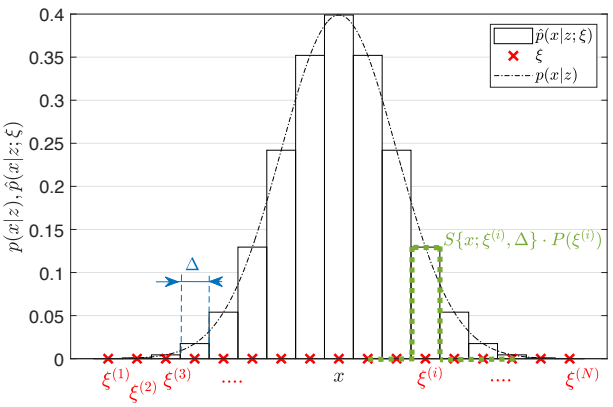

Fig. 2. Point-mass density approximation

The moments, e.g., the mean $\hat{\mathbf{x}}_{k|m} = \mathsf{E}[\mathbf{x}_k|\mathbf{z}^m]$ and covariance matrix $\mathbf{P}_{k|m} = \mathsf{cov}[\mathbf{x}_k|\mathbf{z}^m]$, of the conditional PMD, are not required for the run of the PMF. However, the moments can be readily computed when required [27].

Approximation of the conditional PDF by the PMD at the grid points enables a numerical solution for the BRRs, which leads to the algorithm of the PMF. The PMF, thus, calculates the values of the conditional densities at (discrete) grid points $\boldsymbol{\xi}$.

Although the PMD is the key enabler for the BRRs solution, it has a significant drawback w.r.t. distributed estimation relying on the density transfer between nodes. The description of the PMD is data intensive. For example, to describe and transfer the PMD with N_{pa} points per axis over n_x-dimensional state space, $N_{pa}^{n_x}$ weights $P_{k|k}$ has to be sent between nodes together with the information about the position of the grid and location of the points. The complexity of the grid description depends among others on the grid shape (circular/rectangular) and point placement strategy (equidistant/irregular).

B. Copulas

The theory of *copulas* has been developed for correlation modeling in cases when the marginal distributions of a vector random variable are known [28], [29]. It enables the n_x-dimensional PDF to be split into n_x marginals and one copula describing the correlation between them. Therefore, a (conditional) PDF $p(\mathbf{x})$ can be split as follows

$$p(\mathbf{x}) = c\prod_{i=1}^{n_\mathbf{x}} p(x_i), \tag{10}$$

where $p(x_i)$ is the i-th marginal, and c is the appropriate n_x dimensional copula. Among all the proposed copulas, the Gaussian copula is particularly suitable in filter design [11]. Besides the Gaussian copula, other copulas from the Archimedean family, such as the Clayton, Frank, Gumbel, and Product copula, can be used as well, which might require even fewer memory requirements but at the cost of necessity to specify a set of user-defined parameters.

The Gaussian copula is defined as

$$\overline{c} = \frac{1}{\sqrt{|\mathbf{C}|}}\exp\left(-\frac{1}{2}\mathbf{y}^T(\mathbf{C}^{-1}-\mathbf{I}_{n_x})\mathbf{y}\right), \tag{11}$$

where $\mathbf{C}$ is the correlation matrix

$$\mathbf{C} = \left(\mathrm{diag}(\mathbf{P})\right)^{-\frac{1}{2}}\mathbf{P}\left(\mathrm{diag}(\mathbf{P})\right)^{-\frac{1}{2}}, \tag{12}$$

$\mathrm{diag}(\mathbf{P})$ is the diagonal matrix determined by the diagonal of a covariance matrix $\mathbf{P}$, $|\mathbf{C}|$ denotes the determinant of $\mathbf{C}$, $\mathbf{I}_{n_x}$ is an identity matrix of the indicated dimensions. Moreover

$$\mathbf{y} = \begin{bmatrix} \Phi^{-1}\left(\Phi_1(x(1))\right) \\ \vdots \\ \Phi^{-1}\left(\Phi_{n_x}(x(n_x))\right) \end{bmatrix} \in \mathrm{R}^{n_x}, \tag{13}$$

$\Phi_i(\mathbf{x}_{k+1}(i)))$ stands for the cumulative distribution function (CDF) of the Gaussian PDF with the mean $\hat{\mathbf{x}}_{k+1|k}(i)$ and the variance $\mathbf{P}_{k+1|k}(i,i)$, i.e., of the PDF $\mathcal{N}\{\mathbf{x}_{k+1}(i);\hat{\mathbf{x}}_{k+1|k}(i),\mathbf{P}_{k+1|k}(i,i)\}$, $\Phi^{-1}(\cdot)$ denotes the inverse CDF of the standard Gaussian PDF, and the notation $\mathbf{P}_{k+1|k}(i,i)$ means i-th diagonal element of the matrix $\mathbf{P}_{k+1|k}$. Detailed relations on the PMD split into the marginals and Gaussian copula can be found in [11].

C. Entropy

In the information theory, the *entropy* describes the average level of information that PDF bears [30]. The entropy is, thus, particularly suitable for the determination of the weights of the PMDs to be fused. The entropy is calculated from the arbitrary PMD's weights as follows

$$S = -\sum_{i=1}^{N} P(\boldsymbol{\xi}^{(i)})\ln(P(\boldsymbol{\xi}^{(i)})). \tag{14}$$

The larger the entropy is, the lower amount of information the PMD contains. Thus, the weights for fusion are inversely proportional to the entropy of the given PMD.

D. Distributed PMF Algorithm

The proposed PMF-based efficient distributed fusion algorithm is, for the sake of explanation simplicity, introduced and illustrated for a fusion of two densities. The generalization for the fusion of multiple estimates is straightforward.

Suppose two (conditional) PMDs to be fused, namely $\hat{p}_1(\mathbf{x};\boldsymbol{\xi}_1)$ with weights $P_1(\boldsymbol{\xi}_1)$, and $\hat{p}_2(\mathbf{x};\boldsymbol{\xi}_2)$ with weights $P_2(\boldsymbol{\xi}_2)$. The subscript denotes the index of the PMD to be fused.

The fusion algorithm utilizes a so-called *common grid*. It is a rectangular grid that is supposed to cover the important part of the fused PMD and on which both received, i.e., local, PMDs are interpolated. There is a number of ways this grid can be set up:

1) Calculate the first two moments of the fused estimate using the CI or LE algorithms. Construct the common

grid around the fused mean with the size given by the fused variance.

2) Circumscribe all points $\boldsymbol{\xi}_1$ and $\boldsymbol{\xi}_2$ of local PMDs, for which the PMD weight is greater than a threshold, by a rectangular grid.

In both cases, local estimators can be designed so that local grids have their boundaries aligned with the state-space axis, their grid steps $\boldsymbol{\Delta}_k$ are the same in each time step k, and they are shifted against each other by only integer multiples of $\boldsymbol{\Delta}_k$. Then, the grid points $\boldsymbol{\xi}_1$, and $\boldsymbol{\xi}_2$ will overlap exactly. Therefore no interpolation will be necessary, which will result in a lower computational complexity.

The full PMD fusion algorithm can fuse the whole PMDs, at a price of considerable data transfer. All PMD weights have to be transferred along with the information about underlying grid points (based on the grid design). The data transfer complexity is therefore $\mathcal{O}(N_{pa}^{n_x})$, where N_{pa} is a number of the points per axis.

The data transfer efficient version of the fusion for PMDs has to transfer only the $n_x N_{pa}$ marginal weights, marginal grid point values, and N_{pa} copula parameters, therefore the number of data transferred is significantly decreased. The data transfer complexity is therefore $\mathcal{O}(N_{pa})$ only. The PMD is sent from one node to another in a marginal form alongside its copula parameters and assembled back to the full PMD. Based on the performed simulations, the approximation of the PMD by the copula-based decomposition and the subsequent fusion has negligible impact on the fused estimate accuracy. For completeness, the proposed approach for the distributed PMF is summarized in Algorithm 2. The data exchange can be carried out in any part of the PMF algorithm.

Algorithm 2: Distributed Point-Mass Filter Data Exchange

1) (Optional but recommended step) Each node splits the PMD as in (10) into n_x marginals and copula parameters and sends them to the second node (or others). The second node reconstructs the full (original) PMD approximation.
2) Both PMD's weights are interpolated on a common grid $\boldsymbol{\xi}_c$.
3) Entropy S is calculated for both PMDs.
4) The fusion weights are set as $w_1 = \frac{1}{S_1}$ and $w_2 = \frac{1}{S_2}$ indirectly related to the entropy, and re-scaled so that $w_1 + w_2 = 1$.
5) The PMD weights are fused on the common grid
$$P_{\text{fused}}(\boldsymbol{\xi}_c^{(i)}) = P_1(\boldsymbol{\xi}_c^{(i)})^{w_1} P_2(\boldsymbol{\xi}_c^{(i)})^{w_2}, \forall i. \quad (15)$$
6) The fused PMD is normalized.

VI. Numerical illustration

The performance of the proposed dPMF is numerically illustrated and compared with the moment-based CI and LE methods. Moreover, two implementations of the dPMF are considered, namely dPMF with full PMD transfer (i.e., without copula-based PMD approximation) and dPMF with reduced PMD transfer (i.e., with copula-based approximation). The former is further denoted as *PMF Full*, whereas the latter is as *PMF Cop*.

For numerical evaluation we assume the state-space model [13]

$$\mathbf{x}_{k+1} = \mathbf{x}_k + \mathbf{u}_k + \mathbf{w}_k, \quad (16)$$

$$\mathbf{z}_k = \begin{bmatrix} \sqrt{(\mathbf{x}_k(1)^2 + \mathbf{x}_k(2)^2)} \\ \arctan\left(\frac{\mathbf{x}_k(2)}{\mathbf{x}_k(1)}\right) \end{bmatrix} + \mathbf{v}_k. \quad (17)$$

The two-dimensional state vector $\mathbf{x}_k$ consists of vehicle horizontal position and the horizontal constant shift vector $\mathbf{u}_k = [-10, -10]^T$ can be known from e.g., inertial navigation system or odometer. The measurement $\mathbf{z}_k$ consist of the range and angle from the radar to the vehicle. The noise $\mathbf{v}_k$ is described as

$$\mathcal{N}\{\mathbf{v}_k; \begin{bmatrix}0\\0\end{bmatrix} \begin{bmatrix}20 & 10\\10 & 20\end{bmatrix}\}, \quad (18)$$

and the noise $\mathbf{w_k}$ is a Gaussian mixture PDF with five components

$$p_{\mathbf{w}}(\mathbf{x}) = \sum_{g=1}^{5} \alpha_g \mathcal{N}\{\mathbf{x}; \hat{\mathbf{x}}_g, \mathbf{P}_g\}, \quad (19)$$

where the particular weights, means, and covariance matrices are given as follows

$$\begin{bmatrix}\alpha_1\\\alpha_2\\\alpha_3\\\alpha_4\\\alpha_5\end{bmatrix} = \begin{bmatrix}0.0108\\0.3280\\0.2956\\0.1556\\0.2101\end{bmatrix}, \begin{bmatrix}\hat{\mathbf{x}}_1\\\hat{\mathbf{x}}_2\\\hat{\mathbf{x}}_3\\\hat{\mathbf{x}}_4\\\hat{\mathbf{x}}_5\end{bmatrix} = \begin{bmatrix}0 & 0\\-50 & -0.02\\50 & 0.02\\-100 & 0.06\\250 & -0.1\end{bmatrix}, \quad (20)$$

$$\mathbf{P}_1 = \mathbf{P}_2 = \mathbf{P}_3 = \mathbf{P}_4 = \mathbf{P}_5 \begin{bmatrix}1500 & 0\\0 & 0.001\end{bmatrix}. \quad (21)$$

The results are compared using three criteria:

- RMSE (root mean square error)
$$\sqrt{\frac{1}{M(K+1)} \sum_{i=1}^{M} \sum_{k=0}^{K} (\mathbf{x}_k^{[i]} - \hat{\mathbf{x}}_{k|k}^{[i]})^T (\cdot)}, \quad (22)$$
- ANEES[4] (normalized estimation error squared)
$$\frac{1}{M(K+1)n_x} \sum_{i=1}^{M} \sum_{k=0}^{K} (\mathbf{x}_k^{[i]} - \hat{\mathbf{x}}_{k|k}^{[i]})^T (\mathbf{P}_{k|k}^{[i]})^{-1} (\cdot), \quad (23)$$
- Transfer is the number of transferred values between nodes,

where $M = 1000$ is the number of Monte-Carlo simulations, $K = 20$ is the number of time steps, $(\cdot)$ denotes the repetition

[4] Averaged over time

of the bracket with difference, $\mathbf{x}_k^{[i]}$ is true state at time k at i-the MC simulation, $\mathbf{x}_{k|k}^{[i]} = \mathsf{E}[\mathbf{x}_k|\mathbf{z}^k]$ its filtering estimate, and $(\mathbf{P}_{k|k}^{[i]})^{-1} = (\mathsf{cov}[\mathbf{x}_k|\mathbf{z}^k])^{-1}$ is the inversion of the respective covariance matrix.

The ANESS should be around 1, if it is higher, the estimation algorithm is said to be too optimistic, if it is lower, the algorithm is consistent but too conservative. As the cross-correlation between estimates is unknown in this case, the ANEES should be ideally lower to account for that.

The results can be found in Table I. The number of transferred values for KF's comes from the covariance matrix (3 values) and mean value (2 values). For PMF Full it is PDF values ($50 \times 50 = 2500$ values), corner point of the grid (2 values), and $\mathbf{\Delta}$ (2 values), as the grid is equidistantly spaced, and the axes are aligned with the state space axes. For PMF Cop the marginal PDF values ($50 + 50 = 100$ values), the corner point (2 values), the $\mathbf{\Delta}$ (2 values), and the copula correlation parameter (3 values).

The results indicate that the proposed copula-based dPMF provides more accurate results than standard moment-based fusors with a significant reduction of the data transfer compared to the dPMF communicating the whole PMDs among the nodes.

TABLE I
RESULTS

	RMSE	ANEES	Transfer
EKF IC	15.813	0.5247	5
EKF LE	16.111	0.8396	5
PMF Full	14.462	0.48033	2504
PMF Cop	15.167	0.49259	107

VII. CONCLUDING REMARKS

This paper dealt with distributed nonlinear state estimation by the point-mass filter. In particular, the distributed PMF (dPMF) with a data transfer reduction step based on the copula theory was proposed. The developed dPMF reduces the data transfer between the nodes (i.e., particular estimators) from exponential to linear growth. The numerical illustration confirmed that the copula-based efficient distributed point-mass filter results in significantly reduced data transfer for the price of a small accuracy decrease. In future research, the stress will be laid on the automatic selection of the copula type and its parameter determination, which further increase the accuracy and reduce the computational complexity.

REFERENCES

[1] H. W. Sorenson, "On the development of practical nonlinear filters," *Information Sciences*, vol. 7, pp. 230–270, 1974.

[2] U. Orguner, P. Skoglar, D. Tornqvist, and F. Gustafsson, "Combined point-mass and particle filter for target tracking," in *Proceedings of the 2010 IEEE Aerospace Conference*, April 2010.

[3] J. Lim and C. G. Park, "Rbppff for robust tan," *IET Radar, Sonar & Navigation*, vol. 12, 2019.

[4] M. Vesely, J. Dunik, P. Hotmar, and T. Beda, "Point-mass filter in land and air terrain-aided navigation: Performance evaluation," in *Proceedings of the 2022 25th International Conference on Information Fusion (FUSION)*, July 2022, pp. 1–8.

[5] J. Matoušek, "Point-mass method in state estimation and navigation," Ph.D. dissertation, University of West Bohemia, Czech Republic, 10.13140/RG.2.2.30208.87044, 2020.

[6] S. J. Julier and J. K. Uhlmann, "General decentralized data fusion with covariance intersection," in *Handbook of Multisensor Data Fusion: Theory and Practice*, M. Liggins, D. Hall, and J. Llinas, Eds. Boca Raton, FL, USA: CRC Press, 2009, ch. 14.

[7] ——, "A non-divergent estimation algorithm in the presence of unknown correlations," in *Proceedings of the 1997 American Cont. Conf.*, Albuquerque, NM, USA, Jun. 1997, pp. 2369–2373.

[8] M. B. Hurley, "An information theoretic justification for covariance intersection and its generalization," in *Proceedings of the5th IEEE Int. Conf. on Inf. Fusion*, Annapolis, MD, USA, Jul. 2002.

[9] R. Forsling, Z. Sjanic, F. Gustafsson, and G. Hendeby, "Consistent distributed track fusion under communication constraints," in *Proceedings of the 22nd IEEE Int. Conf. on Inf. Fusion*, Ottawa, Canada, Jul. 2019.

[10] ——, "Optimal linear fusion of dimension-reduced estimates using eigenvalue optimization," in *Proceedings of the 25th IEEE Int. Conf. on Inf. Fusion*, Linköping, Sweden, Jul. 2022.

[11] J. Duník, O. Straka, J. Matoušek, and E. Blasch, "Copula-based convolution for fast point-mass prediction," *Signal Processing*, vol. 192, p. 108367, 2022.

[12] B. D. O. Anderson and J. B. Moore, *Optimal Filtering*. Prentice Hall, New Jersey, 1979.

[13] Y. Bar-Shalom, X. R. Li, and T. Kirubarajan, *Estimation with Applications to Tracking and Navigation: Theory Algorithms and Software*. John Wiley & Sons, 2001.

[14] R. Forsling, A. Hansson, F. Gustafsson, Z. Sjanic, J. Löfberg, and G. Hendeby, "Conservative linear unbiased estimation under partially known covariances," *IEEE Trans. on Sig. Proc.*, jun 2022.

[15] M. Reinhardt, B. Noack, P. O. Arambel, and U. D. Hanebeck, "Minimum covariance bounds for the fusion under unknown correlations," *IEEE Sig. Process. Lett.*, vol. 22, no. 9, pp. 1210–1214, Sep. 2015.

[16] N. R. Ahmed, "Decentralized gaussian mixture fusion through unified quotient approximations." arXiv:1907.04008, 2019.

[17] A. R. Benaskeur, "Consistent fusion of correlated data sources," in *Proceedings of the 28th Annual Conference of the IEEE Industrial Electronics Society*, Sevilla, Spain, Nov. 2002, pp. 2652–2656.

[18] J. Sijs, M. Lazar, and P. P. J. v. d. Bosch, "State fusion with unknown correlation: Ellipsoidal intersection," in *Proceedings of the 2010 American Cont. Conf.*, Baltimore, MD, USA, Jun. 2010, pp. 3992–3997.

[19] F. Gustafsson, *Statistical Sensor Fusion*. Sweden: Studentlit., 2018.

[20] B. Noack, J. Sijs, and U. D. Hanebeck, "Algebraic analysis of data fusion with ellipsoidal intersection," in *Proceedings of theIEEE International Conference on Multisensor Fusion and Integration*, Baden-Baden, Germany, Sep. 2016, pp. 365–370.

[21] G. Koliander, Y. El-Laham, P. M. Djurić, and F. Hlawatsch, "Fusion of probability density functions." arXiv, 2022. [Online]. Available: https://arxiv.org/abs/2202.11633

[22] T. Ghirmai, "Distributed particle filter for target tracking: With reduced sensor communications," *Sensors*, vol. 16, no. 9, 2016.

[23] D. Gu, "Distributed particle filter for target tracking," in *Proceedings of the 2007 IEEE Int. Conf. on Rob. and Auto.*, 2007, pp. 3856–3861.

[24] M. Coates, "Distributed particle filters for sensor networks," in *Proceedings of the 3rd international symposium on Information Processing in Sensor Networks - IPSN'04*. California: ACM Press, 2004.

[25] C. J. Bordin and M. G. S. Bruno, "Distributed Rao-Blackwellized point mass filter for blind equalization in receiver networks," in *Proceedings of 2015 23rd European Signal Processing Conference (EUSIPCO)*, France, 2015, pp. 2186–2190.

[26] N. Bergman, "Recursive Bayesian estimation: Navigation and tracking applications," Ph.D. dissertation, Linköping University, Sweden, 1999.

[27] J. Duník, M. Soták, M. Veselý, O. Straka, and W. J. Hawkinson, "Design of Rao-Blackwellised point-mass filter with application in terrain aided navigation," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 55, no. 1, pp. 251–272, 2019.

[28] I. Žežula, "On multivariate Gaussian copulas," *Journal of Statistical Planning and Inference*, vol. 139, no. 11, pp. 3942–3946, 2009.

[29] R. Nelsen, *An Introduction to Copulas*, ser. Springer Series in Statistics. Springer New York, 2007.

[30] R. M. Gray, *Entropy and Information Theory*. Springer US, 2011.