POINT-MASS FILTER WITH DECOMPOSITION OF TRANSIENT DENSITY

Petr Tichavský

Institute of Information Theory and Automation,
The Czech Academy of Sciences,
Pod Vodárenskou věží 4, 182 00 Praha,
Czech Republic

Ondřej Straka, and Jindřich Duník

Department of Cybernetics,
University of West Bohemia
Univerzitní 8, 306 14 Pilsen,
Czech Republic

ABSTRACT

The paper deals with the state estimation of nonlinear stochastic dynamic systems with special attention on a grid-based numerical solution to the Bayesian recursive relations, the point-mass filter (PMF). In the paper, a novel functional decomposition of the transient density describing the system dynamics is proposed. The decomposition is based on a non-negative matrix factorization and separates the density into functions of the future and current states. Such decomposition facilitates a thrifty calculation of the convolution, which is a bottleneck of the PMF performance. The PMF estimate quality and computational costs can be efficiently controlled by choosing an appropriate rank of the decomposition. The performance of the PMF with the transient density decomposition is illustrated in a terrain-aided navigation scenario.

Index Terms— State estimation, filtering, nonlinear systems, point-mass method, non-negative matrix factorization

1. INTRODUCTION

State estimation of nonlinear discrete-time stochastic dynamic systems from noisy measurements has been a subject of considerable research interest for many decades and plays an indispensable role in fields such as navigation, speech and image processing, fault diagnosis, and adaptive control.

Within the Bayesian framework, a general solution to the state estimation problem is given by the Bayesian recursive relations (BRRs) inferring the probability density functions (PDFs) of the state conditioned on the measurements. The PDFs fully describe the immeasurable state of a possibly nonlinear or non-Gaussian stochastic dynamic system. The relations are analytically tractable for a limited set of models such as linear Gaussian models. In other cases, approximate solutions to the BRRs have to be employed, which offer various levels of approximation. The Gaussian filters assuming the joint state and measurement prediction PDF being Gaussian are attractive for mildly nonlinear models, for which they offer computational efficiency and reasonable estimate quality. For strongly nonlinear or non-Gaussian models one usually resorts to more complex (and thus computationally demanding) filters such as the particle filter [1] or the point-mass filter (PMF) [2].

This paper considers the PMF [3], [4], [5]. It is based on a numerical solution to the BRRs using *deterministic* grid-based numerical integration rules and computes the conditional PDFs at the grid points only. A suitable selection of the grid points is critical as it affects the PMF estimate accuracy and computational complexity. The bottleneck of the PMF limiting the number of the grid points (and thus hindering the PMF application for higher dimensional states) is the predictive step. This step involves an evaluation of a convolution called the Chapman-Kolmogorov equation, where the grids for two consecutive time instants are combined through the transient PDF, which is the essence of the bottleneck.

In this paper, we propose to handle the transient PDF in the form of its decomposition. The decomposition is based on the non-negative matrix factorization[1] (NNMF) and symmetric NNMF, see [8] and references therein. It is specific for the given grid and has to be done once, prior to the estimation itself. In this way, it is possible to handle larger grids while maintaining affordable computational complexity.

The paper is organized as follows. Section 2 deals with introduction of the PMF-based Bayesian state estimation. In Section 3, the decomposition of the transient PDF is proposed. Performance analysis of the decomposition itself and the PMF with the decomposition are provided in Section 4 and, in Section 5, the conclusion remarks are drawn.

2. POINT-MASS FILTER

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

$$\mathbf{x}_{k+1} = \mathbf{f}_k(\mathbf{x}_k, \mathbf{u}_k) + \mathbf{w}_k, \qquad k = 0, 1, 2, \ldots, T, \tag{1}$$

$$\mathbf{z}_k = \mathbf{h}_k(\mathbf{x}_k) + \mathbf{v}_k, \qquad k = 0, 1, 2, \ldots, T, \tag{2}$$

The work was supported by the Czech Science Foundation under grant 22-11101S

[1]The NNMF originally known as non-negative rank factorization or positive matrix factorization has been subject to intensive research for more than three decades [6], [7].

where $\mathbf{x}_k \in \mathbb{R}^{n_x}$, $\mathbf{u}_k \in \mathbb{R}^{n_u}$, 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 $\mathbf{f}_k : \mathbb{R}^{n_x \times n_u} \rightarrow \mathbb{R}^{n_x}$ and $\mathbf{h}_k : \mathbb{R}^{n_x} \rightarrow \mathbb{R}^{n_z}$ 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 $p(\mathbf{x}_0)$.

2.1. Bayesian State Estimation and Recursive Relations

The goal of the state estimation is to find the conditional PDF $p(\mathbf{x}_k|\mathbf{z}^k), \forall k$ conditioned on all measurements $\mathbf{z}^k = [\mathbf{z}_0, \mathbf{z}_1, \ldots, \mathbf{z}_k]$ up to the time instant k, called filtering PDF. The general solution to the state estimation is given by the BRRs for the conditional PDFs[2] computation [9]

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

$$p(\mathbf{x}_{k+1}|\mathbf{z}^k) = \int p(\mathbf{x}_{k+1}|\mathbf{x}_k)p(\mathbf{x}_k|\mathbf{z}^k)d\mathbf{x}_k, \tag{4}$$

where $p(\mathbf{x}_{k+1}|\mathbf{z}^k)$ is the one-step predictive PDF computed by the Chapman-Kolmogorov (CK) equation (4) and $p(\mathbf{x}_k|\mathbf{z}^k)$ is the filtering PDF computed by the Bayes' rule (3). The transient PDF $p(\mathbf{x}_{k+1}|\mathbf{x}_k) = p_{\mathbf{w}_k}(\mathbf{x}_{k+1} - \mathbf{f}_k(\mathbf{x}_k, \mathbf{u}_k))$ and the measurement PDF $p(\mathbf{z}_k|\mathbf{x}_k) = p_{\mathbf{v}_k}(\mathbf{z}_k - \mathbf{h}_k(\mathbf{x}_k))$ are the state transient PDF obtained from (1) and the measurement PDF obtained from (2), 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 (3), (4) starts from $p(\mathbf{x}_0|\mathbf{z}^{-1}) = p(\mathbf{x}_0)$.

2.2. Point-Mass Density Approximation

Assume for convenience a scalar state $x_k \in \mathbb{R}$. The PMF is based on an approximation of a conditional PDF $p(x_k|\mathbf{z}^m)$, where $m = k$ for the filtering PDF and $m = k-1$ for the predictive PDF, by a *piece-wise constant* point-mass density $\Xi_{k|m}$ defined at the set of the discrete grid points $\xi_k = \{\xi_k^{(i)}\}_{i=1}^N, \xi_k^{(i)} \in \mathbb{R}$, as follows

$$\Xi_{k|m} \triangleq \sum_{i=1}^{N} P_{k|m}^{(i)} S\{x_k; \xi_k^{(i)}, \Delta_k\}, \tag{5}$$

with

- $P_{k|m}^{(i)} = c_k \tilde{P}_{k|m}^{(i)}$, where $\tilde{P}_{k|m}^{(i)} = p(\xi_k^{(i)}|\mathbf{z}^m)$ is the conditional PDF $p(x_k|\mathbf{z}^m)$ evaluated at the i-th grid point $\xi_k^{(i)}$, $c_k = \Delta_k \sum_{i=1}^N \tilde{P}_{k|m}^{(i)}$ is a normalization constant.
- Δ_k is a neighborhood of a grid point $\xi_k^{(i)}$, where the PDF $p(x_k|\mathbf{z}^m)$ is assumed to be constant and has value $P_{k|m}^{(i)}$,
- $S\{x_k; \xi_k^{(i)}, \Delta_k\}$ is the *selection* function defined as

$$S\{x_k; \xi_k^{(i)}, \Delta_k\} = \begin{cases} 1, \text{if } |x_k - \xi_k^{(i)}(j)| \leq \frac{\Delta_k}{2}, \\ 0, \text{otherwise.} \end{cases} \tag{6}$$

2.3. Point-Mass Filter Summary

The basic algorithm of the PMF can be summarized as [4]:

Algorithm 1: Point-Mass Filter

1. *Initialization*: Set $k = 0$, construct the initial grid of points $\{\xi_0^{(i)}\}_{i=1}^N$, and define the initial point-mass PDF $\Xi_{0|-1}$ of form (5) approximating the initial PDF.
2. *Meas. update*: Compute the filtering point-mass PDF $\Xi_{k|k}$ of the form (5) where the PDF value at i-th grid point is $P_{k|k}^{(i)} = \frac{p(\mathbf{z}_k|x_k=\xi_k^{(i)})P_{k|k-1}^{(i)}}{\sum_{i=1}^N p(\mathbf{z}_k|x_k=\xi_k^{(i)})P_{k|k-1}^{(i)}\Delta_k}$.
3. *Grid construction:* Construct the new[3] grid $\{\xi_{k+1}^{(j)}\}_{j=0}^N$.
4. *Time update*: Compute the predictive point-mass PDF $\Xi_{k+1|k}$ of the form (5) at the new grid of points, where the value of the predictive PDF at j-th grid point is $P_{k+1|k}^{(j)} = \sum_{i=1}^N p_{x_{k+1}|x_k}(\xi_{k+1}^{(j)}|x_k = \xi_k^{(i)})P_{k|k}^{(i)}\Delta_k$.
5. Set $k = k+1$ and go to the step 2.

The PMF provides estimates in the form of the point-mass density $\hat{p}(\mathbf{x}_k|\mathbf{z}^m; \boldsymbol{\xi}_k)$ approximating $p(\mathbf{x}_k|\mathbf{z}^m)$. The point estimate, i.e., the mean $\hat{\mathbf{x}}_{k|m} = \mathsf{E}[\mathbf{x}_k|\mathbf{z}^m]$ of the conditional PDF is not required for the run of the PMF. However, it can readily be computed if required [10].

The solution to the CK equation (4) represents a PMF bottleneck as it requires evaluation of the transient PDF $p(\mathbf{x}_{k+1}|\mathbf{x}_k)$ for all combinations of the grid points $\{\xi_{k+1}^{(j)}\}_{j=1}^N$ and $\{\xi_k^{(i)}\}_{i=1}^N$. Therefore, a transient PDF decomposition is proposed to reduce the computational costs of the PMF.

3. TRANSIENT DENSITY DECOMPOSITION

Assume that the state transient PDF can be decomposed as

$$p(\mathbf{x}_{k+1}|\mathbf{x}_k) \approx \sum_{r=1}^{R} \mathcal{F}_{1r}(\mathbf{x}_{k+1})\mathcal{F}_{2r}(\mathbf{x}_k) \tag{7}$$

where $\mathcal{F}_{1r}(\cdot), \mathcal{F}_{2r}(\cdot)$, $r = 1, \ldots, R$ are suitable (non-negative) functions, known in advance, and R is the order of the approximation called *rank*.

[2] Considering the model (1), (2), the BRRs (3), (4) should be conditioned also on *available* sequence of the input $\mathbf{u}_k, \forall k$. However, for the sake of notation 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 $p(\mathbf{x}_{k+1}|\mathbf{z}^k) = p(\mathbf{x}_{k+1}|\mathbf{z}^k; \mathbf{u}^k)$.

[3] Usually, the number of grid points N is kept constant $\forall k$ to ensure constant (and predictable) computational complexity of the PMF.

Now, the CK equation (4) can be written as

$$p(\mathbf{x}_{k+1}|\mathbf{z}^k) \approx \sum_{r=1}^{R} \mathcal{F}_{1r}(\mathbf{x}_{k+1}) \int \mathcal{F}_{2r}(\mathbf{x}_k) p(\mathbf{x}_k|\mathbf{z}^k) d\mathbf{x}_k \, . \quad (8)$$

The simplification is that R integrals in (8) can be computed at the grid points $\{\xi_k^{(i)}\}_{i=1}^N$ only, without the need of considering the Kronecker product of the grid points $\{\xi_{k+1}^{(j)}\}_{j=1}^N$ for $\mathbf{x}_{k+1}$ and grid points $\{\xi_k^{(i)}\}_{i=1}^N$ for $\mathbf{x}_k$.

Relation (7) can be seen as a special case of the functional tensor decomposition [11] where the functions $\mathcal{F}_{1r}$ and $\mathcal{F}_{2r}$ are not decomposed further to functions of single elements of $\mathbf{x}_{k+1}$ and $\mathbf{x}_k$. The decomposition in (7) can be found in advance, if the grid for $\mathbf{x}_{k+1}$ and the grid for $\mathbf{x}_k$ are fixed. The task is equivalent to the well known task of NNMF or non-negative tensor factorization.

The decomposition (7) should be constructed prior to the estimation itself. This can be done if the grid points $\{\xi_{k+1}^{(j)}\}_{j=1}^N$ and $\{\xi_k^{(i)}\}_{i=1}^N$ are fixed, which seldom holds. An attractive option is to interpret the transient PDF $p(\mathbf{x}_{k+1}|\mathbf{x}_k) = p_{\mathbf{w}_k}(\mathbf{x}_{k+1} - \mathbf{f}_k(\mathbf{x}_k, \mathbf{u}_k))$ as a function of $\mathbf{x}_{k+1}$ and $\mathbf{f}_k$, where $\mathbf{f}_k = \mathbf{f}_k(\mathbf{x}_k, \mathbf{u}_k)$ denotes the value of the function $\mathbf{f}_k$ for convenience. Then, the value of the transient PDF only depends on the distance of $\mathbf{x}_{k+1}$ and $\mathbf{f}_k$, i.e. $p(\mathbf{x}_{k+1}|\mathbf{x}_k) = p_{\mathbf{w}_k}(\mathbf{x}_{k+1} - \mathbf{f}_k)$. Subsequently, the decomposition (7) in the form

$$p_{\mathbf{w}_k}(\mathbf{x}_{k+1} - \mathbf{f}_k) \approx \sum_{r=1}^{R} \mathcal{F}_{1r}(\mathbf{x}_{k+1}) \mathcal{F}_{2r}(\mathbf{f}_k) \quad (9)$$

needs to be computed only over a region of differences $\mathbf{x}_{k+1} - \mathbf{f}_k$. If the function $p_{\mathbf{w}_k}(\mathbf{x}_{k+1} - \mathbf{f}_k)$ is symmetric (invariant) with respect to permutation of its arguments $\mathbf{x}_{k+1}$ and $\mathbf{f}_k$, we may assume that the decomposition (9) is symmetric as well, i.e., $\mathcal{F}_{1r} = \mathcal{F}_{2r}$ for $r = 1, \ldots, R$.

3.1. Gaussian Transient PDF

This section demonstrates the decomposition for the Gaussian transient PDF. For convenience, consider first the scalar case, i.e, $n_x = 1$, and process noise variance $\mathrm{var}[w_k] = \sigma^2 = 1$

$$p(x_{k+1}|x_k) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}(x_{k+1} - f_k)^2} \, . \quad (10)$$

Consider a region Ω for the decomposition computation be given by the ranges $x_{k+1} \in [-L, L]$, $f_k \in [-L, L]$, $L = 10$, with granularity $1/10$ to obtain a $D \times D$ grid Ψ for the region Ω quantization, $D = 201$. Then, the transient PDF (10) is evaluated at the grid Ψ points, which results in the matrix $\mathbf{M} \in \mathbb{R}^{D \times D}$. This matrix is subject to a NNMF $\mathbf{M} = \mathbf{WH}$ or, better, a symmetric NNMF $\mathbf{M} = \mathbf{WW}^T$, where $\mathbf{W} \in \mathbb{R}^{D \times R}$ and $\mathbf{H} \in \mathbb{R}^{R \times D}$ are matrices with non-negative elements and R is the rank. Here, due to the symmetricity, the algorithm of [8] is utilized.

A suitable rank R of the approximation $\mathbf{WW}^T$ can be deduced from the number of leading eigenvalues of $\mathbf{M}$. Note that the matrix rank is a lower bound on the *non-negative rank* of the matrix, in general (the non-negative rank can be sometimes higher). For (10), the original matrix $\mathbf{M}$ and 40 of its largest eigenvalues are plotted in Fig. 1. The eigenvalues suggest that a good approximation could be obtained for $R \geq 20$.

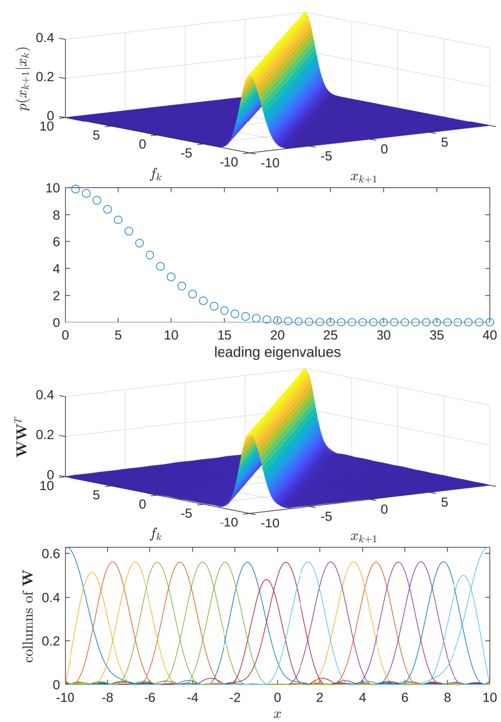

Fig. 1. Original transient PDF, leading eigenvalues of $\mathbf{M}$, an approximate transient PDF for $R = 20$, and columns of $\mathbf{W}$ for $R = 20$ obtained by NNMF.

The curves depicting columns of $\mathbf{W}$ have the Gaussian bell-curve shapes, roughly uniformly distributed[4] within the interval $[-L, L]$. Hence, the columns of $\mathbf{W}$ will be modeled by the functions $\mathcal{F}_{1r} = \mathcal{F}_{2r} = \mathcal{F}_r$ as

$$\mathcal{F}_r(x) = h \cdot e^{-\frac{1}{2}(x - m_r)^2 / w^2} \quad (11)$$

parameterized by the peak position m_r, width w, and height h. Given the range $2L$, rank R, and uniformity of the peak distribution, the distance of the adjacent peak positions will be fixed to $m_r - m_{r-1} = d$, $r = 2, \ldots, R$. The other two parameters w and h will be obtained by a numerical optimization to minimize the average absolute error between the true transient PDF and its approximation (10). The results of the optimization are listed in Table 1. From the table it follows

[4] A minor difference can be spotted in the middle, and close to the margins. The margins are rather unimportant in this problem; they are related to the ranges of x_{k+1} and f_k, which can be easily adjusted to make sure that the values x_{k+1} and f_k will remain in this range.

R	d	w	h	E
10	2	0.8701	0.6411	1.72×10^{-2}
15	1.33	0.7326	0.6266	4.70×10^{-3}
20	1.05	0.7088	0.5768	9.6077×10^{-4}
25	0.82	0.7071	0.5109	5.5380×10^{-5}
30	0.68	0.7071	0.4653	3.8459×10^{-6}
35	0.58	0.7071	0.4297	3.9440×10^{-6}
40	0.52	0.7071	0.4068	4.2067×10^{-6}

Table 1. Rank, peak distances, widths, heights, and the mean absolute error of the true and approximate transient PDF decomposition obtained by numerical optimization.

that the error E decreases with increasing rank. However, for $R > 30$ the shape of the curves deviates significantly from the Gaussian bell-curve shape, and the error does not decrease any more. The result (11) can be generalized for arbitrary variance σ^2 and arbitrary range L. Then, the appropriate rank R', distance of the adjacent peak positions d', and the height h' and width w' of the terms in (11) would be

$$R' = \left\lfloor \frac{2L}{\sigma} \right\rfloor, \quad d' = d\sigma, \quad h' = \frac{h}{\sqrt{\sigma}}, \quad w' = w\sigma. \qquad (12)$$

3.2. Gaussian Transient PDF in higher dimension

In this section, the assumption of a scalar state is dropped and $n_x > 1$ is assumed. Then, for a zero-mean noise $\mathbf{w}_k$ with covariance matrix $\mathsf{cov}[\mathbf{w}_k] = \mathbf{Q}$

$$p(\mathbf{x}_{k+1}|\mathbf{x}_k) = \frac{1}{\sqrt{(2\pi)^{n_x}|\mathbf{Q}|}} e^{-\frac{1}{2}(\mathbf{x}_{k+1}-\mathbf{f}_k)^T\mathbf{Q}^{-1}(\mathbf{x}_{k+1}-\mathbf{f}_k)},$$

where $\mathbf{f}_k = \mathbf{f}_k(\mathbf{x}_k, \mathbf{u}_k)$. Now, if $\mathbf{Q}$ is a diagonal matrix with elements Q^i, $i = 1, \ldots, n_x$ on its diagonal, then

$$p(\mathbf{x}_{k+1}|\mathbf{x}_k) = \prod_{i=1}^{n_x} \frac{1}{\sqrt{(2\pi)Q^i}} e^{-\frac{1}{2}(\mathbf{x}^i_{k+1}-\mathbf{f}^i_k)^2/Q^i}, \qquad (13)$$

where $\mathbf{x}^i_{k+1}$ and $\mathbf{f}^i_k$ are i-th elements of $\mathbf{x}_{k+1}$ and $\mathbf{f}_k$, respectively. From (13) it follows that for the Gaussian process noise $\mathbf{w}_k$ with uncorrelated elements it is possible obtain the decomposition of the transient PDF as a product of the decompositions introduced in the previous section for the scalar case applied to each element $\mathbf{x}^i_{k+1}$ and $\mathbf{f}^i_k$, $i = 1, \ldots, n_x$.

For the process noise with generally correlated elements, one could resort to finding a decomposition of the form (9). Alternatively the variables $\mathbf{x}_{k+1}$ and $\mathbf{f}_k$ can be transformed to achieve diagonal covariance matrix.

4. NUMERICAL ILLUSTRATION

Performance of the PMF with the proposed transient PDF decomposition is illustrated using a terrain-aided navigation scenario [12]. Let a state-space model (1), (2) be considered where $T = 200$, $\mathbf{f}_k(\mathbf{x}_k, \mathbf{u}_k) = \mathbf{x}_k + \mathbf{u}_k$, $\mathbf{x}_k$ is a

Table 2. PMF estimation performance.

N		$\mathrm{PMF}_{\mathrm{STD}}$	$\mathrm{PMF}_{\mathrm{D}}$		
			$R = 10$	$R = 20$	$R = 25$
20^2	IE	30×10^{-3}	130×10^{-3}	31×10^{-3}	30×10^{-3}
	τ	96×10^{-4}	12×10^{-4}	13×10^{-4}	14×10^{-4}
50^2	IE	49×10^{-4}	1147×10^{-4}	62×10^{-4}	56×10^{-4}
	τ	229×10^{-3}	31×10^{-3}	33×10^{-3}	33×10^{-3}

two-dimensional state vector describing the vehicle horizontal position in north and east directions, $\mathbf{u}_k = [300, 300]^T$ is an available shift vector provided, e.g., by the inertial navigation system or odometer, and the state noise characterizing uncertainty in the shift vector is described as[5] $p(\mathbf{w}_k) = \mathcal{N}\{\mathbf{w}_k; [0\,0]^T, 10^2\mathbf{I}_2\}$. The measurement z_k is the terrain altitude below the vehicle[6] and $h_k(\cdot)$ denotes a terrain map connecting the sought horizontal position and the available altitude. The measurement noise v_k includes sensor reading uncertainty and map error and is described by $p(v_k) = \mathcal{N}\{v_k; 0, 8^2\}$. Performance of the following PMF algorithms was analyzed:

- $\mathrm{PMF}_{\mathrm{TRUE}}$ with a high number of grid points $N = 150^2$ providing "almost true" state estimate $p(\mathbf{x}_k|\mathbf{z}^k)$,
- $\mathrm{PMF}_{\mathrm{ST}}$ with the *standard* convolution computation $N = 20^2$ and $N = 50^2$ providing an approximate conditional PDF $\hat{p}_{\mathrm{ST}}(\mathbf{x}_k|\mathbf{z}^k; \boldsymbol{\xi}_k)$,
- $\mathrm{PMF}_{\mathrm{D}}$ with the convolution involving the *proposed* transient PDF decomposition with $N = 20^2$ and $N = 50^2$ providing the approximate conditional PDF $\hat{p}_{\mathrm{D}}(\mathbf{x}_k|\mathbf{z}^k; \boldsymbol{\xi}_k)$,

using (*i*) the normalized filtering PDF integral error[7] IE $= \frac{1}{T+1}\sum_{k=0}^{T} \int \frac{1}{2}|p(\mathbf{x}_k|\mathbf{z}^k) - \hat{p}(\mathbf{x}_k|\mathbf{z}^k; \boldsymbol{\xi}_k)| d\mathbf{x}_k$ and (*ii*) convolution execution time (τ). The results can be found in Table 2. They indicate that for $R = 20$, the estimation quality of the $\mathrm{PMF}_{\mathrm{D}}$ with the proposed convolution computation is very close to that of the $\mathrm{PMF}_{\mathrm{ST}}$ with the standard convolution computation, while the computational costs are almost by an order of magnitude lower.

5. CONCLUDING REMARKS

The state estimation of nonlinear stochastic dynamic systems by the point-mass filter was treated. The paper proposed a non-negative functional decomposition of the transient density, through which the convolution in the PMF can efficiently be calculated. With an appropriate rank of the decomposition, significant computational costs savings can be achieved with only negligible loss of PMF estimate quality.

[5] Notation $\mathbf{I}_2$ stands for the identity matrix of the indicated dimension.

[6] Terrain altitude can be based on the barometric altimeter, radar altimeter, or their combination depending on the type of vehicle.

[7] The PDF integral error is normalized so that IE$\in [0, 1]$.

6. REFERENCES

[1] A. Doucet, N. De Freitas, and N. Gordon, Eds., *Sequential Monte Carlo Methods in Practice*, Springer, 2001, (Ed. Doucet A., de Freitas N., and Gordon N.).

[2] M. Šimandl, J. Královec, and T. Söderström, "Anticipative grid design in point-mass approach to nonlinear state estimation," *IEEE Transactions on Automatic Control*, vol. 47, no. 4, 2002.

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

[4] M. Šimandl, J. Královec, and T. Söderström, "Advanced point-mass method for nonlinear state estimation," *Automatica*, vol. 42, no. 7, pp. 1133–1145, 2006.

[5] J. Matoušek, J. Duník, and O. Straka, "Point-mass filter: Density specific grid design and implementation," in *15th European Workshop on Advanced Control and Diagnosis*, Bologna, Italy, Nov. 2019.

[6] Ji-Cheng Chen, "The nonnegative rank factorizations of nonnegative matrices," *Linear Algebra and its Applications*, vol. 62, pp. 207–217, 1984.

[7] P. Paatero and U. Tapper, "Positive matrix factorization: A non-negative factor model with optimal utilization of error estimates of data values," *Environmetrics*, vol. 5, no. 2, pp. 111–126, 1994.

[8] K. Huang, N. D. Sidiropoulos, and A. Swami, "Nonnegative matrix factorization revisited: Uniqueness and algorithm for symmetric decomposition," *IEEE Transactions on Signal Processing*, vol. 62, no. 1, pp. 211–224, 2014.

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

[10] 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.

[11] A. A. Gorodetsky, S. Karaman, and Y. M. Marzouk, "A continuous analogue of the tensor-train decomposition," arXiv:1510.09088 [math.NA], 2018.

[12] N. Bergman, *Recursive Bayesian Estimation: Navigation and Tracking Applications*, Ph.D. thesis, Linköping University, Sweden, 1999.