Density Difference Grid Design in a Point-Mass Filter

Jakub Matoušek [†], Jindř ich Duník *,[†] and **Ondřej Straka [†]**

Department of Cybernetics, University of West Bohemia, Univerzitní 8, 306 14 Pilsen, Czech Republic;
matoujak@kky.zcu.cz (J.M.); straka30@kky.zcu.cz (O.S.)
* Correspondence: dunikj@kky.zcu.cz; Tel.: +420-377-632-554
† These authors contributed equally to this work.

Abstract: The paper deals with the Bayesian state estimation of nonlinear stochastic dynamic systems. The stress is laid on the point-mass filter, solving the Bayesian recursive relations for the state estimate conditional density computation using the deterministic grid-based numerical integration method. In particular, the grid design is discussed and the novel density difference grid is proposed. The proposed grid design covers such regions of the state-space where the conditional density is significantly spatially varying, by the dense grid. In other regions, a sparse grid is used to keep the computational complexity low. The proposed grid design is thoroughly discussed, analyzed, and illustrated in a numerical study.

Keywords: state estimation; nonlinear filtering; nonlinear system; point-mass method

1. Introduction

State estimation of nonlinear discrete-time stochastic dynamic systems from noisy or incomplete measurements has been a subject of considerable research interest. The topic plays an important role in signal processing, control, and fault detection of both small and large scale systems in virtually all fields of industry [1] including power generation and transmission, battery live prediction [2–8], automotive [9,10], navigation [11,12], speech and image processing etc. State estimation methods has also become important in non-technical fields including, for example, economics, biology, or weather forecasting.

In the Bayesian framework, a general solution to the state estimation problem is given by the Bayesian recursive relations (BRRs), which compute the probability density functions (PDFs) of the state conditioned on the measurements. The conditional PDFs fully describe the estimate of the immeasurable state of a nonlinear or non-Gaussian stochastic dynamic system. Besides a set of linear models leading e.g., to the Kalman filter (KF), the BRRs are not analytically tractable and an approximation in their solution must be used. The resulting approximate methods can be divided with respect to the validity of the resulting estimates into global and local filters [13,14].

The local filters provide computationally efficient estimates in the form of the conditional mean and covariance matrix with potentially limited performance due to inherent underlying Gaussian assumption, which is not always realistic (as the first two moments usually do not represent a full description of the immeasurable state) . These filters are represented by the extended Kalman filter, the unscented Kalman filter, or the stochastic integration filter [14–18].

As opposed to the local filters, the global filters provide estimates in the form of conditional PDFs without any assumption on the conditional distribution family. The global filters are capable of estimating the state of a strongly nonlinear or non-Gaussian system but usually at the cost of higher computational demands. Among these, the Gaussian sum filter [19], the particle filter [20], and the point-mass filter [21–24] have attracted a considerable attention.

The point-mass filter (PMF) considered in this paper is based on a numerical solution to the BRRs using deterministic grid-based numerical integration rules, which means the BRRs are solved and the conditional PDF is computed at the grid points only. Although a conceptual design of the PMF is known from the seventies [13], an application in the terrain-aided navigation (TAN) [22] is reported as late as the nineties mainly because of the computational complexity of the filter. Therefore, since the development of the PMF, extensive attention is devoted to the efficient filter design reducing the computational complexity. In the literature, two main directions of reduction can be found;

- Design of the Rao-Blackwellized (or marginalized) PMF [23,25],
- Efficient convolution in the PMF prediction step [21,22,24,26–29].

The former approach is based on the decomposition of the state vector into two parts; nonlinearly and linearly modeled. As a consequence, just the nonlinearly modeled part of the state is estimated by the computationally expensive PMF and the remaining part of the state vector is estimated by a set of computationally cheap KFs. The Rao–Blackwellized PMF is thus suitable for possibly high-dimensional models with specific conditionally linear structure and Gaussian noises. The latter approach typically does not restrict the class of the allowed state-space models, but it rather focuses on computationally effective implementation of the convolution, as the most demanding step, in the prediction of the PMF.

Recently, a novel method for the efficient convolution, based on the the density-specific grid (DSG) design, was proposed [24]. Compared to other methods (typically assuming one equidistant grid), the DSG assumes two different grids (each with different properties);

- Denser grid supporting symmetrical vicinity of the conditional PDF mean (where a higher volume of the conditional PDF is expected),
- Sparser grid supporting conditional PDF tail region (where a marginal volume of the conditional PDF is expected).

The DSG brings significant improvement to the estimation quality especially for unimodal and rather symmetric conditional PDFs with light tails. If the density is multimodal, heavy-tailed, or heavy-skewed, the benefit of the DSG is marginal.

The goal of the paper is, thus, to design an advanced version of the DSG, further denoted as the density difference grid (DDG). The proposed DDG is based on the differentiation of the conditional PDF at the sparse grid, which allows accurate determination of the regions with a significant spatial change of the PDF. In these regions, the denser grid is created to reasonably well capture the PDF shape. As a consequence, the DDG has an ability to better approximate asymmetric, heavy-tailed, and multi-modal PDFs, i.e., the PDFs with dominant higher-order moments, with a negligible increase of the overall PMF computational complexity.

The rest of the paper is organized as follows. In Section 2, the system description and a brief introduction to the Bayesian state estimation is presented. Then, in Section 3, the points-mass filter is overviewed. This overview is accompanied with a short recapitulation of the state-of-the-art grid designs. Section 4 proposes a novel DDG design and Section 5 shows the DDG performance in a numerical study. Finally, concluding remarks are drawn in Section 6.

2. System Description and Bayesian State Estimation

Let the following discrete-time state-space model of a nonlinear stochastic dynamic system

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

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

be 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 known general 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 $p(\mathbf{x}_0)$.

The goal of the state estimation, particularly of the filtering, is to find an estimate 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. In the Bayesian framework, the state estimate in the form of the conditional PDF $p(\mathbf{x}_k|\mathbf{z}^k), \forall k$, is sought. Note that, considering the model (1), (2), the state transition PDF and the BRRs (3)–(5) 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 $p(\mathbf{x}_{k+1}|\mathbf{z}^k) = p(\mathbf{x}_{k+1}|\mathbf{z}^k; \mathbf{u}^k)$.

The Bayesian state estimator design stems from the solution of the Bayesian recursive relations (BRRs) for the recursive computation of the conditional PDFs [15]

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

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

where $p(\mathbf{x}_k|\mathbf{z}^k)$ is the filtering PDF computed by the Bayes' rule (3) and $p(\mathbf{x}_k|\mathbf{z}^{k-1})$ is the one-step predictive PDF computed by the Chapman–Kolmogorov Equation (4). 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 system dynamics (1) and the measurement PDF obtained from measurement Equation (2), respectively. The PDF

$$p(\mathbf{z}_k|\mathbf{z}^{k-1}) = \int p(\mathbf{x}_k|\mathbf{z}^{k-1})p(\mathbf{z}_k|\mathbf{x}_k)d\mathbf{x}_k \tag{5}$$

is the one-step predictive PDF of the measurement. The recursion (3), (4) starts from the initial PDF, i.e., $p(\mathbf{x}_0|\mathbf{z}^{-1}) = p(\mathbf{x}_0)$.

3. Point-Mass Filter Summary and Grid Designs

The PMF is based on the numerical solution to BRRs using the deterministic integration rules. As a consequence, a significant part of the (continuous) state-space is approximated by the grid of isolated points in which the conditional PDFs are computed. The PDF evaluated at the grid points is denoted as the point-mass PDF or, alternatively, as the point-mass density (PMD).

3.1. Point-Mass PDF Approximation

Approximation of a conditional PDF $p(\mathbf{x}_k|\mathbf{z}^m)$, where $m = k$ for the filtering PDF (3) and $m = k-1$ for the predictive PDF (4), by a piece-wise constant PMD $\hat{p}(\mathbf{x}_k|\mathbf{z}^m; \boldsymbol{\xi}_k)$ defined at the set of the discrete grid points $\boldsymbol{\xi}_k = \{\boldsymbol{\xi}_k^{(i)}\}_{i=1}^N, \boldsymbol{\xi}_k^{(i)} \in \mathbb{R}^{n_x}$, is as follows

$$p(\mathbf{x}_k|\mathbf{z}^m) \approx \hat{p}(\mathbf{x}_k|\mathbf{z}^m; \boldsymbol{\xi}_k) \triangleq \sum_{i=1}^N P_{k|m}(\boldsymbol{\xi}_k^{(i)})S\{\mathbf{x}_k; \boldsymbol{\xi}_k^{(i)}, \boldsymbol{\Delta}_k\}, \tag{6}$$

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 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 normalization constant, and δ_k is a volume of the i-th point neighbourhood defined below,
- $\boldsymbol{\Delta}_k = [\boldsymbol{\Delta}_k(1), \boldsymbol{\Delta}_k(2), \ldots, \boldsymbol{\Delta}_k(n_x)]^T$ 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

$$S\{\mathbf{x}_k; \boldsymbol{\xi}_k^{(i)}, \boldsymbol{\Delta}_k\} = 1, \tag{7}$$

if $\mathbf{x}_k(j) \in [\boldsymbol{\xi}_k^{(i)}(j) - \frac{\boldsymbol{\Delta}_k(j)}{2}, \boldsymbol{\xi}_k^{(i)}(j) + \frac{\boldsymbol{\Delta}_k(j)}{2}]$ for $j = 1, 2, \ldots, n_x$, and

$$S\{\mathbf{x}_k; \boldsymbol{\xi}_k^{(i)}, \boldsymbol{\Delta}_k\} = 0, \tag{8}$$

otherwise, so that

$$\int S\{\mathbf{x}_k; \boldsymbol{\xi}_k^{(i)}\} d\mathbf{x}_k = \boldsymbol{\Delta}_k(1) \times \boldsymbol{\Delta}_k(2) \times \ldots \times \boldsymbol{\Delta}_k(n_x)$$

$$= \delta_k. \tag{9}$$

Note that in (6)–(9) the notation $\mathbf{x}(j)$ meaning the *j*-th element of the vector $\mathbf{x}$ was used. Illustration of the point-mass PDF approximation (6) with omitted time indices is shown in Figure 1. The point-mass density (6) can be, thus, interpreted as a sum of weighted uniform distributions. Another interpretations, such as the a sum of the Dirac delta functions, can be found in [22,26].
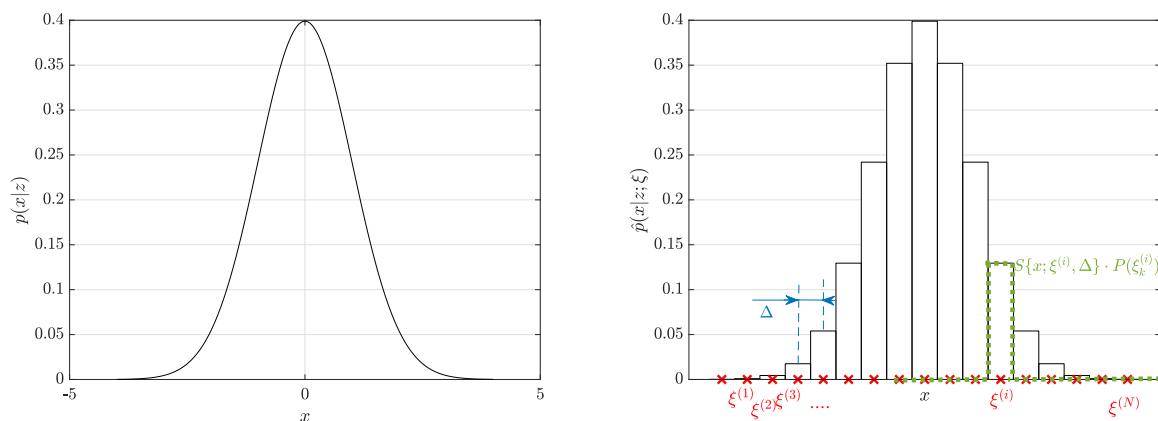

Figure 1. Illustration of point-mass probability density function (PDF) approximation (grid points—red; grid point neighbourhood—blue; scaled selection function—green).

3.2. Point-Mass Filter Algorithm

The basic algorithm of the point-mass filter (PMF) can be summarised by the following Algorithm 1 [26,30]:

Algorithm 1: Point-Mass Filter

---

1. **Initialization**: Set $k = 0$, construct the initial grid of points $\{\boldsymbol{\xi}_0^{(i)}\}_{i=0}^N$, and define the initial point-mass PDF

$$\hat{p}(\mathbf{x}_0|\mathbf{z}^{-1};\boldsymbol{\xi}_0) = \sum_{i=1}^N P_{0|-1}(\boldsymbol{\xi}_0^{(i)}) S\{\mathbf{x}_0;\boldsymbol{\xi}_0^{(i)},\boldsymbol{\Delta}_0\}, \tag{10}$$

approximating the initial PDF $p(\mathbf{x}_0|\mathbf{z}^{-1})$;

2. **Measurement update**: Compute the filtering point-mass PDF

$$\hat{p}(\mathbf{x}_k|\mathbf{z}^k;\boldsymbol{\xi}_k) = \sum_{i=1}^N P_{k|k}(\boldsymbol{\xi}_k^{(i)}) S\{\mathbf{x}_k;\boldsymbol{\xi}_k^{(i)},\boldsymbol{\Delta}_k\}, \tag{11}$$

where

$$P_{k|k}(\boldsymbol{\xi}_k^{(i)}) = \frac{p(\mathbf{z}_k|\mathbf{x}_k=\boldsymbol{\xi}_k^{(i)})P_{k|k-1}(\boldsymbol{\xi}_k^{(i)})}{\sum_{i=1}^N p(\mathbf{z}_k|\mathbf{x}_k=\boldsymbol{\xi}_k^{(i)})P_{k|k-1}(\boldsymbol{\xi}_k^{(i)})\delta_k}. \tag{12}$$

is the value of the filtering PDF at i-th grid point;

3. **Grid construction:** Based on the filtering estimate (11) and the state Equation (1), construct the new grid of points $\{\boldsymbol{\xi}_{k+1}^{(j)}\}_{j=1}^N$;

4. **Time update**: Compute the predictive point-mass PDF at the new grid of points according to

$$\hat{p}(\mathbf{x}_{k+1}|\mathbf{z}^k;\boldsymbol{\xi}_{k+1}) = \sum_{j=1}^N P_{k+1|k}(\boldsymbol{\xi}_{k+1}^{(j)}) S\{\mathbf{x}_{k+1};\boldsymbol{\xi}_{k+1}^{(j)},\boldsymbol{\Delta}_{k+1}\}, \tag{13}$$

where the value of the predictive PDF at j-th grid point is

$$P_{k+1|k}(\boldsymbol{\xi}_{k+1}^{(j)}) = \sum_{i=1}^N p(\boldsymbol{\xi}_{k+1}^{(j)}|\mathbf{x}_k = \boldsymbol{\xi}_k^{(i)}) P_{k|k}(\boldsymbol{\xi}_k^{(i)})\delta_k. \tag{14}$$

5. Set $k = k + 1$ and go to the step 2);

For the sake of clarity, a detailed derivation of the PMF can be found in Appendix A. Note that the PMF filter provides estimates in the form of the PMD $\hat{p}(\mathbf{x}_k|\mathbf{z}^m;\boldsymbol{\xi}_k)$ approximating the conditional PDF $p(\mathbf{x}_k|\mathbf{z}^m)$. Assuming small grid point vicinity $\boldsymbol{\Delta}_k$, the conditional moments, e.g., the mean and covariance matrix, can be easily computed according to [23]

$$\hat{\mathbf{x}}_{\mathrm{PMD},k|m} = \sum_{i=1}^N \delta_k P_{k|m}(\boldsymbol{\xi}_k^{(i)})\boldsymbol{\xi}_k^{(i)} \approx \hat{\mathbf{x}}_{k|m} = \mathsf{E}[\mathbf{x}_k|\mathbf{z}^m], \tag{15}$$

$$\mathbf{P}_{\mathrm{PMD},k|m} = \sum_{i=1}^N \delta_k P_{k|m}(\boldsymbol{\xi}_k^{(i)}) \left(\boldsymbol{\xi}_k^{(i)} - \hat{\mathbf{x}}_{k|m}\right)\left(\boldsymbol{\xi}_k^{(i)} - \hat{\mathbf{x}}_{k|m}\right)^T \approx \mathbf{P}_{k|m} = \mathrm{cov}[\mathbf{x}_k|\mathbf{z}^m], \tag{16}$$

when required. The moments are, however, generally not needed for the PMF run.

3.3. Grid Design

A crucial part of the PMF design that is significantly affecting the filter estimation performance is a specification of the grid points before time update step (13) and, by extension, in the initialization (10).

When designing a grid, the number of grid points should be decided first. The more grid points N are considered, the better the accuracy of the density approximation can be achieved but at the

cost of higher computational demands. Because of the convolution (13), (14), the computational complexity grows exponentially with N. Thus, the number of grid points is solely driven by the available computational capacity and required precision of the estimate (an analysis of ideal grid resolution can be found in [31]).

The grid should be designed to cover a significant region of the conditional PDFs support sufficiently well. Therefore, it is necessary to specify two grid characteristics

1. Region of the PDF support to be covered by the grid points,
2. Locations of grid points inside the region.

The former characteristic specification is common for all single-grid design techniques. The region $\mathcal{R}$ for the predictive PDF $p(\mathbf{x}_{k+1}|\mathbf{z}^k)$ is (hyper-)rectangular and centred around the mean $\hat{\mathbf{x}}_{k+1|k} = \mathsf{E}[\mathbf{x}_{k+1}|\mathbf{z}^k]$ and is selected so that

$$\int_{\mathcal{R}} p(\mathbf{x}_{k+1}|\mathbf{z}^k) > P_{\text{thr}}, \tag{17}$$

where P_{thr} is a user-defined threshold. Unfortunately, except for the initial condition, neither the predictive moments nor the predictive PDF is known at "Grid construction" step of Algorithm 1 (the PDF is to be computed in the subsequent step "Time update"). A reasonable approach is to compute the first two (approximate) predictive moments of the state

$$\hat{\mathbf{x}}_{\text{A},k+1|k} = \mathsf{E}[\mathbf{x}_{k+1}|\mathbf{z}^k], \tag{18}$$

$$\mathbf{P}_{\text{A},k+1|k} = \text{cov}[\mathbf{x}_{k+1}|\mathbf{z}^k], \tag{19}$$

and to construct an approximate Gaussian predictive PDF

$$p_{\text{A}}(\mathbf{x}_{k+1}|\mathbf{z}^k) = \mathcal{N}\{\mathbf{x}_{k+1}; \hat{\mathbf{x}}_{\text{A},k+1|k}, \mathbf{P}_{\text{A},k+1|k}\}, \tag{20}$$

where the notation $\mathcal{N}\{\mathbf{x}; \hat{\mathbf{x}}, \mathbf{P}\}$ stands for the Gaussian PDF of the random variable $\mathbf{x}$ with the mean $\hat{\mathbf{x}}$ and the covariance matrix $\mathbf{P}$ (for simplicity, the random variable has same notation as its realization). Then, the sought region $\mathcal{R}$ can be found analogously to (17), where the approximate predictive PDF $p_{\text{A}}(\mathbf{x}_{k+1}|\mathbf{z}^k)$ (20) is used instead of the exact one $p(\mathbf{x}_{k+1}|\mathbf{z}^k)$. The predictive moments (18) and (19) are computed on the basis of known filtering moments (given by (15), (16), and (11)) and state Equation (1). If the dynamics is nonlinear, approximate predictive moments can be computed e.g., by the unscented transformation or various numerical integration rules [16,32]. Note that, the approximate PDF (20) is used for the region $\mathcal{R}$ determination only.

The latter characteristic is specific for each grid design technique.

3.3.1. Equidistant Grid Design

The equidistant grid design [22] is a standard grid design technique, which covers the region $\mathcal{R}$ with equidistantly located points, i.e., the neighbourhood $\mathbf{\Delta}_k$ is the same for all grid points.

3.3.2. Density Specific Grid Design

The DSG design [24], is based on an unequally spaced placement of the fixed number of points over the PDF support $\mathcal{R}$ to be approximated. The main idea is to approximate the PDF, where the volume of the PDF is significant or significantly varying by a denser grid (as in this area, a slope of the PDF is expected to be significant) and the tail of the PDF by a sparser grid (as the tail is flat). Thus, the DSG design determines, besides the region $\mathcal{R}$, also the hyper-rectangular subregion $\mathcal{R}_{\text{in}}$, where the grid is denser. The determination of the subregion again depends on the computed approximate predictive moments (18), (19).

As can be seen in [24], the DSG design significantly outperforms the standard equidistant grid design if the underlying conditional PDF is unimodal, rather symmetric, and light-tailed. If any of the mentioned properties are violated, the DSG benefit is marginal.

3.4. Goal of the Paper

The goal of the paper is to propose a novel unequally spaced grid design, called the density difference grid design (further abbreviated as DDG), which provides improved PMF estimation performance independently of the conditional PDF properties.

4. Density Difference Grid Design

The DDG design is based on the idea that such part of the support, where the conditional PDF is significantly spatially varying, should be covered by the dense grid (to sufficiently well capture PDF behavior). On the other hand, the remaining part of the support, i.e., the parts where the conditional PDF is almost flat, can be reasonably well approximated by the sparse grid. To realize this basic idea of the DDG, the following three steps should be performed:

1. Design of the sparse (outer) grid and evaluation of the predictive PMD at these grid points,
2. Differentiation of the "sparse" predictive PMD,
3. Design of the dense (inner) grids in the vicinity of such "sparse" grid points, which are associated with a large value of the difference.

Compared to the DSG design discussed in [24], the DDG design utilizes full information about the shape of the predictive PMD (and not only the moments) and also the denser (inner) grid need not be hyper-rectangular, which further improves the fit of the PDF and its grid support. The steps are particularized below.

4.1. Sparse Grid Design

The sparse grid design takes advantage of the PMF time-update, that the PDF value of a single point $\boldsymbol{\xi}_{k+1}^{(j)}$ can be counted independently of any other grid points. The sparse grid design is given by Algorithm 2:

Algorithm 2: DDG Sparse Grid

---

1. Define a threshold p_{thr} below which the probability at a grid point is considered to be negligible and can be neglected. The threshold setting can be related to the computation precision of a given platform;
2. State Similarly as in the DSG design [24], compute the predictive moments $\hat{x}_{\text{A},k+1|k}$ (18) and $P_{\text{A},k+1|k}$ (19) and set the number of standard deviations (STDs) n_σ of the marginal distributions of (20), which are used for specification of an a priori (hyper-)rectangular region $\mathcal{R}_{\text{A}}$ to be covered by the grid points;
3. At the boundaries of the region $\mathcal{R}_{\text{A}}$ select a small set of grid points and evaluate the predictive probability at these points. If the probability of the point is (significantly) below the threshold p_{thr}, then the particular bound can be shifted towards the mean. On the other hand, if the probability of the point is (significantly) over the threshold p_{thr}, then the particular bound can be shifted in an opposite direction from the mean. The shifted points are, then, re-evaluated;
4. When all boundary points are properly set, create the smallest possible hyper-rectangle defining $\mathcal{R}$ including all the boundary points. As a consequence, the resulting region $\mathcal{R}$ need not be symmetrical with respect to the mean;
5. Define a cardinality of the sparse grid N_{sparse} and dense grid $N_{\text{dense}} = N - N_{\text{sparse}}$. Fill the region $\mathcal{R}$ by equidistantly placed N_{sparse} hyper-rectangular grid points $\{\boldsymbol{\xi}_{\text{sparse},k+1}^{(j)}\}_{j=1}^{N_{\text{sparse}}}$ (each sparse grid point is associated with a neighbourhood $\Delta_{\text{sparse},k+1}$). A reasonable choice $N_{\text{sparse}} = \frac{1}{3}N$;

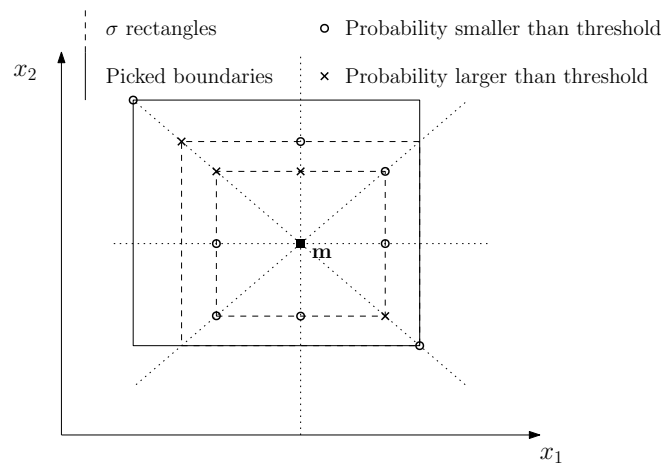This algorithm is illustrated in Figure 2 for the two dimensional state-space.

Figure 2. Illustration of search for sparse grid boundaries.

4.2. Differentiation of Sparse Grid PMD

To find the state-space regions, where the conditional predictive PDF $p(\mathbf{x}_{k+1}|\mathbf{z}^k)$ is expected to significantly spatially vary (i.e., the regions to be covered by the dense grid), the PMD evaluated at the sparse grid points is differenced according to Algorithm 3:

Algorithm 3: Sparse Grid PMD Differentiation

---

1. Compute the PMD for N_{sparse} sparse grid points according to the convolution (14);
2. For each sparse grid point $\boldsymbol{\xi}^{(j)}_{\text{sparse},k+1}$ compute the divided difference vector $\mathbf{d}_{k+1} \in \mathbb{R}^{2^{n_x}}$ [14,33] according to

$$\mathbf{d}^{(j)}_{k+1} = \left[\frac{P_{k+1|k}(\boldsymbol{\xi}^{(j)}_{\text{sparse},k+1}) - P_{k+1|k}(\boldsymbol{\xi}^{(\iota_1)}_{\text{sparse},k+1})}{\|\boldsymbol{\xi}^{(j)}_{\text{sparse},k+1} - \boldsymbol{\xi}^{(\iota_1)}_{\text{sparse},k+1}\|_2}, \dots, \frac{P_{k+1|k}(\boldsymbol{\xi}^{(j)}_{\text{sparse},k+1}) - P_{k+1|k}(\boldsymbol{\xi}^{(\iota_{2^{n_x}})}_{\text{sparse},k+1})}{\|\boldsymbol{\xi}^{(j)}_{\text{sparse},k+1} - \boldsymbol{\xi}^{(\iota_{2^{n_x}})}_{\text{sparse},k+1}\|_2} \right], \quad (21)$$

   where $\|\cdot\|_2$ denotes the Euclidean norm and the index vector $\iota = [\iota_1, \iota_2, \dots, \iota_{2^{n_x}}]$ is selected so that the closest neighbouring points of the j-th point $\boldsymbol{\xi}^{(j)}_{\text{sparse},k+1}$ are chosen. It means two neighbouring points are selected for $n_x = 1$, four neighbouring points for $n_x = 2$, eight neighbouring points for $n_x = 3$, etc.;
3. Compute the grid points rating as a norm of the vector $\mathbf{d}_{k+1} \in \mathbb{R}^{2^{n_x}}$ (21) according to

$$d^{(j)}_{k+1} = \|\mathbf{d}^{(j)}_{k+1}\|_2, \forall j, \quad (22)$$

   which characterizes the overall spatial variability of the PDF in the vicinity of the j-th sparse grid point $\boldsymbol{\xi}^{(j)}_{\text{sparse},k+1}$.

This algorithm is illustrated in Figure 3 for the two dimensional state-space.
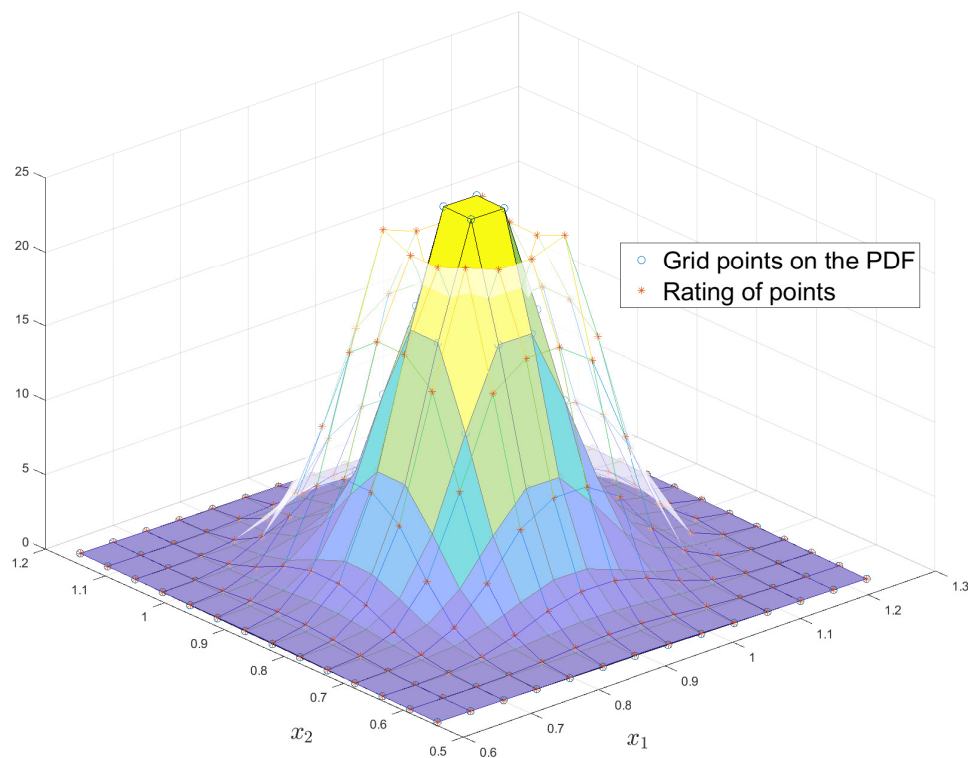
Figure 3. Illustration of sparse points difference ratings.

4.3. Design of Dense Grid

Having defined sparse grid points (Algorithm 2) and computed their difference rating (Algorithm 3), which in fact precisely denotes the state-space regions with significant change or variability of the conditional PDF, it is possible to construct the dense grid according to Algorithm 4:

Algorithm 4: Dense Grid Design

---

1. Sort the sparse grid points $\{\boldsymbol{\xi}_{\mathrm{sparse},k+1}^{(j)}\}_{j=1}^{N_{\mathrm{sparse}}}$ in a descending order according to their rating $d_{k+1}^{(j)}$ (22);
2. Set the splitting ratio r (similarly as in the DSG design in [24]), which determines number of dense grid points r^{n_x} used for covering the vicinity of the j-th sparse grid $\boldsymbol{\xi}_{\mathrm{sparse},k+1}^{(j)}$. Reasonable choice is $r = \{2, 3\}$, but note that, the ratio need not be constant and it can depend on the rating $d_{k+1}^{(j)}$ (22); the higher rating, the higher ratio;
3. Recursively split the vicinity of the ordered grid points until the number of the dense grid points N_{dense} is reached;

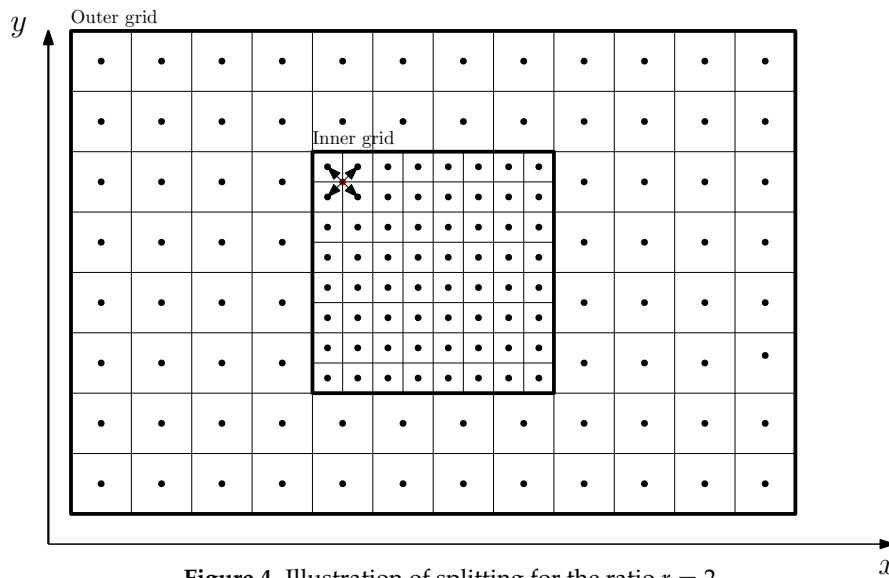This algorithm is illustrated in Figure 4 for the two dimensional state-space.

Figure 4. Illustration of splitting for the ratio $r = 2$.

4.4. Summary of DDG Design and Properties

The DDG design is given by the three steps realized by Algorithms 2–4. An example of the final DDG for the PMD shown in Figure 3 can be found in Figure 5.

Compared to the recently proposed DSG, the DDG dense (inner) grid need not be rectangular and, thus, it more realistically covers the support of the conditional PDF with the significant volume without any requirement on computation and utilization of a set of higher-order moments. The DDG design also reduces the user interaction and minimizes the number of the user-defined parameters (compared to the standard grid design, the DDG design requires specification of three additional parameters only; the number of sparse grid points N_{sparse} and the threshold p_{thr} in Algorithm 2 and splitting ration r in Algorithm 4). On the other, the DDG design is slightly more computationally intensive as illustrated in the following section.
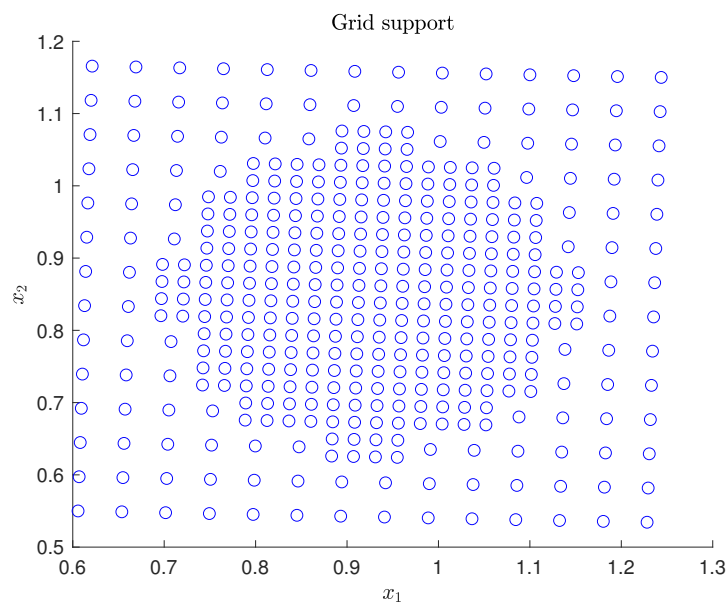

Figure 5. Example of grid support.

5. Numerical Illustration

The influence of the grid design is illustrated in two examples; static point-mass approximation and dynamic state estimation using a PMF with three different grid designs:

- Standard equidistant (EQ) grid design ([22], Section 3.3.1),
- DSG design ([24], Section 3.3.2),
- Proposed DDG design (Section 4).

5.1. PDF Point-Mass Approximation Quality

In the first example, let a true two-dimensional, i.e., $n_x = 2$, Gaussian mixture PDF with four components

$$p(\mathbf{x}) = \sum_{g=1}^{4} \omega_g \mathcal{N}\{\mathbf{x}; \hat{\mathbf{x}}_g, \mathbf{P}_g\} \tag{23}$$

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

$$\begin{bmatrix} \omega_1 & \omega_2 & \omega_3 & \omega_4 \end{bmatrix} = \begin{bmatrix} 0.1 & 0.3 & 0.4 & 0.2 \end{bmatrix}, \tag{24}$$

$$\begin{bmatrix} \hat{\mathbf{x}}_1 & \hat{\mathbf{x}}_2 & \hat{\mathbf{x}}_3 & \hat{\mathbf{x}}_4 \end{bmatrix} = \begin{bmatrix} 1 & 2 & 4 & 5 \\ 4 & 1 & 3.5 & -3.5 \end{bmatrix}, \tag{25}$$

$$\mathbf{P}_1 = \mathbf{P}_2 = \mathbf{P}_3 = \mathbf{P}_4 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}. \tag{26}$$

The PDF is illustrated in Figure 6. The PDF (23) is approximated by the PMD $\hat{p}(\mathbf{x}; \boldsymbol{\xi})$ (6) with the number of grid points $N = \{50, 200, 500\}$ for all three considered grid designs with $\mathcal{R}$ covering three STDs. The DSG and DDG designs used such user-defined parameters, which provided the best performance, namely the PMD with the DSG used $\mathcal{R}_{in}$ covering two STDs and the PMD with the DDG used parameters $r = 2$ and $N_{\text{sparse}} = 0.4N$.
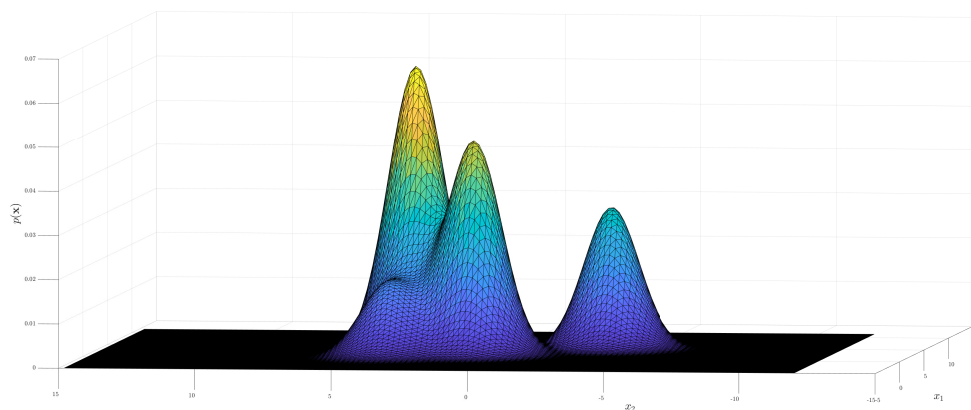

Figure 6. Static example: illustration of Gaussian mixture PDF.

The quality of the PMD is evaluated using the criteria

- Integral sum error: ISE $= 1 - \int \hat{p}(\mathbf{x}; \boldsymbol{\xi}) d\mathbf{x}$,
- Mean error: ME $= \sum_{i=1}^{n_{\mathbf{x}}} |\mathbf{x}(i) - \hat{\mathbf{x}}_{\text{PMD}}(i)|$.

The results can be found in Table 1. It can be seen that the proposed DDG provides the best approximation quality among the considered grids with the same number of points. A degraded

performance of the DSG-based design for the approximation of multimodal distributions can also be noted.

For the sake of completeness, a layout of the tested grids is shown in Figure 7.

Table 1. Approximation errors of the point-mass density (PMD) with three different grid designs.

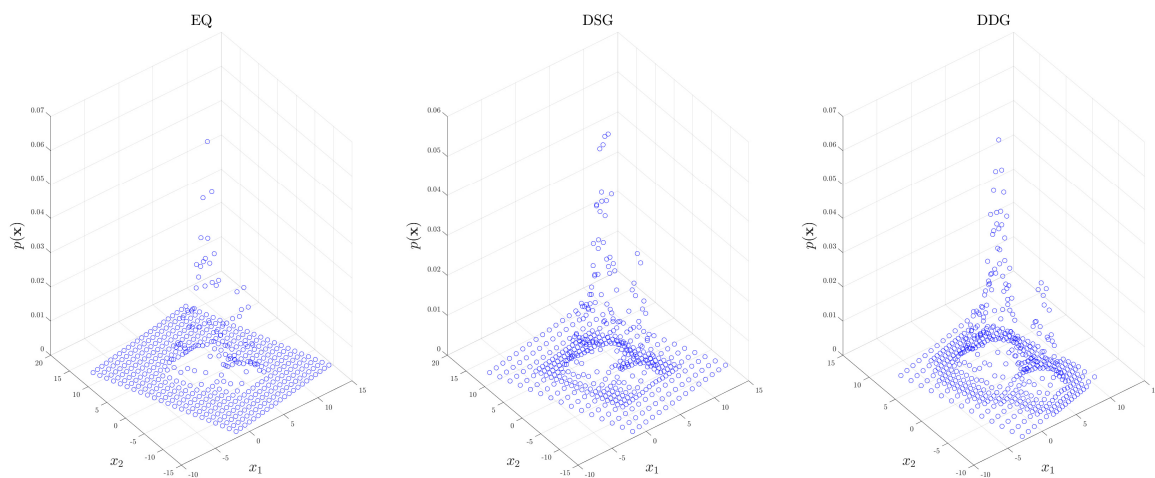|		N	50	200	500
EQ	ISE		0.0468	0.0062	5.5585×10^{-8}
	ME		1.5848	0.0606	2.4479×10^{-5}
DSG	ISE		0.0733	6.3488×10^{-5}	3.8642×10^{-5}
	ME		0.7567	6.1405×10^{-5}	3.1102×10^{-4}
DDG	ISE		1.1102×10^{-15}	2.2204×10^{-16}	2.2204×10^{-16}
	ME		0.2439	2.5314×10^{-5}	6.7043×10^{-6}

Figure 7. Static example: grids with 500 grid points.

5.2. PMF in Terrain-Aided Navigation with Spherical Coordinates

The second example considers the terrain-aided navigation (TAN) scenario. The TAN system is a navigation system suitable for GNSS-denied environments, where the vehicle horizontal position is determined by the correlation of the measured vertical terrain profile and the pre-recorded terrain map [23]. The models used in the TAN are highly nonlinear, thus, they are often used as benchmark examples for illustration of the state estimation techniques performance. In this section, the considered the TAN scenario is described by the state-space model [34]

$$\mathbf{x}_{k+1} = \begin{bmatrix} \phi_{k+1} \\ \lambda_{k+1} \end{bmatrix} = \begin{bmatrix} \phi_k + (\varphi(\lambda_{k+1}) - \varphi(\lambda_k)) \tan(K) \\ \lambda_k + \frac{\Delta T \|\mathbf{V}\|_2}{R(\lambda_k)} \cos(K) \end{bmatrix} + \mathbf{w}_k, \tag{27}$$

$$z_k = h_k(\mathbf{x}_k) + v_k, \tag{28}$$

where $k = 0, 1, 2, \ldots, 10$, ϕ_k is the sought vehicle latitude in degrees, λ_k is the longitude in degrees, $\varphi(\lambda_k) = \log\left(\frac{1+\sin(\lambda_k)}{\cos(\lambda_k)}\right)$, $\Delta T = 1$ is the sampling period in minutes, $K = 0.4363$ is the constant heading relative to the north in radians (measurement provided e.g., by the compass), $\mathbf{V} = [V_{\text{north}}, V_{\text{east}}]^T$ is the known constant velocity vector (measurement provided e.g., by an inertial navigation system) in meters/sec, and

$$R(\lambda_k) = \sqrt{\frac{\left(r_1^2 \cos(\lambda_k)\right)^2 + \left(r_2^2 \sin(\lambda_k)\right)^2}{\left(r_1 \cos(\lambda_k)\right)^2 + \left(r_2 \sin(\lambda_k)\right)^2}} \tag{29}$$

is Earth radius at latitude λ_k. Constants $r_1 = 6,378,137$ and $r_2 = 6,356,800$ are radii of the Earth at the equator and pole, respectively, in meters. The measurement function $h_k(\cdot)$ is the the map in the Geographic coordinate system. For the simulation, the maps provided by the Shuttle Radar Topography Mission (SRTM) were used (https://www2.jpl.nasa.gov/srtm/). The noises $\mathbf{w}_k$ and v_k are described as

$$p(\mathbf{w}_k) = \mathcal{N}\left\{\mathbf{w}_k; \begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 1.8 \cdot 10^{-5} & 0 \\ 0 & 1.8 \cdot 10^{-5} \end{bmatrix}\right\}, \tag{30}$$

$$p(v_k) = \mathcal{N}\{v_k; 0, 3\}. \tag{31}$$

The PMF was configured with three grid designs with $N = 100$, all with $\mathcal{R}$ covering ± 4 STDs in each axis (i.e., in latitude and longitude). Namely, the following implementations were compared

- PMF with equidistant (EQ) grid design,
- PMF with DSG grid design, where the inner denser grid covers $\mathcal{R}_{in}$ with ± 1 STD in each axis,
- PMF with the proposed DDG design with the user-defined parameters $p_{\text{thr}} = 10^{-4}$, $N_{\text{sparse}} = 50$ and $r = 3$. Thus, six sparse points with the highest rating $d_{k+1}^{(j)}$ (22) were split into $r^2 \times 6 = 54$ dense grid points (each sparse point were split into nine dense points). In total, this implementation used $N = 44 + 54 = 98$ grid points.

Performance of the filters were compared using the following criteria

- Root-mean-square deviation [*deg*]: RMSE $= \sum_{i=1}^{n_x} (\mathbf{x}_k(i) - \hat{\mathbf{x}}_{\text{PMD},k|k}(i))^2$,
- Average standard deviation [*deg*]: ASTD $= \frac{1}{n_x} \text{tr}\left(\sqrt{\text{diag}(\mathbf{P}_{\text{PMD},k|k})}\right)$,
- time [*sec*],

in $M = 100$ Monte–Carlo simulations, where the notation $\text{tr}(\mathbf{A})$ is the trace of the matrix $\mathbf{A}$ and $\text{diag}(\mathbf{A})$ stands for the diagonalization of the matrix $\mathbf{A}$, i.e., it zeroes all non-diagonal elements. The results averaged over the time are given in Table 2. It can be seen that the DDG method is better than the equidistant and DSG grid methods, while keeping similar computational complexity. The PMF estimate with the DDG is approximately about 11% more accurate. Note that transformation of the angular longitude and latitude error into the Cartesian coordinate system results in the improvement of the several meters (only by better selection of the grid points and with almost the same computational complexity).

Table 2. Estimation results.

	EQ	DSG	DDG
RMSE [*deg*]	0.015675	0.015521	0.013853
ASTD [*deg*]	0.012741	0.012466	0.012999
Time [*sec*]	0.26061	0.26131	0.26656

6. Concluding Remarks

The paper focused on the point-mass filter for the nonlinear state estimation with the stress on the grid design as a key filter design step. In particular, a new density difference grid design was proposed. The proposed design was based on evaluation of the predictive PDF in a set of the spare grid points. Subsequently, such point-mass density was differentiated and the state-space regions with significantly varying density were located. These regions are, finally, covered by the dense grid and the detailed point-mass density was completed. The density difference grid properties were analyzed, discussed, and illustrated. It was shown that the point-mass density and the point-mass filter with the proposed grid provides consistently better performance independently of the conditional probability density function shape (e.g., multimodal, heavy-tailed) than with other grid design strategies, while the

computational complexity is preserved. The performance was illustrated in two examples, including terrain-aided navigation inspired scenario.

Author Contributions: Conceptualization: J.D., J.M.; Methodology: J.D., O.S.; Investigation: J.M., J.D.; Software: J.M.; Validation: O.S., J.D.; writing: J.D.; J.M., O.S. All authors have read and agreed to the published version of the manuscript.

Conflicts of Interest: The authors declare no conflict of interest.

Abbreviations

The following abbreviations are used in this manuscript:

ASTD	Average STD
BRRs	Bayesian recursive relations
DDG	Density difference grid
DSG	Density specific grid
KF	Kalman filter
PDF	Probability density function
PF	Particle filter
PMD	Point-mass density
PMF	Point-mass filter
RMSE	Root mean square error
STD	Standard deviation

Appendix A. PMF Derivation

The PMF design is based on the numerical solution to the BRRs (3), (4) assuming the PMD form (6) of the conditional PDF. Below, the PMF design is briefly reviewed and a relation to the particle filter is provided. Further details can be found e.g., in [22,26].

Appendix A.1. Initialization

The estimation recursion starts by definition of the initial PDF $p(\mathbf{x}_0|\mathbf{z}^{-1})$ in the form of the PMD $\hat{p}(\mathbf{x}_0|\mathbf{z}^{-1}; \boldsymbol{\xi}_0)$ (6) and setting $k = 0$.

Appendix A.2. Derivation of Measurement-Update

Derivation of a PMF measurement-update starts with the Bayes' rule (3) for the posterior density computation, which can be written as

$$p(\mathbf{x}_k|\mathbf{z}^k) = \tilde{c}_k^{-1} p(\mathbf{z}_k|\mathbf{x}_k) p(\mathbf{x}_k|\mathbf{z}^{k-1}), \tag{A1}$$

where $p(\mathbf{x}_k|\mathbf{z}^{k-1})$ is the predictive density from the last step or from initialization, $\tilde{c}_k$ is a normalization constant and

$$p(\mathbf{z}_k|\mathbf{x}_k) = p_{\mathbf{v}_k}(\mathbf{z}_k - \mathbf{h}_k(\mathbf{x}_k)) \tag{A2}$$

is the PDF of a measurement, which is given by (2).

However, only the point-mass approximation of the predictive PDF $p(\mathbf{x}_k|\mathbf{z}^{k-1})$ is known, i.e., the PMD

$$\hat{p}(\mathbf{x}_k|\mathbf{z}^{k-1}; \boldsymbol{\xi}_k) = \sum_{i=1}^{N} \mathcal{P}_{k|k-1}(\boldsymbol{\xi}_k^{(i)}) S\{\mathbf{x}_k : \boldsymbol{\xi}_k^{(i)}, \boldsymbol{\Delta}_k\}. \tag{A3}$$

Substitution of the predictive PMD (A3) into (A1) can be further treated

$$\hat{p}(\mathbf{x}_k | \mathbf{z}^k; \boldsymbol{\xi}_k) \approx \hat{c}_k^{-1} p(\mathbf{z}_k | \mathbf{x}_k) \sum_{i=1}^{N} \mathcal{P}_{k|k-1}(\boldsymbol{\xi}_k^{(i)}) S\{\mathbf{x}_k : \boldsymbol{\xi}_k^{(i)}, \boldsymbol{\Delta}_k\}$$

$$\approx \sum_{i=1}^{N} \hat{c}_k^{-1} p(\mathbf{z}_k | \mathbf{x}_k = \boldsymbol{\xi}_k^{(i)}) \mathcal{P}_{k|k-1}(\boldsymbol{\xi}_k^{(i)}) \mathcal{P}_{k|k-1}(\boldsymbol{\xi}_k^{(i)}) S\{\mathbf{x}_k : \boldsymbol{\xi}_k^{(i)}, \boldsymbol{\Delta}_k\}$$

$$= \sum_{i=1}^{N} \mathcal{P}_{k|k}(\boldsymbol{\xi}_k^{(i)}) S\{\mathbf{x}_k : \boldsymbol{\xi}_k^{(i)}, \boldsymbol{\Delta}_k\}, \tag{A4}$$

where the evaluation of the filtering PDF at $\boldsymbol{\xi}_k^{(i)}$ is

$$\mathcal{P}_{k|k}(\boldsymbol{\xi}_k^{(i)}) = \hat{c}_k^{-1} p(\mathbf{z}_k | \mathbf{x}_k = \boldsymbol{\xi}_k^{(i)}) \mathcal{P}_{k|k-1}(\boldsymbol{\xi}_k^{(i)}). \tag{A5}$$

and the (point-mass) approximate normalization constant is

$$\hat{c}_k = \int \sum_{i=1}^{N} \mathcal{P}_{k|k}(\boldsymbol{\xi}_k^{(i)}) S\{\mathbf{x}_k : \boldsymbol{\xi}_k^{(i)}, \boldsymbol{\Delta}_k\} d\mathbf{x}_k$$

$$= \sum_{i=1}^{N} \mathcal{P}_{k|k}(\boldsymbol{\xi}_k^{(i)}) \int S\{\mathbf{x}_k : \boldsymbol{\xi}_k^{(i)}, \boldsymbol{\Delta}_k\} d\mathbf{x}_k$$

$$= \sum_{i=1}^{N} \mathcal{P}_{k|k}(\boldsymbol{\xi}_k^{(i)}) \delta_k. \tag{A6}$$

Note that the support grid stays the same during measurement-update.

Appendix A.3. Derivation of Time-Update

Predictive probability is given by the Chapman–Kolmogorov Equation (4), where $p(\mathbf{x}_k | \mathbf{z}^k)$ is the filtration density and

$$p(\mathbf{x}_{k+1} | \mathbf{x}_k) = p_{\mathbf{w}_k}(\mathbf{x}_{k+1} - \mathbf{f}_k(\mathbf{x}_k)) \tag{A7}$$

is a transition PDF of the state, which is given by the dynamics (1). However, only the approximate filtering PMD (A4) is known and its substitution into (4) reads

$$\hat{p}(\mathbf{x}_{k+1} | \mathbf{z}^k; \boldsymbol{\xi}_k) = \int p(\mathbf{x}_{k+1} | \mathbf{x}_k) \sum_{i=1}^{N} \mathcal{P}_{k|k}(\boldsymbol{\xi}_k^{(i)}) S\{\mathbf{x}_k : \boldsymbol{\xi}_k^{(i)}, \boldsymbol{\Delta}_k\} d\mathbf{x}_k. \tag{A8}$$

In order to solve this relation in accord to the PMF design philosophy, a new grid $\{\boldsymbol{\xi}_{k+1}^{(j)}\}_{j=1}^{N}$ is defined. This grid should cover the part of the state-space where the predictive probability is expected to be. The location and properties of the new grid can be found on the basis of the moment-based prediction as discussed in Section 3.3 or in [22,24]. Having the new grid, i.e., the grid for the time instant $k+1$, the transition PDF (A7) can be approximated by the PMD as follows

$$\hat{p}(\mathbf{x}_{k+1} | \mathbf{x}_k; \boldsymbol{\xi}_{k+1}) = \sum_{j=1}^{N} p(\boldsymbol{\xi}_{k+1}^{(j)} | \mathbf{x}_k) S\{\mathbf{x}_{k+1} : \boldsymbol{\xi}_{k+1}^{(j)}, \boldsymbol{\Delta}_{k+1}\}. \tag{A9}$$

By substitution of (A9) into (A8), the predictive PMD is of the form

$$
\hat{p}(\mathbf{x}_{k+1}|\mathbf{z}^k) \approx \int \hat{p}(\mathbf{x}_{k+1}|\mathbf{x}_k) \sum_{i=1}^{N} \mathcal{P}_{k|k}(\xi_k^{(i)}) S\{\mathbf{x}_k : \xi_k^{(i)}, \Delta_k\} d\mathbf{x}_k
$$

$$
= \int \sum_{j=1}^{N} p(\xi_{k+1}^{(j)}|\mathbf{x}_k) S\{\mathbf{x}_{k+1} : \xi_{k+1}^{(j)}, \Delta_{k+1}\} \sum_{i=1}^{N} \mathcal{P}_{k|k}(\xi_k^{(i)}) S\{\mathbf{x}_k : \xi_k^{(i)}, \Delta_k\} d\mathbf{x}_k
$$

$$
\approx \int \sum_{j=1}^{N} \sum_{i=1}^{N} \mathcal{P}_{k|k}(\xi_k^{(i)}) p(\xi_{k+1}^{(j)}|\mathbf{x}_k = \xi_k^{(i)}) S\{\mathbf{x}_k : \xi_k^{(i)}, \Delta_k\} S\{\mathbf{x}_{k+1} : \xi_{k+1}^{(j)}, \Delta_{k+1}\} d\mathbf{x}_k
$$

$$
= \sum_{i=1}^{N} \sum_{j=1}^{N} \mathcal{P}_{k|k}(\xi_k^{(i)}) p(\xi_{k+1}^{(j)}|\mathbf{x}_k = \xi_k^{(i)}) S\{\mathbf{x}_{k+1} : \xi_{k+1}^{(j)}, \Delta_{k+1}\} \underbrace{\int S\{\mathbf{x}_k : \xi_k^{(i)}, \Delta_k\} d\mathbf{x}_k}_{\delta_k}
$$

$$
= \sum_{j=1}^{N} \underbrace{\sum_{i=1}^{N} \mathcal{P}_{k|k}(\xi_k^{(i)}) p(\xi_{k+1}^{(j)}|\mathbf{x}_k = \xi_k^{(i)}) \delta_k}_{\mathcal{P}_{k+1|k}(\xi_{k+1}^{(j)})} S\{\mathbf{x}_{k+1} : \xi_{k+1}^{(j)}, \Delta_{k+1}\}
$$

$$
= \sum_{j=1}^{N} \mathcal{P}_{k+1|k}(\xi_{k+1}^{(j)}) S\{\mathbf{x}_{k+1} : \xi_{k+1}^{(j)}, \Delta_{k+1}\}, \tag{A10}
$$

where the predictive probability value at a new grid point $\xi_{k+1}^{(j)}$ is

$$
\mathcal{P}_{k+1|k}(\xi_{k+1}^{(j)}) = \sum_{i=1}^{N} p(\xi_{k+1}^{(j)}|\mathbf{x}_k = \xi_k^{(i)}) \mathcal{P}_{k|k}(\xi_k^{(i)}) \delta_k, \forall j. \tag{A11}
$$

The expression (A10), commonly denoted as the convolution, is the most computationally demanding step of the PMF. Note that the value of $\mathcal{P}_{k+1|k}(\xi_{k+1}^{(j)})$ is independent over *j*, therefore parallel programming and computing tools can be used to speed the algorithm up.

Appendix A.4. Relation of Point-Mass and Particle Filters

The point-mass filter and particle filter (PF) are based on numerical, i.e., point-based, solution to the BRRs (3), (4). Whereas, the PMF solves the BRRs using the deterministic integration rule, the PF uses the stochastic Monte-Carlo integration techniques [20]. The main difference can be, thus, found in the generation of the points. In the PMF, the points are generated in, possibly equally spaced, grid covering (approximating) a significant part of the state space and the conditional PDF is computed in the grid points only. On the other hand, in the PF, the conditional PDF is approximated by the set of samples, which means, that in parts of the state space with "higher values" of the conditional PDF, a higher number of the samples can be expected. Consequently, the PF does not include the convolution, as the PMF, and, thus, it is computationally cheaper. As an advantage of the PMF over the PF, its deterministic behavior can be mentioned. Contrary to the PMF, the PF may not provide the same estimates while repetitively processing the same data on different platforms. This property may be limiting for certain safety-critical applications [23].

References

1. Simon, D. *Optimal State Estimation: Kalman, H Infinity, and Nonlinear Approaches*; Wiley-Interscience: New York, NY, USA, 2004.
2. Abur, A.; Expósito, A. *Power System State Estimation: Theory and Implementation*; CRC Press: Boca Raton, FL, USA, 2004.
3. Hering, P.; Janeček, P.; Janeček, E. On-line Ampacity Monitoring from Phasor Measurements. In *Proceedings of the 19th IFAC World Congress*; Elsevier: Cape Town, South Africa, 2014.

4.   Guo, P.; Infield, D.   Wind Turbine Tower Vibration Modeling and Monitoring by the Nonlinear State Estimation Technique (NSET). *Energies* **2012**, *7*, 5279–5293. [CrossRef]

5.   Wang, D.; Guan, X.; Liu, T.; Gu, Y.; Shen, C.; Xu, Z. Ext. Distrib. State Estim. A Detect. Method Tolerable False Data Inject. Attacks Smart Grids. *Energies* **2014**, *7*, 1517–1538. [CrossRef]

6.   Antončič, M.; Papič, I.; Blažič, B. Robust Fast State Estim. Poorly-Obs. Low Volt. Distrib. Networks Based Kalman Filter Algorithm. *Energies* **2019**, *7*, 4457. [CrossRef]

7.   Yu, Z.; Huai, R.; Xiao, L. State- Estim. Lithium-Ion Batter. Using A Kalman Filter Based Local Linearization. *Energies* **2015**, *8*, 7854–7873. [CrossRef]

8.   Zhang, R.; Xia, B.; Li, B.; Cao, L.; Lai, Y.; Zheng, W.; Wang, H.; Wang, W. State Art Lithium-Ion Battery SOC Estim. Electr. Veh. *Energies* **2018**, *11*, 1820. [CrossRef]

9.   Simon, D. *Adaptive Filtering and Change Detection*; John Wiley & Sons: Hoboken, NJ, USA, 2000.

10. Lindfors, M.; Hendeby, G.; Gustafsson, F.; Karlsson, R. Vehicle Speed Tracking Using Chassis Vibrations. In Proceedings of the 2016 IEEE Intelligent Vehicles Symposium, Gothenburg, Sweden, 19–22 June 2016.

11. Bar-Shalom, Y.; Li, X.R.; Kirubarajan, T.   *Estimation with Applications to Tracking and Navigation: Theory Algorithms and Software*; John Wiley & Sons: Hoboken, NJ, USA, 2001.

12. Groves, P.D. *Principles of GNSS, Inertial, and Multisensor Integrated Navigation Systems*; Artech House: Norwood, MA, USA, 2008.

13. Sorenson, H.W. On the Development of Practical Nonlinear Filters. *Inf. Sci.* **1974**, *7*, 230–270. [CrossRef]

14. Šimandl, M.; Duník, J. Derivative-free Estimation Methods: New Results and Performance Analysis. *Automatica* **2009**, *45*, 1749–1757. [CrossRef]

15. Anderson, B.D.O.; Moore, J.B. *Optimal Filtering*; Prentice Hall: Englewood Cliffs, NJ, USA, 1979.

16. Julier, S.J.; Uhlmann, J.K. Unscented Filtering and Nonlinear Estimation. *IEEE Proc.* **2004**, *92*, 401–421. [CrossRef]

17. Duník, J.; Straka, O.; Šimandl, M. Stochastic Integration Filter. *IEEE Trans. Autom. Control* **2013**, *58*, 1561–1566. [CrossRef]

18. Särkkä, S. *Bayesian Filtering and Smoothing*; Cambridge University Press: Cambridge, UK, 2013.

19. Sorenson, H.W.; Alspach, D.L. Recursive Bayesian Estimation Using Gaussian Sums. *Automatica* **1971**, *7*, 465–479. [CrossRef]

20. Doucet, A.; De Freitas, N.; Gordon, N. (Eds.)   *Sequential Monte Carlo Methods in Practice*; Springer: Berlin/Heidelberg, Germany, 2001.

21. Šimandl, M.; Královec, J.; Söderström, T. Anticipative Grid Design in Point-Mass Approach to Nonlinear State Estimation. *IEEE Trans. Autom. Control* **2002**, *47*, 699–702. [CrossRef]

22. Bergman, N. Recursive Bayesian Estimation: Navigation and Tracking Applications.   Ph.D. Thesis, Linköping University, Linköping, Sweden, 1999.

23. Duník, J.; Soták, M.; Veselý, M.; Straka, O.; Hawkinson, W.J. Design of Rao-Blackwellised Point-Mass Filter with Application in Terrain Aided Navigation. *IEEE Trans. Aerosp. Electron. Syst.* **2019**, *55*, 251–272. [CrossRef]

24. Matoušek, J.; Duník, J.; Straka, O. Point-Mass Filter: Density Specific Grid Design and Implementation. In Proceedings of the 15th European Workshop on Advanced Control and Diagnosis, Bologna, Italy, 21–22 November 2019.

25. Šmídl, V.; Gašperin, M. Rao-Blackwellized Point Mass Filter for Reliable State Estimation. In Proceedings of the 16th International Conference on Information Fusion, Istanbul, Turkey, 9–12 July 2013.

26. Šimandl, M.; Královec, J.; Söderström, T. Advanced Point-mass Method for Nonlinear State Estimation. *Automatica* **2006**, *42*, 1133–1145. [CrossRef]

27. Sirola, N. Nonlinear Filtering with Piecewise Probability Densities. Ph.D. Thesis, Tampere University of Technology, Tampere, Finland, 2007.

28. Yoo, Y.M.; Park, C.G. Improvement of Terrain Referenced Navigation using a Point Mass Filter with Grid Adaptation. *Int. J. Control. Autom. Syst.* **2015**, *13*, 1173–1181. [CrossRef]

29. Peng, D.; Zhou, T.; Xu, C.; Zhang, W.; Shen, J. Marginalized Point Mass Filter with Estimating Tidal Depth Bias for Underwater Terrain-Aided Navigation. *J. Sens.* **2019**, *2019*. [CrossRef]

30. Bergman, N. Bayesian Approach to Terrain-Aided Navigation.   In *11th IFAC Symposium on System Identification*; Elsevier: Kitakyushu/Fukuoka, Japan, 1997; pp. 1531–1536.

31. Jeon, H.C.; Park, W.J.; Park, C.G. Grid Design for Efficient and Accurate Point Mass Filter-Based Terrain Referenced Navigation. *IEEE Sens. J.* **2018**, *18*, 1731–1737. [CrossRef]

32. Duník, J.; Straka, O.; Šimandl, M.; Blasch, E. Random-Point-Based Filters: Analysis and Comparison in Target Tracking. *IEEE Trans. Aerosp. Electron. Syst.* **2015**, *51*, 303–308. [CrossRef]

33. Nørgaard, M.; Poulsen, N.K.; Ravn, O. New Developments in State Estimation for Nonlinear Systems. *Automatica* **2000**, *36*, 1627–1638. [CrossRef]

34. Musso, C.; Bresson, A.; Bidel, Y.; Zahzam, N.; Dahia, K.; Allard, J.; Sacleux, B. Absolute Gravimeter for Terrain-aided Navigation. In Proceedings of the 20th International Conference on Information Fusion, Xi'an, China, 10–13 July 2017.