Sigma-Point Set Rotation for Derivative-Free Filters in Target Tracking Applications

JINDŘICH DUNÍK
ONDŘEJ STRAKA
MIROSLAV ŠIMANDL
ERIK BLASCH

The paper focuses on the state estimation of the nonlinear discrete-time stochastic dynamic systems by the derivative-free filters. In particular the impact of the σ-point set rotation on the performance of the unscented transform and the unscented Kalman filter (UKF) is analysed. It is shown that the σ-point set rotation is an additional user-defined parameter closely tied with the covariance matrix decomposition technique used in σ-point computation that significantly affects the estimation performance. Analysis, algorithms, and recommendations for computations of the optimal σ-point set rotation are provided to determine either the rotation prior to the estimation experiment (off-line) or during the estimation experiment (on-line). Further, two approaches for a reduction of optimisation computational costs are presented. The proposed algorithms, namely the on-line adaptive-sigma-point-set-UKF (AUKF) and off-line trained-sigma-point-set-UKF (TUKF), are illustrated and verified in a numerical study considering two static and two dynamic examples. The TUKF improves the UKF performance, while the computational complexity is preserved. The AUKF further improves the estimate accuracy with increased computational burden.

Manuscript received February 16, 2015; revised December 22, 2015; released for publication April 2, 2016.

Refereeing of this contribution was handled by U. Hanebeck.

Authors' addresses: J. Duník, O. Straka, and M. Šimandl, Department of Cybernetics and European Centre of Excellence NTIS, Faculty of Applied Sciences, University of West Bohemia, Univerzitn´ 8, 306 14 Pilsen, Czech Republic, tel.: +420377632549, fax: +420377632502 (E-mail: {dunikj,straka30,simandl}@kky.zcu.cz), E. Blasch, Air Force Research Lab, Information Directorate, Rome, NY 13441 (E-mail: erik.blasch.1@us.af.mil).

This work was supported by the grant GA15-12068S by the Czech Science Foundation (GACR).

LIST OF ABBREVIATIONS

AUKF	Adaptive-sigma-point-set Unscented Kalman Filter
BRRs	Bayesian Recursive Relations
CKF	Cubature Kalman Filter
DDFs	Divided Difference Filters
KF	Kalman Filter
MC	Monte-Carlo
TUKF	Trained-sigma-point-set Unscented Kalman Filter
PDF	Probability Density Function
RMSE	Root Mean Square Error
STD	Standard Deviation
SVD	Singular Value Decomposition
TE	Taylor series Expansion
UKF	Unscented Kalman Filter
UT	Unscented Transformation

1. INTRODUCTION

State estimation of nonlinear discrete-time stochastic systems plays an important role in many fields such as adaptive and optimal control, fault detection, and signal processing in many applications such as navigation and target tracking.

The state estimation can be solved by various techniques among them the Bayesian and optimisation approaches are widely preferred. The Bayesian approach stems from the solution to the Bayesian recursive relations (BRRs) computing the probability density functions (PDFs) of the state conditioned by the available measurements. The conditional PDF provides a complete description of the immeasurable state, which is valid almost over the whole state space. Therefore, the BRRs-based methods are usually called global. In contrast, the solution to the estimation problem provided by the optimisation approach is in the form of conditional moments of the state, which do not represent a complete description of the estimated state. Therefore, the optimisation-based methods are usually called local[1] as the estimate is valid in a small vicinity of the working point only. The local methods are based on specification and (usually) minimisation of a criterion (most often the mean squared error) under assumption of a certain estimator structure.

The closed-form solution to the state estimation problem is available only for a limited set of the systems. Among these systems, the linear ones are the most important [21], [1]. In other cases [5], [29], [30], if the closed-form solution is not available, approximate methods have to be used. The approximate global methods are based on various approximations to the BRRs solution and are represented for example by the particle filter [5], point-mass method [29], the Gaussian sum method [30], or their combinations [19]. The application

[1] In literature besides the term "local filters," terms "Gaussian filters" or "Kalman filters" can be found as well.

of the global methods is usually limited by their computational complexity especially for high-dimensional systems. Therefore, in most practical applications the approximate local methods are often preferred.

The approximate local filters use the algorithm structure of the Kalman filter[2] (KF) for solving the state estimation problem of nonlinear systems. The local filters can further be divided into two groups; derivative and derivative-free. Derivative filters approximate nonlinear functions in a system description by derivative-based expansions, for example by the Taylor or the Fourier-Hermite series expansions which lead to the extended Kalman filter, the second-order filter [1], or the Fourier-Hermite Kalman filter [27]. Derivative-free local filters are based on differential polynomial interpolations, the unscented transform, or various numerical integration rules. These approximations might be viewed as approximations of the state estimate description by a set of weighted points while the nonlinear functions in the system description remains unaffected. The filters within this group are represented by the divided difference filters (DDFs) based on the Stirling polynomial interpolation [23], the unscented Kalman filter (UKF) based on various versions of the unscented transformation (UT) [16], or the quadrature [13], cubature [2], and stochastic integration [9] based filters utilizing deterministic and stochastic integration rules.[3]

Derivative-free filters (contrary to the derivative ones) evaluate the nonlinear functions in the system description at multiple points (often called σ-points). Placement of the σ-points in the state-space is determined by i) inherent parameters, which are the current mean and associated covariance matrix of the estimate error (i.e., the approximation point) and ii) user-defined parameters affecting the quality of the approximation and subsequently the filter performance. The user-defined parameters might include specification of the σ-point set rotation (also tied to the selection of the covariance matrix decomposition) or the σ-point set scaling (if applicable).

Scaling of the σ-point set by specification of a scaling parameter or parameters has been widely studied in the last decade and recommendations for both fixed and adaptive parameter settings have been proposed. Recommendations for fixed parameters setting stem from a term-by-term comparison of the Taylor series expansion (TE) of the true mean and covariance matrix of a random variable transformed through a nonlinear function with the TE of the point-approximated statistics [23], [17]. In this case, the scaling parameter is a function of the state-space dimension only. On the other hand, an adaptive parameter setting takes an advantage of possibly different parameter values over time instants reflecting the actual working (or approximation) point. Various off-line and on-line adaptive strategies were proposed e.g., in [26], [10], [36], [11], resulting in a non-negligible estimation performance improvement.

Rotation of the σ-point set has been recently directly analysed in [4], [7], [6] and indirectly via the covariance matrix decomposition[4] in [24], [34]. In [7], an adaptive selection of the σ-point set rotation has been studied and illustrated using a simplistic bearings-only tracking example. It was shown that the impact of the σ-point set rotation on the transformation and subsequently on the filter performance is comparable with the σ-point set scaling, which is worth for a deeper analysis.

The goal of the paper is therefore twofold; to summarise the recent results related to the σ-point set rotation in a unified local filter design framework and to thoroughly analyse and explain the impact of the σ-point set rotation on a filter performance. Special emphasis is also placed on a numerical illustration and on reduction of the computational costs of the optimal σ-point set rotation specification.

The rest of the paper is organised as follows. In Section 2 the system description and the introduction to the state estimation with the stress on the UKF are given. Rotation of the σ-point set, its influence, parametrisation and optimisation are discussed in Section 3. Section 4 focuses on two techniques to reduce optimisation costs of σ-point set rotation. In Section 5 a general algorithm for the UKF with rotated σ-point set is summarised. Section 6 compares the proposed algorithms in a numerical study. Conclusions are drawn in Section 7.

2. SYSTEM DESCRIPTION AND STATE ESTIMATION BY UKF

A discrete-time nonlinear stochastic dynamic system is given as

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

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

where the vectors $\mathbf{x}_k \in \mathbb{R}^{n_x}$ and $\mathbf{z}_k \in \mathbb{R}^{n_z}$ represent the immeasurable state of the system and measurement at time instant k, respectively, $\mathbf{f}_k : \mathbb{R}^{n_x} \to \mathbb{R}^{n_x}$ and $\mathbf{h}_k : \mathbb{R}^{n_x} \to \mathbb{R}^{n_z}$ are known vector functions, and $\mathbf{w}_k \in \mathbb{R}^{n_x}$ and $\mathbf{v}_k \in \mathbb{R}^{n_z}$ are the state and measurement white noises. The noises are assumed to be zero-mean with known

[2]The Kalman filter is an optimal (in the mean squared error sense) linear estimator for linear systems.

[3]Although the approximations used in the derivative-free filters originate from quite different basic ideas, the resulting filter algorithms are in many cases identical. As examples, the analysis and analytic comparison of the DDFs and UKF are given in [28], cubature and quadrature filters and UKF in [14], cubature and stochastic integration filters in [9], and the stochastic integration filter and Monte-Carlo filter are in [8].

[4]The particular matrix decompositions lead to the same σ-point sets up to their rotation (or reflection), i.e., all the sets lies on the surface of a hyper-ellipsoid determined by the covariance matrix [34].

covariance matrices $\mathbf{\Sigma}_k^w = \mathrm{cov}[\mathbf{w}_k]$ and $\mathbf{\Sigma}_k^v = \mathrm{cov}[\mathbf{v}_k]$, respectively. The first two moments of the initial state $\mathbf{x}_0$ are assumed to be known as well, i.e., $\mathsf{E}[\mathbf{x}_0] = \bar{\mathbf{x}}_0$, $\mathrm{cov}[\mathbf{x}_0] = \mathbf{P}_0$.

The local state estimation methods provide the point estimate $\hat{\mathbf{x}}_{k|k}$ approximating the conditional mean $\mathsf{E}[\mathbf{x}_k \mid \mathbf{z}^k]$, in which $\mathbf{z}^k = [\mathbf{z}_0, \mathbf{z}_1, \ldots, \mathbf{z}_k]$, and the corresponding covariance matrix of the estimation error $\mathbf{P}_{k|k}^{\mathbf{xx}}$.

The first two moments can be realised as a Gaussian approximation of the conditional PDF, i.e., $p(\mathbf{x}_k \mid \mathbf{z}^k) \approx \mathcal{N}\{\mathbf{x}_k : \hat{\mathbf{x}}_{k|k}, \mathbf{P}_{k|k}^{\mathbf{xx}}\}$ [2], [9], depending on the type of employed approximation.

For calculation of predictive statistics of the state and measurement the UKF utilises the UT.

2.1. Unscented transformation

The UT [16] is a tool for approximate computation of the mean, covariance matrix, and cross-covariance matrix of a transformed random variable

$$\mathbf{y} = \mathbf{g}(\mathbf{x}) = [g_1(\mathbf{x}), \ldots, g_{n_y}(\mathbf{x})]^{\mathrm{T}}, \tag{3}$$

where $\mathbf{x} \in \mathbb{R}^{n_x}$ and $\mathbf{y} \in \mathbb{R}^{n_y}$, under the assumption of known vector function $\mathbf{g}(\cdot)$ and known mean $\hat{\mathbf{x}} = \mathsf{E}[\mathbf{x}]$ and covariance matrix $\mathbf{P}^{\mathbf{xx}} = \mathrm{cov}[\mathbf{x}]$ of $\mathbf{x}$. The UT is based on computation of a symmetric set of deterministic σ-points $\{\mathcal{X}_i\}_{i=0}^{2n_x}$ with appropriate weights $\{\mathcal{W}_i\}_{i=0}^{2n_x}$ according to

$$\mathcal{X}_0 = \hat{\mathbf{x}}, \qquad \mathcal{W}_0 = \frac{\kappa}{n_x + \kappa}, \tag{4}$$

$$\mathcal{X}_j = \hat{\mathbf{x}} + \sqrt{(n_x + \kappa)}\mathbf{S}^{\mathbf{xx}}\mathbf{e}_j, \quad \mathcal{W}_j = \frac{1}{2(n_x + \kappa)}, \tag{5}$$

$$\mathcal{X}_{n_x+j} = \hat{\mathbf{x}} - \sqrt{(n_x + \kappa)}\mathbf{S}^{\mathbf{xx}}\mathbf{e}_j, \quad \mathcal{W}_{n_x+j} = \mathcal{W}_j, \tag{6}$$

where $j = 1, \ldots, n_x$, term $\mathbf{e}_j$ is the jth column of the n_x-dimensional identity matrix $\mathbf{I}_{n_x}$, and $\mathbf{S}^{\mathbf{xx}}$ is a decomposition of the covariance matrix $\mathbf{P}^{\mathbf{xx}}$ such that $\mathbf{S}^{\mathbf{xx}}(\mathbf{S}^{\mathbf{xx}})^{\mathrm{T}} = \mathbf{P}^{\mathbf{xx}}$. The variable κ is a scaling parameter. To get an approximate characteristic of $\mathbf{y}$, each point is transformed through the nonlinear function

$$\mathcal{Y}_i = \mathbf{g}(\mathcal{X}_i), \quad \forall i. \tag{7}$$

The resulting approximate characteristics calculated by the UT are given by

$$\hat{\mathbf{y}}^{\mathrm{UT}} = \sum_{i=0}^{2n_x} \mathcal{W}_i \mathcal{Y}_i, \tag{8}$$

$$\mathbf{P}^{\mathbf{yy},\mathrm{UT}} = \sum_{i=0}^{2n_x} \mathcal{W}_i (\mathcal{Y}_i - \hat{\mathbf{y}}^{\mathrm{UT}})(\cdot)^{\mathrm{T}}, \tag{9}$$

$$\mathbf{P}^{\mathbf{xy},\mathrm{UT}} = \sum_{i=0}^{2n_x} \mathcal{W}_i (\mathcal{X}_i - \hat{\mathbf{x}})(\mathcal{Y}_i - \hat{\mathbf{y}}^{\mathrm{UT}})^{\mathrm{T}}, \tag{10}$$

where the notation $(\mathbf{a})(\cdot)^{\mathrm{T}}$ stands for $(\mathbf{a})(\mathbf{a})^{\mathrm{T}}$. Now, having the UT introduced, the UKF algorithm can be stated.

2.2. Unscented Kalman filter

The UKF algorithm has the following structure [16]:

ALGORITHM 1: *Unscented Kalman Filter*

Step 1: (initialisation) Set the time instant $k = 0$ and define a priori initial condition by the predictive mean $\hat{\mathbf{x}}_{0|-1} = \mathsf{E}[\mathbf{x}_0] = \bar{\mathbf{x}}_0$ and the predictive covariance matrix $\mathbf{P}_{0|-1}^{\mathbf{xx}} = \mathrm{cov}[\mathbf{x}_0] = \mathbf{P}_0$. Set the scaling parameter κ and compute the σ-point weights

$$\mathcal{W}_0 = \frac{\kappa}{n_x + \kappa}, \tag{11}$$

$$\mathcal{W}_i = \frac{1}{2(n_x + \kappa)}, \quad i = 1, \ldots, 2n_x. \tag{12}$$

Step 2: (filtering step) The state predictive estimate is updated with respect to the last measurement $\mathbf{z}_k$ according to

$$\hat{\mathbf{x}}_{k|k} = \hat{\mathbf{x}}_{k|k-1} + \mathbf{K}_k(\mathbf{z}_k - \hat{\mathbf{z}}_{k|k-1}), \tag{13}$$

$$\mathbf{P}_{k|k}^{\mathbf{xx}} = \mathbf{P}_{k|k-1}^{\mathbf{xx}} - \mathbf{K}_k \mathbf{P}_{k|k-1}^{\mathbf{zz}} \mathbf{K}_k^{\mathrm{T}}, \tag{14}$$

where $\mathbf{K}_k = \mathbf{P}_{k|k-1}^{\mathbf{xz}}(\mathbf{P}_{k|k-1}^{\mathbf{zz}})^{-1}$ is the filter gain,

$$\hat{\mathbf{z}}_{k|k-1} = \sum_{i=0}^{2n_x} \mathcal{W}_i \mathcal{Z}_{i,k|k-1}, \tag{15}$$

$$\mathbf{P}_{k|k-1}^{\mathbf{zz}} = \sum_{i=0}^{2n_x} \mathcal{W}_i (\mathcal{Z}_{i,k|k-1} - \hat{\mathbf{z}}_{k|k-1})(\cdot)^{\mathrm{T}} + \mathbf{\Sigma}_k^v, \tag{16}$$

$$\mathbf{P}_{k|k-1}^{\mathbf{xz}} = \sum_{i=0}^{2n_x} \mathcal{W}_i (\mathcal{X}_{i,k|k-1} - \hat{\mathbf{x}}_{k|k-1})(\mathcal{Z}_{i,k|k-1} - \hat{\mathbf{z}}_{k|k-1})^{\mathrm{T}}, \tag{17}$$

$$\mathcal{Z}_{i,k|k-1} = \mathbf{h}_k(\mathcal{X}_{i,k|k-1}), \tag{18}$$

and the predictive state σ-points are computed according to

$$\mathcal{X}_{0,k|k-1} = \hat{\mathbf{x}}_{k|k-1}, \tag{19}$$

$$\mathcal{X}_{j,k|k-1} = \hat{\mathbf{x}}_{k|k-1} + \sqrt{(n_x + \kappa)}\mathbf{S}_{k|k-1}^{\mathbf{xx}}\mathbf{e}_j, \tag{20}$$

$$\mathcal{X}_{n_x+j,k|k-1} = \hat{\mathbf{x}}_{k|k-1} - \sqrt{(n_x + \kappa)}\mathbf{S}_{k|k-1}^{\mathbf{xx}}\mathbf{e}_j, \tag{21}$$

with $j = 1, \ldots n_x$ and $\mathbf{S}_{k|k-1}^{\mathbf{xx}}$ being a decomposition of $\mathbf{P}_{k|k-1}^{\mathbf{xx}}$.

Step 3: (predictive step) The predictive statistics are given by the relations

$$\hat{\mathbf{x}}_{k+1|k} = \sum_{i=0}^{2n_x} \mathcal{W}_i \mathcal{X}_{i,k+1|k}, \tag{22}$$

$$\mathbf{P}_{k+1|k}^{\mathbf{xx}} = \sum_{i=0}^{2n_x} \mathcal{W}_i (\mathcal{X}_{i,k+1|k} - \hat{\mathbf{x}}_{k+1|k})(\cdot)^{\mathrm{T}} + \mathbf{\Sigma}_k^w, \tag{23}$$

$$\mathcal{X}_{i,k+1|k} = \mathbf{f}_k(\mathcal{X}_{i,k|k}), \tag{24}$$

where the filtering state σ-points are computed according to

$$\mathcal{X}_{0,k|k} = \hat{\mathbf{x}}_{k|k}, \tag{25}$$

$$\mathcal{X}_{j,k|k} = \hat{\mathbf{x}}_{k|k} + \sqrt{(n_x + \kappa)}\mathbf{S}^{\mathbf{xx}}_{k|k}\mathbf{e}_j, \tag{26}$$

$$\mathcal{X}_{n_x+j,k|k} = \hat{\mathbf{x}}_{k|k} - \sqrt{(n_x + \kappa)}\mathbf{S}^{\mathbf{xx}}_{k|k}\mathbf{e}_j, \tag{27}$$

with $j = 1,\dots n_x$ and $\mathbf{S}^{\mathbf{xx}}_{k|k}$ being a decomposition of $\mathbf{P}^{\mathbf{xx}}_{k|k}$.

Let $k = k + 1$. The algorithm continues by **Step 2**.

2.3. σ-point set

The σ-point set of the UT (4)–(6) ((19)–(21) and (25)–(27) in the UKF) is determined by the inherent parameters given by the mean and covariance matrix of $\mathbf{x}$ (i.e., by $\hat{\mathbf{x}}$, $\mathbf{P}^{\mathbf{xx}}$) and by the user-defined parameters. The latter include the *scaling parameter* κ influencing the set scaling (the area of the state-space covered by the points) and the *decomposition technique* which is inevitably tied with the σ-point set rotation [34]. However, independently of the selected user-defined parameter, the weighted σ-point set has at least the same mean and covariance matrix as the original to-be-transformed random variable, i.e.,

$$\hat{\mathbf{x}}^{\mathrm{UT}} = \sum_{i=0}^{2n_x} \mathcal{W}_i\mathcal{X}_i = \hat{\mathbf{x}}, \tag{28}$$

$$\begin{aligned}\mathbf{P}^{\mathbf{xx},\mathrm{UT}} &= \sum_{i=0}^{2n_x} \mathcal{W}_i(\mathcal{X}_i - \hat{\mathbf{x}}^{\mathrm{UT}})(\mathcal{X}_i - \hat{\mathbf{x}}^{\mathrm{UT}})^{\mathrm{T}} \\ &= \mathbf{S}^{\mathbf{xx}}[\mathbf{S}^{\mathbf{xx}}]^{\mathrm{T}} = \mathbf{P}^{\mathbf{xx}}.\end{aligned} \tag{29}$$

This is the ultimate property of any σ-point set.

2.4. Scaling parameter of σ-point set

In literature, three main options for the selection of the scaling parameter can be identified, namely:

- constant parameter (specified prior to the estimation experiment),
- off-line computed time-varying parameter,
- on-line computed time-varying parameter.

The standard setting of the constant scaling parameter follows the recommendation $\kappa = 3 - n_x$ [16], [17], which minimises the error of the TE of the true mean and its UT approximation.[5]

The strategies for time-varying setting of the scaling parameter take into account not only the state dimension but also the actual filter working point (mean and covariance matrix) and the particular nonlinear functions in the state-space model. The strategies for off-line setting can be found e.g., in [26] and for on-line setting in [11], [35], [33]. The on-line strategies compute κ (usually in the filtering step) at every time instant to minimise or maximise a chosen criterion.

2.5. Covariance matrix decomposition

The choice of the covariance matrix decomposition technique was discussed in [34], where it was shown that all decomposition techniques provide the same σ-point set up to the σ-point set rotation or reflection.[6] The consequence is that there are infinitely many decompositions which can be parametrised by a rotation matrix. The analysis also indicated that the impact of the rotation is increasing for covariance matrices with non-negligible off-diagonal elements and with difference in magnitude of the matrix eigenvalues. Illustration of three different σ-point sets based on the Cholesky decomposition, singular value decomposition (SVD), and matrix square root and three different sets obtained by the SVD and consequently rotated by 0, 30, and 60 degrees can be found in Fig. 1.

In [4], [6], and [7] the discussion was further extended by introducing arbitrary rotation and reflection matrices in the σ-point set computation and a significant impact of the rotation was illustrated by a set of numerical examples without a theoretical analysis. As a consequence, the provided recommendations for the σ-point set rotation were rather ad hoc without theoretical justification.

The aim of the following section is to provide a thorough theoretical analysis of the impact of the σ-point rotation on the UT and subsequently on the UKF performance and the justification for on-line and off-line rotation matrix optimisation. The theoretical results are then illustrated by static and dynamic numerical examples.

3. OPTIMISATION OF σ-POINT SET ROTATION IN UT

The σ-points lie on an hyper-ellipsoid with its size determined by the scaling parameter and the semi-axes given by eigenvectors of the covariance matrix $\mathbf{P}^{\mathbf{xx}}$. Specific position of the σ-points on the hyper-ellipsoid are determined by the chosen decomposition. Different decompositions can be obtained by right-multiplying a decomposition by a rotation matrix $\mathbf{C}$. To avoid confusion, in the rest of the paper the SVD will be used to generate the decomposition $\mathbf{S}^{\mathbf{xx}}$ if not stated otherwise. The SVD decomposes the symmetric and positive definite covariance matrix $\mathbf{P}^{\mathbf{xx}}$ as $\mathbf{P}^{\mathbf{xx}} = \mathbf{U}\mathbf{D}\mathbf{U}^{\mathrm{T}}$, where $\mathbf{U}$ is a unitary matrix and $\mathbf{D}$ is a diagonal matrix. Then, the decomposition $\mathbf{S}^{\mathbf{xx}}$ is given by $\mathbf{S}^{\mathbf{xx}} = \mathbf{U}\sqrt{\mathbf{D}}$, where $\sqrt{\mathbf{D}}$ is a diagonal matrix with elements given by square root of elements on the diagonal of $\mathbf{D}$. Note that using the SVD generates the σ-points on the principal axes of a hyper-ellipsoid.

[5] Note that, if $n_x > 3$, the covariance matrices may lose positive semi-definiteness as $\kappa < 0$. Then, it is better to choose $\kappa = 0$ [17], which ensures positive definiteness of covariance matrices. For such a choice, the UKF converts into the cubature Kalman filter (CKF) [2].

[6] Contrary to the rotation, the reflection changes the order of the points besides their rotation. If a symmetric σ-point set is considered, the rotation and reflection are interchangeable.

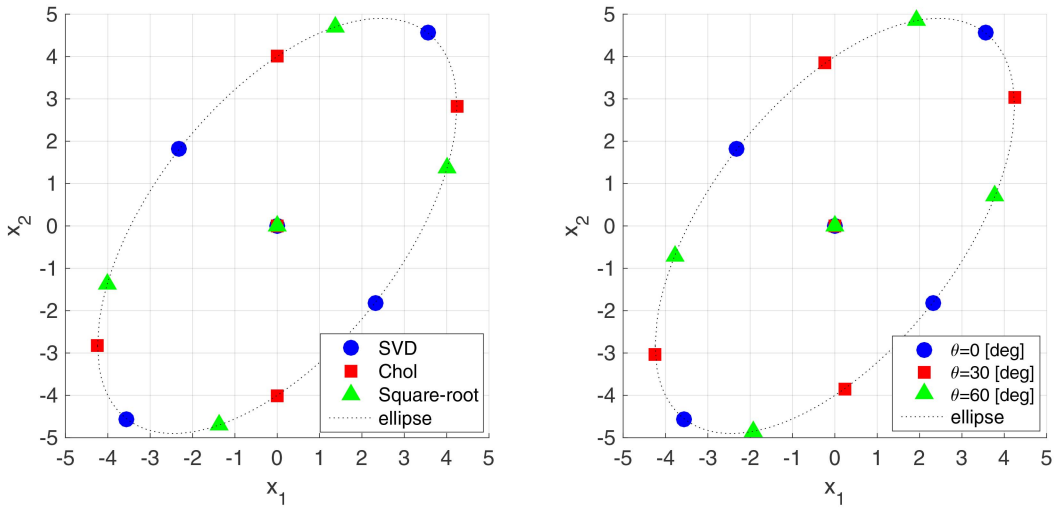


Fig. 1. Illustration of σ-point sets based on different decompositions and with different rotations.

3.1. Rotation of UT σ-point set

The σ-point set is typically computed by (4)–(6). The σ-point set with explicit consideration of the rotation can be computed according to

$$\mathcal{X}_0^r = \hat{\mathbf{x}}, \tag{30}$$

$$\mathcal{X}_j^r = \hat{\mathbf{x}} + \sqrt{(n_x+\kappa)}\mathbf{S}^{\mathbf{xx},r}\mathbf{e}_j, \tag{31}$$

$$\mathcal{X}_{n_x+j}^r = \hat{\mathbf{x}} - \sqrt{(n_x+\kappa)}\mathbf{S}^{\mathbf{xx},r}\mathbf{e}_j, \tag{32}$$

where $j = 1,\dots n_x$ and $\mathbf{S}^{\mathbf{xx},r}$ is a decomposition of $\mathbf{P}^{\mathbf{xx}}$ parametrised by a rotation matrix $\mathbf{C}$

$$\mathbf{S}^{\mathbf{xx},r} = \mathbf{S}^{\mathbf{xx}}\mathbf{C}, \tag{33}$$

where $\mathbf{C}$ is a rotation matrix in n_x dimensional space. Because of the orthogonal property of the rotation matrix, the rotated covariance matrix decomposition $\mathbf{S}^{\mathbf{xx},r}$ also produces the covariance matrix $\mathbf{P}^{\mathbf{xx}}$ as well as the original decomposition $\mathbf{S}^{\mathbf{xx}}$:

$$\mathbf{P}^{\mathbf{xx}} = \mathbf{S}^{\mathbf{xx},r}[\mathbf{S}^{\mathbf{xx},r}]^{\mathrm{T}} = \mathbf{S}^{\mathbf{xx}}\mathbf{C}\mathbf{C}^{\mathrm{T}}[\mathbf{S}^{\mathbf{xx}}]^{\mathrm{T}} = \mathbf{S}^{\mathbf{xx}}\mathbf{I}_{n_x}[\mathbf{S}^{\mathbf{xx}}]^{\mathrm{T}}. \tag{34}$$

As the rotated σ-point set (30)–(32) is still symmetric and the relation (34) is valid, the rotated set preserves the mean and covariance matrix of the original variable $\mathbf{x}$ similarly to the "unrotated" set in (28), (29). The σ-point set rotation, therefore, does not affect the weights $\{\mathcal{W}_i\}_{i=0}^{2n_x}$.

Hence, the rotation matrix $\mathbf{C}$ can be seen as another design parameter of the σ-point set (besides the scaling parameter).

3.2. Effect of σ-points rotation

Let us focus on the UT based approximation of $\hat{\mathbf{y}} = \mathsf{E}[\mathbf{y}]$ by $\hat{\mathbf{y}}^{\mathrm{UT}}$ in (8). The approximation uses the weights $\mathcal{W}_i$ and σ-points $\mathcal{Y}_i(\mathbf{C}) = \mathbf{g}(\mathcal{X}_i^r)$, where the notation emphasises dependence of the σ-points on the rotation matrix $\mathbf{C}$.

The error of approximation denoted as $\tilde{\mathbf{y}}(\mathbf{C})$ is given by

$$\tilde{\mathbf{y}}(\mathbf{C}) = \hat{\mathbf{y}} - \hat{\mathbf{y}}^{\mathrm{UT}}(\mathbf{C}). \tag{35}$$

The optimum rotation will be sought to minimise the weighted squared error of approximation

$$J(\mathbf{C}) = \tilde{\mathbf{y}}(\mathbf{C})^{\mathrm{T}}\mathbf{W}\tilde{\mathbf{y}}(\mathbf{C}), \tag{36}$$

where $\mathbf{W}$ can be any symmetric positive definite matrix weighting individual elements of $\tilde{\mathbf{y}}(\mathbf{C})$. The weight matrix $\mathbf{W}$ can for example be chosen as $\mathbf{W} = (\mathbf{P}^{\mathbf{yy}})^{-1}$, in which case the criterion puts stress on the elements of $\tilde{\mathbf{y}}(\mathbf{C})$ corresponding to small diagonal values of $\mathbf{P}^{\mathbf{yy}}$.

The true mean $\hat{\mathbf{y}}$ can be expressed using the TE of $\mathbf{g}$ around $\hat{\mathbf{x}}$ as

$$\begin{aligned}\hat{\mathbf{y}} = \mathsf{E}[\mathbf{y}] = \mathsf{E}[\mathbf{g}(\mathbf{x})] &= \mathsf{E}[\sum_{i=0}^{\infty}\frac{1}{i!}\mathbf{g}^{(i)}(\hat{\mathbf{x}})(\mathbf{x}-\hat{\mathbf{x}})^{\otimes i}] \\ &= \sum_{i=0}^{\infty}\frac{1}{i!}\mathbf{g}^{(i)}(\hat{\mathbf{x}})\mathbf{M}^{\mathbf{x}(i)},\end{aligned} \tag{37}$$

where $\otimes$ is the Kronecker product and the jth row of $\mathbf{g}^{(i)}(\hat{\mathbf{x}}) \in \mathbb{R}^{n_y\times(n_x)^i}$ is given by

$$\mathbf{g}_j^{(i)}(\hat{\mathbf{x}}) = \sum_{l_1,\dots l_i}^{n_x,\dots,n_x}\frac{\partial^i g_j(\mathbf{x})}{\partial\mathbf{x}_{l_1}\cdots\partial\mathbf{x}_{l_i}}\bigg|_{\mathbf{x}=\hat{\mathbf{x}}}(\mathbf{e}_{l_1}\otimes\mathbf{e}_{l_2}\otimes\cdots\otimes\mathbf{e}_{l_i}) \tag{38}$$

with $\mathbf{x}_{l_i}$ being the l_ith element of $\mathbf{x}$. Further, $(\mathbf{x}-\hat{\mathbf{x}})^{\otimes i} = (\mathbf{x}-\hat{\mathbf{x}})\otimes(\mathbf{x}-\hat{\mathbf{x}})^{\otimes i-1}$, $(\mathbf{x}-\hat{\mathbf{x}})^{\otimes 0} = \mathbf{1}_{n_x\times 1}$, $\mathbf{M}^{\mathbf{x}(i)} = \mathsf{E}[(\mathbf{x}-\hat{\mathbf{x}})^{\otimes i}]$, and $\mathbf{M}^{\mathbf{x}(i)} \in \mathbb{R}^{(n_x)^i\times 1}$ is the ith central moments of $\mathbf{x}$ stacked column-wise.

Similarly, by expansion of each transformed σ-point $\mathcal{Y}_i(\mathbf{C}) = \mathbf{g}(\mathcal{X}_i^r)$ around the mean $\hat{\mathbf{x}}$, the mean $\hat{\mathbf{y}}^{\mathrm{UT}}(\mathbf{C})$ can

be written as

$$\hat{\mathbf{y}}^{\mathrm{UT}}(\mathbf{C}) = \sum_{j=0}^{2n_x} \mathcal{W}_j \sum_{i=0}^{\infty} \frac{1}{i!} \mathbf{g}^{(i)}(\hat{\mathbf{x}})(\mathcal{X}_j^r - \hat{\mathbf{x}})^{\otimes i}$$
$$= \sum_{i=0}^{\infty} \frac{1}{i!} \mathbf{g}^{(i)}(\hat{\mathbf{x}}) \sum_{j=0}^{2n_x} \mathcal{W}_j (\mathcal{X}_j^r - \hat{\mathbf{x}})^{\otimes i}$$
$$= \sum_{i=0}^{\infty} \frac{1}{i!} \mathbf{g}^{(i)}(\hat{\mathbf{x}}) \mathbf{M}^{\mathbf{x}(i),\mathrm{UT}}(\mathbf{C}), \tag{39}$$

where $\mathbf{M}^{\mathbf{x}(i),\mathrm{UT}}(\mathbf{C})$ is an UT approximation of $\mathbf{M}^{\mathbf{x}(i)}$. Note that

$$\mathbf{M}^{\mathbf{x}(1)} = \mathbf{M}^{\mathbf{x}(1),\mathrm{UT}}(\mathbf{C}) = \mathbf{0}_{n_x \times 1}, \tag{40}$$

$$\mathbf{M}^{\mathbf{x}(2)} = \mathbf{M}^{\mathbf{x}(2),\mathrm{UT}}(\mathbf{C}) = \mathrm{vec}(\mathbf{P}^{\mathbf{xx}}), \tag{41}$$

where $\mathrm{vec}(\mathbf{A})$ is a column vector obtained by stacking the columns of $\mathbf{A}$. Hence, the first two moments are independent of the σ-point rotation $\mathbf{C}$.

Assuming $\mathbf{x}$ being Gaussian and the σ-point set being symmetric, the odd terms are zero $\mathbf{M}^{\mathbf{x}(2i+1)} = \mathbf{0}$, $\mathbf{M}^{\mathbf{x}(2i+1),\mathrm{UT}} = \mathbf{0}$, $\forall i \in \mathbb{Z}^*$, and considering (40) and (41), the error of approximation is equal to

$$\tilde{\mathbf{y}}(\mathbf{C}) = \sum_{i=2}^{\infty} \frac{1}{(2i)!} \mathbf{g}^{(2i)}(\hat{\mathbf{x}})(\mathbf{M}^{\mathbf{x}(2i)} - \mathbf{M}^{\mathbf{x}(2i),\mathrm{UT}}(\mathbf{C})). \tag{42}$$

This means that the first non-zero term of the error $\tilde{\mathbf{y}}(\mathbf{C})$ expansion (42) is the fourth one, i.e., $\frac{1}{4!}\mathbf{g}^{(4)}(\hat{\mathbf{x}})(\mathbf{M}^{\mathbf{x}(4)} - \mathbf{M}^{\mathbf{x}(4),\mathrm{UT}}(\mathbf{C}))$. For a scalar function $\mathbf{g}$ the fourth term may be zeroed by a choice of the scaling parameter κ [32] and the first non-zero term is the sixth term. Nevertheless, for general vector functions $\mathbf{g}$ the fourth term is non-zero and depends on the rotation matrix $\mathbf{C}$. This implies that the value of the criterion (36) also depends on the rotation matrix $\mathbf{C}$, of which appropriate choice may lead to a minimum value of the criterion and consequently to the best $\hat{\mathbf{y}}^{\mathrm{UT}}$.

3.3. Selection of the optimum rotation

A direct utilisation of the TE of the mean error (42) for the rotation matrix $\mathbf{C}$ optimisation purposes is impossible due to the infinite number of terms. Assuming the magnitude of the ith term in (42) is decreasing with $i \to \infty$, the criterion (36) can be replaced by

$$J^4(\mathbf{C}) = [\tfrac{1}{4!}\mathbf{g}^{(4)}(\hat{\mathbf{x}})\tilde{\mathbf{M}}^{\mathbf{x}(4)}(\mathbf{C})]^{\mathrm{T}}\mathbf{W}[\tfrac{1}{4!}\mathbf{g}^{(4)}(\hat{\mathbf{x}})\tilde{\mathbf{M}}^{\mathbf{x}(4)}(\mathbf{C})]. \tag{43}$$

where $\tilde{\mathbf{M}}^{\mathbf{x}(4)}(\mathbf{C}) = \mathbf{M}^{\mathbf{x}(4)} - \mathbf{M}^{\mathbf{x}(4),\mathrm{UT}}(\mathbf{C})$. Hence, $J^4(\mathbf{C})$ in (43) contains the first non-zero term of (42).

The criterion $J^4(\mathbf{C})$ can be used in the UT for finding an optimal rotation matrix $\mathbf{C}$. However, its utilisation is not very practical as the criterion requires calculation of fourth derivatives of $\mathbf{g}$, which might be limiting for the UT and subsequently for the UKF designed to be derivative-free.

An alternative of (36) more suitable for optimisation of the rotation matrix $\mathbf{C}$ to achieve the smallest error is based on a replacement of the mean $\hat{\mathbf{y}}$ by a sample $\mathbf{y}_s$ of the variable $\mathbf{y}$. This leads to the criterion

$$J^{\mathrm{S}}(\mathbf{C}) = [\mathbf{y}_s - \hat{\mathbf{y}}^{\mathrm{UT}}(\mathbf{C})]^{\mathrm{T}}\mathbf{W}[\mathbf{y}_s - \hat{\mathbf{y}}^{\mathrm{UT}}(\mathbf{C})], \tag{44}$$

which is suitable for an application in the UKF.

Note that the criterion (36) and its consequent simplifications (43) and (44) aim to minimise the approximation error of the mean $\hat{\mathbf{y}}^{\mathrm{UT}}$. Other criteria also can be proposed for minimising approximation error of other quantities, such as the covariance matrix $\mathbf{P}^{\mathbf{yy},\mathrm{UT}}$. In that case a modification analogous to (43) can be proposed while no modification analogous to (44) exists as there is no sample of $\mathbf{P}^{\mathbf{yy}}$ available.

3.4. Choice of the weight matrix

It has been suggested above to choose the weight matrix $\mathbf{W}$ as the inverse of the covariance matrix $\mathbf{P}^{\mathbf{yy}}$ to properly weight individual elements of the mean error. This choice is however usually impossible as the covariance matrix $\mathbf{P}^{\mathbf{yy}}$ is usually unknown and is approximated by the UT. Its approximation also depends on the choice of the rotation matrix $\mathbf{C}$. So it may be replaced by its estimate $\mathbf{P}^{\mathbf{yy},\mathrm{UT}}(\mathbf{C})$, in which case the criterion would be

$$J^{\mathrm{S}}(\mathbf{C}) = [\mathbf{y}_s - \hat{\mathbf{y}}^{\mathrm{UT}}(\mathbf{C})]^{\mathrm{T}}(\mathbf{P}^{\mathbf{yy},\mathrm{UT}}(\mathbf{C}))^{-1}[\mathbf{y}_s - \hat{\mathbf{y}}^{\mathrm{UT}}(\mathbf{C})]. \tag{45}$$

Hence, by minimising (45), the difference $\mathbf{y}_s - \hat{\mathbf{y}}^{\mathrm{UT}}(\mathbf{C})$ and $(\mathbf{P}^{\mathbf{yy},\mathrm{UT}}(\mathbf{C}))^{-1}$ are minimised and thus $\mathbf{P}^{\mathbf{yy},\mathrm{UT}}(\mathbf{C})$ is maximised.

If $\mathbf{W}$ in (44) is selected as $\mathbf{W} = (\mathbf{P}^{\mathbf{yy},\mathrm{UT}}(\mathbf{C}))^{-1}$, then $J^{\mathrm{S}}(\mathbf{C})$ is χ^2 distributed random variable with n_y degrees of freedom with the mean $\mathsf{E}[J^{\mathrm{S}}(\mathbf{C})] = n_y$. Also note that $\mathsf{E}[\mathbf{y}_s - \hat{\mathbf{y}}^{\mathrm{UT}}(\mathbf{C})] = \hat{\mathbf{y}} - \hat{\mathbf{y}}^{\mathrm{UT}}(\mathbf{C})$. Consequently, the criterion can be formulated as

$$J^{\mathrm{MS}}(\mathbf{C}) = |[\mathbf{y}_s - \hat{\mathbf{y}}^{\mathrm{UT}}(\mathbf{C})]^{\mathrm{T}}(\mathbf{P}^{\mathbf{yy},\mathrm{UT}}(\mathbf{C}))^{-1} \times [\mathbf{y}_s - \hat{\mathbf{y}}^{\mathrm{UT}}(\mathbf{C})] - n_y|. \tag{46}$$

Minimising (46) results in $J^{\mathrm{S}}(\mathbf{C})$ (45), computed on the basis of predicted measurement statistics, which is closest to its mean value, i.e., to n_y.

3.5. Rotation parametrisation

Minimisation of the criterion (44) w.r.t. the rotation matrix $\mathbf{C}$ directly is quite difficult and should be avoided. The reason is twofold; the rotation matrix has to be an orthogonal matrix, thus it is difficult to perform minimisation while the structure is respected, and n_x^2 elements need to be found. Instead the rotation matrix could be parametrised by a set of

$$n_\theta = n_x(n_x - 1)/2 \tag{47}$$

rotation angles $\theta_1, \ldots, \theta_{n_\theta}$. Therefore, if $J(\mathbf{C})$ is minimised with respect to the rotation matrix parametrised by the rotation angles, then n_θ is the dimension of the optimisation space. Additionally, the rotation angles correspond to a minimal representation of a general rotation in $\mathbb{R}^{n_x}$.

As an example, parametrisation of a rotation matrix by six consecutive rotations $\theta_1, \theta_2, \ldots, \theta_6$ in four dimensional space is given by [22], [12]

$$\begin{aligned}\mathbf{C} = &\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & \cos(\theta_6) & -\sin(\theta_6) \\ 0 & 0 & \sin(\theta_6) & \cos(\theta_6) \end{bmatrix} \\ &\times \begin{bmatrix} \cos(\theta_5) & 0 & 0 & -\sin(\theta_5) \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ \sin(\theta_5) & 0 & 0 & \cos(\theta_5) \end{bmatrix} \\ &\times \begin{bmatrix} \cos(\theta_4) & 0 & -\sin(\theta_4) & 0 \\ 0 & 1 & 0 & 0 \\ \sin(\theta_4) & 0 & \cos(\theta_4) & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \\ &\times \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos(\theta_3) & 0 & -\sin(\theta_3) \\ 0 & 0 & 1 & 0 \\ 0 & \sin(\theta_3) & 0 & \cos(\theta_3) \end{bmatrix} \\ &\times \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos(\theta_2) & -\sin(\theta_2) & 0 \\ 0 & \sin(\theta_2) & \cos(\theta_2) & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \\ &\times \begin{bmatrix} \cos(\theta_1) & -\sin(\theta_1) & 0 & 0 \\ \sin(\theta_1) & \cos(\theta_1) & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}, \end{aligned} \tag{48}$$

where the last right-hand side matrix rotates a vector through an angle θ_1 in the $x-y$ plane (about the axis perpendicular to the plane) with $\mathbf{x} = [x, y, u, z]^{\mathrm{T}} \in \mathbb{R}^4$.

4. REDUCTION OF OPTIMISATION COSTS

Due to quadratic dependence of the number of rotation angles n_θ (47) on the state dimension it is convenient to pay attention to a possible costs reduction. In this section two approaches to optimisation costs reduction will be described. The reduction is achieved by decreasing the dimension n_θ of the optimisation space.

The first approach is applicable only for a class of functions $\mathbf{g}$ and offers a costs reduction that does not affect the smallest attainable approximation error of $\hat{\mathbf{y}}^{\mathrm{UT}}(\mathbf{C})$ and $\mathbf{P}^{\mathbf{yy},\mathrm{UT}}(\mathbf{C})$. Hence, the approach will be called lossless.

The second approach is applicable for general functions $\mathbf{g}$ and suggests optimising over the angles with a significant effect only. Application of this approach leads a suboptimal solution, i.e., to the approximation error probably higher than if the optimisation over the full optimisation space is done. Hence, the approach will be called lossy.

4.1. Lossless optimisation space dimension decreasing

If the variable $\mathbf{y}$ is a function of only some elements of $\mathbf{x}$, then the optimisation over some rotation angles does not affect the mean approximation $\hat{\mathbf{y}}^{\mathrm{UT}}(\mathbf{C})$ and the covariance matrix approximation $\mathbf{P}^{\mathbf{yy},\mathrm{UT}}(\mathbf{C})$. Hence, these rotations are useless in the optimisation and the dimension of the optimisation space can be decreased. This is for example the case of tracking applications where only some elements (usually positional) of the state $\mathbf{x}_k$ are directly observed by the measurement $\mathbf{z}_k$, while other elements (velocity, acceleration) are not.

Suppose, the variable $\mathbf{x}$ can be split into two parts, $\mathbf{x}_a \in \mathbb{R}^{n_a}$ and $\mathbf{x}_b \in \mathbb{R}^{n_b}$, $n_b \geq 2$, i.e., $\mathbf{x} = [\mathbf{x}_a^{\mathrm{T}}, \mathbf{x}_b^{\mathrm{T}}]^{\mathrm{T}}$ with $n_a + n_b = n_x$, where only the part $\mathbf{x}_a$ is directly observable through $\mathbf{y}$. More specifically,

$$\begin{aligned}\mathbf{y} &= \mathbf{g}\left(\left[\begin{array}{c} \mathbf{x}_a \\ \hline \mathbf{x}_{b,1} \end{array}\right]\right) \\ &= \mathbf{g}\left(\left[\begin{array}{c} \mathbf{x}_a \\ \hline \mathbf{x}_{b,2} \end{array}\right]\right), \quad \forall \mathbf{x}_{b,1} \in \mathbb{R}^{n_b}, \quad \mathbf{x}_{b,2} \in \mathbb{R}^{n_b}. \end{aligned} \tag{49}$$

To show, that some rotation matrices do not affect the σ-points $\mathcal{Y}_i = \mathbf{g}(\mathcal{X}_i^r)$, $i = 0, \ldots 2n_x$, it is convenient to consider $\mathbf{S}^{\mathbf{xx}}$ being obtained by the Cholesky decomposition for which the independence will be proven. Later, the rotations that do not affect the σ-points $\mathcal{Y}_i$ will be specified for a general decomposition.

The decomposition $\mathbf{S}^{\mathbf{xx}}$ obtained by the Cholesky decomposition has the following form

$$\mathbf{S}_{\mathrm{Ch}}^{\mathbf{xx}} = \left[\begin{array}{c|c} \mathbf{L}_a & \mathbf{0}_{n_a \times n_b} \\ \hline \mathbf{F} & \mathbf{L}_b \end{array}\right], \tag{50}$$

where $\mathbf{L}_a \in \mathbb{R}^{n_a \times n_a}$ and $\mathbf{L}_b \in \mathbb{R}^{n_b \times n_b}$ are lower triangular matrices and $\mathbf{F} \in \mathbb{R}^{n_b \times n_a}$ is a full-rank matrix. Note that the notation $\mathbf{S}_{\mathrm{Ch}}^{\mathbf{xx}}$ emphasises that it is the Cholesky decomposition. The σ-points $\mathcal{X}_{i,a}^r$ related to $\mathbf{x}_a$ that do affect the σ-points $\mathcal{Y}_i$ are given by

$$\mathcal{X}_{0,a}^r = \hat{\mathbf{x}}_a, \tag{51}$$

$$\mathcal{X}_{j,a}^r = \hat{\mathbf{x}}_a + \sqrt{(n_x + \kappa)}[\mathbf{L}_a, \mathbf{0}_{n_a \times n_b}]\mathbf{e}_j, \tag{52}$$

$$\mathcal{X}_{n_x+j,a}^r = \hat{\mathbf{x}}_a - \sqrt{(n_x + \kappa)}[\mathbf{L}_a, \mathbf{0}_{n_a \times n_b}]\mathbf{e}_j, \tag{53}$$

where $j = 1, \ldots, n_x$, $\mathbf{e}_j$ is the jth column of $\mathbf{I}_{n_x}$ and $\hat{\mathbf{x}} = [\hat{\mathbf{x}}_a^{\mathrm{T}}, \hat{\mathbf{x}}_b^{\mathrm{T}}]^{\mathrm{T}}$.

Now, consider an arbitrary n_x-dimensional rotation matrix $\mathbf{C}'$ that rotates in $\mathbb{R}^{n_b}$ while the space $\mathbf{R}^{n_a}$ is unaffected. Such rotation matrix can be written as

$$\mathbf{C}' = \left[\begin{array}{c|c} \mathbf{I}_{n_a} & \mathbf{0}_{n_a \times n_b} \\ \hline \mathbf{0}_{n_b \times n_a} & \mathbf{C}_b \end{array}\right]. \tag{54}$$

Note that the rotation matrix of this form is the first rotation matrix in (48) with rotation angle θ_6.

The Cholesky decomposition $\mathbf{S}_{\mathrm{Ch}}^{\mathbf{xx},r}$ rotated by $\mathbf{C}'$ is given by

$$\mathbf{S}_{\mathrm{Ch}}^{\mathbf{xx},r} = \mathbf{S}_{\mathrm{Ch}}^{\mathbf{xx}}\mathbf{C}' = \left[\begin{array}{c|c} \mathbf{L}_a & \mathbf{0}_{n_a\times n_b} \\ \hline \mathbf{F} & \mathbf{L}_b \end{array}\right] \left[\begin{array}{c|c} \mathbf{I}_{n_a} & \mathbf{0}_{n_a\times n_b} \\ \hline \mathbf{0}_{n_b\times n_a} & \mathbf{C}_b \end{array}\right] = \left[\begin{array}{c|c} \mathbf{L}_a & \mathbf{0}_{n_a\times n_b} \\ \hline \mathbf{F} & \mathbf{L}_b\mathbf{C}_b \end{array}\right]. \quad (55)$$

From (55) it can be seen that the rotation matrix $\mathbf{C}'$ in (54) does not affect the first n_a rows of $\mathbf{S}_{\mathrm{Ch}}^{\mathbf{xx}}$. Hence, the rotation matrix $\mathbf{C}'$ does not affect the σ-points $\mathcal{X}_{i,a}^r$ and consequently the σ-points $\mathcal{Y}_i$, the mean $\hat{\mathbf{y}}^{\mathrm{UT}}(\mathbf{C}')$ and the covariance matrix $\mathbf{P}^{\mathbf{yy},\mathrm{UT}}(\mathbf{C}')$. This implies that rotation matrices in the form (54) do not have to be considered for optimisation. This gives $n_b(n_b-1)/2$ rotation angles that do not have to be optimised. Hence, the dimension of the optimisation space is now reduced to

$$n_\theta = [n_x(n_x-1)/2] - [n_b(n_b-1)/2]. \quad (56)$$

A general decomposition $\mathbf{S}^{\mathbf{xx}}$ can be written as $\mathbf{S}^{\mathbf{xx}} = \mathbf{S}_{\mathrm{Ch}}^{\mathbf{xx}}\bar{\mathbf{C}}$, where $\bar{\mathbf{C}}$ is an n_x dimensional rotation matrix. Then, the rotation matrices unnecessary for optimisation are given by $\bar{\mathbf{C}}^{\mathrm{T}}\mathbf{C}'$ as the rotated decomposition $\mathbf{S}^{\mathbf{xx},r}$ is given by

$$\mathbf{S}^{\mathbf{xx},r} = \mathbf{S}^{\mathbf{xx}}\bar{\mathbf{C}}^{\mathrm{T}}\mathbf{C}' = \mathbf{S}_{\mathrm{Ch}}^{\mathbf{xx}}\bar{\mathbf{C}}\bar{\mathbf{C}}^{\mathrm{T}}\mathbf{C}' = \mathbf{S}_{\mathrm{Ch}}^{\mathbf{xx}}\mathbf{I}_{n_x}\mathbf{C}', \quad (57)$$

in which case again the σ-points $\mathcal{X}_{i,a}^r$ are unaffected by the rotation $\bar{\mathbf{C}}^{\mathrm{T}}\mathbf{C}'$.

However, for the considered decrease of n_θ, a direct usage of the Cholesky decomposition $\mathbf{S}_{\mathrm{Ch}}^{\mathbf{xx}}$ is convenient due to parametrisation of the rotation matrix $\mathbf{C}$ by angles $\boldsymbol{\theta}$ as is shown in Section 3.5.

The proposed lossless decrease of optimisation space dimension does not affect the smallest error of $\hat{\mathbf{y}}^{\mathrm{UT}}(\mathbf{C})$ and $\mathbf{P}^{\mathbf{yy},\mathrm{UT}}(\mathbf{C})$ attainable by the optimisation. However, it does affect the cross-covariance matrix $\mathbf{P}^{\mathbf{xy}}$ (10). Therefore, when applied in the UKF, the predictive state σ-points $\mathcal{X}_{i,k|k-1}^r$ affect through $\mathbf{P}_{k|k-1}^{\mathbf{xz}}$ the gain $\mathbf{K}_k$ of the filter.

4.2. Lossy optimisation space dimension decreasing

For reduction of the computational costs, it is advisable to consider only the rotation angles that significantly affect the criterion value. For such analysis the usage of the criterion (43) is opportune. To set up an order of elements of $\mathbf{x}$ according to their influence on (43), the value $J_i^4(\mathbf{C})$ is calculated for $i = 1,\ldots,n_x$

$$J_i^4(\mathbf{C}) = [\tfrac{1}{4!}\mathbf{g}^{(4)}(\hat{\mathbf{x}})\tilde{\mathbf{M}}^{\mathbf{x}(4)}(\mathbf{C})\mathbf{J}_i]^{\mathrm{T}}\mathbf{W}[\tfrac{1}{4!}\mathbf{g}^{(4)}(\hat{\mathbf{x}})\tilde{\mathbf{M}}^{\mathbf{x}(4)}(\mathbf{C})\mathbf{J}_i] \quad (58)$$

where $\mathbf{J}_i = \mathrm{diag}[(\mathbf{e}_i)^{\otimes 4}]$ is a matrix selecting the element of $\tilde{\mathbf{M}}^{\mathbf{x}(4)}(\mathbf{C})$ corresponding to the ith element of $\mathbf{x}$. Then, the values $J_i^4(\mathbf{C})$, $i = 1,\ldots,n_x$ are sorted and the order and individual values provide information of magnitude of influence of $\mathbf{x}$ elements on the criterion. Using this information the user can decide over which rotation angles the chosen criterion should be optimised.

Note that for calculation of (58) any rotation matrix $\mathbf{C}$ can be used, e.g., $\mathbf{C} = \mathbf{I}$.

5. UKF WITH ROTATED σ-POINT SET

The concept of the σ-point set rotation can directly be extended from the UT to the UKF. The UKF with the rotated σ-point set has the same structure as the UKF given by Algorithm 1, where the filtering state σ-points and the predictive state σ-points respect the optimal rotation matrices.

When finding an optimal rotation matrix for the predictive state σ-points in the filtering step both criteria (43) and (44) can be used as the measurement $\mathbf{z}_k$ is available and can be used in place of the sample $\mathbf{y}_s$ in (44). To find an optimal rotation matrix for the filtering state σ-points in the predictive step, only the criterion (43) can be used.

The UKF with optimised rotation of the σ-point set is given by Algorithm 2.

ALGORITHM 2: *Unscented Kalman Filter with Rotated σ-point Set*

Step 1: Identical with **Step 1** of Algorithm 1.
Step 2: Find an optimal rotation matrix $\mathbf{C}_{k|k-1}$.
Step 3: Similar to **Step 2** of Algorithm 1 with $\mathbf{S}_{k|k-1}^{\mathbf{xx}}$ replaced by $\mathbf{S}_{k|k-1}^{\mathbf{xx},r} = \mathbf{S}_{k|k-1}^{\mathbf{xx}}\mathbf{C}_{k|k-1}$.
Step 4: Find an optimal rotation matrix $\mathbf{C}_{k|k}$.
Step 5: Similar to **Step 3** of Algorithm 1 with $\mathbf{S}_{k|k}^{\mathbf{xx}}$ replaced by $\mathbf{S}_{k|k}^{\mathbf{xx},r} = \mathbf{S}_{k|k}^{\mathbf{xx}}\mathbf{C}_{k|k}$.

Let $k = k+1$. The algorithm continues by **Step 2**.

The rotation matrices $\mathbf{C}^{*k|k}$ and $\mathbf{C}^{*k|k-1}$ can be in principle selected on on-line or off-line basis.

5.1. Off-line computed time-invariant and time-varying rotation matrix

The simplest choice of the rotation matrices is $\mathbf{C}_{k|k} = \mathbf{C}_{k|k-1} = \mathbf{I}_{n_x}$, $\forall k$. This corresponds to utilisation of a constant rotation of the σ-points that is determined by the selected covariance matrix decomposition technique.

Alternatively, a constant rotation matrix (or matrices different for the filtering and prediction step) might be found by e.g., a prior analysis of a considered problem as outlined in [6].

Another possibility is to find a sequence (or sequences) of the rotation matrices for a considered scenario based on a prior Monte-Carlo (MC) analysis. This approach is suitable for systems with set-up allowing periodical repetition of their dynamic behaviour. As an example, during the landing phase all aircrafts of a given category follow the same path (trajectory). Nevertheless, once the set-up is changed (either trajectory, sensor location, etc.) the procedure of finding the rotation matrices must be repeated. This is illustrated in numerical simulations in Section 6.

For the prior analysis either criterion (36), (43) or (44) may be used. Note that the rotation matrices $\mathbf{C}_{k|k-1}$, $\mathbf{C}_{k|k}$, if computed off-line, are not conditioned by the

past measurements of considered experiment as the notation indicates. The notation is, however, kept to be sufficiently general to cover also the possibility for the on-line computation of the matrices.

The UKF given by Algorithm 2 with rotation matrices $\mathbf{C}_{k|k-1}$ and $\mathbf{C}_{k|k}$ optimised off-line by means of a prior analysis will be denoted as trained-sigma-point-set-UKF (TUKF).

5.2. On-line computed time-varying rotation matrix

The on-line computation of the rotation matrix (or matrices) respects the current status of the filter. This procedure can be understood as a rotation matrix optimisation and in fact, a similar approach, in principle, has been used for the scaling parameter adaptation [33], [4] and is adopted here.

The rotation matrices used in the filtering and prediction steps, i.e., $\mathbf{C}_{k|k-1}$ and $\mathbf{C}_{k|k}$, can be computed by minimising (43) or (44).

The UKF given by Algorithm 2 with rotation matrices $\mathbf{C}_{k|k-1}$ and $\mathbf{C}_{k|k}$ optimised on-line using either (43) or (44) will be denoted as adaptive-sigma-point-set-UKF (AUKF).

5.3. Notes

Note 1: The on-line optimisation of the rotation matrix is clearly the most computationally demanding part of the algorithm. Moreover, the dimension of the optimisation space n_θ grows quadratically with the state dimension (see (47)). Selection of the rotation angles for optimisation and the optimisation technique is thus crucial. Two techniques for optimisation space dimension n_θ decrease have been proposed in Section 4. A brief discussion regarding the suitable optimisation technique can be found in [10]. In this paper, the grid method is preferred due to its simplicity. The method is summarised in Appendix A.

Note 2: The order of the successive rotations affects the overall rotation matrix. However, if a sufficiently dense grid is assumed, then the impact of the rotation order in the optimisation can be neglected.

Note 3: The rotation matrix optimisation might be performed independently in the predictive and the filtering steps. However, the impact of the adaptation in one step is often dominant and the adaptation in the second step can be skipped to reduce computational costs.

Note 4: The σ-point set rotation does not affect the UT performance if the function $\mathbf{g}(\cdot)$ in (3) is linear. Therefore, it is reasonable to supplement the UKF adaptively rotating of the σ-point set with an algorithm evaluating the severity of the nonlinearities at the actual working points. If the nonlinearity is mild at a given time, the adaptive selection of the rotation matrix might be skipped without any significant impact on the estimation performance. Such nonlinearity measures were proposed and integrated with local filters in [33].

Note 5: The paper considered the system (1) and (2) with additive noises. However, the concept can readily be extended for systems with non-additive noises. Just the optimisation space for rotations would be higher dimensional due to the state augmentation by the noises.

Note 6: Although the concept of the σ-point set rotation was discussed in the UKF framework, the σ-point set rotation adaptation can be used with any deterministic σ-point set based local filter.

6. NUMERICAL ILLUSTRATION

The performance of the UT and the UKF is affected by the σ-point set rotation and scaling. In this section, the UT and UKF are evaluated each using two examples mainly with respect to the σ-point set rotation influence.

6.1. Static example I: Fourth-order polynomial

Consider a random variable nonlinear transformation (as defined by (3)–(10)) with

$$\hat{\mathbf{x}} = \begin{bmatrix} \hat{x}_1 \\ \hat{x}_2 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \quad \mathbf{P}^{\mathbf{xx}} = \begin{bmatrix} 4 & 0.8 \\ 0.8 & 10 \end{bmatrix} \tag{59}$$

and

$$y = g(\mathbf{x}) = \mathbf{x}^{\mathrm{T}}\mathbf{x}\mathbf{x}^{\mathrm{T}}\mathbf{x}. \tag{60}$$

In Fig. 2, the impact of the σ-point set rotation through an angle θ on the approximate characteristics of y are illustrated. The characteristics are computed using the UT (8)–(10) with $\mathbf{S}^{\mathbf{xx}}$ computed by the SVD,[7] scaling parameters $\kappa = 1$ and $\kappa = 2$, and the rotation matrix used in (30)–(33) of the form

$$\mathbf{C} = \begin{bmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{bmatrix}. \tag{61}$$

Besides these characteristics, the true ones and the ones approximated using the UT with

- a recommended fixed $\kappa = 1$,
- a fixed $\kappa = 2$,

both with $\theta = 0$ [deg], are plotted in Fig. 2. Here, it should be noted, that the true and fixed UT characteristics are not functions of the rotation parameter θ, they are plotted for the whole range of the parameter just for ease of comparison. The figure shows a significant dependency of the UT performance on the σ-point set rotation. As shown in (42) under assumption of the Gaussian PDF of $\mathbf{x}$, the set rotation affects fourth- and higher-order even terms of the TE of the UT-based mean computation. Considering the mean value computation, i.e., $\hat{y}^{\mathrm{UT}}$ (8), in this polynomial example, the fourth-order term of the TE is the only term depending on the rotation. Therefore, the error (42) of $\hat{y}^{\mathrm{UT}}$ is equal to

$$\tilde{y}(\mathbf{C}) = \tfrac{1}{4!}\mathbf{g}^{(4)}(\hat{\mathbf{x}})(\mathbf{M}^{\mathbf{x}(4)} - \mathbf{M}^{\mathbf{x}(4),\mathrm{UT}}(\mathbf{C})), \tag{62}$$

[7] Throughout this section, the SVD of the covariance matrix $\mathbf{P}^{\mathbf{xx}}$ is considered. In this case, the σ-points lie on the principal axes of the covariance matrix ellipsoid [34].

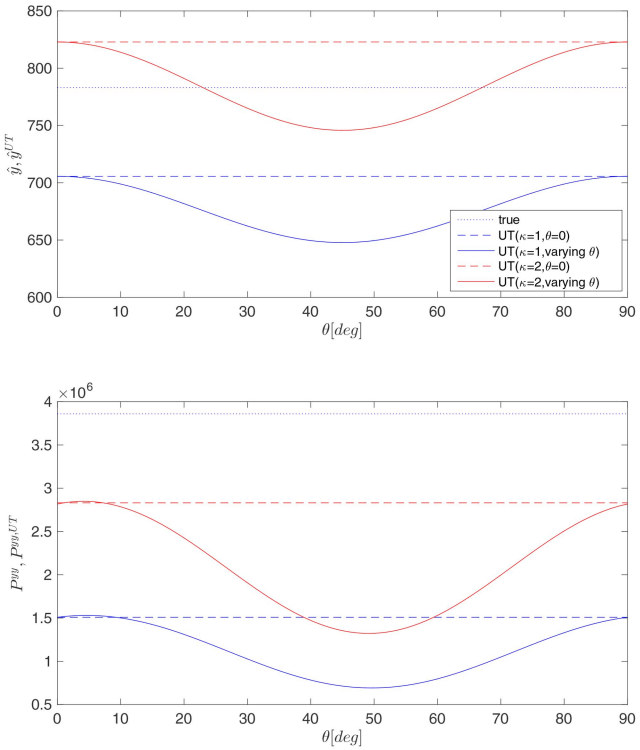


Fig. 2. Static example I: Dependence of the UT-based approximate characteristics on the rotation of σ-point sample set.

where

$$\mathbf{g}^{(4)}(\hat{\mathbf{x}}) = \left[\frac{\partial^4 g(\mathbf{x})}{\partial x_1^4}, \frac{\partial^4 g(\mathbf{x})}{\partial x_1^3 \partial x_2}, \frac{\partial^4 g(\mathbf{x})}{\partial x_1^3 \partial x_2}, \frac{\partial^4 g(\mathbf{x})}{\partial x_1^2 \partial x_2^2}, \frac{\partial^4 g(\mathbf{x})}{\partial x_1^3 \partial x_2}, \dots \right.$$
$$\frac{\partial^4 g(\mathbf{x})}{\partial x_1^2 \partial x_2^2}, \frac{\partial^4 g(\mathbf{x})}{\partial x_1^2 \partial x_2^2}, \frac{\partial^4 g(\mathbf{x})}{\partial x_1 \partial x_2^3}, \frac{\partial^4 g(\mathbf{x})}{\partial x_1^3 \partial x_2}, \frac{\partial^4 g(\mathbf{x})}{\partial x_1^2 \partial x_2^2}, \dots$$
$$\left. \frac{\partial^4 g(\mathbf{x})}{\partial x_1^2 \partial x_2^2}, \frac{\partial^4 g(\mathbf{x})}{\partial x_1 \partial x_2^3}, \frac{\partial^4 g(\mathbf{x})}{\partial x_1^2 \partial x_2^2}, \frac{\partial^4 g(\mathbf{x})}{\partial x_1 \partial x_2^3}, \frac{\partial^4 g(\mathbf{x})}{\partial x_1 \partial x_2^3}, \frac{\partial^4 g(\mathbf{x})}{\partial x_2^4} \right]$$
$$= [24,0,0,8,0,8,8,0,0,8,8,0,8,0,0,24]. \quad (63)$$

Hence the criterion (43) is a function of the fourth moments only and $\mathbf{g}^{(4)}(\hat{\mathbf{x}})$. Thus, it is equal to (36).

Value of the criterion (43) with $\mathbf{W} = 1$ as a function of the rotation θ is shown in Fig. 3 for both setting of κ.

6.2. Static example II: Arctangent

Arctangent is a nonlinear transformation appearing in tracking[8] applications for a conversion between Cartesian and polar coordinates. Contrary to the previous polynomial function, arctangent cannot by expressed by the TE with a finite number of terms.

Let a random variable nonlinear transformation with

$$\hat{\mathbf{x}} = \begin{bmatrix} \hat{x}_1 \\ \hat{x}_2 \end{bmatrix} = \begin{bmatrix} 10 \\ 1 \end{bmatrix}, \quad \mathbf{P}^{\mathbf{xx}} = \begin{bmatrix} 4 & 0.8 \\ 0.8 & 10 \end{bmatrix} \quad (64)$$

and

$$y = g(\mathbf{x}) = \operatorname{atan}(x_2/x_1). \quad (65)$$

[8]The four-quadrant inverse tangent function (atan2) is considered.

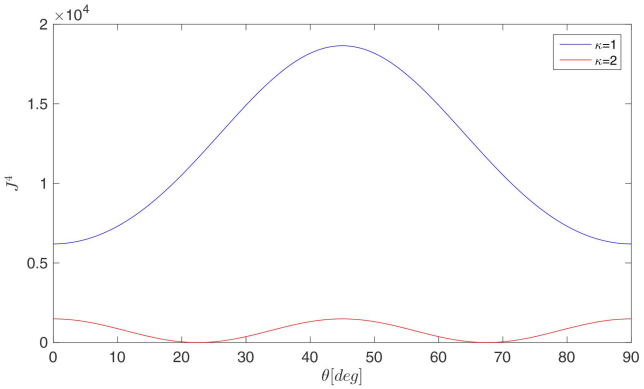


Fig. 3. Static example I: Dependence of the UT-based mean error on the rotation of σ-point sample set (for $\kappa = 1$).

be considered. In Fig. 4, the impact of both parameters affecting the UT performance, i.e., κ and θ, is illustrated for the mean $\hat{y}^{\mathrm{UT}}$ and variance $P^{yy,\mathrm{UT}}$ computation together with the true characteristics of y and the ones approximated using the UT with fixed $\kappa = 1$, and $\theta = 0$ [deg].

Fig. 4 indicates that both user-defined parameters (scaling and rotation) heavily impact the UT-approximated mean value $\hat{y}^{\mathrm{UT}}$ and the variance $P^{yy,\mathrm{UT}}$. By a suitable rotation, it is possible to get the true values of the mean and variance of y (for mean $\theta \approx 31$ and 79 [deg], for variance $\theta \approx 10$ and 58 [deg]; the values are highlighted by the vertical red lines). On the other hand, in this example it is not possible to get the true values of the statistics by any selection of the scaling parameter κ.

The value of the criterion $J(\mathbf{C}(\theta))$ (36) with $\mathbf{W} = 1$ is plotted in Fig. 5. The value is slightly different from the fourth-order moment based criterion $J^4(\mathbf{C}(\theta))$ (43) as, in this case, the rotation affects not only the fourth-order term of the TE but also all remaining higher-order even terms. However, the criterion $J^4(\mathbf{C}(\theta))$ (which can be computed without the knowledge of the true mean $\hat{y}$) still represents a reasonable approximation of the criterion $J(\mathbf{C}(\theta))$ (which cannot be computed without the knowledge of the true mean $\hat{y}$).

6.3. Dynamic example I: Bearings-only tracking

The impact of the σ-point set rotation either fixed or adaptive on the UKF performance is illustrated using the bearings-only tracking example where a manoeuvring object is tracked by a radar platform [25]. The object follows a course of -140 [deg] (the angles are referenced clockwise positive to the y axis) starting 12.2 [km] away from the platform at a constant speed of 4 [knots]. The platform follows a course of 140 [deg] at a constant speed of 5 [knots] and at the time interval $k = \langle 13,17 \rangle$ executes a manoeuvre to reach a new course of 18 [deg]. The initial positions are [12,2] [km] for the object and [0,0] [km] for the platform. The geometry of the motion is depicted in Fig. 6. The object

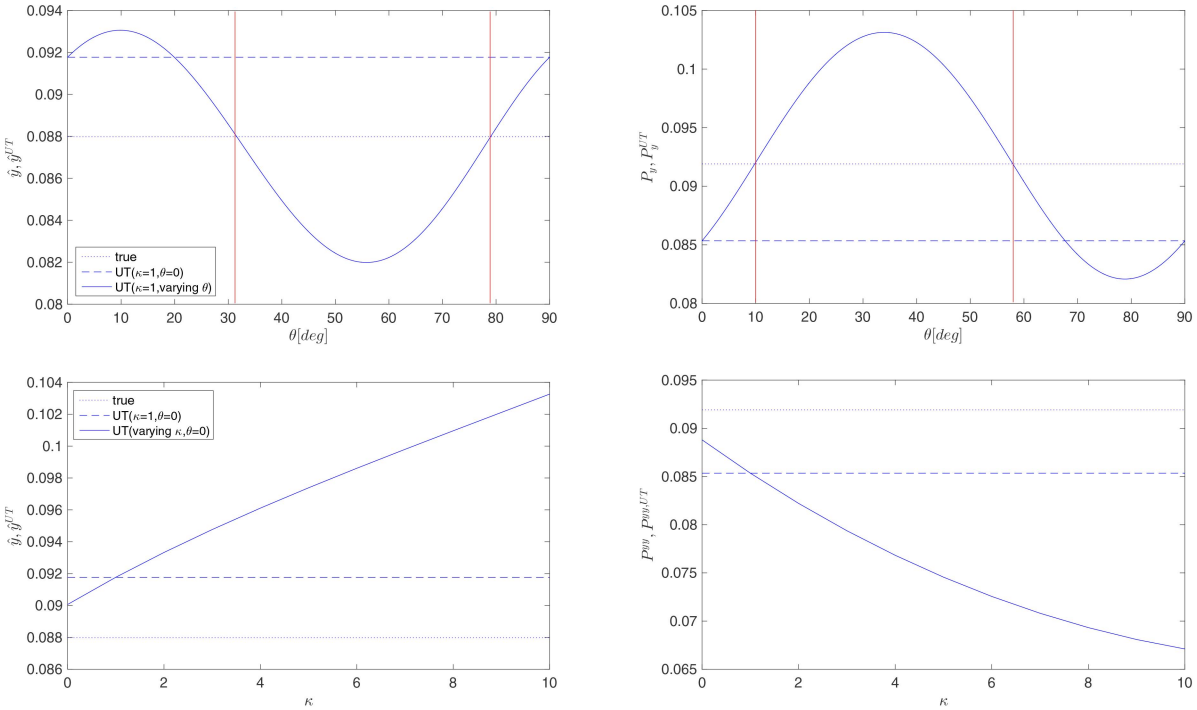


Fig. 4. Static example II: Dependence of the UT-based approximate characteristics on the rotation and scaling of σ-point sample set (together with the true and fixed UT characteristics).

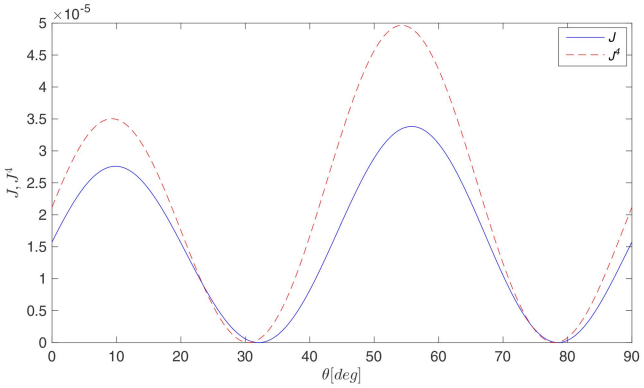


Fig. 5. Static example II: Dependence of the UT-based mean error on the rotation of σ-point sample set (for $\kappa = 1$).

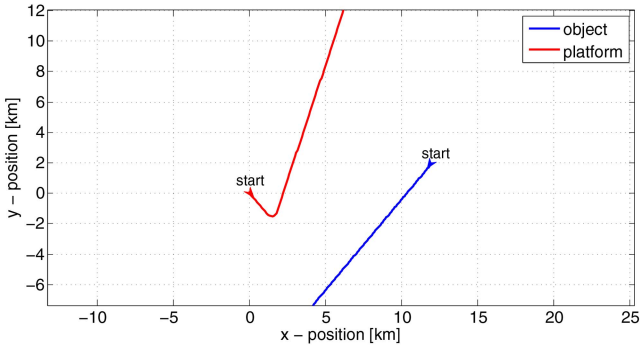


Fig. 6. Geometry of bearings-only tracking example.

motion (relative to the platform) is modelled by a continuous white noise acceleration model [3]. The state of the model $\mathbf{x}_k = [x_{1,k}, x_{2,k}, x_{3,k}, x_{4,k}]^{\mathrm{T}}$ consists of the positions in x and y-directions ($[x_{1,k}, x_{2,k}]$) and respective velocities ($[x_{3,k}, x_{4,k}]$) which evolves as

$$\mathbf{x}_{k+1} = \mathbf{F}\mathbf{x}_k + \mathbf{G}\mathbf{w}_k, \tag{66}$$

with

$$\mathbf{F} = \begin{bmatrix} 1 & 0 & T & 0 \\ 0 & 1 & 0 & T \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}, \quad \mathbf{G} = \begin{bmatrix} 0.5T^2 & 0 \\ 0 & 0.5T^2 \\ T & 0 \\ 0 & T \end{bmatrix}, \tag{67}$$

where $T = 1$ [min] is the sampling interval, $k = 0, 1, \ldots, K = 100$, and $\mathbf{w}_k$ is the Gaussian zero-mean state noise with covariance matrix $\mathbf{\Sigma}^w$ with $\mathbf{\Sigma}^w = 10^{-4}\mathbf{I}_2$ $\times[\mathrm{km}^2/\mathrm{sec}^4]$.

The measurement z_k providing the relative angle of the object w.r.t. the platform at time k is

$$z_k = \arctan \frac{x_{1,k} - x^p_{1,k}}{x_{2,k} - x^p_{2,k}} + v^\theta_k, \tag{68}$$

where $[x^p_{1,k}, x^p_{2,k}]$ is the known platform position and the variance of the measurement noise is $\Sigma^v = (3\ [\mathrm{deg}])^2$.

In total $M = 10^3$ MC simulations were carried out and the UKF with different settings of the user-defined parameters were simulated. In particular, the following UKFs and CKFs are considered

- UKF(svd)—UKF with $\mathbf{S}^{\mathbf{xx}}$ computed by the SVD,
- UKF(chol)—UKF with $\mathbf{S}^{\mathbf{xx}}$ computed by the Cholesky decomposition,

- AUKF(θ_{12},$J^4(\mathbf{C})$)—AUKF with the Cholesky decomposition and rotation matrix computed on the basis of θ_{12} with the criterion $J^4(\mathbf{C})$ (43),
- AUKF(θ_{12})—AUKF with the Cholesky decomposition and rotation matrix computed on the basis of θ_{12} with the criterion $J^{\mathrm{MS}}(\mathbf{C})$ (46),
- AUKF(θ_{34})—AUKF with the Cholesky decomposition and rotation matrix computed on the basis of θ_{34} with the criterion $J^{\mathrm{MS}}(\mathbf{C})$ (46),
- AUKF($\boldsymbol{\theta}_{12,23,34}$)—AUKF with the Cholesky decomposition and rotation matrix computed on the basis of $\boldsymbol{\theta}_{12,23,34}$ with the criterion $J^{\mathrm{MS}}(\mathbf{C})$ (46),
- AUKF($\boldsymbol{\theta}_{all}$)—AUKF with the Cholesky decomposition and rotation matrix computed on the basis of $\boldsymbol{\theta}_{all}$ with the criterion $J^{\mathrm{MS}}(\mathbf{C})$ (46),
- TUKF—UKF with the Cholesky decomposition and off-line identified sequence of rotation matrices,
- CKF5(svd)—CKF of the fifth-order proposed in [15] with $\mathbf{S}^{\mathbf{xx}}$ computed by the SVD,
- CKF5(rot)—CKF5 with the SVD and fixed rotation of the σ-points through 30 [deg],
- ACKF5(θ_{12})—CKF with the SVD and rotation matrix computed on the basis of θ_{12} with the criterion $J^{\mathrm{MS}}(\mathbf{C})$ (46),

where θ_{12} denotes the rotation in the plane of the first two state vector elements, i.e., the respective rotation matrix is of the form

$$\mathbf{C}(\theta_{12}) = \begin{bmatrix} \cos(\theta_{12}) & -\sin(\theta_{12}) & 0 & 0 \\ \sin(\theta_{12}) & \cos(\theta_{12}) & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \tag{69}$$

etc., and $\boldsymbol{\theta}_{all} = [\theta_{23},\theta_{12},\theta_{13},\theta_{14},\theta_{24}]$. It means that rotation θ_{12} rotates the part of sigma points relative to the position, where as θ_{34} rotates relative to the velocity components. Note that all the considered filters were run with the recommended scaling parameter $\kappa = 0$.

Hence, the filters AUKF(θ_{12},$J^4(\mathbf{C})$), AUKF(θ_{12}), AUKF(θ_{34}), and AUKF($\boldsymbol{\theta}_{12,23,34}$) used the lossy decrease of n_θ proposed in Section 4.2 with $n_\theta = 1$, $n_\theta = 1$, $n_\theta = 1$, and $n_\theta = 3$, respectively. The filter AUKF($\boldsymbol{\theta}_{all}$) used the lossless decrease of n_θ proposed in Section 4.1 and thus ignored rotation with the angle θ_{34}.

Note that the filter AUKF(θ_{34}) was used to demonstrate that the rotation computed on the basis of θ_{34}, i.e., in the plane defined by the state elements that are not directly measurable, has almost no effect on the estimation quality.

The AUKFs use the grid optimisation technique [10] with the grid defined as $\theta_{ij} = \{0,15,\ldots,75\}$ [deg], $\forall i,j$ which results in 6 different values used in optimisation for the AUKF(θ_{12}), AUKF(θ_{34}), in $6^3 = 216$ values for the AUKF($\theta_{12,23,34}$), and in $6^5 = 7776$ for the AUKF($\boldsymbol{\theta}_{all}$). The grid optimisation method is further discussed in Appendix A.

The TUKF takes advantage of the optimal rotation θ_{12} found prior to the estimation experiment on the basis of a set of MC simulations. In particular 10^3 simulations of the AUKF(θ_{12}) were carried out and the average (sample-based) optimal rotation angle was computed as

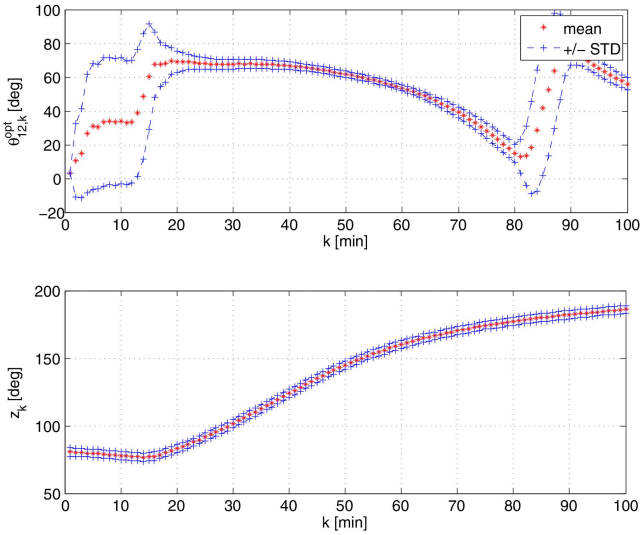


Fig. 7. Dynamic example I: Averaged optimal rotation angle θ_{12} used in TUKF and the averaged measurement sequence.

$$\theta^{\mathrm{opt}}_{12,k|k-1} = \frac{1}{10^3}\sum_{i=1}^{10^3}\theta^{(i)*}_{12,k|k-1}, \quad \forall k, \tag{70}$$

where $\theta^{(i)*}_{12,k|k-1}$ is the optimal rotation at time k of ith MC simulation. The average optimal rotation was used for the rotation matrix $\mathbf{C}^{\mathrm{opt}}_{k|k-1}(\theta_{12})$ computation in the TUKF. The optimal rotation angle is plotted in Fig. 7 together with the respective sample standard deviation (STD). In Fig. 7 the sample statistics of the measurement z_k, i.e., the mean and STD, over the MC simulations are given as well.

The filters were initialised according to [25] with initial range pdf $p(r) = \mathcal{N}\{r;\sqrt{12^2+2^2},4^2\}$ and a speed pdf $p(s) = \mathcal{N}\{s;\bar{s},4^2\}$, where $\bar{s}$ is the true speed.

The filter results were compared using the root mean square error (RMSE) defined as

$$\mathrm{RMSE}^{\mathrm{pos}}_k = \sqrt{\frac{1}{M}\sum_{m=1}^{M}(\hat{x}^{(m)}_{1,k|k} - x^{(m)}_{1,k})^2 + (\hat{x}^{(m)}_{2,k|k} - x^{(m)}_{2,k})^2} \tag{71}$$

for the position error and

$$\mathrm{RMSE}^{\mathrm{vel}}_k = \sqrt{\frac{1}{M}\sum_{m=1}^{M}(\hat{x}^{(m)}_{3,k|k} - x^{(m)}_{3,k})^2 + (\hat{x}^{(m)}_{4,k|k} - x^{(m)}_{4,k})^2} \tag{72}$$

for the velocity error and using the non-credibility index (NCI) [20] defined as

$$\mathrm{NCI}_k = \frac{1}{M}\sum_{m=1}^{M}[10\log_{10}((\tilde{\mathbf{x}}^{(m)}_{k|k})^{\mathrm{T}}(\mathbf{P}^{\mathbf{xx}(m)}_{k|k})^{-1}\tilde{\mathbf{x}}^{(m)}_{k|k}) - 10\log_{10}((\tilde{\mathbf{x}}^{(m)}_{k|k})^{\mathrm{T}}\boldsymbol{\Sigma}^{-1}_k\tilde{\mathbf{x}}^{(m)}_{k|k})], \tag{73}$$

TABLE I
Dynamic example I: RMSE, NCI, and time results of bearings-only tracking.

	UKF(svd)	UKF(chol)	AUKF(θ_{12})	AUKF(θ_{12}, $J^4(\mathbf{C})$)	AUKF(θ_{34})	AUKF($\theta_{12,13,24}$)	AUKF($\boldsymbol{\theta}_{all}$)	TUKF
RMSE^{pos} [km]	3.30	3.40	2.90	2.80	3.30	2.70	2.70	3.10
RMSE^{vel} [km/sec]	0.16	0.16	0.15	0.15	0.15	0.15	0.14	0.16
NCI	4.60	4.80	3.80	3.80	4.60	3.40	2.90	4.40
time [msec]	12.8	9.5		86		3175	104191	10.6

	CKF5(svd)	CKF5(rot)	ACKF5(θ_{12})
RMSE^{pos} [km]	3.10	3.50	2.80
RMSE^{vel} [km/sec]	0.15	0.16	0.14
NCI	4.15	5.10	3.50
time [msec]	13.30		93

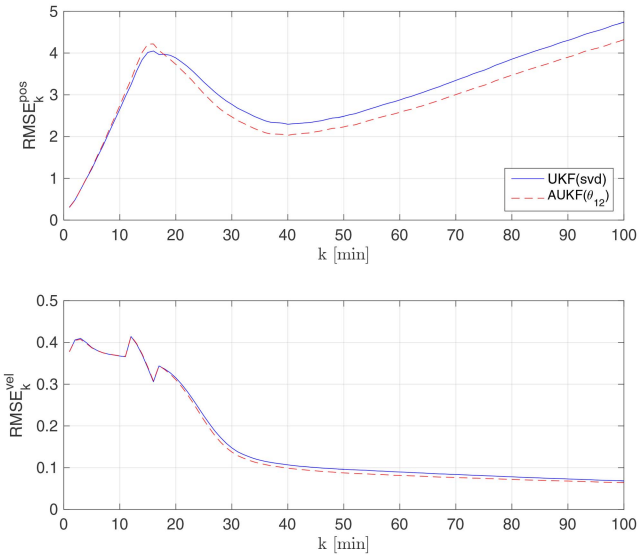


Fig. 8. Dynamic example I: Time behaviour of the RMSE for UKF(svd) and the AUKF(θ_{12}).

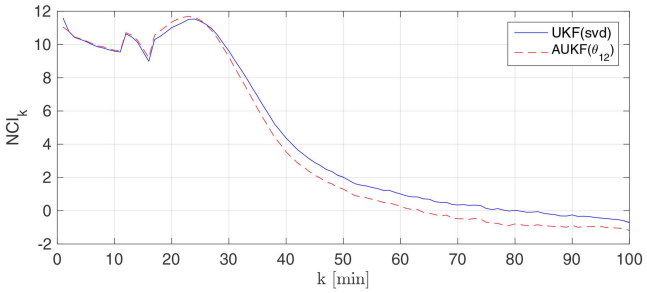


Fig. 9. Dynamic example I: Time behaviour of the NCI for UKF(svd) and the AUKF(θ_{12}).

where $\tilde{\mathbf{x}}_{k|k}^{(m)} \triangleq (\mathbf{x}_k^{(m)} - \hat{\mathbf{x}}_{k|k}^{(m)})$, $\mathbf{x}_k^{(m)}$ is the true state, $\hat{\mathbf{x}}_{k|k}^{(m)}$ and $\mathbf{P}_k^{\mathbf{xx}(m)}$ are the filtering mean and covariance matrix of the estimate provided by the filter at the mth MC run and $\boldsymbol{\Sigma}_k$ is the mean square error. Whereas the RMSE compares just the quality of the mean estimate (the lower RMSE value, the better performance), the NCI assesses the credibility of the estimate, i.e., whether the error covariance matrix of the filter corresponds to the real mean square error of the state estimate. The NCI value should be ideally zero. Negative NCI values indicate pessimistic estimates of the filter, positive NCI values, on the other hand, imply optimistic estimates. The averaged values of the RMSE and the NCI over all time instants are given in Table I. Illustration of time behaviour of the RMSE and the NCI for the UKF(svd) and the AUKF(θ_{12}) is depicted in Fig. 8 and Fig. 9.

It can be seen that the UKF(svd) slightly outperforms the UKF(chol) at the cost of mild increase of the computational complexity. The further improvement can be reached by using the off-line found averaged optimal rotation $\theta_{12,k|k-1}^{\text{opt}}$ with negligible computational requirements increase.

The results for the AUKFs confirm the theoretical analysis of Section 4.1 that the largest impact on the estimation quality is tied with the rotation in the planes defined by the state components which are "directly" measured, i.e., the positions in this example. The major improvement is caused by the rotation in the position plane, i.e., through the angle θ_{12}. Then, the additional rotations through the angles θ_{23} and θ_{13} reduces the RMSE and the NCI further. Adding the adaptive selection of the optimal rotation also in the remaining directions (five in total) brings almost no benefit. This is also confirmed by the results of the AUKF(θ_{34}) which indicates that the rotation in the plane defined by the velocities in x and y directions does not lead to any performance improvement.

Concerning the optimisation criteria, there is no significant difference in using the criterion $J^4(\mathbf{C})$ (43) or the criterion $J^{\text{MS}}(\mathbf{C})$ (46). With respect to the fact that the criterion $J^4(\mathbf{C})$ requires computation and evaluation of the fourth-order derivatives, the criterion $J^{\text{MS}}(\mathbf{C})$ seems to be more suitable in the considered set-up.

From the results it is evident that the lossy decrease of n_θ may save a considerable amount of computational costs. If only the directly observable state elements are considered in the lossy decrease of n_θ, the decrease of estimate quality is negligible in this illustration.

TABLE II
Dynamic example II: RMSE and NCI of two-dimensional example.

	UKF	AUKF(15[deg])	AUKF(30[deg])
RMSE	14.70	8.90	8.95
NCI	11.10	11.00	11.10

For the sake of completeness, the fifth-order CKF, the CKF5 [15], was tested as well. Compared to the UKF, the CKF5 processes $2n_x^2 + 1$ σ-points. Therefore, it provides estimates of better quality, in this example, in both monitored criteria at the costs of slightly higher computational complexity. The performance of the CKF5 is, however, still affected by the rotation of the σ-point set. Fixed σ-point set rotation may improve but also worsen the estimation performance (as can be seen for CKF5(rot)). In principle, the same rotation adaptation criteria may be used as for the UKF. The CKF5 with the adaptation, the ACKF5, according to the criterion (46), then offers further improvement to the estimation performance.

6.4. Dynamic example II: Nonlinear two-dimensional system

The second dynamic example considers a nonlinear system introduced in [37] described by

$$\mathbf{x}_{k+1} = \begin{bmatrix} x_{1,k+1} \\ x_{2,k+1} \end{bmatrix} = \begin{bmatrix} 3\sin(5x_{2,k}^2) \\ x_{1,k} + e^{-0.05x_{2,k}} + 10 \end{bmatrix} + \mathbf{w}_k, \tag{74}$$

$$z_k = \cos(x_{1,k}) + x_{2,k}^2 + v_k, \tag{75}$$

with $p(\mathbf{w}_k) = \mathcal{N}\{\mathbf{w}_k : \mathbf{0}, 6\mathbf{I}_2\}$, $p(v_k) = \mathcal{N}\{v_k : 0, 1\}$, $p(\mathbf{x}_0) = \mathcal{N}\{\mathbf{x}_0 : [-0.7, 1]^{\mathrm{T}}, \mathbf{I}_2\}$, and $k = 0, 1, \ldots, 100$.

Three filters, namely

- UKF(svd),
- AUKF(15 [deg])—AUKF with the SVD with the criterion $J^{\mathrm{MS}}(\mathbf{C})$ (46) and grid-based optimisation with step of 15 [deg],
- AUKF(30 [deg])—AUKF with the SVD with the criterion $J^{\mathrm{MS}}(\mathbf{C})$ (46) and grid-based optimisation with step of 30 [deg],

all with $\kappa = 1$, were compared in terms of the averaged RMSE (71) and NCI (73). The simulation results are summarised in Table II.

Adaptation in the σ-point set rotation significantly improves the UKF performance in terms of the RMSE. The NCI remains almost unaffected, which means that the optimisation reduces the estimation error and estimated covariance matrix proportionally. The results also reveal that the AUKF is not heavily dependent on the optimisation grid density in this example. Doubling the optimisation grid points does not have almost and impact on the estimation performance while the computational requirements of AUKF(30[deg]) are half of the AUKF(15[deg]) requirements.

7. CONCLUDING REMARKS

The paper dealt with an analysis of the σ-point set rotation in the derivative-free approximations used in the local filter design, namely in the unscented transform being a cornerstone of the unscented Kalman filter. It was shown that the covariance matrix decomposition, used in σ-point computation, can be multiplied by an arbitrary rotation matrix. The matrix can be then viewed as the user-defined parameter significantly impacting the filter performance. In principle, two different approaches for selection of the appropriate rotation were proposed; off-line approach determining the optimal rotation prior to the estimation experiment (in certain recurrent scenarios even time-varying) leading to the TUKF, or an on-line approach computing the optimal rotation during the experiment respecting the actual conditions leading to the AUKF. To avoid excessive increase of computational costs due to the optimisation of the rotation angles, two techniques were proposed to decrease dimension of the optimisation space. The proposed approaches for rotation matrix optimisation were illustrated and compared in terms of the estimation performance and the computational complexity in several numerical examples.

Note that algorithms of the UKF with rotated σ-point set are part of the Nonlinear Estimation Framework available at http://nft.kky.zcu.cz/.

APPENDIX A GRID-BASED OPTIMISATION METHOD USED IN AUKF

Computationally efficient adaptation of the σ-point set rotation matrix is the key enabler allowing effective usage of the AUKFs.

Any optimisation technique used for the solution to (43) and (44) requires specification of the range for the parameters being optimised as the task is generally nonlinear. If the rotation matrix parameterisation by a sequence of the subsequent rotations is selected, then the simplest way is to define the range for each particular rotation (omitting time indices) as

$$\theta_i = \langle 0, 360 \rangle, \quad i = 1, 2, \ldots, n_\theta, \tag{76}$$

where n_θ is the number of subsequent rotations (47). However, because of the symmetry of the considered σ-point set, the interval can be significantly reduced to

$$\theta_i = \langle 0, 90 \rangle, \quad \forall i, \tag{77}$$

without the loss of generality. This is illustrated for a two- and a three-dimensional case below assuming the grid-based optimisation method used in the filtering step.

The grid-based optimisation method simply covers the interval θ_i (77) by a grid of equidistantly placed points, i.e.,

$$\theta_i^{\mathrm{grid}} = \{\theta_i^{(1)}, \theta_i^{(2)}, \ldots, \theta_i^{(G)}\}, \quad \forall i. \tag{78}$$

If $n_\theta > 1$, then the final (multidimensional) grid is given by the Cartesian product of the particular grids as

$$\boldsymbol{\theta}^{\text{grid}} = \theta_1^{\text{grid}} \times \theta_2^{\text{grid}} \times \cdots \times \theta_{n_\theta}^{\text{grid}}. \tag{79}$$

The cardinality of the grid-points set (79) is

$$n_{\text{grid}} = G^{n_\theta}, \tag{80}$$

thus exponentially growing with the number of the rotation angles θ_i.

The criterion function is then evaluated at all points of the set $\boldsymbol{\theta}^{\text{grid}}$ (79) and the optimum point is chosen according to

$$\boldsymbol{\theta}^* = \arg\min_{\boldsymbol{\theta}^{\text{grid}}} J(\mathbf{C}(\boldsymbol{\theta})). \tag{81}$$

The chosen vector $\boldsymbol{\theta}^*$ is subsequently used for the rotation matrix computation as illustrated by (48).

1.1. Two dimensional case

Considering $n_x = 2$, $\hat{\mathbf{x}} = \mathbf{0}_{2\times 1}$, $\mathbf{P}^{\mathbf{xx}} = \mathbf{I}_2 = [\mathbf{e}_1, \mathbf{e}_2]$, and $(n_x + \kappa) = 1$, the σ-point set computed according to (4)–(6) is

$$\mathcal{X}_0 = \mathbf{0}_{2\times 1},\ \mathcal{X}_1 = \mathbf{e}_1,\ \mathcal{X}_2 = \mathbf{e}_2,\ \mathcal{X}_3 = -\mathbf{e}_1,\ \mathcal{X}_4 = -\mathbf{e}_2, \tag{82}$$

where $\mathbf{0}_{n_x\times 1}$ is the zero matrix of the indicated dimension. The number of available rotation angles (47) is $n_\theta = 1$, therefore $\boldsymbol{\theta}^{\text{grid}} = \theta_1^{\text{grid}}$, and the respective covariance matrix $\mathbf{C}$ is computed according to (61).

Then, it is not difficult to see that the σ-point set rotation through an angle $\theta_h \in \langle 90, 180)$ [deg] results in virtually the same σ-point set as for the rotation $\theta_l = (\theta_h - 90) \in \langle 0, 90)$ [deg]. It means that

$$\mathcal{X}_0^r = \mathbf{C}(\theta_h)\mathcal{X}_0 = \mathbf{C}(\theta_l)\mathcal{X}_0 = \mathcal{X}_0, \tag{83}$$

$$\begin{aligned}\mathcal{X}_i^r &= \mathbf{C}(\theta_h)\mathcal{X}_i = \mathbf{C}(\theta_h - 90)\mathbf{C}(90)\mathcal{X}_i = \mathbf{C}(\theta_h - 90)\mathcal{X}_{i+1} \\ &= \mathbf{C}(\theta_l)\mathcal{X}_{i+1},\end{aligned} \tag{84}$$

$$\mathcal{X}_4^r = \mathbf{C}(\theta_h)\mathcal{X}_4 = \mathbf{C}(\theta_l)\mathcal{X}_1, \tag{85}$$

where $\mathbf{C}(90) = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$ and $i = 1, 2, 3$.

As the σ-points $\mathcal{X}_1, \ldots, \mathcal{X}_4$ are weighted equally and the UT (8)–(10) does not reflect the σ-point order, the UT with the rotated σ-points (30)–(32) provides the same results for

$$\mathbf{C}(\theta) = \mathbf{C}(\theta + 90) = \mathbf{C}(\theta + 180) = \mathbf{C}(\theta + 270). \tag{86}$$

assuming θ in [deg]. That means the optimisation of the criterion (81) does not bring any benefit if performed on the interval $\theta \in \langle 0, 360)$ [deg] w.r.t. optimisation on $\theta \in \langle 0, 90)$ [deg].

1.2. Three dimensional case

Considering $n_x = 3$, $\hat{\mathbf{x}} = \mathbf{0}_{3\times 1}$, $\mathbf{P}^{\mathbf{xx}} = \mathbf{I}_3 = [\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3]$, and $(n_x + \kappa) = 1$, the σ-point set computed according to (4)–(6) is

$$\begin{aligned}&\mathcal{X}_0 = \mathbf{0}_{3\times 1},\ \mathcal{X}_1 = \mathbf{e}_1,\ \mathcal{X}_2 = \mathbf{e}_2,\ \mathcal{X}_3 = \mathbf{e}_3,\ \mathcal{X}_4 = -\mathbf{e}_1, \\ &\mathcal{X}_5 = -\mathbf{e}_2,\ \mathcal{X}_6 = -\mathbf{e}_3.\end{aligned} \tag{87}$$

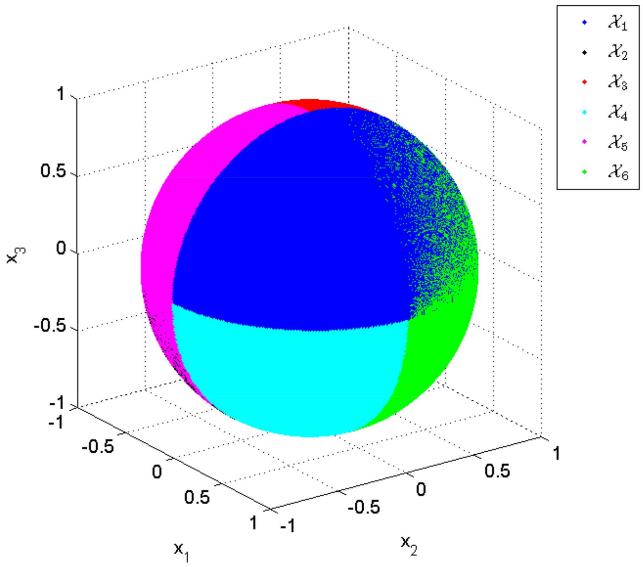


Fig. 10. Illustration of the σ-point set rotation in three dimensional case.

The number of available rotation angles (47) is $n_\theta = 3$, therefore $\boldsymbol{\theta}^{\text{grid}} = \theta_1^{\text{grid}} \times \theta_2^{\text{grid}} \times \theta_3^{\text{grid}}$, and the respective covariance matrix $\mathbf{C}$ is computed according to

$$\begin{aligned}\mathbf{C}(\boldsymbol{\theta}) &= \mathbf{C}_3(\theta_3)\mathbf{C}_2(\theta_2)\mathbf{C}_1(\theta_1) \\ &= \begin{bmatrix} 1 & 0 & 1 \\ 0 & \cos(\theta_3) & -\sin(\theta_3) \\ 0 & \sin(\theta_3) & \cos(\theta_3) \end{bmatrix} \begin{bmatrix} \cos(\theta_2) & 0 & -\sin(\theta_2) \\ 0 & 1 & 0 \\ \sin(\theta_2) & 0 & \cos(\theta_2) \end{bmatrix} \\ &\quad \times \begin{bmatrix} \cos(\theta_1) & -\sin(\theta_1) & 0 \\ \sin(\theta_1) & \cos(\theta_1) & 0 \\ 0 & 0 & 1 \end{bmatrix}.\end{aligned} \tag{88}$$

Similarly to the two dimensional case, the σ-point set rotated through the angle(s) greater than 90° is analysed and it is shown that such rotation results again in virtually the same σ-point set rotated just within the range from 0° to 90°. This is illustrated by the rotation $\theta_{1,h} \in \langle 90, 180)$ [deg] as

$$\mathcal{X}_0^r = \mathbf{C}_{3,2}\mathbf{C}_1(\theta_{1,h})\mathcal{X}_0 = \mathbf{C}_{3,2}\mathbf{C}_1(\theta_{1,l})\mathcal{X}_0 = \mathcal{X}_0, \tag{89}$$

$$\begin{aligned}\mathcal{X}_i^r &= \mathbf{C}_{3,2}\mathbf{C}_1(\theta_{1,h})\mathcal{X}_i = \mathbf{C}_{3,2}\mathbf{C}_1(\theta_{1,h} - 90)\mathbf{C}_1(90)\mathcal{X}_i \\ &= \mathbf{C}_{3,2}\mathbf{C}(\theta_{1,h} - 90)\mathcal{X}_j = \mathbf{C}_{3,2}\mathbf{C}(\theta_{1,l})\mathcal{X}_j,\end{aligned} \tag{90}$$

where $\theta_{1,l} \in \langle 0, 90)$ [deg], $\mathbf{C}_{3,2} = \mathbf{C}_3(\cdot)\mathbf{C}_2(\cdot)$,

$$\mathbf{C}_1(90) = \begin{bmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \tag{91}$$

$i,j = 1,2,\ldots,6$, and

$$\mathbf{C}_1(90)\mathcal{X}_1 = [0,1,0]^{\mathrm{T}} = \mathcal{X}_2, \tag{92}$$

$$\mathbf{C}_1(90)\mathcal{X}_2 = [-1,0,0]^{\mathrm{T}} = \mathcal{X}_4, \tag{93}$$

$$\mathbf{C}_1(90)\mathcal{X}_3 = [0,1,0]^{\mathrm{T}} = \mathcal{X}_3, \tag{94}$$

$$\mathbf{C}_1(90)\mathcal{X}_4 = [0,-1,0]^{\mathrm{T}} = \mathcal{X}_5, \tag{95}$$

$$\mathbf{C}_1(90)\mathcal{X}_5 = [1,0,0]^{\mathrm{T}} = \mathcal{X}_1, \tag{96}$$

$$\mathbf{C}_1(90)\mathcal{X}_6 = [0,0,-1]^{\mathrm{T}} = \mathcal{X}_6. \tag{97}$$

The σ-points $\mathcal{X}_1,\ldots,\mathcal{X}_6$ are again weighted equally and the UT (8)–(10) does not reflect the σ-point order, therefore, the optimisation for an interval $\theta_i \in \langle 0,\theta_{i,max}\rangle$ [deg] with $\theta_{i,max} > 90^\circ$ cannot improve the results.

For illustration, the set of σ-points (87) rotated for the Cartesian product $\boldsymbol{\theta}^{\mathrm{grid}}$ with $\theta_i^{\mathrm{grid}} = \{0,1,\ldots,89\}$, $i = 1,2,3$, is shown in Fig. 10.

REFERENCES

[1] B. D. O. Anderson and J. B. Moore
Optimal Filtering.
New Jersey: Prentice Hall Ins.: Englewood Cliffs, 1979.

[2] I. Arasaratnam and S. Haykin
Cubature Kalman filters.
IEEE Transactions on Automatic Control, vol. 54, no. 6 (2009), 1254–1269.

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

[4] L. Chang, B. Hu, and F. Qin
Unscented type Kalman filter: Limitation and combination.
IET Signal Processing, vol. 7, no. 3 (2013), 167–176.

[5] A. Doucet, N. De Freitas, and N. Gordon
Sequential Monte Carlo Methods in Practice.
Springer, 2001.

[6] J. Duník, O. Straka, and M. Šimandl
On sigma-point set rotation in derivative-free filters.
In *Proceedings of the 17th International Conference on Information Fusion*, Salamanca, Spain, July 2014, 1–8.

[7] J. Duník, O. Straka, and M. Šimandl
Sigma-point set rotation in unscented Kalman filter: Analysis and adaptation.
In *Proceedings of the 19th IFAC World Congress*, Cape Town, South Africa, September 2014, 5951–5956.

[8] J. Duník, O. Straka, M. Šimandl, and E. Blasch
Random-point-based filters: Analysis and comparison in target tracking.
IEEE Transactions on Aerospace and Electronic Systems, vol. 51, no. 2 (2015), 1403–1421.

[9] J. Duník, O. Straka, and M. Šimandl
Stochastic integration filter.
IEEE Transactions on Automatic Control, vol. 58, no. 6 (2013), 1561–1566.

[10] J. Duník, M. Šimandl, and O. Straka
Adaptive choice of scaling parameter in derivative-free local filters.
In *Proceedings of the 2010 International Conference on Information Fusion*, Edinburgh, Great Britain, 2010, 1–8.

[11] J. Duník, M. Šimandl, and O. Straka
Unscented Kalman filter: Aspects and adaptive setting of scaling parameter.
IEEE Transactions on Automatic Control, vol. 57, no. 9 (2012), 2411–2416.

[12] P. D. Groves
Principles of GNSS, Inertial, and Multisensor Integrated Navigation Systems (2nd Edition).
Artech House, 2013.

[13] K. Ito and K. Xiong
Gaussian filters for nonlinear filtering problems.
IEEE Transactions on Automatic Control, vol. 45, no. 5 (2000), 910–927.

[14] B. Jia, M. Xin, and Y. Cheng
Sparse Gauss-Hermite quadrature filter with application to spacecraft attitude estimation.
Journal of Guidance, Control, and Dynamics, vol. 34, no. 2 (2011), 367–379.

[15] B. Jia, M. Xin, and Y. Cheng
High-degree cubature Kalman filter.
Automatica, vol. 49, no. 2 (2013), 510–518.

[16] S. J. Julier and J. K. Uhlmann
Unscented filtering and nonlinear estimation.
IEEE Review, vol. 92, no. 3 (2004), 401–421.

[17] S. J. Julier, J. K. Uhlmann, and H. F. Durrant-Whyte
A new method for the nonlinear transformation of means and covariances in filters and estimators.
IEEE Transactions on Automatic Control, vol. 45, no. 3 (2000), 477–482.

[18] T. Kollo and D. Von Rosen
Advanced Multivariate Statistics with Matrices, volume 579 of Mathematics and Its Applications.
Springer, 2005.

[19] J. H. Kotecha and P. M. Djurić
Gaussian sum particle filtering.
IEEE Transactions on Signal Processing, vol. 51, no. 10 (2003), 2602–2612.

[20] X. R. Li, Z. Zhao
Measuring estimator's credibility: Noncredibility index.
In *Proceedings of the 2006 International Conference on Information Fusion*, Florence, Italy, 2006.

[21] F. L. Lewis
Optimal Estimation.
New York: John Wiley & Sons, 1986.

[22] C. D. Meyer
Matrix Analysis and Applied Linear Algebra.
SIAM, 2000.

[23] M. Nørgaard, N. K. Poulsen, and O. Ravn
New developments in state estimation for nonlinear systems.
Automatica, vol. 36, no. 11 (2000), 1627–1638.

[24] M. Rhudy, Y. Gu, J. Gross, and M. R. Napolitano
Evaluation of matrix square root operations for UKF within a UAV GPS/INS sensor fusion application.
International Journal of Navigation and Observation, vol. 2011 (2011).

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

[26] A. Sakai and Y. Kuroda
Discriminatively trained unscented Kalman filter for mobile robot localization.
Journal of Advanced Research in Mechanical Engineering, vol. 1, no. 3 (2010), 153–161.

[27] J. Sarmavuori and S. Särkkä
Fourier-Hermite Kalman filter.
IEEE Transactions on Automatic Control, vol. 57, no. 6 (2012), 1511–1515.

[28] M. Šimandl and J. Duník
Derivative-free estimation methods: New results and performance analysis.
Automatica, vol. 45, no. 7 (2009), 1749–1757.

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

[30] H. W. Sorenson and D. L. Alspach
Nonlinear bayesian estimation using Gaussian sum approximations.
IEEE Transactions of Automatic Control, no. 4 (1972), 439–448.

[31] J. Steinbring and U. Hanebeck
S2KF: The smart sampling Kalman filter.
In *Proceedings of the 16th International Conference on Inference Fusion*, Istanbul, Turkey, July 2013, 2089–2096.

[32] O. Straka, J. Duník, and M. Šimandl
Scaling parameter in unscented transform: Analysis and specification.
In *Proceedings of the 2012 American Control Conference*, 2012, 5550–5555.

[33] O. Straka, J. Duník, and M. Šimandl
Unscented Kalman filter with advanced adaptation of scaling parameter.
Automatica, vol. 5, no. 10 (2014), 2657–2664.

[34] O. Straka, J. Duník, M. Šimandl, and J. Havlík
Aspects and comparison of matrix decompositions in unscented Kalman filter.
In *Proceedings of the 2013 American Control Conference*, Washington, DC, USA, 2013, 3075–3080.

[35] O. Straka, J. Duník, and M. Šimandl
Scaling parameter in unscented transformation: Analysis and specification.
In *Proceedings of the 2012 American Control Conference*, Montreal, Canada, June 2012, 5550–5555.

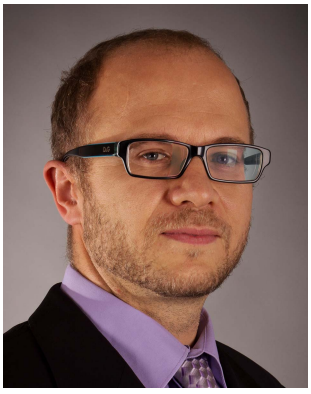
[36] R. Turner and C. E. Rasmussen
Model based learning of sigma points in unscented Kalman filter.
Neurocomputing, vol. 80 (2012), 47–53.

[37] X. Wang, Y. Liang, Q. Pan, and F. Yang
A Gaussian approximation recursive filter for nonlinear systems with correlated noises.
Automatica, vol. 48, no. 9 (2012), 2290–2297.

Jindřich Duník is a scientist at the Department of Cybernetics, Faculty of Applied Sciences, University of West Bohemia (UWB), Czech Republic. He received Ing. (M.Sc.) and Ph.D. degrees in Automatic Control in 2003 and 2008, respectively, both from the UWB. Until 2010, he was with the Department of Cybernetics focusing on the state estimation methods. Then, in the period 2010–2013, he was with Honeywell International, Aerospace Advanced Technology Europe (ATE), working in the areas of inertial and satellite based navigation systems, integrity monitoring methods, and advanced nonlinear estimation techniques. From 2013, he is with the European Centre of Excellence NTIS and Department of Cybernetics, UWB, focusing on the state estimation and system identification. From 2013, he is with Honeywell International, ATE, for part-time.

He is the author or co-author of more than 50 technical papers and patent applications devoted to the nonlinear filtering, system identification, and navigation information integrity monitoring. The papers were published at prestigious conferences (IFAC, IEEE, ISIF, ION) and journals (e.g., *Automatica*, *IEEE Transactions on Automatic Control*, *IEEE Transactions on Aerospace and Electronic Systems*). He is a guest area editor for the *Journal of Advances in Information Fusion* for the special issue on Nonlinear Derivative-Free Filters: Theory and Application and IPC Co-chair of the 12th European Workshop on Advanced Control and Diagnosis. He has also participated in a number of projects of fundamental and applied research (Czech Science Foundation, EU SESAR, ESA).

Jindřich received several Honeywell Bravo Awards and the Technology Achievement Award in 2016 for aircraft navigation system design and Werner von Siemens Excellence Award in 2014 for the basic research New Approaches and Methods of Nonlinear State Estimation and Optimal Decision Making under Uncertainty.



Ondřej Straka is head of Research Group Identification and Decision-Making in the European Centre of Excellence NTIS—New Technologies for the Information Society, where he focuses on nonlinear state estimation, stochastic systems, and system identification.

Dr. Straka focuses his research interests on both local and global nonlinear state estimation methods, performance evaluation and fault detection in navigation systems. He has published over fifty journal and conference papers in journals such as *Automatica*, *IEEE Transactions on Automatic Control*, *IEEE Transactions on Aerospace and Electronic Systems*, *Signal Processing* and at international conferences such as American Control Conference, World Congresses and Symposia of the International Federation of Automatic Control, IFAC, and FUSION Conferences. He has participated in a number of projects of fundamental research and in several project of applied research (e.g., GNSS-based safe train localization, attitude and heading reference system). He was involved in development of several software frameworks for nonlinear state estimation.

Dr. Straka received his Master's degree in Cybernetics and Control Engineering (in 1998) and his Ph.D. degree in Cybernetics (in 2004) both from University of West Bohemia. Since 2015 he has been associate professor at the Department of Cybernetics, University of West Bohemia.

Miroslav Šimandl was born in Ledce, Czechoslovakia, in 1954. He received the M.Sc. (Dipl.Ing.) degree in control engineering in 1978 and the Ph.D. (C.Sc.) degree in technical cybernetics in 1984, both from the Institute of Technology in Pilsen, Czechoslovakia.

In the period 1978–1992, he held various research and teaching positions at the Institute of Technology in Pilsen. In the period 1993–2001, he was associate professor, and since 2002 he has held the position of professor at the Department of Cybernetics, Faculty of Applied Sciences, University of West Bohemia, Pilsen, Czech Republic. Since 2011, he has held the position of vice-rector for research and development at the University of West Bohemia. Currently, he is also with the European Centre of Excellence—New Technologies for the Information Society (NTIS). Within IFAC, he served as the co-chair of the International Programme Committee of the 16th International Federation of Automatic Control (IFAC) World Congress in Prague, 2005, and he has served as a member of two technical committees of IFAC.

Prof. Šimandl is the author or co-author of many journal and conference technical papers. His main research interests are in the fields of nonlinear filtering, active fault detection, nonlinear identification, and adaptive dual control. In 2014 he received the Werner von Siemens Excellence Award for the basic research New Approaches and Methods of Nonlinear State Estimation and Optimal Decision Making under Uncertainty.

Erik P. Blasch is a principal scientist at the the United States Air Force Research Laboratory (AFRL) in the Information Directorate at Rome, NY, USA. From 2009–2012, he was an exchange scientist to Defence Research and Development Canada (DRDC) at Valcartier, Quebec. From 2000–2009, Dr. Blasch was the Information Fusion Evaluation Tech Lead for the AFRL Sensors Directorate—COMprehensive Performance Assessment of Sensor Exploitation (COMPASE) Center supporting design evaluations in Dayton, OH. Dr. Blasch has been an Adjunct Electrical Engineering Professor at Wright State University, the Air Force Institute of Technology, and the University of Dayton teaching signal processing, target tracking, and information fusion. He has also held various assignements as a reserve military officer supporting space situational awareness, photonics design, and dynamic data driven application systems.

Dr. Blasch was a founding member of the International Society of Information Fusion (ISIF) (www.isif.org) in 1998, held various leadership roles for the Fusion conferences, a Board of Governors (BoG) member (1999–2010), and the 2007 ISIF President. For the Institute of Electrical and Electronics Engineers (IEEE), he has served as a member of the Aerospace and Electronics Systems Society (AESS) BoG (2011–2016), the AESS International Chapters Chair (2012–2016), representative to AIAA (2013–2016), and an AESS Distinguished lecturer (2013–2016). He has focused on information fusion, target tracking, pattern recognition, and robitcs research compiling 600+ scientific papers and book chapters. He holds 8 patents, presented over 30 tutorials, and is an associate editor of three academic journals. His books include High-Level Information Fusion Management and Systems Design (Artech House, 2012) and Context Enhanced Information Fusion (Springer, 2016).

Dr. Blasch received his B.S. in Mechanical Engineering from the Massachusetts Institute of Technology in 1992 and Master's Degrees in Mechanical ('94), Health Science ('95), and Industrial Engineering (Human Factors) ('95) from Georgia Tech and attended the University of Wisconsin for a MD/PhD in Neurosciences/Mech. Eng until being called to military service in 1996 to the United States Air Force. He completed an MBA ('98), MSEE ('98), MS Econ('99), and a PhD ('99) in Electrical Engineering from Wright State University and is a graduate of Air War College ('08). He is the recipeint of the IEEE Bio-Engineering Award (Russ-2008), IEEE AESS Magazine Best Paper Award (Mimno-2012), and Military Sensing Symposium Ledaership in Data Fusion Award (Mignogna-2014). He is a Fellow of SPIE, Associate Fellow of AIAA, and a senior member of IEEE.