Preprints of the 19th World Congress
The International Federation of Automatic Control
Cape Town, South Africa. August 24-29, 2014

Sigma-Point Set Rotation in Unscented Kalman Filter: Analysis and Adaptation ⋆

Jindřich Duník, Ondřej Straka, and Miroslav Šimandl

Department of Cybernetics & European Centre of Excellence, NTIS - New Technologies for the Information Society, Faculty of Applied Sciences
University of West Bohemia, Univerzitní 8, 306 14 Pilsen, Czech Republic
e-mails: {dunikj,straka30,simandl}@kky.zcu.cz (J. Duník, O. Straka, M. Šimandl)

Abstract: The paper deals with analysis and illustration of the impact of the σ-point set rotation on the approximation quality of the unscented transformation and the estimation performance of the unscented Kalman filter. It is shown that the covariance matrix factor, used in σ-point computation, can be multiplied by an arbitrary rotation matrix which moves the σ-points along the surface of a hyper-ellipsoid related to the covariance matrix. The rotation matrix can be thus considered as another user-defined parameter (in addition to the scaling parameter) and the unscented Kalman filter with adaptive selection of both user-defined parameters is proposed. The impact of fixed or adaptively selected parameters on the performance of the unscented Kalman filter is illustrated by a numerical study.

1. INTRODUCTION

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

General solution to the recursive state estimation problem is given by the Bayesian recursive relations (BRRs). The BRRs provide probability density functions (PDFs) of the state conditioned by the measurements representing a full description of the state, which itself cannot be measured. A closed-form solution to the BRRs is available only for a few special cases such as a linear Gaussian system [Anderson and Moore, 1979] for which the resulting algorithm corresponds to the well-known Kalman filter (KF). In other cases, approximate methods, either global or local, must be used.

The global methods are based on a certain type of BRRs solution approximation and thus, they generate the conditional PDF of the state. As an example of these methods the particle filter [Doucet et al., 2001], point-mass method [Simandl et al., 2006], and the Gaussian sum method [Anderson and Moore, 1979] can be mentioned. The local methods often approximate the system description so that the KF design technique can be used even for nonlinear systems. In consequence, such methods provide state estimate in the form of the conditional mean and covariance matrix with local validity.

Local [1] methods can be divided into derivative and derivative-free local methods. Derivative methods approximate nonlinear functions in a system description by derivative-based expansions. As an example the Taylor or Fourier-Hermite series can be mentioned which lead to e.g., the extended Kalman filter, second order filter, or Fourier-Hermite Kalman filter [Anderson and Moore, 1979, Sarmavuori and Sarkka, 2012].

Derivative-free methods to the local filter design are based on differential-based polynomial interpolations, the unscented transformation, or various numerical integration rules. These methods are represented by the divided difference filters (DDFs) based on the Stirling polynomial interpolation [Norgaard et al., 2000], unscented Kalman filter (UKF) based on the unscented transformation (UT) [Julier and Uhlmann, 2004], or the quadrature, cubature, and stochastic integration filters utilizing deterministic and stochastic integration rules [Ito and Xiong, 2000, Arasaratnam and Haykin, 2009, Duník et al., 2013]. It is worth to note that although the approximations used in derivative-free methods come from quite different basic ideas, the resulting filter algorithms are in many cases identical. Among others, similarity analyses of the DDFs and UKF [Simandl and Dunik, 2009], cubature and quadrature filters and UKF [Jia et al., 2011], cubature and stochastic integration filters [Duník et al., 2013] can be pointed out.

Derivative-free methods, contrary to the derivative ones, evaluate the nonlinear functions in the system description in multiple points (often called σ-points). Placement of the σ-points in the state-space is determined by i) the actual estimated mean and covariance matrix (i.e., the linearisation point) and ii) σ-point set design parameters. The parameters are user-defined and affect the quality of the approximation and subsequently the filter estimation performance. The parameters might include selection of the covariance matrix decomposition technique and specification of the scaling parameter(s) (if applicable).

In the literature, vast majority of the research interest related to the user-defined parameters specification has been devoted to the scaling parameter specification in the context of the UKF. A fundamental recommendation on how to select the scaling parameter was provided in [Julier and Uhlmann, 2004]. The recommendation stems form a term-by-term comparison of the Taylor series expansion of the true mean and covariance matrix of a random variable transformed through a nonlinear func-

⋆ This work was supported by the Czech Science Foundation, project no. GACR P103-13-07058J, and by the project EXLIZ CZ.1.07/2.3.00/30.0013, which is co-financed by the European Social Fund and the state budget of the Czech Republic.

[1] In literature instead of the term 'local methods', terms like 'Gaussian filters' or 'Kalman filters' can be found.

tion with the Taylor series expansion of the UT-approximated statistics. In this case, the scaling parameter is a function of the state-space dimension only. In [Sakai and Kuroda, 2010], a technique for the off-line scaling parameter selection using a training procedure was proposed. Contrary to the previous procedure, adaptive techniques for the parameter selection has been of interest in past few years [Duník et al., 2010, Turner and Rasmussen, 2012, Duník et al., 2012]. Such techniques take an advantage of possibly different parameter value at each time instant resulting in significant increase of the estimation performance. Regarding the covariance matrix decomposition in the scope of the local filters, several decompositions were studied in [Rhudy et al., 2012] and thoroughly analysed and simulated in [Straka et al., 2013].

The goal of the paper is to analyse the impact of the σ-point set rotation on the performance of the UT or UKF[2]. It is shown that different covariance matrix decompositions provide the same σ-point set except of its rotation. Then, relations for an arbitrarily rotated set are proposed and an algorithm of the UKF with adaptive σ-point set rotation selection is studied.

The rest of the paper is organised as follows. In Section II the system is defined and in Section III the state estimation by the UKF is summarized. Then, the rotation of the σ-point set in the UT is analysed and novel UKF with adaptive selection of the user-defined parameters is proposed in Sections IV and V, respectively. Finally, a numerical study is the content of Section VI and the conclusion is given in Section VII.

2. SYSTEM DEFINITION

Let the discrete-time nonlinear stochastic system be considered

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

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

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 PDFs of the noises are supposed to be Gaussian with zero means and known covariance matrices $\mathbf{\Sigma}_{w,k}$ and $\mathbf{\Sigma}_{v,k}$, i.e., $p_{\mathbf{w}_k}(\mathbf{w}_k) = \mathcal{N}\{\mathbf{w}_k : \mathbf{0}, \mathbf{\Sigma}_{w,k}\}$ and $p_{\mathbf{v}_k}(\mathbf{v}_k) = \mathcal{N}\{\mathbf{v}_k : \mathbf{0}, \mathbf{\Sigma}_{v,k}\}$, respectively. The PDF of the initial state is Gaussian and known as well, i.e., $p_{\mathbf{x}_0}(\mathbf{x}_0) = \mathcal{N}\{\mathbf{x}_0 : \bar{\mathbf{x}}_0, \mathbf{P}_0\}$, and independent of the noises.

3. STATE ESTIMATION BY UKF

The aim of the local state estimation methods is to compute the first two moments of the state conditioned by the measurements, namely, the conditional mean $\hat{\mathbf{x}}_{k|k} = \mathsf{E}[\mathbf{x}_k|\mathbf{z}^k]$ and covariance matrix $\mathbf{P}_{k|k} = \mathrm{cov}[\mathbf{x}_k|\mathbf{z}^k]$ in which $\mathbf{z}^k = [\mathbf{z}_0, \mathbf{z}_1, \dots, \mathbf{z}_k]$. The moments can be understood as a Gaussian approximation of the conditional PDF, i.e., $p(\mathbf{x}_k|\mathbf{z}^k) \approx \mathcal{N}\{\mathbf{x}_k : \hat{\mathbf{x}}_{k|k}, \mathbf{P}_{k|k}\}$ [Arasaratnam and Haykin, 2009, Duník et al., 2013], depending on the type of used approximation.

All the local filter algorithms follow the structure of the generic local filter algorithm which can be summarised as [Simandl and Dunik, 2009, Arasaratnam and Haykin, 2009]:

Algorithm 1: Generic Local Filter

Step 1: 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 $\mathbf{P}_{0|-1} = \mathrm{cov}[\mathbf{x}_0] = \mathbf{P}_0$.

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

$$\hat{\mathbf{x}}_{k|k} = \hat{\mathbf{x}}_{k|k-1} + \mathbf{K}_{k|k}(\mathbf{z}_k - \hat{\mathbf{z}}_{k|k-1}), \quad (3)$$

$$\mathbf{P}_{k|k} = \mathbf{P}_{k|k-1} - \mathbf{K}_{k|k}\mathbf{P}_{z,k|k-1}\mathbf{K}_{k|k}^T, \quad (4)$$

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

$$\hat{\mathbf{z}}_{k|k-1} = \mathsf{E}[\mathbf{z}_k|\mathbf{z}^{k-1}] = E[\mathbf{h}_k(\mathbf{x}_k)|\mathbf{z}^{k-1}], \quad (5)$$

$$\begin{aligned}\mathbf{P}_{z,k|k-1} &= \mathsf{E}[(\mathbf{z}_k - \hat{\mathbf{z}}_{k|k-1})(\mathbf{z}_k - \hat{\mathbf{z}}_{k|k-1})^T|\mathbf{z}^{k-1}] = \\ &= \mathsf{E}[(\mathbf{h}_k(\mathbf{x}_k) - \hat{\mathbf{z}}_{k|k-1})\times \\ &\quad \times (\mathbf{h}_k(\mathbf{x}_k) - \hat{\mathbf{z}}_{k|k-1})^T|\mathbf{z}^{k-1}] + \mathbf{\Sigma}_{v,k}, \quad (6)\end{aligned}$$

$$\mathbf{P}_{xz,k|k-1} = \mathsf{E}[(\mathbf{x}_k - \hat{\mathbf{x}}_{k|k-1})(\mathbf{z}_k - \hat{\mathbf{z}}_{k|k-1})^T|\mathbf{z}^{k-1}]. \quad (7)$$

Step 3: The predictive statistics are given by the relations

$$\hat{\mathbf{x}}_{k+1|k} = \mathsf{E}[\mathbf{x}_{k+1}|\mathbf{z}^k] = E[\mathbf{f}_k(\mathbf{x}_k)|\mathbf{z}^k], \quad (8)$$

$$\begin{aligned}\mathbf{P}_{k+1|k} &= \mathsf{E}[(\mathbf{x}_{k+1} - \hat{\mathbf{x}}_{k+1|k})(\mathbf{x}_{k+1} - \hat{\mathbf{x}}_{k+1|k})^T|\mathbf{z}^k] = \\ &= \mathsf{E}[(\mathbf{f}_k(\mathbf{x}_k) - \hat{\mathbf{x}}_{k|k})(\mathbf{f}_k(\mathbf{x}_k) - \hat{\mathbf{x}}_{k|k})^T|\mathbf{z}^k] + \mathbf{\Sigma}_{w,k}. \quad (9)\end{aligned}$$

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

Applying the UT for the approximate solution to the measurement and state predictive statistics (5)–(7) and (8)–(9) leads to the UKF [Julier and Uhlmann, 2004]. For the sake of simplicity and completeness, the basic version[3] of the UT is introduced below.

The UT can be viewed as a tool for approximate computation of the mean, covariance and cross-covariance matrices of a transformed random variable $\mathbf{y} = \mathbf{g}(\mathbf{x})$ under the assumption of the known function $\mathbf{g}(\cdot)$ and the known mean $\hat{\mathbf{x}} = \mathsf{E}[\mathbf{x}]$ and covariance matrix $\mathbf{P}_x = \mathrm{cov}[\mathbf{x}]$ of $\mathbf{x}$. The UT is based on computation of the set of deterministic σ-points $\{\mathcal{X}_i\}_{i=0}^{2n_x}$ with appropriate weights $\{\mathcal{W}_i\}_{i=0}^{2n_x}$ according to

$$\mathcal{X}_0 = \hat{\mathbf{x}}, \qquad \mathcal{W}_0 = \frac{\kappa}{n_x + \kappa}, \quad (10)$$

$$\mathcal{X}_j = \hat{\mathbf{x}} + \sqrt{(n_x + \kappa)}\mathcal{S}_j, \mathcal{W}_j = \frac{1}{2(n_x + \kappa)}, \quad (11)$$

$$\mathcal{X}_{n_x+j} = \hat{\mathbf{x}} - \sqrt{(n_x + \kappa)}\mathcal{S}_j, \mathcal{W}_{n_x+j} = \mathcal{W}_j, \quad (12)$$

where $j = 1, \dots, n_x$, term $\mathcal{S}_j$ is the j-th column of the matrix $\mathbf{S}_x$ which is a factor of the covariance matrix $\mathbf{P}_x$ so that $\mathbf{S}_x(\mathbf{S}_x)^{\mathrm{T}} = \mathbf{P}_x$, and variable κ is the scaling parameter. To get approximate characteristic of $\mathbf{y}$, each point is transformed via the nonlinear function

$$\mathcal{Y}_i = \mathbf{g}(\mathcal{X}_i), \forall i. \quad (13)$$

Resulting UT-based characteristics[4] are given by

$$\hat{\mathbf{y}}^{\mathrm{UT}} = \sum_{i=0}^{2n_x} \mathcal{W}_i\mathcal{Y}_i, \quad (14)$$

$$\mathbf{P}_y^{\mathrm{UT}} = \sum_{i=0}^{2n_x} \mathcal{W}_i(\mathcal{Y}_i - \hat{\mathbf{y}}^{\mathrm{UT}})(\mathcal{Y}_i - \hat{\mathbf{y}}^{\mathrm{UT}})^T, \quad (15)$$

[2] Note that although the paper is focused on the UKF, the results can be directly applied to other derivative-free filters as well.

[3] Quite a few other advanced UT versions have been proposed, e.g., higher order, simplified, scaled, orthogonally transformed, or the smart sampling, differing in the σ-point set computation [Julier and Uhlmann, 2004, Steinbring and Hanebeck, 2013, Chang et al., 2013].

[4] The characteristics are generally approximate. Exact values are obtained for a few special functions $\mathbf{g}(\cdot)$ only, e.g., for a linear one.

$$\mathbf{P}_{xy}^{\mathrm{UT}} = \sum_{i=0}^{2n_x} \mathcal{W}_i(\mathcal{X}_i - \hat{\mathbf{x}})(\mathcal{Y}_i - \hat{\mathbf{y}}^{\mathrm{UT}})^T. \quad (16)$$

4. ROTATION IN SIGMA-POINT SET COMPUTATION

The ultimate property of any σ-point set is the equality of its first (at least) two moments with the moments of the original random variable, i.e.,

$$\hat{\mathbf{x}}^{\mathrm{UT}} = \sum_{i=0}^{2n_x} \mathcal{W}_i\mathcal{X}_i = \hat{\mathbf{x}}, \quad (17)$$

$$\mathbf{P}_x^{\mathrm{UT}} = \sum_{i=0}^{2n_x} \mathcal{W}_i(\mathcal{X}_i - \hat{\mathbf{x}}^{\mathrm{UT}})(\mathcal{X}_i - \hat{\mathbf{x}}^{\mathrm{UT}})^T = \mathbf{S}_x\mathbf{S}_x^T = \mathbf{P}_x, \quad (18)$$

regardless of the choice of the scaling parameter and covariance matrix decomposition.

4.1 Sigma-Point set with arbitrary rotation

Let a specific decomposition of the covariance matrix $\mathbf{P}_x$ be considered, the covariance matrix factor $\mathbf{S}_x$ be computed, and n_x-dimensional rotation matrix $\mathbf{C}$ be specified. Then, the rotated σ-point set (with preserved weights) is

$$\mathcal{X}_0^r = \hat{\mathbf{x}}, \quad (19)$$

$$\mathcal{X}_j^r = \hat{\mathbf{x}} + \sqrt{(n_x+\kappa)}\mathcal{S}_j^r, \quad (20)$$

$$\mathcal{X}_{n_x+j}^r = \hat{\mathbf{x}} - \sqrt{(n_x+\kappa)}\mathcal{S}_j^r, \quad (21)$$

where $\mathcal{S}_j^r$ is the j-th column of the matrix

$$\mathbf{S}_x^r = \mathbf{S}_x\mathbf{C}. \quad (22)$$

Because of the orthogonal property of the rotation matrix, the rotated covariance matrix decomposition forms also the covariance matrix $\mathbf{P}_x$, i.e.,

$$\mathbf{P}_x = \mathbf{S}_x^r(\mathbf{S}_x^r)^T = \mathbf{S}_x\mathbf{C}\mathbf{C}^T\mathbf{S}_x^T = \mathbf{S}_x\mathbf{I}\mathbf{S}_x^T, \quad (23)$$

where $\mathbf{I}$ is the identity matrix of appropriate dimension.

As the rotated σ-point set (19)–(21) is still symmetric and relation (23) is valid, the rotated set preserve the mean and covariance matrix of the original variable $\mathbf{x}$ similarly to 'unrotated' set, see (17), (18).

For illustration, three σ-point sets are shown in Figure 1; unrotated (standard) set, rotated set through $\theta = 30$ degrees, and rotated set through $\theta = 60$ degrees. It can be seen that the points belong to the ellipsoid (in $\mathbb{R}^{n_x}$ to hyper-ellipsoid) with the principal axes given by the orthogonal matrix U and diagonal matrix D computed by the singular value decomposition [5] (SVD) of the scaled covariance matrix $(n_x+\kappa)\mathbf{P}_x$ as

$$(n_x+\kappa)\mathbf{P}_x = (n_x+\kappa)\mathbf{U}\mathbf{D}\mathbf{U}^T. \quad (24)$$

Note 1: Two-dimensional (counter-clockwise) rotation matrix is of the form

$$\mathbf{C}_2 = \begin{bmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{bmatrix}. \quad (25)$$

Rotation matrix computation of arbitrary dimension is discussed e.g., in [Meyer, 2000]. Generally, a rotation matrix of dimension n_x, $\mathbf{C}_{n_x}$, requires specification of $n_x(n_x-1)/2$ rotation angles, i.e., $\boldsymbol{\theta} = [\theta_1, \ldots, \theta_{n_x(n_x-1)/2}]^T$.

[5] For symmetric positive definite matrix SVD reduces to the spectral decomposition.

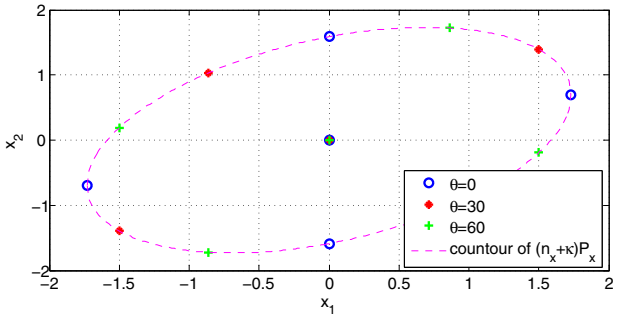

Fig. 1. Illustration of σ-point set rotation.

Note 2: Note also that with respect to the form of (23), any orthogonal [6] matrix can be used as $\mathbf{C}$.

4.2 Matrix decomposition: A specific σ-point set transformation

Based on the analysis in [Straka et al., 2013], it can be shown that various covariance matrix decompositions are the same except of the matrix $\mathbf{C}$.

As an example, comparison of the SVD and Cholesky decomposition of $\mathbf{P}_x$ is introduced. Singular value decomposition of $\mathbf{P}_x$ (24) results in the covariance matrix factor

$$\mathbf{S}_x = \mathbf{S}_x^{\mathrm{SVD}} = \mathbf{U}\sqrt{\mathbf{D}}. \quad (26)$$

Let the orthogonal-triangular QR decomposition of the matrix $\left(\mathbf{U}\sqrt{\mathbf{D}}\right)^T$, i.e.,

$$[\mathbf{Q}, \mathbf{R}] = \mathrm{qr}\left(\left(\mathbf{U}\sqrt{\mathbf{D}}\right)^T\right), \quad (27)$$

be supposed, where $\mathbf{Q}$ and $\mathbf{R}$ is the orthogonal and (upper-)triangular matrix, respectively. Hence, $\mathbf{Q}\mathbf{R} = (\mathbf{U}\sqrt{\mathbf{D}})^T$. Then, selecting

$$\mathbf{C} = \mathbf{Q}, \quad (28)$$

it holds that

$$\mathbf{P}_x = \mathbf{S}_x^{\mathrm{SVD}}\mathbf{C}(\mathbf{S}_x^{\mathrm{SVD}}\mathbf{C})^T = \mathbf{U}\sqrt{\mathbf{D}}\mathbf{Q}\left(\mathbf{U}\sqrt{\mathbf{D}}\mathbf{Q}\right)^T = \\ = \mathbf{R}^T\mathbf{R} = \mathbf{L}\mathbf{L}^T, \quad (29)$$

as

$$\mathbf{R} = \mathbf{Q}^T(\mathbf{U}\sqrt{\mathbf{D}})^T = (\mathbf{U}\sqrt{\mathbf{D}}\mathbf{Q})^T = \left((\mathbf{Q}\mathbf{R})^T\mathbf{Q}\right)^T. \quad (30)$$

The lower triangular matrix $\mathbf{L}$ in (29) then corresponds to

$$\mathbf{L} = \mathbf{U}\sqrt{\mathbf{D}}\mathbf{Q} = \mathbf{S}_x^{\mathrm{Chol}} \quad (31)$$

which is the Cholesky decomposition based covariance matrix factor $\mathbf{S}_x$. The notation $\sqrt{\mathbf{D}}$ stands for the square-root of all diagonal elements of the matrix $\mathbf{D}$.

Therefore, the SVD- and Cholesky-based covariance matrix factors $\mathbf{S}_x^{\mathrm{SVD}}$ (26) and $\mathbf{S}_x^{\mathrm{Chol}}$ (31) are uniquely related through the orthogonal matrix $\mathbf{Q}$. Note that the factors are the same if $\mathbf{P}_x$ is diagonal; in this case matrix $\mathbf{Q}$ is the identity matrix.

4.3 Numerical illustration

Selection of the scaling parameter is often considered as a major user-defined parameter affecting the approximation quality

[6] Orthogonal matrix is a unitary matrix which preserves inner products of two vectors. Orthogonal matrices can be further divided into rotations (having determinant equal to 1) and reflections or their combinations (having determinant equal to -1) [Meyer, 2000]. Both operations belong into the set of isometric transformations. Within this paper the special focus is laid on the rotations.

characteristics (14)–(16) and subsequently, impacting the filter performance. However, as it shown in this illustration, selection of the σ-point set rotation is another factor heavily impacting the approximation quality.

This fact is illustrated using a nonlinear random variable transformation example (as defined in Section 3) where

$$\hat{\mathbf{x}} = \begin{bmatrix} \hat{x}_1 \\ \hat{x}_2 \end{bmatrix} = \begin{bmatrix} 10 \\ 2 \end{bmatrix}, \mathbf{P}_x = \begin{bmatrix} 6 & 4 \\ 4 & 30 \end{bmatrix} \tag{32}$$

and

$$\mathbf{y} = \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \mathbf{g}(\mathbf{x}) = \begin{bmatrix} \sqrt{x_1^2 + x_2^2} \\ \text{atan}(x_2/x_1) \end{bmatrix}. \tag{33}$$

This example represents transformation of Cartesian to polar coordinate system [7].

In Figure 2, the impact of the σ-point set rotation through the angle $\theta \in \langle 0, 90\rangle\ [deg]$ and σ-point set scaling through the parameter $\kappa \in \langle 0, 10\rangle$ on the resulting approximate characteristics of y is shown. Namely, the mean value $\hat{\mathbf{y}}^{\text{UT}} = [\hat{y}_1^{\text{UT}}, \hat{y}_2^{\text{UT}}]^T$ (14) and the determinant of the covariance matrix $\mathbf{P}_y^{\text{UT}}$ (15) are studied. Except those characteristics, the true ones and ones approximated using the UT with fixed parameters [8] $\theta = 0[deg]$ and $\kappa = 1$ are plotted in Figure 2. Here it should be noted, that the true and fixed UT characteristics are not functions of the parameters θ or κ, they are plotted for the whole range of the parameters just for ease of comparison.

From the figure, it can be seen that the rotation of the σ-point set (i.e., the UT parameter θ) heavily impacts both the mean value and covariance matrix. In the case of the mean value, it is possible to get the true value of $\hat{y}_1^{\text{UT}}$ by proper selection of the rotation ($\theta \approx 11, 64[deg]$) and of $\hat{y}_2^{\text{UT}}$ by $\theta \approx 41, 83[deg]$, whereas by any setting of κ it is not possible to get the true mean. On the other hand, the true determinant can be reached by proper selection either of $\theta \approx 52, 82[deg]$ or $\kappa \approx 3.5$. It is also worth mentioning, that the 'optimum' parameters (in terms of their closeness to the true values) differ for the particular characteristics.

It can be seen that the performance of the UT (and as will be shown, the performance of the UKF) is affected not only by the selection of the scaling parameter but also by the rotation of the σ-points. Thus, the rotation can be viewed as another user-defined parameter (extending the possibilities imposed by selection of the covariance matrix decomposition method).

5. UKF WITH ADAPTIVE SELECTION OF SCALING PARAMETER AND SIGMA-POINT SET ROTATION

In [Duník et al., 2012], the UKF with adaptive setting of the scaling parameter (UKF-A(κ)) was proposed. Its algorithm follows the generic local filter algorithm with only one exception which is the computation of κ prior to the filtering step. The scaling parameter can be computed according to the several criteria, e.g., maximizing the likelihood function [Duník et al., 2012, Chang et al., 2013] or posterior probability [Straka et al., 2012]. As an example, computation of κ according to the latter is recapitulated below;

$$\kappa_k^* = \arg\max_{\kappa} p(\hat{\mathbf{x}}_{k|k}(\kappa)|\mathbf{z}^k; \kappa), \tag{34}$$

where $p(\mathbf{x}_k|\mathbf{z}^k;\kappa) = p(\mathbf{z}_k|\mathbf{x}_k)p(\mathbf{x}_k|\mathbf{z}^{k-1}) \left(p(\mathbf{z}_k|\mathbf{z}^{k-1};\kappa)\right)^{-1}$, $p(\mathbf{x}_k|\mathbf{z}^{k-1}) \approx \mathcal{N}\{\mathbf{x}_k : \hat{\mathbf{x}}_{k|k-1}, \mathbf{P}_{k|k-1}\}$ is the predictive estimate PDF, $p(\mathbf{z}_k|\mathbf{z}^{k-1};\kappa) \approx \mathcal{N}\{\mathbf{z}_k : \hat{\mathbf{z}}_{k|k-1}(\kappa), \mathbf{P}_{z,k|k-1}(\kappa)\}$ is the likelihood function emphasizing its dependence on κ at given time k, and $p(\mathbf{z}_k|\mathbf{x}_k)$ is the measurement PDF obtained from (1). Relation (34) leads, after a few rearrangements, to the following minimization

$$\begin{aligned}\kappa_k^* = \arg\min_{\kappa} \Big\{ & \tilde{\mathbf{z}}_k(\kappa)^T [\mathbf{L}_{k|k-1}(\kappa) - \mathbf{P}_{z,k|k-1}(\kappa)] \tilde{\mathbf{z}}_k(\kappa) \\ & + [\mathbf{z}_k - \mathbf{h}_k(\hat{\mathbf{x}}_{k|k}(\kappa))]^T \boldsymbol{\Sigma}_{v,k} [\mathbf{z}_k - \mathbf{h}_k(\hat{\mathbf{x}}_{k|k}(\kappa))] \\ & - \tfrac{1}{2} \log |\mathbf{P}_{z,k|k-1}(\kappa)| \Big\}\end{aligned} \tag{35}$$

where $\tilde{\mathbf{z}}_k(\kappa) = \mathbf{z}_k - \hat{\mathbf{z}}_{k|k-1}(\kappa)$ and

$$\mathbf{L}_{k|k-1} = \mathbf{P}_{z,k|k-1}^{-1} \mathbf{P}_{xz,k|k-1}^T \mathbf{P}_{k|k-1} \mathbf{P}_{xz,k|k-1} \mathbf{P}_{z,k|k-1}^{-1}.$$

In [Duník et al., 2012], several optimisation schemes were discussed. Within this paper, the grid method is selected for its simplicity. The method covers a feasible domain $\langle \kappa_{\min}, \kappa_{\max} \rangle$ within κ is sought by an equally spaced grid of points in which the function (35) is evaluated.

Considering the rotation matrix (more specifically the respective rotation angles $\boldsymbol{\theta}$) as a free parameter, the optimal rotation angles $\boldsymbol{\theta}^*$ can be computed analogously to κ^* (35) at every time instant k. Then, the resulting algorithm of the UKF with adaptive setting of $\boldsymbol{\theta}$ and κ (UKF-A($\boldsymbol{\theta}, \kappa$)) is as follows.

Algorithm 2: UKF with adaptive setting of parameters

Step 1: Initialisation is the same as in Algorithm 1.
Step 2: Parameters $\boldsymbol{\theta}_k^*$ and κ_k^* are computed according to

$$[\boldsymbol{\theta}_k^*, \kappa_k^*] = \arg\max_{\kappa, \boldsymbol{\theta}} p(\hat{\mathbf{x}}_{k|k}(\boldsymbol{\theta}, \kappa)|\mathbf{z}^k; \boldsymbol{\theta}, \kappa). \tag{36}$$

Step 3: Filtering step is the same as in Algorithm 1 but the measurement predictive statistics are computed on the basis of σ-point set with $\boldsymbol{\theta}_k^*$ and κ_k^*.
Step 4: Prediction step is the same as in Algorithm 1 where the σ-point set uses either $\boldsymbol{\theta}_k^*$ and κ_k^* or other values stemming for example from recommendation in [Julier and Uhlmann, 2004].
Let $k = k + 1$ and algorithm continues by **Step 2**.

Note 3: The algorithm is conditioned by $\boldsymbol{\theta}$- and κ-dependent optimisation function (36). This is not fulfilled for e.g., linear function in measurement equation (2) for which the filtering estimate is the same for any $\boldsymbol{\theta}$ and κ.

Note 4: Clear limitation of this adaptive algorithm is the computational complexity increasing with the state dimension. If κ is optimized only, it is one dimensional optimisation problem for $\forall n_x$. However, if $\boldsymbol{\theta}$ is optimized as well, the searched space is of dimension $\frac{n_x(n_x+1)}{2} + 1$. Therefore, selection of the proper optimisation method is crucial decision. Another possibility is to perform sets of off-line simulations for considered scenarios and then, to select the suitable parameters accordingly.

6. NUMERICAL EXAMPLE

The impact of the user-defined parameters values on the UKF performance is evaluated in this section together with the performance of the proposed UKF-A($\boldsymbol{\theta}, \kappa$). The evaluation is based on the bearings only tracking [Duník et al., 2012] described by the following mathematical model

[7] Four-quadrant inverse tangent function (atan2) is considered.
[8] Fixed parameters were set on the basis of recommendation given in [Julier and Uhlmann, 2004].

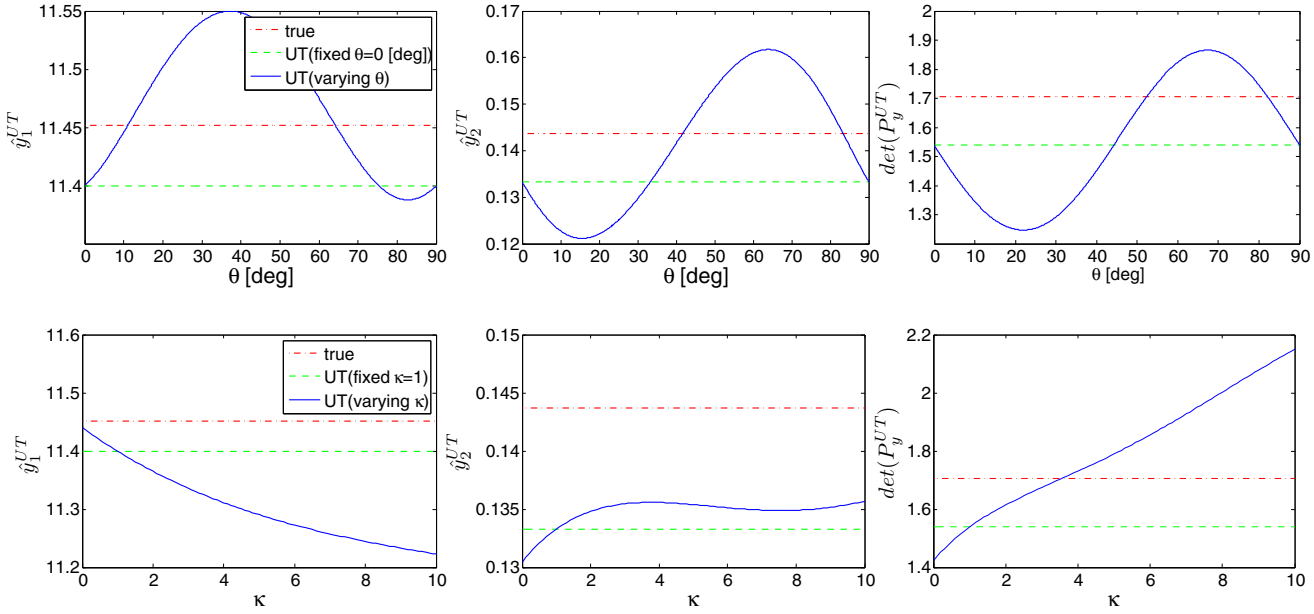

Fig. 2. Dependence of the UT-based approximate characteristics on the rotation θ of σ-point sample set and the scaling parameter κ (together with the true and fixed UT characteristics).

$$\mathbf{x}_{k+1} = \begin{bmatrix} 0.9 & 0 \\ 0 & 1 \end{bmatrix} \mathbf{x}_k + \mathbf{w}_k, \tag{37}$$

$$z_k = \operatorname{atan}\left(\frac{x_{2,k} - \sin(k)}{x_{1,k} - \cos(k)} \right) + v_k, \tag{38}$$

where $k = 0, 1, \ldots, N$, $N = 51$, $p(\mathbf{x}_0) = \mathcal{N}\{\mathbf{x}_0 : [20, 5]^T, 0.1\mathbf{I}_2\}$, $\mathbf{\Sigma}_{w,k} = \begin{bmatrix} 0.1 & 0.01 \\ 0.01 & 0.1 \end{bmatrix}$, $\Sigma_{v,k} = 0.025, \forall k$.

The UKF (with various but fixed values of θ and κ) and UKF-A (with different settings) are studied using a set of $M = 1000$ Monte Carlo (MC) simulations and three performance metrics;

- *Root Mean Square Error (RMSE)* defined as

$$\mathrm{RMSE} = \sqrt{\frac{1}{MNn_x} \sum_{i=1}^{M} \sum_{k=1}^{N} (\hat{x}^i_{1,k} - x^i_{1,k})^2 + (\hat{x}^i_{2,k} - x^i_{2,k})^2},$$

  where $(x^i_{1,k}, x^i_{2,k})$ and $(\hat{x}^i_{1,k}, \hat{x}^i_{2,k})$ denote true and estimated state components at the i-th MC run,
- *Averaged Normalized Estimation Error Squared (ANEES)* Blasch et al. [2006] defined as

$$\mathrm{ANEES} = \frac{\sum_{i=1}^{M} \sum_{k=1}^{N} \left((\mathbf{x}^i_k - \hat{\mathbf{x}}^i_k)^{\mathrm{T}} (\mathbf{P}^i_{k|k})^{-1} (\mathbf{x}^i_k - \hat{\mathbf{x}}^i_k) \right)}{n_x MN},$$

  where $\mathbf{P}^i_{k|k}$ is the covariance matrix of the estimate provided by the filter at the i-th MC run,
- *Time* indicating an average time per one MC simulation.

The RMSE metric provides an evaluation of the estimate error expressed as the Euclidean distance between the true state and its estimate. The value of the RMSE provides an absolute evaluation of the estimate error; the closer RMSE to zero, the better. The ANEES metric, on the other hand, provides evaluation of a relative estimation error (self-assessment of a filter); the closer ANEES to one, the better.

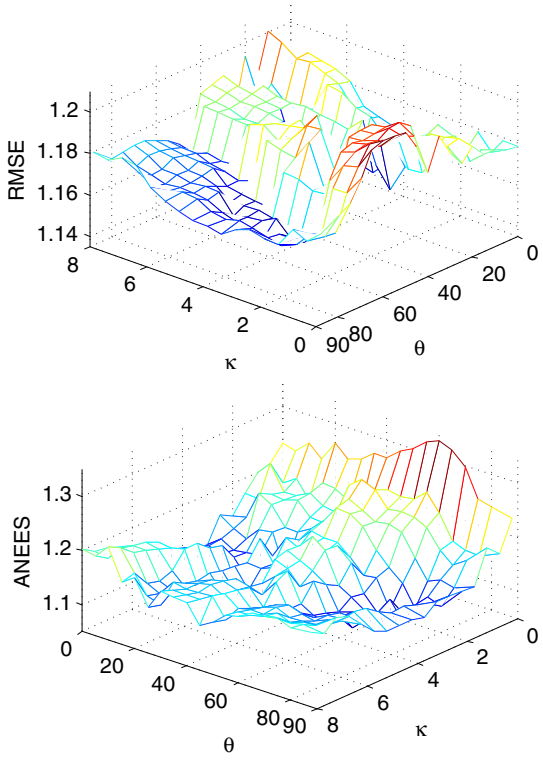

Fig. 3. RMSE and ANEES of UKF with different values of θ and κ.

In Figure 3, the estimation performance of the UKF with the fixed parameters is shown for different values of θ and κ. It can be seen that the RMSE varies between 1.1560 and 1.2104 (approx. 5% change) and ANEES between 1.0782 and 1.3322 (approx. 20% change). From the perspective of the RMSE

Table 1. RMSE, ANEES, and average time for various UKF and UKF-A algorithms.

	fixed par.	RMSE	ANEES	Time [msec]
UKF-A(θ,κ)	-	**1.1484**	**1.0175**	81.1
UKF-A(κ)	$\theta = 0[deg]$	1.1602	1.1041	15.3
UKF-A(κ)	$\theta = 30[deg]$	1.1587	1.0852	
UKF-A(κ)	$\theta = 50[deg]$	**1.1496**	1.0387	
UKF-A(κ)	$\theta = 75[deg]$	1.1598	1.0610	
UKF-A(θ)	$\kappa = 0$	1.1719	1.1614	12.7
UKF-A(θ)	$\kappa = 1$	1.550	1.0598	
UKF-A(θ)	$\kappa = 2$	1.1519	1.0345	
UKF-A(θ)	$\kappa = 3$	1.1521	**1.0306**	
UKF	$\kappa = 1, \theta = 0$	1.1791	1.2050	2.4
UKF	$\kappa = 2, \theta = 70$	1.1569	1.0782	

the optimum fixed UKF parameters are $\theta^*_{fix} = 65[deg]$ and $\kappa^*_{fix} = 3$ and for ANEES are $\theta^*_{fix} = 70[deg]$ and $\kappa^*_{fix} = 2$. Note that based on the analysis in [Julier and Uhlmann, 2004], the scaling parameter should be $\kappa = 3 - n_x = 1$ without any explicit consideration of σ-point set rotation.

Simulation results of the UKF-A are summarized in Table 1. Namely, the following UKF-As are considered:

- UKF-A with adaptation of both parameters, UKF-A(θ,κ),
- UKF-A with adaptation of κ and fixed θ, UKF-A(κ),
- UKF-A with adaptation of θ and fixed κ, UKF-A(θ).

Adaptation of the parameter(s) (equivalently, selection of such parameter(s) minimizing (36)) is performed using a grid method with the grid points $\{0, 15, 30, \ldots, 75\}[deg]$ for θ and $\{0, 1, \ldots, 8\}$ for κ. For completeness, performance of the UKF with recommended and optimum fixed parameters is given as well.

The simulation results illustrates the estimation quality improvement of the UKF-As over fixed parameters based UKFs. The best performance is reached by the UKF-A(θ,κ) in both considered criteria. Regarding single parameter UKF-A, the UKF-A(κ) provides lower RMSE, but the UKF-A(θ) lower ANEES. Considering the time complexity, the UKF-A is naturally more computationally demanding because of the employed optimisation technique for (36). To reduce the complexity, an alternative optimisation technique should be used. This topic is, however, out of the scope of this study.

7. CONCLUDING REMARKS

The paper dealt with analysis of the σ-point set rotation within the UT and UKF. It was proven that the covariance matrix factor, used in σ-point computation, can be multiplied by an arbitrary rotation matrix. Selected rotation matrix subsequently (and significantly) affects performance of the UT and UKF and it can be viewed as a user-defined parameter similarly to the scaling parameter. Therefore, the UKF with adaptive setting of the scaling parameter was reviewed and modified to adapt not only the scaling performance but also σ-point set rotation. The superior performance of the UKF with adaptation of the scaling parameter and set rotation was illustrated using a numerical study.

Note that the algorithm of the UKF-A is a part of the Nonlinear Estimation Framework available at http://nft.kky.zcu.cz/.

REFERENCES

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

I. Arasaratnam and S. Haykin. Cubature Kalman filters. *IEEE Transactions on Automatic Control*, 54(6):1254–1269, 2009.

E. P. Blasch, A. Rice, and C. Yang. Nonlinear track evaluation using absolute and relative metrics. In *Proceedings of the International Society for Optical Engineering*. SPIE, 2006.

L. Chang, B. Hu, and F. Qin. Unscented type Kalman filter: Limitation and combination. *IET Signal Processing*, 7(3): 167–176, 2013.

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

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.

J. Duník, M. Šimandl, and O. Straka. Unscented Kalman filter: Aspects and adaptive setting of scaling parameter. *IEEE Transactions on Automatic Control*, 57(9):2411–2416, 2012.

J. Duník, O. Straka, and M. Šimandl. Stochastic integration filter. *IEEE Trans. on AC*, 58(6):1561–1566, 2013.

K. Ito and K. Xiong. Gaussian filters for nonlinear filtering problems. *IEEE Trans. on AC*, 45(5):910–927, 2000.

B. Jia, M. Xin, and Y. Cheng. Sparse Gauss–Hermite quadrature filter with application to spacecraft attitude estimation. *Journal of Guidance, Control, and Dynamics*, 34(2), 2011.

S. J. Julier and J. K. Uhlmann. Unscented filtering and nonlinear estimation. *IEEE Review*, 92(3):401–421, 2004.

C. D. Meyer. *Matrix Analysis and Applied Linear Algebra*. SIAM, 2000.

M. Norgaard, N. K. Poulsen, and O. Ravn. New developments in state estimation for nonlinear systems. *Automatica*, 36 (11):1627–1638, 2000.

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*, 2012.

A. Sakai and Y. Kuroda. Discriminatively trained unscented Kalman filter for mobile robot localization. *J. of Adv. Res. in Mech. Eng.*, 1(3):153–161, 2010.

J. Sarmavuori and S. Sarkka. Fourier-Hermite Kalman filter. *IEEE Trans. on AC*, 57(6):1511–1515, 2012.

M. Simandl and J. Dunik. Derivative-free estimation methods: New results and performance analysis. *Automatica*, 45(7): 1749–1757, 2009.

M. Simandl, J. Kralovec, and T. Soderstrom. Advanced point – mass method for nonlinear state estimation. *Automatica*, 42 (7):1133–1145, 2006.

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.

O. Straka, J. Duník, and M. Šimandl. Scaling parameter in unscented transformation: Analysis and specification. In *Proc. of the 2012 ACC*, Montreal, Canada, June 2012.

O. Straka, J. Dunik, M. Simandl, and J. Havlik. Aspects and comparison of matrix decompositions in unscented Kalman filter. In *Proceedings of the 2013 American Control Conference*, Washington, DC, USA, 2013.

R. Turner and C. E. Rasmussen. Model based learning of sigma points in unscented Kalman filter. *Neurocomputing*, 80:47–53, 2012.