The structure of the KF.

(KFF), performs some approximations of the model and adopts the concept, referred here to as the Kalman filtering framework for generally nonlinear and non-Gaussian systems, mean squared error sense for linear Gaussian systems. To estimate the state, the Kalman filter (KF) represents the best estimator in the problem [4].

The randomized unscented Kalman filter is introduced. The filter utilizes the randomized unscented transform, which is based on a degree 3 stochastic integration rule. The paper discusses several aspects of the randomized unscented transform and its relation to the unscented transform, which can be seen as a special deterministic version. The aspects are demonstrated on polar to Cartesian coordinate transformation. The randomized unscented Kalman filter is illustrated in a numerical example concerning a reentry problem.

Keywords: state estimation, nonlinear filtering, target tracking.

I. INTRODUCTION

State estimation of dynamic stochastic systems represents a key component of both single and multiple target tracking applications [1], [2]. Its aim, in general, is to find an estimate of system quantities, called state, using a set of measurements (observations). In target tracking applications the state usually consists of kinematic variables (such as the position of the target, its velocity and acceleration) and variables characterizing the target (such as aerodynamic properties, target signal-to-noise ratio, spectral characteristics, and so forth).

The system behavior is represented by a state-space model given by an equation expressing state evolution and an equation characterizing the relation between the state and measurement. The system description is considered to be stochastic to account for possible uncertainties in the state or measurement. Due to stochastic nature of the system, the estimate is given in the form of a conditional probability density function (pdf) of the state conditioned by the measurement, which is the complete solution to the estimation problem as it contains all statistical information given by the available measurements and initial condition [3]. Calculation of the conditional pdf for nonlinear or non-Gaussian models is an intricate functional problem [4].

Thus, usually a simpler concept providing only point estimates of the state is preferred. The concept stems from the Kalman filter (KF), which represents the best estimator in the mean squared error sense for linear Gaussian systems. To estimate state of generally nonlinear and non-Gaussian systems, the concept, referred here to as the Kalman filtering framework (KFF), performs some approximations of the model and adopts the structure of the KF.

The first used approximations of the model were based on the Taylor series expansions of the nonlinear functions in the model description which lead to the extended Kalman filter (EKF) or the second-order extended Kalman filter [5].

In the last decade a number of new filters within the KFF have appeared. Instead of the Taylor series based linearization they utilize stochastic and polynomial linearizations. The filters are called sigma-point or derivative-free filters and their main advantage over the traditional methods is that they do not require the Jacobi matrix of the nonlinear functions in the state and measurement equations of the model [6].

The idea of the stochastic linearization is to approximate a random variable by a set of points which are transformed through nonlinear functions (i.e. functions in the model) [7]–[9]. The unscented transform (UT) used in the unscented Kalman filter (UKF) belongs to stochastic linearization methods [10].

The idea of the polynomial linearization is to approximate a nonlinear function by a first or second order polynomial interpolation [8], [11]. The divided difference filter (DDF) is the main representative of the filters utilizing the polynomial linearization.

Recently, several nonlinear filters have been proposed based on numerical calculation of the predictive statistics within the KFF. For the calculation the Gauss-Hermite quadrature [8] or cubature integration rules [12] were used.

In [13] a new filter called the randomized unscented Kalman filter (RUKF) has been proposed. Its core is the randomised UT (RUT) which possesses some favorable properties as apposed to the UT.

The aim of the paper is to introduce the RUKF, discuss its properties and particularly focus on its application to target tracking.

The paper is organized as follows: System specification and nonlinear state estimation by means of a generic filter of the KFF are introduced in Section II. The UT and RUT are presented in Section III and properties of the RUT are discussed and analysed in Section IV. Section V deals with the RUKF algorithm and its efficient implementation. In Section VI the RUKF is applied to a target tracking numerical example and concluding remarks are drawn in Section VII.
II. SYSTEM SPECIFICATION AND GENERIC FILTER

Let the discrete-time nonlinear stochastic system be considered

\[ x_{k+1} = f_k(x_k) + w_k, \quad k = 0, 1, 2, \ldots, \quad (1) \]
\[ z_k = h_k(x_k) + v_k, \quad k = 0, 1, 2, \ldots, \quad (2) \]

where the vectors \( x_k \in \mathbb{R}^{n_x} \) and \( z_k \in \mathbb{R}^{n_z} \) represent the immeasurable state of the system and measurement at time instant \( k \), respectively. \( f_k : \mathbb{R}^{n_x} \rightarrow \mathbb{R}^{n_x} \) and \( h_k : \mathbb{R}^{n_x} \rightarrow \mathbb{R}^{n_z} \) are known vector functions, and \( w_k \in \mathbb{R}^{n_z}, v_k \in \mathbb{R}^{n_z} \) are the state and measurement white noises. The pdf’s of the noises are supposed to be Gaussian with zero means and known covariance matrices \( Q_k \) and \( R_k \), i.e., \( p_{w_k}(w_k) = \mathcal{N}(0, Q_k) \) and \( p_{v_k}(v_k) = \mathcal{N}(0, R_k) \), respectively. The pdf of the initial state is supposed to be Gaussian and known as well, i.e., \( p_{x_0}(x_0) = \mathcal{N}(x_0, P_0) \). The initial state is independent of the noises.

As has been mentioned above, the generic filter of the KFF provides only point estimate of the state \( \hat{x}_{k|k} = \mathbb{E}[x_k|z^k] \), where \( z^k = [z_0^T, z_1^T, \ldots, z_k^T]^T \) together with a measure of its accuracy given by the conditional covariance matrix \( P_{k|k} = \text{cov}[x_k|z^k] \). Its algorithm can be written in the following form.

Algorithm 1: generic filter

Step 1: (initialization) Set the time instant \( k = 0 \) and define a priori initial condition by its first two moments

\[ \hat{x}_{0|0} = \mathbb{E}[x_0] = \hat{x}_0, \]
\[ P_{0|0} = \mathbb{E}[(x_0 - \hat{x}_{0|0})(x_0 - \hat{x}_{0|0})^T] = P_0. \]

Step 2: (filtering, measurement update, correction) The filtering mean \( \hat{x}_{k|k} \) and covariance matrix \( P_{k|k} \) are computed by means of

\[ \hat{x}_{k|k} = \hat{x}_{k|k-1} + K_k(z_k - \hat{z}_{k|k-1}), \]
\[ P_{k|k} = P_{k|k-1} - K_kP_{z|k-1}K_k^T, \]

where

\[ K_k = P_{z|k-1}(1) = P_{z|k-1}^{-1}(P_{k|k-1} - K_kH_k^T) \]

is the filter gain and the measurement prediction \( \hat{z}_{k|k-1} \) is given by

\[ \hat{z}_{k|k-1} = \mathbb{E}[z_k|z^{k-1}]|z^{k-1}. \]

The predictive covariance matrices \( P_{k|k-1} \) and \( P_{z|k-1} \) are computed as

\[ P_{z|k-1} = \mathbb{E}[(z_k - \hat{z}_{k|k-1})(z_k - \hat{z}_{k|k-1})^T|z^{k-1}] \]
\[ = \mathbb{E}[(h_k(x_k) - \hat{z}_{k|k-1}) \times (h_k(x_k) - \hat{z}_{k|k-1})^T|z^{k-1}] + R_k, \]
\[ P_{k|k-1} = \mathbb{E}[(x_k - \hat{x}_{k|k-1})(z_k - \hat{z}_{k|k-1})^T|z^{k-1}] \]

Step 3: (prediction, time update) The predictive mean \( \hat{x}_{k+1|k} \) and covariance matrix \( P_{k+1|k} \) are given by

\[ \hat{x}_{k+1|k} = \mathbb{E}[x_{k+1}|z^k] = \mathbb{E}[f_k(x_k, u_k)|z^k], \]
\[ P_{k+1|k} = \mathbb{E}[(f_k(x_k, u_k) - \hat{x}_{k+1|k}) \times (f_k(x_k, u_k) - \hat{x}_{k+1|k})^T|z^k] + Q_k. \]

Let \( k = k + 1 \). The algorithm then continues by Step 2.

The measurement predictive statistics \((8)\) and state predictive statistics \((11, 12)\) can be evaluated in a few special cases such as when the functions \( f_k \) and \( h_k \) are linear. In other cases, an approximate calculation of the statistics has to be applied.

The crucial difference between particular filters is given by the approximation. For example, the EKF [3] approximates the nonlinear function \( h_k \) by the first two terms of its Taylor expansion at \( \hat{x}_{k|k-1} \) as

\[ h_k(x_k) \approx h_k(\hat{x}_{k|k-1}) + h'_k(x_k - \hat{x}_{k|k-1}), \]

where \( h_k = \frac{\partial h_k}{\partial x_k} \) is the Jacobi matrix. Similarly, the nonlinear mapping \( f_k \) is approximated by the first two terms of its Taylor expansion at \( \hat{x}_{k|k} \) as

\[ f_k(x_k) \approx f_k(\hat{x}_{k|k}, u_k) + f'_k(x_k - \hat{x}_{k|k}, u_k), \]

where the Jacobi matrix is \( f_k = \frac{\partial f_k}{\partial x_k} \) evaluated at \( \hat{x}_{k|k} \).

Then the terms \( \hat{z}_{k|k-1}, P_{z|k-1}^{zz} \) and \( P_{k|k-1} \) for the filtering step and \( \hat{x}_{k+1|k} \) and \( P_{k+1|k}^{xx} \) for the prediction step are computed according to the following relations:

\[ \hat{z}_{k|k-1} = h_k(\hat{x}_{k|k-1}) + \hat{v}_k \]
\[ P_{z|k-1}^{zz} = H_kP_{k|k-1}H_k^T + R_k, \]
\[ P_{k|k-1} = P_{k|k-1} - K_kH_k^T, \]
\[ \hat{x}_{k+1|k} = f_k(\hat{x}_{k|k}, u_k) + \hat{w}_k, \]
\[ P_{k+1|k}^{xx} = F_kP_{k|k}F_k^T + Q_k. \]

III. UNSCENTED TRANSFORMATION AND RANDOMIZED UNSCENTED TRANSFORMATION

This section introduces the UT and RUT which are used in the UKF and RUKF to calculate the predictive statistics of the measurement \((8)\) and state \((11, 12)\).

A. Unscented transform

The basic idea of the UT can be illustrated by an example of transformation of a random variable through a nonlinear function \([6], [14]\).

Let \( x \in \mathbb{R}^{n_x} \) and \( y \in \mathbb{R}^{n_y} \) be random vector variables related through a known nonlinear function

\[ y = g(x) = [g_1(x), \ldots, g_{n_y}(x)]^T. \]

The variable \( x \) is given by the first two moments, i.e., the mean \( \hat{x} \) and the covariance matrix \( P_x \). The aim is to calculate the
mean and the covariance matrix of $y$, and the cross-covariance matrix $P_{xy}$, i.e.,

$$
\hat{y} = E[y] = E[g(x)],
$$

$$
P_y = \text{cov}[y] = E[(y - \hat{y})(y - \hat{y})^T],
$$

$$
P_{xy} = E[(x - \hat{x})(y - \hat{y})^T].
$$

The solution to the problem by means of the UT is based on approximation of the random variable $x$ by a set of deterministically chosen points, so called $\sigma$-points, $\{\mathcal{X}_i\}$ with corresponding weights $\{W_i\}$

$$
\mathcal{X}_0 = \hat{x}, W_0 = \frac{\kappa}{n_x + \kappa},
$$

$$
\mathcal{X}_i = \hat{x} + \sqrt{(n_x + \kappa)P_x}i, \quad W_i = \frac{1}{2(n_x + \kappa)},
$$

$$
\mathcal{X}_{n_x+i} = \hat{x} - \sqrt{(n_x + \kappa)P_x}i, \quad W_{n_x+i} = W_i,
$$

where $i = 1, \ldots, n_x$, the term $\sqrt{(n_x + \kappa)P_x}i$ represents the $i$-th column of the matrix $(n_x + \kappa)P_x = \sqrt{(n_x + \kappa)S_x}$. The matrix $S_x$ is the square-root of the covariance matrix $P_x$, so that $P_x = S_xS_x^T$. Then, each $\sigma$-point is transformed via the nonlinear function as

$$
y_i = g(\mathcal{X}_i), \quad \forall i,
$$

and the resulting characteristics are given as

$$
\hat{y}^{\text{UT}} = \sum_{i=0}^{2n_x} W_i y_i,
$$

$$
P^{\text{UT}}_y = \sum_{i=0}^{2n_x} W_i (y_i - \hat{y}^{\text{UKF}})(y_i - \hat{y}^{\text{UKF}})^T,
$$

$$
P^{\text{UT}}_{xy} = \sum_{i=0}^{2n_x} W_i (\mathcal{X}_i - \hat{x})(y_i - \hat{y}^{\text{UKF}})^T.
$$

Note that these results are only approximations of the true mean and the covariance and cross-covariance matrices which cannot generally be computed.

The variable $\kappa$ is the scaling parameter influencing accuracy of the approximation. The recommended setting of the scaling parameter is $\kappa = 3 - n_x$ [14].

The UT (24–30) represents a basic algorithm which suffers from a major weakness, namely the possible loss of positive semi-definiteness of $P_x$ (29) for multi-dimensional variable $x$ due to negative $\kappa$ (if $n_x > 3$, then $\kappa < 0$). Thus, several improved algorithms have been proposed which differ mainly in the $\sigma$-point computation, such as the scaled UT [7]. Another possibility is to choose $\kappa = 0$ for $n_x > 3$.

It is important to note that although the UT was designed to calculate moments of a random variable $y = g(x)$ given the moments of $x$, it can in fact be thought of as an integration rule to calculate integrals of the form

$$
\int_{\mathbb{R}^n} \phi(x)N(x; \hat{x}, P)\,dx
$$

with $\phi(x) = g(x)$ (calculation of the mean $\hat{y}$), $\phi(x) = (g(x) - \hat{y})(g(x) - \hat{y})^T$ (calculation of the covariance matrix $P_y$) and $\phi(x) = (x - \hat{x})(g(x) - \hat{y})^T$ (calculation of the cross-covariance matrix $P_{xy}$). This integration rule can be shown to provide only approximate value of the integrals [14].

### B. Randomized Unscented Transformation

The RUT introduced in [13] is a special case of a degree 3 stochastic integration rule (SIR3) proposed in [15], [16]. The SIR3 aims at evaluating an integral of the form

$$
\mu = \int_{\mathbb{R}^n} \phi(x)(2\pi)^{-n/2}e^{-\frac{1}{2}(x-\hat{x})^TP^{-1}(x-\hat{x})}\,dx,
$$

where $\phi(\cdot)$ is an arbitrary function. Relation (31) can be interpreted as computation of the expected value of the function $\phi(x)$ where $x$ is a random variable with $p(x) = N[\hat{x}, P]$, i.e. $\mu = E[\phi(x)]$. The SIR3 proposed in [15] for solution to (31) is given by the following algorithm

#### Algorithm 2: Degree 3 stochastic integration rule

**Step 1:** Choose an error tolerance $\varepsilon$ a maximum number of iterations $N_{\text{max}}$.

**Step 2:** Set the number of iterations $N = 0$, initial value of the integral $\hat{\mu} = 0_{n_x \times 1}$, initial square error of the integral estimate $V = 0_{n_x \times n_x}$ and calculate a square root $\sqrt{P_x}$ of $P_x$ such that $P_x = \sqrt{P_x}\sqrt{P_x}^T$. Note that $0_{a \times b}$ denotes a $a \times b$ matrix of zeros.

**Step 3:** Repeat (until $N = N_{\text{max}}$ or $\|V\| < \varepsilon$) the following loop:

a) Set $N = N + 1$.

b) Generate a uniformly random orthogonal matrix $U$ of dimension $n_x \times n_x$ and generate a random number $\rho$ from $\rho \sim \text{Chi}(n_x + 2)$.

c) Compute a set of points $\chi_i$ and appropriate weights $\omega_i$ according to

$$
\chi_0 = \hat{x}, \quad \omega_0 = 1 - \frac{n_x}{\rho},
$$

$$
\chi_i = \hat{x} - \rho U(\sqrt{P_x}), \quad \omega_i = \frac{1}{2\rho^2},
$$

$$
\chi_{n_x+i} = \hat{x} + \rho U(\sqrt{P_x}), \quad \omega_{n_x+i} = \omega_i,
$$

where $i = 1, 2, \ldots, n_x$.

d) Compute the value $S$ of the integral at current iteration

$$
S = \sum_{i=0}^{2n_x} \omega_i \phi(\chi_i),
$$

and use it to update the approximate value $\hat{\mu}$ of the integral and square error of the integral estimate $V$ as

$$
D = (S - \hat{\mu})/N,
$$

$$
\hat{\mu} = \hat{\mu} + D,
$$

$$
V = (N - 2)V/N + DD^T.
$$

**Step 4:** The approximate value of the integral $\mu$ is $\hat{\mu}$. 

The SIR3 can be thought of as a blend of a cubature rule and Monte Carlo method. Contrary to the UT the SIR3 provides asymptotically exact estimates of integral, i.e. $E[\hat{\mu}] = \mu$ [15]. The SIR3 consists of the spherical integration rule, which generates a set of $2n_x$ points located on an unit $n_x$-sphere, and
the radial integration rule, which governs spread of the points within the \( \mathbb{R}^n \). In [13] it was shown that the location of points generated by the UT matches location of points generated by the SIR3 when choosing a trivial orthogonal matrix \( U = I \) and the parameter \( \rho \) as \( \rho^2 = \kappa + n_x \). This fact lead to coining the term randomized UT in [13] for the SIR3 used to estimate moments of a random variable.

Utilization of the RUT in the filtering or prediction steps of the generic filter requires computation of three statistics of the random variable \( y \) - the mean \( \tilde{y} \), the covariance matrix \( P_y \) and the cross-covariance matrix \( P_{xy} \). These can be calculated by three runs of the SIR3 algorithm i) by considering \( \phi(x) = g(x) \), ii) \( \phi(x) = (g(x) - \tilde{y})(g(x) - \tilde{y})^T \) and iii) \( \phi(x) = (x - \tilde{x})(g(x) - \tilde{y})^T \).

IV. ASPECTS OF THE RANDOMIZED UNSCENTED TRANSFORMATION

This section first focuses on some important aspects of the RUT and then these aspects are illustrated in a numerical example involving polar to Cartesian transformation, which is ubiquitous in target tracking applications.

- To obtain the desired statistics of \( y \), the RUT introduced in the previous section requires three runs of the SIR3 algorithm, which leads to high computational costs. Instead it is possible to obtain estimates of the statistics in a single run of the SIR3 algorithm. The idea is to start first with simultaneous computing of the mean \( \tilde{y} \), second raw moment \( \mathbb{E} [yy^T] \) by considering \( \phi(x) = g(x)g(x)^T \) and \( \mathbb{E} [xy^T] \) by considering \( \phi(x) = xg(x)^T \). Then, the covariance matrix \( P_y \) is given by \( P_y = \mathbb{E} [yy^T] - \tilde{y}\tilde{y}^T \) and the cross-covariance matrix \( P_{xy} \) by \( P_{xy} = \mathbb{E} [xy^T] - \tilde{x}\tilde{y}^T \).

- The RUT requires a pdf of \( x \) to be Gaussian, as opposed to the UT which itself requires the first two moments of \( x \) only, regardless of \( x \) distribution. This might be seen as drawback of the RUT however, if no other information concerning the random variable \( x \) distribution besides its first two moments is known, the Gaussian pdf is a convenient distribution to be considered. The reason is that the Gaussian distribution has the largest entropy of all distributions with a given mean and covariance matrix (i.e. it represents the worst case for the given first two moments) [17]. If the distribution of \( x \) is known and non-Gaussian, the importance sampling technique can be used to cope with this fact [16].

- Now, the number of iterations of the RUT will be discussed. In [15] it was shown that for the function \( \phi(\cdot) \) in (31) in the form of a polynomial of order less than 3 the result of the SIR3 is exact in one iteration, i.e. \( \hat{\mu} = \mu \) for \( N_{max} = 1 \). The functions in state space description are seldom of this kind. However, due to the bell-curve shape of the Gaussian density of \( x \), estimate quality provided by the SIR3 is substantially affected by behavior of the function \( \phi \) only within a certain neighborhood of \( \hat{x} \). The size of the interval is governed by the covariance matrix \( P_y \). If, for example, the behavior within the neighborhood of \( x \) is close to linear, than only one iteration of the RUT algorithm is required. In such a case, performance of the RUT will be close to that of the UT. On the other hand, a function \( \phi \) that is highly nonlinear within the neighborhood will lead to an increased number of iterations.

To illustrate the above mentioned aspects, consider an example of conversion from polar to Cartesian coordinates [7]. This conversion is a usual component of observation models of many sensors. Position of a target in 3D is obtained by a sensor providing range \( r \), bearing \( b \) and elevation \( e \) in a local coordinate frame. The sensor data have to be converted to some global Cartesian coordinate frame

\[
\begin{bmatrix}
x \\
y \\
z \\
\end{bmatrix} =
\begin{bmatrix}
r \cos(b) \sin(e) \\
r \sin(b) \sin(e) \\
r \cos(e) \\
\end{bmatrix} .
\]

(38)

The pdf of \( x = [r, b, e]^T \) is supposed to be

\[
p(x) = \mathcal{N} \{ x : [1, \pi/4, \pi/4]^T, \text{diag}([0.001, 0.01, 0.05]) \}
\]

(39)

The UT and RUT were used to estimate the mean \( \hat{y} = [\hat{x}, \hat{y}, \hat{z}] \) of \( y = [x, y, z]^T \), covariance matrix \( P_y \), and cross-covariance matrix \( P_{xy} \). The values computed by the UT and RUT were copared with the true values obtained by numerical integration. The quality of the mean evaluation was measured by the root mean squared error \( \text{RMSE} = \sqrt{(\hat{x}^{est} - \hat{x}^{true})^2 + (\hat{y}^{est} - \hat{y}^{true})^2 + (\hat{z}^{est} - \hat{z}^{true})^2} \). The quality of the covariance and cross-covariance matrices was measured by an absolute value of the difference between trace of the estimated and true covariance and cross-covariance matrices, i.e. \( \hat{P}_y = | \text{tr}(P_y^{est}) - \text{tr}(P_y^{true}) | \) and \( \hat{P}_{xy} = | \text{tr}(P_{xy}^{est}) - \text{tr}(P_{xy}^{true}) | \).

Three experiments were carried out. As the results achieved by the RUT exhibit random nature, the experiments involved 1000 Monte Carlo (MC) simulations.

The first experiment involved comparison of the UT and the RUT. For the UT, the recommended setting for the scaling parameter is, according to Subsection III-A \( \kappa = 3 - 3 = 0 \). To take into account more possible choices, three values of \( \kappa \) were considered: \( \kappa = 0, 1, 2 \).

The RUT was run with a fixed number of iterations \( N_{max} = 2, 5, 10, 20, \ldots, 100, 200, 500, 1000 \). The results are illustrated in Fig. 1 (RMSE), Fig. 2 (\( \hat{P}_y \)), and Fig. 3 (\( \hat{P}_{xy} \)). Note that the UT provides only a single value of the error for each \( \kappa \), which, for illustrative reasons, was depicted by a horizontal line. From the figures it can be seen that the RUT clearly outperforms the UT in approximately 10 iterations in all monitored errors no matter how \( \kappa \) is specified. From the figures it also is clear that a single value of the scaling parameter \( \kappa \) in the UT does not lead to the smallest error for all statistics. In [18] it was shown that by tuning the scaling parameter \( \kappa \) the UT may achieve much higher estimate quality than the UT with the standard choices of \( \kappa \). The price for the increased quality is high computational costs of the tuning procedure.
Thus, the RUT with a single run of SIR3 will be preferred in 63 percent lower than that of the RUT with multiple SIR3 runs. The computational costs of the RUT with a single SIR run were by approximately 4 percent for the mean estimates and 2 percent for covariance and cross-covariance matrix estimates. The experiments showed that the RUT using only one SIR3 run leads to significant computational savings with a negligible decrease of estimate quality. Moreover, the accuracy of the parameter $\alpha$ clearly illustrates that for higher degree of nonlinearity of $\phi$ the mean estimate quality provided by the RUT deteriorates. The RUT keeps the RMSE approximately constant by increasing the number of iterations.

The above experiments indicate that using the RUT for evaluation of the predictive measurement and state statistics in the generic filter could be more suitable than using the UT. The experiments showed that the RUT using only one SIR3 run leads to significant computational savings with a negligible decrease of estimate quality. Moreover, the accuracy of the RUT can be naturally controlled by the parameter $\epsilon$ which determines accuracy of the statistic estimates. This direct relation between the parameter and the estimate accuracy is much more suitable for controlling quality of the filter estimates than tuning the scaling parameter $\kappa$ in the UKF.

V. RANDOMIZED UNSCENTED KALMAN FILTER

The algorithm of the RUKF corresponds to the Algorithm 1 of the generic filter with the state and measurement predictive statistics calculated by the following RUT algorithm. For simplicity reasons, RUT calculation of only measurement behind see the discussion in the beginning of Section IV. In each simulation the covariance matrix $P_x$ was obtained by multiplying the covariance matrix from the first experiment by a scalar parameter $\alpha = 0.1, 0.3, 0.5, 1, 2, \ldots, 10$, i.e. parameter $\alpha < 1$ made the neighborhood of $\hat{x}$ smaller and $\alpha > 1$ enlarged it.

In Fig 4 dependency of the average number of iterations on the parameter $\alpha$ is depicted. The results confirm the fact that the number of iterations is directly related to the degrees of nonlinearity. Fig 5 showing dependency of the RMSE on the parameter $\alpha$ clearly illustrates that for higher degree of nonlinearity of $\phi$ the mean estimate quality provided by the UT deteriorates. The RUT keeps the RMSE approximately constant by increasing the number of iterations.

The above experiments indicate that using the RUT for evaluation of the predictive measurement and state statistics in the generic filter could be more suitable than using the UT. The experiments showed that the RUT using only one SIR3 run leads to significant computational savings with a negligible decrease of estimate quality. Moreover, the accuracy of the RUT can be naturally controlled by the parameter $\epsilon$ which determines accuracy of the statistic estimates. This direct relation between the parameter and the estimate accuracy is much more suitable for controlling quality of the filter estimates than tuning the scaling parameter $\kappa$ in the UKF.
predictive statistics will be presented. State predictive statistics will be calculated analogously. For clarity purposes the time index $k | k - 1$ will be dropped in the algorithm.

Note that a careful implementation of the SIR3 requires specification of a minimum number of iterations $N_{\text{min}}$.

**Algorithm 3: Randomized unscented transform**

**Step 1:** Choose an error tolerance $\varepsilon$, a maximum number of iterations $N_{\text{max}}$, a minimum number of iterations $N_{\text{min}}$ and set the number of iterations $N = 0$. Set the initial values of the measurement statistics to be calculated

- predictive mean $\hat{z} = 0_{n_z \times 1}$
- predictive covariance matrix $P^{zz} = 0_{n_z \times n_z}$
- predictive cross-covariance matrix $P^{xz} = 0_{n_x \times n_z}$
- initial square error of the mean estimate $V = 0_{n_z \times n_z}$

Calculate a square root $\sqrt{P}$ of the state predictive covariance matrix $P$.

**Step 2:** Repeat (while $N < N_{\text{min}}$ and until $N = N_{\text{max}}$ or $\|V\| < \varepsilon$) the following loop:

a) Set $N = N + 1$.

b) Generate a uniformly random orthogonal matrix $U$ of dimension $n_x \times n_x$ and generate a random number $\rho$ from $\rho \sim \text{Chi}(n_z + 2)$.

c) Compute a set of points $X_i$ and appropriate weights $\omega_i$ according to

$$X_0 = \hat{x}, \quad \omega_0 = 1 - \frac{n_z}{n_x},$$

$$X_i = \hat{x} - \rho U(X_i) \rho, \quad \omega_i = \frac{1}{2 n_x},$$

$$X_{n_z+i} = \hat{x} + \rho U(X_i) \rho, \quad \omega_{n_z+i} = \omega_i,$$

where $i = 1, 2, \ldots, n_x$.

d) Compute the values $S_{zz}$, $S_{pz}$, and $S_{pz}$ of the statistics at current iteration

$$S_z = \sum_{i=0}^{2 n_x} \omega_i \mathbf{h}_k(X_i),$$

$$S_{pz} = \sum_{i=0}^{2 n_x} \omega_i \mathbf{h}_k(X_i)\mathbf{h}_k(X_i)^T,$$

$$S_{pz} = \sum_{i=0}^{2 n_x} \omega_i \mathbf{h}_k(X_i)\mathbf{h}_k(X_i)^T.$$

Update the values of the statistics and square error of the mean estimate $V$ as

$$D = (S_z - \hat{z})/N,$$

$$\hat{z} = \hat{z} + D,$$

$$V = (N - 2)V/N + DD^T,$$

$$P^{zz} = ((N - 1)P^{zz} + S_{pz})/N,$$

$$P^{xz} = ((N - 1)P^{xz} + S_{pz})/N.$$

**Step 3:** Correct the covariance matrices as

$$P^{zz} = P^{zz} - \hat{z} \hat{z}^T$$

$$P^{xz} = P^{xz} - xz^T$$

**VI. APPLICATION OF THE RANDOMIZED UNSCENTED KALMAN FILTER TO REENTRY PROBLEM**

Consider a vehicle entering atmosphere at high altitude with a high speed [7]. The state of the model consists of position of the vehicle ($x_1$ and $x_2$), velocity ($x_3$ and $x_4$) and aerodynamic properties $x_5$. Its dynamics is described by

$$\begin{bmatrix}
\dot{x}_1(t) \\
\dot{x}_2(t) \\
\dot{x}_3(t) \\
\dot{x}_4(t) \\
\dot{x}_5(t)
\end{bmatrix} =
\begin{bmatrix}
x_3(t) \\
x_4(t) \\
D(t) x_3(t) + G(t) x_1(t) + w_1(t) \\
D(t) x_4(t) + G(t) x_2(t) + w_2(t) \\
w_3(t)
\end{bmatrix},$$

where $D(t)$ is the drag-related force term given by

$$D(t) = \beta(t) \exp\left(\frac{R_0 - R(t)}{H_0}\right)V(t)$$

with $\beta(t)$ being the ballistic coefficient $\beta(t) = \beta_0 \exp(x_5(t))$, $R(t)$ being the distance form the center of the earth $R(t) = \sqrt{x_1^2(t) + x_2^2(t)}$ and $V(t)$ being the speed $V(t) = \sqrt{x_3^2(t) + x_4^2(t)}$. The term $G(t)$ denotes the gravity-related term $G(t) = \frac{G_m}{R(t)}$. The typical values characterizing the vehicle and environment are [7] $\beta_0 = -0.59783$, $H_0 = 13.406$, $G_m = 3.9860 \times 10^5$ and $R_0 = 6347$. The accelerations $\dot{x}_3(t)$ and $\dot{x}_4(t)$ and the aerodynamic properties $x_5(t)$ are corrupted by random noises with $\mathbb{E}[w(t)w(\tau)^T] = \text{diag}(10^{-5}, 2.4064 \times 10^{-5}, 0) \delta(t - \tau)$. The initial conditions for the vehicle are given by the mean and covariance matrix

$$\mathbf{x}(0) = \begin{bmatrix}
6500.4 & \\
349.14 & \\
-1.8093 & \\
-6.7967 & \\
0.6932 & \\
\end{bmatrix},$$

$$\mathbf{P}_0 = \begin{bmatrix}
10^{-6} & 0 & 0 & 0 & 0 \\
0 & 10^{-6} & 0 & 0 & 0 \\
0 & 0 & 10^{-6} & 0 & 0 \\
0 & 0 & 0 & 10^{-6} & 0 \\
0 & 0 & 0 & 0 & 0 \\
\end{bmatrix}.\quad (56)$$

The position of the vehicle is observed by a radar located at $[x_{1,r}, x_{2,r}]$ providing the measurement $z_k$ at a frequency of 10 Hz. The measurement $z_k$ consists of range $r_k$ and bearing $b_k$

$$r_k = \sqrt{(x_1(t_k) - x_{1,r})^2 + (x_2(t_k) - x_{2,r})^2 + v_{1,k}},$$

$$b_k = \tan^{-1}\left(\frac{x_2(t_k) - x_{2,r}}{x_1(t_k) - x_{1,r}}\right) + u_{2,k},$$

where $\text{var}(v_{1,k}) = 10^{-3}[km^2]$ and $\text{var}(u_{2,k}) = 17 \times 10^{-3}[rad^2]$ [7].
The state was estimated by means of the EKF, UKF ($\kappa = 0$) and RUKF ($\varepsilon = 10^{-17}$, $N_{\text{max}} = 100$) using 1000 MC simulations. The filters used the discrete time model obtained by Euler approximation of (53) with the integration step 100ms. The initial condition for the filters were

$$
\mathbf{x}_{0|\text{1}} = \begin{bmatrix}
6500.4 \\
349.14 \\
-1.8093 \\
-6.7967 \\
0
\end{bmatrix},
\quad \text{(59)}
$$

$$
\mathbf{P}_{0|\text{1}} = \begin{bmatrix}
10^{-6} & 0 & 0 & 0 & 0 \\
0 & 10^{-6} & 0 & 0 & 0 \\
0 & 0 & 10^{-6} & 0 & 0 \\
0 & 0 & 0 & 10^{-6} & 0 \\
0 & 0 & 0 & 0 & 1
\end{bmatrix}.
\quad \text{(60)}
$$

Note that the true initial conditions (55–56) and the filter initial conditions (59–60) differ in the mean and variance of the aerodynamic properties. For estimation the state noise was given by $\mathbb{E} [\mathbf{w}(t) \mathbf{w}(t)^\top] = \text{diag}(2.4064 \times 10^{-5}, 2.4064 \times 10^{-5}, 10^{-5})\delta(t - \tau)$.

Performance of the filters was measured using the root mean squared error (RMSE) and non-credibility index (NCI). The NCI is defined as

$$
\text{NCI}_k = \frac{1}{M} \sum_{i=1}^{M} \left[ 10 \log_{10} \left( (\mathbf{x}_i^k - \mathbf{x}_{\text{hat}}^k)^\top (\mathbf{P}_{\text{hat}}^k)^{-1} (\mathbf{x}_i^k - \mathbf{x}_{\text{hat}}^k) \right) \\
- 10 \log_{10} \left( (\mathbf{x}_0^k - \mathbf{x}_{\text{hat}}^k)^\top \Sigma_k^{-1} (\mathbf{x}_0^k - \mathbf{x}_{\text{hat}}^k) \right) \right],
$$

where $\mathbf{P}_{\text{hat}}^k$ is the covariance matrix of the estimate provided by the filter at the $i$-th MC run and $\Sigma_k$ is the mean square error matrix of the estimate [19], [20].

The achieved results are depicted on Fig 6 (RMSE) and Fig 7 (NCI). Fig 8 illustrates average number of iterations of the RUKF filtering step. The figures show that the RUKF achieves higher estimate quality in terms of RMSE than the UKF especially between 20 and 80 seconds. This correlates with the position and velocity evolution (not depicted here), which exhibits strongly nonlinear behavior within this time interval. To accommodate this nonlinear behavior the RUKF increased the number of iterations (see Fig 8). Performance of the EKF is quite poor, especially for the position and velocity estimates.

Also in terms of NCI the EKF is overly optimistic whereas the UKF and RUKF stay pessimistic for the whole simulation interval.

VII. CONCLUSION

The paper dealt with state estimation of nonlinear dynamic systems by the randomized unscented Kalman filter. The filter follows the Kalman filtering framework and to compute the measurement and state predictive statistics, the randomized unscented transform, which is based on a degree 3 stochastic integration rule, is utilized. The RUT and the RUKF itself are introduced with a special focus on target tracking. The RUT is analyzed and applied to a problem of polar to...
Cartesian coordinate transformation, which is ubiquitous in target tracking. This application demonstrates advantages of the RUT over the plain UT. Application of the RUKF is then illustrated in a reentry example.

ACKNOWLEDGEMENT

This work was supported by the European Regional Development Fund (ERDF), project NTIS - New Technologies for Information Society, European Centre of Excellence, CZ.1.05/1.1.00/02.0090 and by the Czech Science Foundation, project no. GACR P103/11/1353.

REFERENCES


