The Development of a Randomised Unscented Kalman Filter *

Jindřich Duník, Ondřej Straka and Miroslav Šimandl

Department of Cybernetics & Research Centre Data- Algorithms-Decision Making, 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 state estimation of nonlinear stochastic dynamic systems. Traditional filters providing local estimates of the states, such as the extended Kalman filter, unscented Kalman filter or the cubature Kalman filter, are based on approximations which lead to biased estimates of the state and measurement statistics. The aim of the paper is to propose a new local filter that utilises a randomised unscented transformation which is a special case of stochastic integration rules providing an unbiased estimate of an integral. The new filter provides estimates of higher quality than the traditional filters and renders a randomised version of the unscented Kalman filter. The proposed filter is illustrated in a numerical example.

Keywords: State estimation; Estimation theory; Kalman filtering

1. INTRODUCTION

The problem of recursive state estimation of discrete-time stochastic dynamic systems from noisy or incomplete measurement data has been a subject of considerable research interest for the last several decades (Sorenson, 1974; Anderson and Moore, 1979; Ito and Xiong, 2000; Kushner and Budhiraja, 2000).

The general solution to the estimation problem is given by the Bayesian recursive relations (BRR’s) for computation of probability density functions (pdf’s) of the state conditioned by the measurements (Anderson and Moore, 1979). These pdf’s provide a full description of the immeasurable state. The closed form solution to the BRR’s is available only for a few special cases (Anderson and Moore, 1979), e.g. for linear Gaussian system, which leads to the well-known Kalman filter (KF).

In other cases, it is necessary to apply some approximate methods. These methods can be divided into two groups: local and global methods (Sorenson, 1974; Arasaratnam and Haykin, 2000; Ito and Xiong, 2000; Kushner and Budhiraja, 2000; Quine, 2006; Julier and Uhlmann, 2004; Duník et al., 2005; Šimandl and Duník, 2009; Šimandl and Šimandl, 2009).

The local methods are often based on an approximation of the nonlinear functions in the state or measurement equation so that the KF design technique can be used for the BRR’s solution. This approach leads to description of the conditional pdf’s of the state estimate by only the first two moments, i.e. the mean value and covariance matrix. This rough approximation of the a posteriori estimates induces local validity of the state estimates and consequently impossibility to generally ensure convergence of the local filter estimates. Thus, estimates generated by the local filters are suitable mainly for point estimation. On the other hand, the advantage of the local methods can be found in a relative simplicity of the BRR’s solution and acceptable computational demands.

The global methods are based on a certain type of approximation of the conditional pdf of the state estimate to accomplish better state estimates. Unfortunately, they have higher computational demands than the local methods (Šimandl and Duník, 2009). The global methods are not considered in this paper.

The standard local methods approximate nonlinear functions in the state or the measurement equation by the Taylor expansion up to the first or second order. The BRR’s solution based on this approximation leads to e.g. the extended Kalman filter or the second order filter (Anderson and Moore, 1979).

In the last decade, novel approaches to local filter design based on the polynomial interpolation (Nørgaard et al., 2000; Schei, 1997; Ito and Xiong, 2000; Duník et al., 2005; Šimandl and Duník, 2009) or on the unscented transformation (UT) (Julier et al., 2000; Ito and Xiong, 2000; Kushner and Budhiraja, 2000; Quine, 2006; Julier and Uhlmann, 2004; Duník et al., 2005; Šimandl and Duník, 2009; Arasaratnam and Haykin, 2009), have been published. The approximation of the nonlinear functions by means of Stirling’s polynomial interpolation leads to the divided difference filters (DDF’s). Further, the DDF’s can be classified into the divided difference filter first order (DD1) and the divided difference filter second order (DD2) which are based on the first and second order Stirling’s interpolation formula, respectively (Nørgaard et al., 2000). Instead of the direct substitution of the nonlinear functions in the system description by an approximation, an approximation of the pdf’s representing state estimates by a set of deterministically chosen weighted points (so called σ-points) can be utilised as a base for the local filters. This transformation is known as the UT. The unscented Kalman filter (UKF) (Julier et al., 2000), the Gauss-Hermite filter (Ito and Xiong, 2000) or the cubature Kalman Filter (Arasaratnam and Haykin, 2009) exemplify this approach. It is very important to mention that for the UKF and the DDF’s (and their variants) common features can be found although they come from quite different basic ideas (Nørgaard et al., 2000; Lefebvre et al., 2002; Šimandl and Duník, 2009).

* The work was supported by the Ministry of Education, Youth and Sports of the Czech Republic, project no. 1M0572 and by the Czech Science Foundation, project No. P103/11/1353.
Therefore, these local filters are often jointly referred to as the σ-point KFs or the derivative-free Kalman (local) filters.

All the local filters follow the structure of the KF algorithm and can be written in a unified framework (Šimandl and Duník, 2009). The only difference between particular local filters is the employed approximation in the filter design, more specifically the approximation used in computation of integrals. The currently used approximations (i.e. those based on the Taylor or Stirling expansion, the UT etc.) suffer, however, from one major weakness; their provide an approximate solution to an integral containing a systematic error which has subsequently an adverse impact on the local filter estimation performance.

The goal of the paper is to propose an improved version of the UT allowing computation of integrals without the systematic error and to employ the improved UT in the local filter design.

The paper is organised as follows: System specification and Bayesian estimation is introduced in Section 2, the UT and UKF are presented and analysed in sections 3 and 4. The randomised UT is described in Section 5 and the new filter is proposed in Section 6. In Section 7 a numerical illustration of the new filter is given and concluding remarks are drawn in Section 8.

2. SYSTEM SPECIFICATION AND BAYESIAN ESTIMATION

Let the discrete-time nonlinear stochastic system be considered

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

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

where the vectors \( x_k \) ∈ \( \mathbb{R}^n \) and \( z_k \) ∈ \( \mathbb{R}^m \) represent the immeasurable state of the system and measurement at time instant \( k \), respectively, \( f_k : \mathbb{R}^n \to \mathbb{R}^n \) and \( h_k : \mathbb{R}^n \to \mathbb{R}^m \) are the known vector mappings, and \( w_k \) ∈ \( \mathbb{R}^n \), \( v_k \) ∈ \( \mathbb{R}^m \) 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. \( p_w(w_k) = \mathcal{N}(w_k; 0, Q_k) \) and \( p_v(v_k) = \mathcal{N}(v_k; 0, R_k) \), respectively. The pdf of the initial state is supposed to be Gaussian and known as well, i.e. \( p_{x_0}(x_0) = \mathcal{N}(x_0; \bar{x}_0, P_0) \), and independent of the noises.

The aim of the state estimation is to find a state estimate in the form of the conditional pdf \( p(x_k|z^k) \) in which \( z^k = [z_0, z_1, \ldots, z_k] \). A widely used approach for computation of the conditional pdf is based on the solution of the Bayesian recursive relations (BRR’s) (Anderson and Moore, 1979)

\[ p(x_k|z^k) = \frac{p(z_k|z^{k-1}) p(x_k|x_{k})}{p(z_k|z^{k-1})}, \] \hspace{1cm} (3)

\[ p(x_{k+1}|z^k) = \int p(x_{k+1}|x_k) p(x_k|z^k) dx_k, \] \hspace{1cm} (4)

where \( p(x_{k+1}|z^k) \) represents one-step ahead predictive conditional pdf. The BRR’s are solvable for a few special cases only, involving a linear Gaussian system. In other cases an approximation has to be employed in BRR’s solution which may lead to a recursive computation of the first two conditional moments, i.e. the filtering and predictive means \( \hat{x}_{k|k} = E[x_k|z^k] \) and \( \hat{x}_{k+1|k} = E[x_{k+1}|z^k] \), respectively, and the covariance matrices \( P_{k|k} = \text{cov}[x_k|z^k] \) and \( P_{k+1|k} = \text{cov}[x_{k+1}|z^k] \) only, instead of the conditional pdf’s. Such methods are often denoted as the local filters. Mention that the conditional moments can be understood as a Gaussian approximations of the conditional pdfs, i.e. \( p(x_k|z^k) \approx \mathcal{N}(x_k; \hat{x}_{k|k}, P_{k|k}) \) and \( p(x_{k+1}|z^k) \approx \mathcal{N}(x_{k+1}; \hat{x}_{k+1|k}, P_{k+1|k}) \) (Ito and Xiong, 2000; Arasaratnam and Haykin, 2009).

3. UNSCENTED TRANSFORMATION AND UNSCENTED KALMAN FILTER

The algorithms of the local filters can be expressed using the same structure; the filtering and predictive means and covariance matrices are recursively computed. The only difference arises in computation of predictive estimates of the measurement and the state, where different approximation techniques are employed.

In this section an approximation using the UT is selected and used in the filter design.

3.1 Unscented transformation

The basic idea of the UT can be illustrated by an example of transformation of a random variable through a nonlinear function (Julier et al., 2000; Šimandl and Duník, 2009).

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

\[ y = g(x) = [g_1(x), \ldots, g_n(x)]^T. \] \hspace{1cm} (5)

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.

\[ \hat{y} = E[y] = E[g(x)], \] \hspace{1cm} (6)

\[ P_y = \text{cov}[y] = E[(y - \hat{y})(y - \hat{y})^T], \] \hspace{1cm} (7)

\[ P_{xy} = E[(x - \hat{x})(y - \hat{y})^T]. \] \hspace{1cm} (8)

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 σ-points, \( \{X_i\} \) with corresponding weights \( \{W_i\} \).

\[ X_0 = \hat{x}, \quad W_0 = \frac{\kappa}{n_x + \kappa}, \] \hspace{1cm} (9)

\[ X_i = \hat{x} + \left( \sqrt{(n_x + \kappa)P_x} \right)_i, \quad W_i = \frac{1}{2(n_x + \kappa)}, \] \hspace{1cm} (10)

\[ X_{n_x+i} = \hat{x} - \left( \sqrt{(n_x + \kappa)P_x} \right)_i, \quad W_{n_x+i} = W_i, \] \hspace{1cm} (11)

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

\[ \hat{y}_i = g(X_i), \quad \forall i, \] \hspace{1cm} (12)

and the resulting characteristics are given as

\[ \hat{y}_{\text{UT}} = \sum_{i=0}^{2n_x} W_i \hat{y}_i, \] \hspace{1cm} (13)

\[ P_{y|\text{UT}} = \sum_{i=0}^{2n_x} W_i (\hat{y}_i - \hat{y}_{\text{UT}})(\hat{y}_i - \hat{y}_{\text{UT}})^T, \] \hspace{1cm} (14)

\[ P_{xy|\text{UT}} = \sum_{i=0}^{2n_x} W_i (X_i - \hat{x})(\hat{y}_i - \hat{y}_{\text{UT}})^T. \] \hspace{1cm} (15)
Note that these results are only approximations of the true mean and the 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 $\kappa$ is $\kappa = 3 - n_x$ (Julier et al., 2000).

The UT (9)–(15) represents a basic algorithm which suffers from a major weakness, namely the loss of positive semi-definiteness of $P^\text{UT}_y$ (14) 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, e.g. scaled UT (Julier and Uhlmann, 2004). Another possibility is to choose $\kappa = 0$ for $n_x > 3$.

For the sake of presentation clarity, the computation of the $\sigma$-point set (9)–(11) will be denoted as

$$[[\{X_i\}, \{W_i\}]] = \text{UTpoints}(\hat{x}_k, P_{kk}, \kappa),$$

where $\{X_i\} = \{X_0, \ldots, X_{2n_x}\}$ is the set of all $\sigma$-points with corresponding weights in $\{W_i\}$.

### 3.2 Unscented Kalman filter

The unscented transformation (9)–(15) can be directly used in the design of the UKF (Julier et al., 2000). The resulting algorithm of the UKF can be summarised as follows:

**Step 1:** Set the time instant $k = 0$ and define the a priori initial condition by the predictive mean $\hat{x}_{0|0} = E[x_0] = \hat{x}_0$ and the predictive covariance matrix $P_{0|0} = \text{cov}[x_0] = P_0$.

**Step 2:** The state predictive estimate is updated with respect to the last measurement $z_k$ according to

$${\hat{x}}_{k|k-1} = \hat{x}_{k|k-1} + K_{k|k-1} (z_k - \hat{z}_{k|k-1}),$$

$${P}_{k|k-1} = {P}_{k|k-1} - K_{k|k-1} {P}_{z,k|k-1} P^T_{z,k|k-1},$$

where

$${\hat{z}}_{k|k-1} = E[z_k|z^{k-1}] = E[h_k(x_k)|z^{k-1}] \approx \sum_{i=0}^{2n_x} W_i Z_i x_{i,k|k-1},$$

$${P}_{z,k|k-1} = E[(z_k - \hat{z}_{k|k-1})(z_k - \hat{z}_{k|k-1})^T|z^{k-1}] =$$

$$= E[(h_k(x_k) - \hat{h}_k(x_k))(h_k(x_k) - \hat{h}_k(x_k))^T|z^{k-1}] + R_k \approx \sum_{i=0}^{2n_x} W_i (Z_i x_{i,k|k-1} - \hat{z}_{i,k|k-1})(Z_i x_{i,k|k-1} - \hat{z}_{i,k|k-1})^T + R_k,$$

and $[[\{X_i\}_{i,k|k-1}, \{W_i\}_{i,k|k-1}]] = \text{UTpoints}(\hat{x}_{i,k|k-1}, P_{i,k|k-1}, \kappa)$.

**Step 3:** The predictive statistics are given by

$${\hat{x}}_{k+1|k} = E[x_{k+1}|z^k] = E[f_k(x_k)|z^k] \approx \sum_{i=0}^{2n_x} W_i X_i x_{i,k+1|k},$$

$${P}_{k+1|k} = E[(x_{k+1} - \hat{x}_{k+1|k})(x_{k+1} - \hat{x}_{k+1|k})^T|z^k] \approx \sum_{i=0}^{2n_x} W_i (X_i x_{i,k+1|k} - \hat{x}_{i,k+1|k})(X_i x_{i,k+1|k} - \hat{x}_{i,k+1|k})^T + Q_k,$$

where $[[\{X_i\}_{i,k|k}, \{W_i\}_{i,k|k}]] = \text{UTpoints}(\hat{x}_{i,k|k}, P_{i,k|k}, \kappa)$. Let $k = k+1$ and algorithm continues by **Step 2**.

The parameter $\kappa$ is usually selected in accord with the recommendation given in the case of the UT. Note, however, that the adaptive setting of the scaling parameter may improve estimation performance of the UKF (Dunkert et al., 2010).

### 4. PROPERTIES OF UNSCENTED TRANSFORMATION AND PROBLEM STATEMENT

As was mentioned above the UT (5)–(15) provides an approximate solution to the problem given by relations (5)–(8). In this section an error of the mean $\hat{y}_\text{UKF}$ and the covariance matrices $P^\text{UKF}$ and $P^\text{xy} \text{UKF}$ provided by the UT is analysed.

#### 4.1 Systematic error of results computed using UT

The error of the mean computed using the UT will be considered first. The error will be expressed by means of the Taylor expansion of the true mean $\hat{y}$ (6) and the approximate $\hat{y}\text{UT}$ (13).

The Taylor expansion of the true mean of $y$ is of the form

$$\hat{y} = g(\hat{x}) + \frac{1}{2} \sum_{i=1}^{n_x} \sum_{j=1}^{n_x} P_{i,j}(i, j) \nabla^2_{ij} g +$$

$$+ E \left[ \frac{D^3_{\partial g(x)}(x)}{4!} + \frac{D^5_{\partial g(x)}(x)}{6!} + \ldots \right].$$

where

$$\nabla^2_{ij} g = \frac{\partial^2}{\partial x_i \partial x_j} g(x)|_{x = \tilde{x}},$$

$$\frac{1}{j!} D^j_{\partial g(x)}(x) = \frac{1}{j!} \left( \sum_{i=1}^{n_x} \Delta x_i \frac{\partial}{\partial x_i} \right)^j g(x)|_{x = \tilde{x}},$$

$$\Delta x = x - \tilde{x},$$

$$x_i$$ is the $i$-th element of the vector $x$.

The approximate mean (13) can be simplified to

$$\hat{y}_\text{UT} = g(\hat{x}) + \frac{(n_x + \kappa)}{2(n_x + \kappa)} \sum_{i=1}^{n_x} \sum_{j=1}^{n_x} P_{i,j}(i, j) \nabla^2_{ij} g +$$

$$+ \frac{1}{2(n_x + \kappa)} \sum_{p=1}^{2n_x} \left( \frac{D^3_{\partial g(x_p)}}{4!} + \frac{D^5_{\partial g(x_p)}}{6!} + \ldots \right),$$

where

$$D^j_{\partial g(x_p)}(x) = \frac{1}{j!} \left( \sum_{i=1}^{n_x} (X_p(i) - \hat{x}_i) \frac{\partial}{\partial X_p(i)} \right)^j g(x)|_{X_p = \hat{x}},$$

is the $j$-th term of the Taylor series expansion of the $p$-th $\sigma$-point $g(X_p)$ and $X_p(i)$ is the $i$-th element of $X_p$. Derivation of relation (29) can be found e.g. in the paper by Julier et al. (2000).

The particular error, in this paper called the systematic error, is

$$\varepsilon_\text{UT} = E \left[ \frac{D^3_{\partial g(x)}}{4!} + \frac{D^5_{\partial g(x)}}{6!} + \ldots \right]$$

$$- \frac{1}{2(n_x + \kappa)} \sum_{p=1}^{2n_x} \left( \frac{D^3_{\partial g(x_p)}}{4!} + \frac{D^5_{\partial g(x_p)}}{6!} + \ldots \right).$$
The error $\varepsilon_{\text{UT}}$ is non-zero, except for some cases with the function $g(\cdot)$ being a polynomial of degree $2n_x$.

Analogously, it can be shown that the covariance matrices computed using the UT also contain a systematic error.

In the UKF, the UT is employed in the filtering and predictive step at each time instant. The systematic error in the results produced by the UT thus can have an adverse impact on a local filter estimation performance.

Finally, it should be noted that the systematic error can be found in the approximate results based on e.g. the Taylor expansion, Stirling’s interpolation, the quadrature and cubature rules, used in the local filter design.

4.2 Problem statement

In the previous parts it was mentioned that all the standard and derivative-free filters are based on an approximate calculation of the predictive statistics of the state and measurement, which contains a systematic error. This fact was illustrated by means of the UT.

The goal of the paper is thus to introduce an improved version of the UT without the systematic error and to employ it in the local filter design.

5. RANDOMIZED UNSCENTED TRANSFORMATION

The improved version of the UT providing results without the systematic error is given in this section.

5.1 Stochastic integration rule

The basic idea of the improved version of the UT is based on a stochastic integration rule (SIR) proposed by Genz and Monahan (1998, 1999). The SIR is suitable for solution to an integral of the form

$$
\mu = \int_{\mathbb{R}^n} \varphi(x)(2\pi)^{-n/2} \rho e^{-\frac{1}{2}x^T x} \, dx,
$$

(32)

where $\varphi(\cdot)$ is an arbitrary function. Relation (32) can be interpreted as computation of the expected value of the function $\varphi(x)$ where $x$ is a random variable with $p(x) = \mathcal{N}[x : 0, I]$, i.e.

$$
\mu = E[\varphi(x)].
$$

(33)

The symbols $0$ and $I$ mean the zero vector and the identity matrix of appropriate size, respectively.

The stochastic integration rule proposed in (Genz and Monahan, 1998) for solution to (32) or (33) is given by the following algorithm

**Step 1:** Choose a maximum number of iterations $N_{\text{max}}$

**Step 2:** Set the number of iterations $N = 0$, initial value of the integral $\tilde{\mu} = 0$, and compute the point $x_0 = \varphi(0)$.

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

- **a)** Set $N = N + 1$.
- **b)** Generate uniformly random orthogonal matrix $U \in \mathbb{R}^{n_x \times n_x}$ and a random number $\rho$ from the chi distribution, i.e. $\rho \sim \chi(\eta + 2)$.
- **c)** Compute a set of points $\chi_i$ and appropriate weights $\omega_i$ according to

$$
\chi_i = -\rho U e_i,
$$

(34)

$$
\chi_{n_x + i} = \rho U e_i,
$$

(35)

and appropriate weights as

$$
\omega_0 = 1 - \frac{n_x}{\rho^2},
$$

(36)

$$
\omega_i = \omega_{n_x+i} = \frac{1}{2\rho^2},
$$

(37)

where $i = 1, 2, \ldots, n_x$ and $e_i$ is the $i$-th column of the identity matrix $I$.

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

$$
S = x_0 \omega_0 + \sum_{i=1}^{n_x} (\varphi(\chi_i) + \varphi(\chi_{n_x+i})) \omega_i,
$$

(38)

and use it to update the approximate value $\tilde{\mu}$ of the integral as

$$
\tilde{\mu} = \tilde{\mu} + (S - \tilde{\mu})/N,
$$

(39)

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

The variable $N_{\text{max}}$ represent a maximum number of iterations and is user-defined. The random orthogonal matrix $U$ can be generated using a product of appropriately chosen random reflections (Genz and Monahan, 1998).

If the random variable is given by $p(x) = \mathcal{N}[x : \hat{x}, P_{x}]$ the algorithm for the stochastic integration is the same except relations (34) and (35). The points are then computed as

$$
\chi_i = -\rho S_{x_e} U e_i + \hat{x},
$$

(40)

$$
\chi_{n_x+i} = \rho S_{x_e} U e_i + \hat{x}.
$$

(41)

Finally, for the sake of presentation clarity, computing $\tilde{\mu}$ using the above mentioned stochastic integration algorithm will be denoted as

$$
[\tilde{\mu}] = \text{SISI}(\hat{x}, P_x, \varphi(x), N_{\text{max}}).
$$

(42)

5.2 Properties of SIR and its relation to UT

The SIR can be understood as an improved version of the perfect Monte Carlo (MC) method. The SIR provides a better performance with respect to the perfect MC especially in situations where the function $\varphi(\cdot)$ is not approximately constant (Genz and Monahan, 1998). In such a case the perfect MC may have low accuracy and slow convergence (Genz and Monahan, 1998).

The important property of the SIR is that it comprises the UT without the systematic error and to employ it in the local filter design.

In the light of the previous relations, the SIR can be viewed as the UT with $N_{\text{max}}$ $\sigma$-point sets where the spread and rotation...
6. RANDOMISED UNSCENTED KALMAN FILTER

The RUT as was introduced in the previous section can directly be used in the LF design. The resulting algorithm, denoted as the randomised unscented Kalman filter (RUKF), can be summarised in the following steps:

**Step 1:** Set the time instant \( k = 0 \) and define the a priori initial condition by the predictive mean \( \hat{x}_{0|k} = E[x_0] = \bar{x}_0 \) and the predictive covariance matrix \( P_{0|k} = \text{cov}[x_0] = P_0 \). Define the maximum number of iteration steps \( N_{\text{max}} \) in the SIR.

**Step 2:** The state predictive estimate is updated with respect to the last measurement \( z_k \) according to

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

\[
P_{k|k} = P_{k|k-1} - K_k P_{k|k-1} K_k^T,
\]

where \( K_k = P_{xz,k-1} (P_{z,k-1})^{-1} \) is the filter gain and \( \hat{z}_{k|k-1} = \text{Slalg}(\hat{x}_{k|k-1}, P_{k|k-1}, h_k(x_k), N_{\text{max}}) \).

**Step 3:** The predictive statistics are given by relations

\[
\hat{x}_{k+1|k} = \text{Slalg}(\hat{x}_{k|k}, P_{k|k}, f_k(x_k), N_{\text{max}}),
\]

\[
P_{k+1|k} = \text{Slalg}(\hat{x}_{k|k}, P_{k|k}),
\]

\[
(f_k(x_k) - \hat{x}_{k|k})(f_k(x_k) - \hat{x}_{k|k})^T, N_{\text{max}}) + Q_k
\]

Let \( k = k + 1 \) and algorithm continues by **Step 2**.

7. NUMERICAL ILLUSTRATION

In this section the newly proposed RUKF will be compared with the UKF using a numerical example describing the bearings only tracking (Šimandl et al., 2002). The model is given by the following equations

\[
x_{k+1} = \begin{bmatrix} 0.9 & 0 \\ 0 & 1 \end{bmatrix} x_k + w_k,
\]

\[
z_k = \tan^{-1} \left( \frac{x_{2,k} - \sin(k)}{x_{1,k} - \cos(k)} \right) + v_k,
\]

where \( k = 0, 1, \ldots, N \), \( N = 100 \), \( p(x_0) = N[x_0 : [20, 5]^T, 0.1 I_2] \), \( Q_k = 10^{-3} \begin{bmatrix} 1 & 0.5 \\ 0.5 & 1 \end{bmatrix} \), \( R_k = 0.005 \), \( v_k \), and \( I_n \) is the identity matrix of dimension \( n \).

The filters are compared using the following two criteria providing mean square error (MSE) and variance of the squared error (SE)

\[
\text{MSE} = \frac{1}{M} \sum_{m=1}^{M} \text{SE}_m,
\]

\[
\text{var}[\text{SE}] = \frac{1}{M-1} \sum_{m=1}^{M} (\text{SE}_m - \text{MSE})^2,
\]
Table 1. Simulation results
(a) Standard and derivative-free filters.

<table>
<thead>
<tr>
<th></th>
<th>EKF</th>
<th>UKF (κ = 0)</th>
<th>CKF</th>
<th>UKF (κ = 1)</th>
<th>UKF (ML)</th>
</tr>
</thead>
<tbody>
<tr>
<td>MSE</td>
<td>0.8087</td>
<td>0.1572</td>
<td>0.1161</td>
<td>0.0506</td>
<td></td>
</tr>
<tr>
<td>var[SE]</td>
<td>1.1878</td>
<td>0.7409</td>
<td>0.4646</td>
<td>0.1069</td>
<td></td>
</tr>
<tr>
<td>time [sec]</td>
<td>0.0274</td>
<td>0.0487</td>
<td>0.2281</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

(b) RUKF.

<table>
<thead>
<tr>
<th></th>
<th>Nmax</th>
</tr>
</thead>
<tbody>
<tr>
<td>MSE</td>
<td>0.2898</td>
</tr>
<tr>
<td>var[SE]</td>
<td>0.9427</td>
</tr>
<tr>
<td>time [sec]</td>
<td>0.0566</td>
</tr>
</tbody>
</table>

where $SE_m = \sum_{k=0}^{N} \sum_{i=1}^{n_x} (x_i^{(m)} - \hat{x}_{i,k})^2/((N + 1)n_x)$ and $x_i^{(m)}$ is the $i$-th component of the true state in the $m$-th MC simulation at time $k$, and $\hat{x}_{i,k}$ is its filtering estimate. The number of simulations is $M = 10^3$.

The results comparing the local filters are given in Table 1. In Table 1a the results for the EKF and the UKF’s with $\kappa = 0$ and $\kappa = 1$ are given. Also the results for the UKF with adaptive choice of scaling parameter (Dunik et al., 2010) within the interval $\kappa \in (0, 4)$, denoted as UKF(ML), are presented. Predictably, the EKF provides worse estimation performance than the UKF’s. The UKF(ML) provides better performance than the UKF’s with the fixed scaling parameter. Note that the algorithm of the UKF with the scaling parameter $\kappa = 0$ is the same as the algorithm of the cubature Kalman filter (CKF) (Arasaratnam and Haykin, 2009) based on third degree cubature integration rule.

The newly proposed RUKF generally provides better estimation performance with respect to the EKF and the UKF’s. The estimation quality of the RUKF increases with increasing number of allowed iterations $N_{max}$. The results can be found in Table 1b. Note that for the RUKF the improvement of the estimation quality measured by MSE (55) for $N_{max} > 20$ is rather insignificant with respect to increase of the computational demands. However, value of the criterion $\text{var}[\text{SE}]$ (56) decreases with increasing $N_{max}$. That means, the performance of such estimator does not significantly depends on a particular realisation.

The EKF has lowest computational demands. The computational demands of the UKF and the RUKF with number of iteration $N_{max} = 1$ are almost the same. The RUKF with $N_{max} = 10$ provides computational complexity similar to the UKF(ML), however, with better estimation quality.

8. CONCLUDING REMARKS

In the paper recursive state estimation of nonlinear stochastic dynamic systems was treated. It was shown that the standard filters providing local state estimates suffer from a systematic error induced by the approximation used. To eliminate this flaw the RUKF has been proposed. The RUKF is based on a randomised version of the UT used in the UKF to solve the integrals appearing in computation of the state and measurement statistics. The randomised UT is a special kind of a stochastic integration rule which provides unbiased estimates of the integrals as opposed to the deterministic integration rule represented by the UT. Utilising the randomised UT in a local filter enables the filter to provide estimates of higher quality than the traditional local filters with a slight increase of computational time. Performance of the proposed RUKF was illustrated in a numerical example.

REFERENCES