

SIGMA POINT GAUSSIAN SUM FILTER DESIGN USING SQUARE ROOT UNSCENTED FILTERS

Miroslav Šimandl, Jindřich Duník

*Department of Cybernetics and
Research Centre: Data - Algorithms - Decision
University of West Bohemia in Pilsen,
Univerzitní 8, 306 14 Plzeň, Czech Republic*

Abstract: Local and global estimation approaches are discussed, above all the Unscented Kalman Filter and the Gaussian Sum Filter. The square root modification of the Unscented Kalman Filter is derived and it is used in the Gaussian Sum Filter framework. The new Sigma Point Gaussian Sum Filter is designed and some aspects of the filter are presented. Estimation quality and computational demands of the filter are illustrated in a numerical example. *Copyright © 2005 IFAC*

Keywords: stochastic systems, state estimation, estimation theory, filtering techniques, nonlinear filters

1. INTRODUCTION

The general solution of recursive state estimation problem, based on the Bayesian approach, is given by the functional relations for probability density functions (pdf's) of the state conditioned by the measurements which provide a full description of the hidden state.

The closed-form solution of the Bayesian Recursive Relation (BRR) is available only for a few special cases e.g. for linear Gaussian system (Anderson and Moore, 1979). In other cases it is necessary to apply some approximative methods. These methods are divided with respect to the validity of the resulting estimates into two groups (Sorenson, 1974). The first group of methods provides results with validity within some neighborhood of a point estimate only and thus they are called local methods. The second group of methods provides results valid in almost whole state space and so they are called global methods.

The local methods approximate the pdf's representing state estimate with Gaussian pdf's. The main disadvantage of these methods is local va-

lidity of the state estimate which can causes the divergence of the local filter estimates. The reason for local methods popularity is simplicity of BRR solution. The local filters can be based on approximation of the nonlinear functions in the state and measurement equation so that the analytical solution of the BRR is possible using techniques based on the Kalman Filter approach. The useful approximations are based on the Taylor expansion or Stirling's interpolation which leads e.g. to the Extended Kalman Filter (EKF) or Second Order Filter (Anderson and Moore, 1979) or to the Divide Difference Filter (DDF) (Nørgaard *et al.*, 2000), respectively. Instead of the substitution of nonlinear functions the local filters can be also based on the approximation of the pdf of the state estimate by a set of deterministically chosen weighted points so called σ -points. The Unscented Kalman Filter (UKF) (Julier *et al.*, 2000) or Gauss-Hermite Filter (Ito and Xiong, 2000) exemplify this local approach.

Although the DDF and the UKF rise up from quite different bases, these methods have a few common features. Thus relatively high interest has

been devoted to compare these methods (Lefebvre *et al.*, 2001; Nørsgaard *et al.*, 2000). Unfortunately, comparison of algorithms of these methods is quite difficult due to the fact that the DDF can be understood as square root algorithm, whereas the UKF has “classical non-square root” structure.

Therefore the first aim of this paper is to derive the square root version of the UKF.

The global estimation methods generate the conditional pdf of the state estimate. Comparing to local methods they are more sophisticated and produce the better state estimate, but they have higher computational demands. There are three main approaches to the global filtering method design: analytical approach based on system approximation and Gaussian sum approximation of pdf’s (Sorenson and Alspach, 1971; Šimandl and Královec, 2000), numerical approach using numerical solution of the integrals in the BRR, (Bucy and Senne, 1971; Šimandl *et al.*, 2002) and simulation approach taking advantage of the BRR solution performed using Monte Carlo integration (Van Der Merwe and Wan, 2003; Šimandl and Straka, 2003).

Some global methods are based on a multiple application of local methods, e.g. Gaussian Sum Filter (GSF), consists of a bank of EKF’s, or Unscented Particle Filter and Gaussian Mixture Sigma Point Particle Filter (GMSPPF), where samples are moving by means of the UKF’s to reduce a significant problem of basic algorithm of the Particle Filter (PF) called “sample depletion” (Van Der Merwe and Wan, 2003).

The idea of multiple application of local filters in global filter design is the motivation for the second aim of this paper, to utilize the Square Root UKF, derived in this paper, in the framework of Gaussian sum approach for design of the global Sigma Point Gaussian Sum Filter.

The paper is organized as follows. Section 2 deals with the estimation problem statement, Bayesian state estimation and brief description of the UKF. Subsequently the description and application of the Householder triangularization in the Square Root UKF design is given in Section 3. Section 4 is focused on utilization of the Square Root UKF in the GSF framework. Afterwards, in Section 5 a numerical example is given.

2. PROBLEM STATEMENT

Consider the discrete-time nonlinear non-Gaussian stochastic system

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

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

where the vectors $\mathbf{x}_k \in \mathbb{R}^{n_x}$ and $\mathbf{z}_k \in \mathbb{R}^{n_z}$ represent the unknown state of the system and the measurement at time instant k , respectively, and $\mathbf{w}_k \in \mathbb{R}^{n_x}$ and $\mathbf{v}_k \in \mathbb{R}^{n_z}$ are state and measurement white noises. Both noises are mutually independent and independent on the initial state \mathbf{x}_0 . The vector functions $\mathbf{f}_k : \mathbb{R}^{n_x} \rightarrow \mathbb{R}^{n_x}$, $\mathbf{h}_k : \mathbb{R}^{n_x} \rightarrow \mathbb{R}^{n_z}$ are known. Also the initial condition pdf $p_{\mathbf{x}_0}(\mathbf{x}_0)$ is known as well as the pdf’s of the noises $p_{\mathbf{w}_k}(\mathbf{w}_k)$, $p_{\mathbf{v}_k}(\mathbf{v}_k)$.

The system can be alternatively written as a set of conditional pdf’s called transient pdf’s $p_{\mathbf{x}_{k+1}|\mathbf{x}_k}(\mathbf{x}_{k+1}|\mathbf{x}_k) = p_{\mathbf{w}_k}(\mathbf{x}_{k+1} - \mathbf{f}_k(\mathbf{x}_k))$ and measurement pdf $p_{\mathbf{z}_k|\mathbf{x}_k}(\mathbf{z}_k|\mathbf{x}_k) = p_{\mathbf{v}_k}(\mathbf{z}_k - \mathbf{h}_k(\mathbf{x}_k))$, $\forall k$. For the sake of simplicity all pdf’s will be given with their argument, i.e. $p(\mathbf{x}_{k+1}|\mathbf{x}_k) = p_{\mathbf{x}_{k+1}|\mathbf{x}_k}(\mathbf{x}_{k+1}|\mathbf{x}_k)$.

The aim of the filtering is to find the state estimate in the form of the posterior filtering pdf $p(\mathbf{x}_k|\mathbf{z}^k)$. This pdf is given by the BRR, expressed as

$$p(\mathbf{x}_k|\mathbf{z}^k) = \frac{p(\mathbf{x}_k|\mathbf{z}^{k-1})p(\mathbf{z}_k|\mathbf{x}_k)}{\int p(\mathbf{x}_k|\mathbf{z}^{k-1})p(\mathbf{z}_k|\mathbf{x}_k)d\mathbf{x}_k} \quad (3)$$

where $p(\mathbf{x}_{k+1}|\mathbf{z}^k) = \int p(\mathbf{x}_k|\mathbf{z}^k) p(\mathbf{x}_{k+1}|\mathbf{x}_k)d\mathbf{x}_k$ is a one-step prediction pdf, $\mathbf{z}^k = [\mathbf{z}_0, \dots, \mathbf{z}_k]$ and $p(\mathbf{x}_0|\mathbf{z}^{-1}) = p(\mathbf{x}_0)$.

The local filters the UKF and the DDF have been developed to overcome the well-known disadvantages of the EKF, e.g. as inaccuracy, required derivation of the state and the measurement functions $\mathbf{f}_k(\cdot)$ and $\mathbf{h}_k(\cdot)$ at preservation of computational demands. However, approximation of all pdf descriptions by the Gaussian pdf’s is still assumed. It is also important to mention that due to the common features the local filters UKF, DDF and their variants are often called together as Sigma Point Kalman Filters (SPKF’s) (Van Der Merwe and Wan, 2003).

The basic UKF algorithm for system with additive Gaussian noises $\mathbf{w}_k \sim \mathcal{N}\{\mathbf{w}_k : \mathbf{0}, \mathbf{Q}_k\}$ and $\mathbf{v}_k \sim \mathcal{N}\{\mathbf{v}_k : \mathbf{0}, \mathbf{R}_k\}$ consists of the following steps:

Step 1: Set the time instant $k = 0$ and define a prior Gaussian initial condition $p(\mathbf{x}_k|\mathbf{z}^{k-1}) = \mathcal{N}\{\mathbf{x}_k : \hat{\mathbf{x}}'_k, \mathbf{P}'_k\}$, where $\hat{\mathbf{x}}'_k = E[\mathbf{x}_k|\mathbf{z}^{k-1}]$ and $\mathbf{P}'_k = cov[\mathbf{x}_k|\mathbf{z}^{k-1}]$.

Step 2: Determine the set of predictive σ -points $\mathcal{X}_{i,k|k-1}$ and their weights \mathcal{W}_i (weights are computed only at $k = 0$) at time instant k according to

$$\mathcal{X}_{0,k|k-1} = \hat{\mathbf{x}}'_k, \quad \mathcal{W}_0 = \frac{\kappa}{n_x + \kappa}, \quad (4)$$

$$\mathcal{X}_{i,k|k-1} = \hat{\mathbf{x}}'_k + \left(\sqrt{(n_x + \kappa)\mathbf{P}'_k}\right)_i, \quad (5)$$

$$\mathcal{X}_{j,k|k-1} = \hat{\mathbf{x}}'_k - \left(\sqrt{(n_x + \kappa)\mathbf{P}'_k}\right)_{j-n_x}, \quad (6)$$

where $i = 1, \dots, n_x$, $j = n_x + 1, \dots, 2n_x$ and $\mathcal{W}_i = \mathcal{W}_j = \frac{1}{2(n_x + \kappa)}$, $\forall i, j$. The parameter κ de-

termines spreading of σ -points around the mean and thereby it affects accuracy of the estimation and the term $(\sqrt{(n_x + \kappa)\mathbf{P}_k})_i$ represents the i -th column of the matrix $\sqrt{(n_x + \kappa)\mathbf{P}_k}$.

Step 3: Update the state estimate, i.e. compute the Gaussian approximation of the posterior pdf $p(\mathbf{x}_k|\mathbf{z}^k) \approx \mathcal{N}\{\mathbf{x}_k : \hat{\mathbf{x}}_k, \mathbf{P}_k\}$, according to the last measurement \mathbf{z}_k :

$$\mathcal{Z}_{i,k|k-1} = \mathbf{h}_k(\mathcal{X}_{i,k|k-1}), \forall i, \quad (7)$$

$$\hat{\mathbf{z}}'_k = \sum_{i=0}^{2n_x} \mathcal{W}_i \mathcal{Z}_{i,k|k-1}, \quad (8)$$

$$\begin{aligned} \mathbf{P}'_{z,k} &= \sum_{i=0}^{2n_x} \mathcal{W}_i (\mathcal{Z}_{i,k|k-1} - \hat{\mathbf{z}}'_k) \\ &\quad \times (\mathcal{Z}_{i,k|k-1} - \hat{\mathbf{z}}'_k)^T + \mathbf{R}_k, \end{aligned} \quad (9)$$

$$\begin{aligned} \mathbf{P}'_{xz,k} &= \sum_{i=0}^{2n_x} \mathcal{W}_i (\mathcal{X}_{i,k|k-1} - \hat{\mathbf{x}}'_k) \\ &\quad \times (\mathcal{Z}_{i,k|k-1} - \hat{\mathbf{z}}'_k)^T, \end{aligned} \quad (10)$$

$$\hat{\mathbf{x}}_k = \hat{\mathbf{x}}'_k + \mathbf{P}'_{xz,k} \mathbf{P}'_{z,k}^{-1} (\mathbf{z}_k - \hat{\mathbf{z}}'_k), \quad (11)$$

$$\mathbf{P}_k = \mathbf{P}'_k - \mathbf{P}'_{xz,k} \mathbf{P}'_{z,k}^{-1} \mathbf{P}'_{xz,k}{}^T. \quad (12)$$

Step 4: Compute the filtering σ -points $\mathcal{X}_{i,k|k}$ in accordance with (4)-(6), where the predictive moments $\hat{\mathbf{x}}'_k, \mathbf{P}'_k$ are substituted by the filtering ones $\hat{\mathbf{x}}_k, \mathbf{P}_k$.

Step 5: Calculate the set of predictive σ -points $\mathcal{X}_{i,k+1|k}$ in accord with the system equation

$$\mathcal{X}_{i,k+1|k} = \mathbf{f}_k(\mathcal{X}_{i,k|k}), \forall i. \quad (13)$$

The predictive pdf $p(\mathbf{x}_{k+1}|\mathbf{z}^k)$ is approximated by the Gaussian pdf $\mathcal{N}\{\mathbf{x}_{k+1} : \hat{\mathbf{x}}'_{k+1}, \mathbf{P}'_{k+1}\}$,

$$\hat{\mathbf{x}}'_{k+1} = \sum_{i=0}^{2n_x} \mathcal{W}_i \mathcal{X}_{i,k+1|k}, \quad (14)$$

$$\begin{aligned} \mathbf{P}'_{k+1} &= \sum_{i=0}^{2n_x} \mathcal{W}_i (\mathcal{X}_{i,k+1|k} - \hat{\mathbf{x}}'_{k+1}) \\ &\quad \times (\mathcal{X}_{i,k+1|k} - \hat{\mathbf{x}}'_{k+1})^T + \mathbf{Q}_k. \end{aligned} \quad (15)$$

Let $k \leftarrow k + 1$ and go to the Step 2.

Note that the square root of matrix $\sqrt{(n_x + \kappa)\mathbf{P}_k}$ can be found with assistance of e.g. Cholesky decomposition (Julier *et al.*, 2000). The recommended settings of the scaling parameter κ is $\kappa = n_x - 3$ for Gaussian distribution which was proved in detail in (Julier *et al.*, 2000). This basic algorithm suffers from significant disadvantage particularly in case if $n_x > 3$. Then parameter κ and weight \mathcal{W}_0 can be negative which may cause non-positive semi-definite posterior covariance matrix. In this case a few substitutive ways ensuring positive definiteness of posterior covariance matrix have been proposed e.g. scaled UKF (Van Der Merwe and Wan, 2001), reduced σ -points UKF (Julier & Uhlmann, 2002), Gauss-Hermite Filter (Ito and Xiong, 2000). All these approaches differ only in the way of computing of

σ -points, whereas the equations (7)-(15) remain without any change.

Computation of a set of σ -points is the most computationally demanding operation due to the Cholesky decomposition performed twice at each time instant. Although the σ -point set calculation is based on the square root of the state estimate covariance matrix, the “full” covariance matrixes (12), (15) are computed.

Therefore the improvement of the basic algorithm of the UKF should be proposed, where square root of the covariance would be immediately available.

As it was mentioned above, the local approaches are often used to make the global method properties better, e.g. GMSPPF (Van Der Merwe and Wan, 2003) which improve estimation performance of the PF. Unfortunately, this enhancement induces growth of computational demands.

For that reason the additional aim is to propose the alternative global estimation method which reduces the computational demands and maintains the estimation quality of the GMSPPF.

3. SQUARE ROOT UNSCENTED FILTER

This section deals with the square root modification of the UKF (SRUKF). This improvement should facilitate direct comparison of the SRUKF and the DDF algorithms and it should reduce the computational demands of the UKF. Moreover, it should ensure positive semi-definiteness of the state estimate covariance matrixes and thereby it should improve numerical stability of the UKF.

As a basic tool to the SRUKF derivation the Householder triangularization is used (Nørgaard *et al.*, 2000). The Householder triangularization can be used for transformation of known rectangular matrix $\mathbf{M} \in \mathbb{R}^{n \times m}$ to square matrix $\mathbf{N} \in \mathbb{R}^{n \times n}$ so that the equality $\mathbf{N}\mathbf{N}^T = \mathbf{M}\mathbf{M}^T$ is accomplished.

It should be mentioned that all variables used in this section are defined in accord with the basic UKF algorithm (4)-(15) in previous section. The matrixes $\mathbf{S}_{Q,k} \in \mathbb{R}^{n_x \times n_x}$, $\mathbf{S}_{R,k} \in \mathbb{R}^{n_z \times n_z}$ represent the state and measurement square root covariance matrixes at time instant k which accomplish $\mathbf{Q}_k = \mathbf{S}_{Q,k} \mathbf{S}_{Q,k}^T$, $\mathbf{R}_k = \mathbf{S}_{R,k} \mathbf{S}_{R,k}^T$. The SRUKF algorithm for the system (1), (2) with Gaussian noises is given by the following steps.

Step 1: Set the time instant $k = 0$ and define a prior Gaussian initial condition $p(\mathbf{x}_k|\mathbf{z}^{k-1}) = \mathcal{N}\{\mathbf{x}_k : \hat{\mathbf{x}}'_k, \mathbf{P}'_k\}$. Compute the square root of the covariance $\mathbf{S}'_k = \text{chol}(\mathbf{P}'_k)$, where the function $\text{chol}(\cdot)$ means the Cholesky decomposition.

Step 2: Determine the set of predictive σ -points $\mathcal{X}_{i,k|k-1}$ and their weights \mathcal{W}_i (weights are com-

puted only at $k = 0$) at time instant k according to

$$\mathcal{X}_{0,k|k-1} = \hat{\mathbf{x}}'_k, \mathcal{W}_0 = \frac{\kappa}{n_x + \kappa}, \quad (16)$$

$$\mathcal{X}_{i,k|k-1} = \hat{\mathbf{x}}'_k + (\sqrt{(n_x + \kappa)} \mathbf{S}'_k)_i, \quad (17)$$

$$\mathcal{X}_{j,k|k-1} = \hat{\mathbf{x}}'_k - (\sqrt{(n_x + \kappa)} \mathbf{S}'_k)_{j-n_x}, \quad (18)$$

where $i = 1, \dots, n_x$, $j = n_x + 1, \dots, 2n_x$ and $\mathcal{W}_i = \mathcal{W}_j = \frac{1}{2(n_x + \kappa)}$, $\forall i, j$.

Step 3: Update the state estimate, i.e. compute the Gaussian approximation of the posterior pdf $p(\mathbf{x}_k | \mathbf{z}^k) \approx \mathcal{N}\{\mathbf{x}_k : \hat{\mathbf{x}}_k, \mathbf{S}_k \mathbf{S}_k^T\}$, according to the last measurement \mathbf{z}_k :

$$\mathcal{Z}_{i,k|k-1} = \mathbf{h}_k(\mathcal{X}_{i,k|k-1}), \forall i, \quad (19)$$

$$\hat{\mathbf{z}}'_k = \sum_{i=0}^{2n_x} \mathcal{W}_i \mathcal{Z}_{i,k|k-1}, \quad (20)$$

$$\mathbf{M}'_{z,k} = [\sqrt{\mathcal{W}_0}(\mathcal{Z}_{0,k|k-1} - \hat{\mathbf{z}}'_k), \dots, \sqrt{\mathcal{W}_{2n_x}}(\mathcal{Z}_{2n_x,k|k-1} - \hat{\mathbf{z}}'_k)], \quad (21)$$

$$\mathbf{S}'_{z,k} = [\mathbf{M}'_{z,k}, \mathbf{S}_{R,k}], \quad (22)$$

$$\mathbf{M}'_{x,k} = [\sqrt{\mathcal{W}_0}(\mathcal{X}_{0,k|k-1} - \hat{\mathbf{x}}'_k), \dots, \sqrt{\mathcal{W}_{2n_x}}(\mathcal{X}_{2n_x,k|k-1} - \hat{\mathbf{x}}'_k)], \quad (23)$$

$$\mathbf{P}'_{xz,k} = \mathbf{M}'_{x,k} \mathbf{M}'_{z,k}{}^T, \quad (24)$$

$$\mathbf{K}_k = \mathbf{P}'_{xz,k} (\mathbf{S}'_{z,k} \mathbf{S}'_{z,k}{}^T)^{-1}, \quad (25)$$

$$\hat{\mathbf{x}}_k = \hat{\mathbf{x}}'_k + \mathbf{K}_k (\mathbf{z}_k - \hat{\mathbf{z}}'_k), \quad (26)$$

$$\mathbf{S}_k = ht([\mathbf{M}'_{x,k} - \mathbf{K}_k \mathbf{M}'_{z,k}, \mathbf{K}_k \mathbf{S}_{R,k}]), \quad (27)$$

where function $ht(\mathbf{M})$ is the Householder triangularization applied to rectangular matrix \mathbf{M} .

Step 4: Compute the filtering σ -points $\mathcal{X}_{i,k|k}$ in accordance with (16)-(18), where the predictive characteristics $\hat{\mathbf{x}}'_k$, \mathbf{S}'_k are substituted by the filtering ones $\hat{\mathbf{x}}_k$, \mathbf{S}_k .

Step 5: Calculate the set of predictive σ -points $\mathcal{X}_{i,k+1|k}$ in accord with the system equation

$$\mathcal{X}_{i,k+1|k} = \mathbf{f}_k(\mathcal{X}_{i,k|k}), \forall i. \quad (28)$$

The predictive pdf $p(\mathbf{x}_{k+1} | \mathbf{z}^k)$ is approximated by the Gaussian pdf $\mathcal{N}\{\mathbf{x}_{k+1} : \hat{\mathbf{x}}'_{k+1}, \mathbf{S}'_{k+1} \mathbf{S}'_{k+1}{}^T\}$, where

$$\hat{\mathbf{x}}'_{k+1} = \sum_{i=0}^{2n_x} \mathcal{W}_i \mathcal{X}_{i,k+1|k}, \quad (29)$$

$$\mathbf{S}'_{k+1} = ht([\sqrt{\mathcal{W}_0}(\mathcal{X}_{0,k+1|k} - \hat{\mathbf{x}}'_{k+1}), \dots, \sqrt{\mathcal{W}_{2n_x}}(\mathcal{X}_{2n_x,k+1|k} - \hat{\mathbf{x}}'_{k+1}), \mathbf{S}_{Q,k+1}]). \quad (30)$$

Let $k \leftarrow k + 1$ and go to the Step 2.

Note that the formulas for computation of all means in the UKF and the SRUKF are the same. The differences between these algorithms arise in calculation of covariance matrixes. Some of these terms can be transformed to the square root structure easily and directly, namely $\mathbf{P}'_{z,k}$ (9), $\mathbf{P}'_{xz,k}$ (10) and \mathbf{P}'_{k+1} (15). But the transformation

of the covariance \mathbf{P}_k (12) to the square root form \mathbf{S}_k (27) is slightly complicated due to the minus sign in (12). This problem can be solved by extending the equation (12) to the form

$$\mathbf{P}_k = \mathbf{S}'_k \mathbf{S}'_k{}^T + \mathbf{K}_k \mathbf{S}'_{z,k} \mathbf{S}'_{z,k}{}^T \mathbf{K}_k^T - 2\mathbf{K}_k \mathbf{S}'_{z,k} \mathbf{S}'_{z,k}{}^T \mathbf{K}_k^T.$$

Then replace the terms $\mathbf{K}_k \mathbf{S}'_{z,k} \mathbf{S}'_{z,k}{}^T \mathbf{K}_k^T$ by means of following substitutions

$$\begin{aligned} \mathbf{K}_k \mathbf{S}'_{z,k} \mathbf{S}'_{z,k}{}^T \mathbf{K}_k^T &= \mathbf{P}'_{xz,k} \mathbf{K}_k^T = \mathbf{M}'_{x,k} \mathbf{M}'_{z,k}{}^T \mathbf{K}_k^T, \\ &= \mathbf{K}_k \mathbf{P}'_{xz,k} = \mathbf{K}_k \mathbf{M}'_{z,k} \mathbf{M}'_{x,k}{}^T, \\ &= \mathbf{K}_k (\mathbf{M}'_{z,k} \mathbf{M}'_{z,k}{}^T + \mathbf{S}'_{R,k} \mathbf{S}'_{R,k}{}^T) \mathbf{K}_k^T. \end{aligned}$$

Afterwards, only a few basic matrix operations have to be done to obtain the equation (27).

The structure of the SRUKF algorithm is quite similar to the DDF one which allows superficial comparison between these algorithms. The DDF can be based on the Stirling's interpolation formula first (DD1) or second order (DD2) (Nørgaard *et al.*, 2000). It can be easily shown that the formulas for the predictive mean in the SRUKF and in the DD2 are the same. On the other hand the structure of the SRUKF formulas for the predictive and the filtering covariance matrixes are formally similar to the formulas of the DD1 ones, but the columns of the matrixes are different. It is also important that the square root modification of the UKF has been already done in (Van Der Merwe and Wan, 2001), but this variation is based on different techniques as QR decomposition, Cholesky factor updating or efficient least squares which do not facilitate the straight comparison of the SRUKF with the DDF. Note that all refined versions of basic UKF, which differ only in the σ -point determination, e.g. the Gauss-Hermite Filter (Ito and Xiong, 2000), can be easily modified to the square root structure as well.

4. SIGMA POINT GAUSSIAN SUM FILTER

The main goal of the section is to design the new global Sigma Point Gaussian Sum Filter (SPGSF) as an alternative to the computational demanding GMSPPF which uses the Monte Carlo simulation and the bank of the standard UKF (Van Der Merwe and Wan, 2003).

The main idea of the Gaussian sum approach (Sorenson and Alspach, 1971; Šimandl and Královic, 2000) is based on the approximation of an arbitrary pdf by a Gaussian mixture. To apply this idea for the system (1), (2) it is necessary to assume the prior pdf, the state and the measure noise pdf in the form of Gaussian mixtures. To obtain a close-loop solution of the BRR, the multi-point linearization of nonlinear functions $\mathbf{f}_k(\cdot)$, $\mathbf{h}_k(\cdot)$ has to be performed. Then, the GSF can be

understood as a mixture of the EKF's generating filtering pdf's in the form of Gaussian mixtures.

However, the EKF as the corner stone of the GSF suffers with some above mentioned disadvantages which have been overcome by some newer approaches as the UKF, DDF. Therefore, with usage of bank of SPKF's instead of bank of the EKF's, it should be possible to reach better estimation results than it would be reached by the standard GSF. It seems to be reasonable to choose the SRUKF as the representant of the SPKF's due to small computations demands and numerical stability. Substituting the SRUKF to the GSF structure, the global SPGSF can be obtained.

The SPGSF algorithm can be briefly summarized in the following steps.

Step 1: Set the time instant $k = 0$ and define a prior initial condition in the form of mixture of N Gaussians

$$p(\mathbf{x}_k | \mathbf{z}^{k-1}) = \sum_{j=1}^N \mathbf{w}_{k-1}^{(j)} \mathcal{N}(\mathbf{x}_k : \hat{\mathbf{x}}_k^{(j)}, \mathbf{P}_k^{(j)}), \quad (31)$$

where $\mathbf{w}_{k-1}^{(j)} > 0$ is weight of the j -th Gaussian and $\sum_{j=1}^N \mathbf{w}_{k-1}^{(j)} = 1$.

Step 2: Determine the sets of predictive σ -points $\mathcal{X}_{i,k|k-1}^{(j)}$ and weights $\mathcal{W}_i^{(j)}$ for all Gaussians, i.e. $\forall j$, according to (16) - (18).

Step 3: The filtering pdf is determined in the form

$$p(\mathbf{x}_k | \mathbf{z}^k) = \sum_{j=1}^{N_f} \mathbf{w}_k^{(j)} \mathcal{N}(\mathbf{x}_k : \hat{\mathbf{x}}_k^{(j)}, \mathbf{P}_k^{(j)}), \quad (32)$$

with number of Gaussians $N_f = N \times N_v$, where N_v denotes the number of Gaussians in measurement noise description. The j -th pdf $\mathcal{N}(\mathbf{x}_k : \hat{\mathbf{x}}_k^{(j)}, \mathbf{P}_k^{(j)})$ is derived with respect to (19) - (27), $\forall j$, and relevant filtering weight according to

$$\mathbf{w}_k^{(j)} = \frac{\mathbf{w}_{k-1}^{(j)} \zeta_k^{(j)}}{\sum_{i=1}^{N_f} \mathbf{w}_{k-1}^{(i)} \zeta_k^{(i)}}, \quad (33)$$

where $\zeta_k^{(j)} = \mathcal{N}(\mathbf{z}_k : \hat{\mathbf{z}}_k^{(j)}, \mathbf{S}'_{z,k} \mathbf{S}'_{z,k}{}^T)$.

Step 4: The number of terms in posterior mixture grows exponentially, due to the fact that the state and the measurement noise have a Gaussian mixture form. Therefore, it is necessary to reduce the number of terms. The simplest way of mixture reduction is to retain only a few Gaussians with highest weights $\mathbf{w}_k^{(j)}$.

Step 5: Similarly as in the filtering step the prediction pdf is determined in the form

$$p(\mathbf{x}_{k+1} | \mathbf{z}^k) = \sum_{j=1}^{N_p} \mathbf{w}_k^{(j)} \mathcal{N}(\mathbf{x}_{k+1} : \hat{\mathbf{x}}_{k+1}^{(j)}, \mathbf{P}_{k+1}^{(j)}), \quad (34)$$

where number of Gaussians $N_p = N_f \times N_w$ and N_w denotes number of term in state noise description. The j -th element of $p(\mathbf{x}_{k+1} | \mathbf{z}^k)$ is calculated according to (28) - (30), $\forall j$.

Let $k \leftarrow k + 1$ and go to the Step 2.

The basic structure of the SPGSF algorithm is quite similar to the GSF one. The substantial differences can be found in the relations (32), (34) where the components of filtering and predictive pdf are obtained by means of the SRUKF instead of the EKF in the GSF.

5. NUMERICAL ILLUSTRATION

The properties of the local SRUKF and the global SPGSF will be discussed and illustrated in this section.

The better estimation performance of the newer local approaches (UKF, DDF) over the EKF is well-known (Julier *et al.*, 2000; Van Der Merwe and Wan, 2003). Estimation performance of the derived SRUKF is naturally the same as the UKF, but the computational demands are slightly smaller, because number of numerical operation in the SRUKF algorithm is smaller.

The estimation performance of the global approaches, particularly the SPGSF, can be verified with experiment presented e.g. in (Van Der Merwe and Wan, 2003). Consider the nonlinear non-Gaussian system with one-dimensional state

$$x_{k+1} = \phi_1 x_k + 1 + \sin(\omega \pi k) + w_k \quad (35)$$

with the state noise w_k with Gamma pdf $Ga(3, 2)$, $\forall k$, $\phi_1 = 0.5$, $\omega = 0.04$ are scalar parameters and $k = 1, \dots, 60$. The state is observed by the scalar measurement described by the equation

$$z_k = \begin{cases} \phi_2 x_k^2 + v_k, & k \leq 30, \\ \phi_3 x_k - 2 + v_k, & k > 30. \end{cases} \quad (36)$$

The measurement z_k is influenced by the measurement noise $v_k \sim \mathcal{N}(v_k : 0, 10^{-5})$, $\forall k$, and the scalar parameters are $\phi_2 = 0.2$ and $\phi_3 = 0.5$. The initial condition is given by the mixture of five Gaussians $p(x_0) = \sum_{j=1}^5 \mathbf{w}_{-1}^{(j)} \times \mathcal{N}(x_0 : \hat{x}_0^{(j)}, P_0^{(j)}) = \sum_{j=1}^5 0.2 \times \mathcal{N}(x_0 : j - 3, 10)$. The predictive pdf $p(x_0 | z^{-1})$ is equal to $p(x_0)$.

The SPGSF was compared with three different global filters, specifically with a generic PF with sampling-importance resampling, GMSPPF and "standard" GSF based on the bank of the EKF's. The last two filters and the SPGSF used an approximation of the Gamma pdf in the state equation (35) by the Gaussian mixture $\tilde{p}(w_k) = 0.29 \times \mathcal{N}(w_k : 2.14, 0.72) + 0.18 \times \mathcal{N}(w_k : 7.45, 8.05) + 0.53 \times \mathcal{N}(w_k : 4.31, 2.29)$, $\forall k$ which was obtained with assistance of the EM algorithm. Both particle filters (GMSPPF and generic PF) exploited 500

Table 1 Global filter estimation results.

Algorithm	MSE	Time(s)
PF	1.9262	3.28
GMSPPF	0.0156	4.90
GSF	0.0253	0.91
SPGSF	0.0149	2.08

particles. The GMSPPF used 5 Gaussians in the mixture for the posterior state estimate. Reduction of components in the posterior pdf for the GSF and the SPGSF is accomplished by the rejection of Gaussians with weight $w_k^{(j)} < 0.05, \forall k, j$. The experiment was repeated 100 times with re-initialization for each run in order to calculate Monte Carlo performance estimates for each filter. The results for the global filters are presented in the Table 1, where the mean square error $MSE = \frac{\sum_{i=1}^{100} \sum_{k=1}^{60} (x_k^i - \hat{x}_k^i)^2}{6000}$ of the state estimate and the computational time (in seconds) for each filter are shown. The superscript i determines the order of repetition.

Relatively poor estimation performance of the standard PF is caused by the “sample depletion” due to the small observation noise variance. This problem is solved with Particle Filters which exploit SPKF’s (Van Der Merwe and Wan, 2003). As an representative the GMSPPF was chosen which gives a substantial improvement in estimation performance. However, this improvement leads to the growth of the computational demands. On other hand, the GSF has the smallest computational demands, but the estimation quality is worse. These imperfections are rectified by the SPGSF which provides the highly precious estimates as well as the GMSPPF with computational cost slightly higher than that of the GSF.

6. CONCLUSION

The UKF was described and modified to the square root form which ensures the positive semi-definiteness of state covariance matrixes and slight reduction of computational demands. An interesting attribute of the derived Square Root Unscented Kalman Filter algorithm concerning its structure was discussed. It is in the form similar to the structure of the filter algorithms based on the Stirling’s approximation. Further, the novel global method Sigma Point Gaussian Sum Filter was proposed. It provides the same estimation performance as the GMSPPF with reduced computational demands and theoretical complexity.

ACKNOWLEDGEMENT

The work was supported by the Ministry of Education, Youth and Sports of the Czech Republic, projects No. MSM 235200004 and 1M6798555601.

REFERENCES

- Anderson, B. D. O. and S. B. Moore (1979). *Optimal filtering*. Englewood Cliffs.
- Bucy, R. S. and K. D. Senne (1971). Digital synthesis of non-linear filters. *Automatica* **7**, 287–298.
- Ito, K. and K. Xiong (2000). Gaussian filters for nonlinear filtering problems. *IEEE Transactions On Automatic Control* **45**(5), 910–927.
- Julier, S. J. and J. K. Uhlmann (2002). Reduced sigma point filters for the propagation of means and covariances through nonlinear transformations. In: *Proceedings of the IEEE American Control Conference*, pp. 887–892, Anchorage AK, USA.
- Julier, S. J., J. K. Uhlmann and H. F. Durrantwhite (2000). A new method for the nonlinear transformation of means and covariances in filters and estimators. *IEEE Transactions On Automatic Control* **45**(3), 477–482.
- Lefebvre, T., H. Bruyninckx and J. De Schutter (2001). Kalman filters for nonlinear systems: a comparison of performance. Technical report. Dept. of Mechanical Eng., Katholieke Universiteit Leuven. Belgium.
- Nørgaard, M., N. K. Poulsen and O. Ravn (2000). New developments in state estimation for nonlinear systems. *Automatica* **36**(11), 1627–1638.
- Sorenson, H. W. (1971). On the development of practical nonlinear filters. *Inf. Sci.* **7**, 230–270.
- Sorenson, H. W. and D. L. Alspach (1971). Recursive bayesian estimation using gaussian sums. *Automatica* **7**, 465–479.
- Van Der Merwe, R. and E. A. Wan (2003). Sigma-point particle filters for sequential probabilistic inference in dynamic state-space models. In: *Proc. of the Int. Conf. On Acoustics, Speech, and Signal Proc.* IEEE. Hong Kong.
- Van Der Merwe, R. and E. Wan (2001). The square-root unscented Kalman filter for state and parameter-estimation. In: *International Conference on Acoustics, Speech, and Signal Processing*. Salt Lake City, Utah.
- Šimandl, M. and J. Královec (2000). Filtering, prediction and smoothing with gaussian sum representation. In: *Proceedings of the 12th IFAC Symposium on System Identification*. Santa Barbara, California: Elsevier Science.
- Šimandl, M. and O. Straka (2003). Nonlinear filtering methods: Some aspects and performance evaluation. In: *Proceedings of the 22nd IASTED International Conference On Modelling, Identification and Control*. Innsbruck: Acta Press.
- Šimandl, M., J. Královec and T. Söderström (2002). Anticipative grid design in point-mass approach to nonlinear state estimation. *IEEE Transactions On Automatic Control* **47**(4), 699–702.