A new method for generating sigma points and weights for nonlinear filtering

Rahul Radhakrishnan, Ajay Yadav, Paresh Date, and Shovan Bhaumik

Abstract—In this paper, a new method termed as new sigma point Kalman filter (NSKF), is proposed for generating sigma points and weights for estimating the states of a stochastic nonlinear dynamic system. The sigma points and their corresponding weights are generated such that the points nearer to the mean (in inner product sense) have a higher probability of occurrence, and the mean vector and covariance matrix are matched exactly. Performance of the new algorithm is compared with the existing unscented Kalman filter (UKF), the cubature Kalman filter (CKF), the cubature quadrature Kalman filter (CQKF) and higher order unscented filter (HOUF) for two different problems. Comparison is done by calculating the root mean square error (RMSE), relative computational time and track-loss. From simulation results, it can be concluded that the proposed algorithm performs with superior estimation accuracy when compared to the UKF, CKF, CQKF and HOUF.

Index Terms—Kalman filtering, Stochastic systems, Estimation.

I. Introduction

MANY real-life problems require estimation of state of a dynamic system from noisy measurement data. For linear systems with additive white Gaussian noise, Kalman filter (KF) [1] is the optimal estimator. However, most of the real life problems are nonlinear and they can be represented with the help of process and measurement equations given by

$$\text{Process equation:} \quad x_{k+1} = \phi(x_k) + \eta_k, \tag{1}$$

$$\text{Measurement equation:} \quad z_k = \gamma(x_k) + v_k, \tag{2}$$

where $x_k \in \Re^n$ is the n-dimensional states of the system, $\phi(x_k)$ and $\gamma(x_k)$ are the nonlinear functions of x_k and k, $z_k \in \Re^d$ is the measurement, η_k and v_k are zero mean white Gaussian noise with covariance Q_k and R_k respectively.

In the Bayesian filtering approach, at each time instant the posterior density $p(x_k|z_k)$ gives the complete description of the states x_k. This posterior density can be determined recursively by the following two steps

Prediction: It is given by Chapman-Kolmogorov equation

$$p(x_k|z_{1:k-1}) = \int p(x_k|x_{k-1})p(x_{k-1}|z_{1:k-1})dx_{k-1}, \tag{3}$$

Update: It is given by Bayes' rule

$$p(x_k|z_{1:k}) = \frac{p(z_k|x_k)p(x_k|z_{1:k-1})}{p(z_k|z_{1:k-1})}, \tag{4}$$

R. Radhakrishnan, A. Yadav and S. Bhaumik are with the Department of Electrical Engineering, Indian Institute of Technology Patna, India. e-mail: (rahul.pee13, ajay.pee15, shovan.bhaumik)@iitp.ac.in

P. Date is with the Department of Mathematics, College of Engineering, Design and Physical Sciences, Brunel University, United Kingdom. e-mail: Paresh.Date@brunel.ac.uk

where $p(z_k|z_{1:k-1}) = \int p(z_k|x_k)p(x_k|z_{1:k-1})dx_k$.

As the equations (3) and (4) are intractable for nonlinear systems, it is impossible to determine the posterior density $p(x_k|z_k)$ in closed form. The simplest way to deal with this type of nonlinear estimation problem is to assume that the posterior density is Gaussian. Initially, extended Kalman Filter (EKF) [1], [2] was introduced, which linearize the nonlinear function via Taylor series expansion. However, it fails to track or converge when the process and measurement models are highly nonlinear.

Other popular filters which give better estimation accuracy than the EKF are the unscented Kalman filter (UKF) [3], [4], particle filter (PF) [5], central difference filter (CDF) [6], cubature Kalman filter (CKF) [7], cubature quadrature Kalman filter (CQKF) [8], Gauss-Hermite filter (GHF) [6] and sparse-grid quadrature filter (SGQF) [9]. But, due to low computational cost and ease of implementation, UKF is still the most popular among practitioners and have been implemented in many real-life applications such as target tracking [10], [11].

There are many variants of the unscented transformation which are available in literature. To improve accuracy, various modifications to the conventional UKF has been done. Julier, who proposed the UKF came up with variants such as the simplex sigma set and spherical simplex sigma set [12], [13]. Recently, [14] proposed a new variant of the sigma point set. In that work, it was mentioned that [12], [13] fail to match the mean and covariance of the prior random variable exactly. Few variants with more number of sigma points than UKF exist such as the higher order unscented filter (HOUF) [15] and higher order sigma point filter (HoSPF) [16]. For increasing the robustness of UKF along with accuracy, marginalised iterated UKF (MIUKF) [17] and risk-sensitive UKF (RSUKF) [18] were also proposed. A detailed description of UKF variants is available in [19]. Even though many variants of the unscented transformation is proposed, there is still scope of improvement.

In this work, we propose a new method for generating the sigma points and weights for estimating the states of a nonlinear system. The conventional unscented transformation results in a small set of points which assigns exactly the same probability weights (except for the mean itself). In the proposed new sigma point transformation, a set of new sigma points and weights are generated such that the sigma points nearer to the mean in inner product sense have a higher probability of occurrence, and the mean and covariance are still matched exactly.

II. UNSCENTED KALMAN FILTER

In UKF [4], unscented transformation (UT) is used where a set of deterministic sample points and their corresponding weights are generated to approximate the probability densities. The sample points, also called as sigma points, are propagated through the known nonlinear functions to capture the posterior mean and the covariance.

Let us consider a mean vector μ and covariance matrix $\mathbf{P}$, a discrete distribution $(\mathbb{Z}_i, w_i)$ for a random variable $\mathbb{Z}$ can be defined such that $\mathbb{P}(\mathbb{Z} = \mathbb{Z}_i) = w_i$,

$$\mathbb{E}[\mathbb{Z}] = \sum_i w_i \mathbb{Z}_i = \mu \quad \text{and}$$

$$\mathbb{E}\left[(\mathbb{Z} - \mu)(\mathbb{Z} - \mu)^T\right] = \sum_i w_i (\mathbb{Z}_i - \mu)(\mathbb{Z}_i - \mu)^T = \mathbf{P}.$$

The discrete distribution with $2n + 1$ support points for a random variable $\mathbb{Z}$ is given by

$$\mathbb{P}(\mathbb{Z} = \mu) = \frac{\kappa}{(n + \kappa)},$$

$$\mathbb{P}\left(\mathbb{Z} = \mu \pm \sqrt{(n+\kappa)\mathbf{P}_i}\right) = \frac{1}{2(n+\kappa)}, \quad i = 1, \cdots, n$$

where $\mu \pm \sqrt{(n+\kappa)\mathbf{P}_i}$ are the sigma points and the probabilities associated with it are the weights. Here κ is a user defined scaling parameter which takes value according to the relation $n + \kappa = 3$ [3]. Apart from this transformation, several modifications to the basic algorithm have been suggested, see [12], [13], [14] for example.

III. NEW SIGMA POINT KALMAN FILTER

In this section, we shall first address the reason as to what prompted us to explore for a new sigma point generation algorithm.

A. Motivation

To motivate our discussion on the new sigma point generation algorithm, note that matching only the first two moments of a continuous distribution can generate misleading discrete distributions. As a simple example, consider the two point distribution for the random variable $\mathbb{Z}$, $\mathbb{P}(\mathbb{Z} = -1) = \mathbb{P}(\mathbb{Z} = 1) = 0.5$. This has the same first three moments as the standard normal distribution, although it is a very poor approximation for the same. To overcome this at least partially, we propose an algorithm which achieves the following, in addition to matching the first two moments:

- It's unimodal with its mode coinciding with its mean. This is obviously the case with Gaussian distribution, which we are trying to approximate.
- Consider a covariance matrix $\mathbf{P}$, whose positive semidefinite square root is given by $\begin{bmatrix} U_1 & U_2 & \cdots & U_n \end{bmatrix}$, where column vectors $U_1 \cdots U_{n-1}$ are successively 'closer' to μ in the sense that $< U_i, \mu > \; < \; < U_{i+1}, \mu >$, $i = 1, 2, \cdots, n-1$, where $< \cdot, \cdot >$ is the inner product. Suppose that we are approximating the multivariate distribution by a discrete distribution whose support points include (possibly scaled versions of) $U_1, U_2, \cdots, U_n$, as is the case

in any filtering algorithm based on sigma points. Suppose further that we are assuming that the vector realizations closer to the mean are more likely than those away from the mean. This should clearly be the case if we are approximating a unimodal continuous distribution with the mode at its mean, such as the normal distribution. If $\mathbb{Z}$ represents the underlying random variable, it clearly makes intuitive sense to allocate probability weights such that $\mathbb{P}(\mathbb{Z} = U_i) < \mathbb{P}(\mathbb{Z} = U_{i+1})$. In the traditional UKF, all these vectors will be allocated the same probability weights. Our algorithm orders the weights of U_i in proportion with their inner product with the mean vector, while still matching the covariance matrix exactly. It also suggests a simple, optimization-free way of potentially matching other properties of the distribution being approximated.

B. Generation of sigma points and weights

The following steps represent the proposed new unscented transformation:

- Find any real U such that $UU^T = \mathbf{P}$, and U_i and $\mathbf{P}_i$ denote the ith column of U and $\mathbf{P}$, respectively.
- Calculate $\alpha_i = \frac{|<\mu, \mathbf{P}_i>|}{\|\mu\|_2 \|\mathbf{P}_i\|_2}$.
- Choose a real constant $m \in (0.5, 1)$.
- Choose a real constant β such that $\beta > \{\frac{1}{4} \max(m\alpha_i) - \frac{1}{2}\sum_{i=1}^n \alpha_i\}$.

Now define a discrete distribution for the random variable $\mathbb{Z}$ as

$$\mathbb{P}(\mathbb{Z} = \mu) = 1 - \frac{\sum_{i=1}^n \alpha_i}{2(\sum_{i=1}^n \alpha_i + \beta)},$$

$$\mathbb{P}\left(\mathbb{Z} = \mathbb{Z}_i = \mu \pm \sqrt{\frac{\sum_{i=1}^n \alpha_i + \beta}{m\alpha_i}} U_i\right) = \frac{m\alpha_i}{4(\sum_{i=1}^n \alpha_i + \beta)},$$

$$\mathbb{P}\left(\mathbb{Z} = \mathbb{Z}_{i+n} = \mu \pm \sqrt{\frac{\sum_{i=1}^n \alpha_i + \beta}{(1-m)\alpha_i}} U_i\right) = \frac{(1-m)\alpha_i}{4(\sum_{i=1}^n \alpha_i + \beta)},$$

for $i = 1, \cdots, n$. Hence we have a total of $4n + 1$ sigma points. This discrete distribution satisfies $\mathbb{E}(\mathbb{Z}) = \mu$ and $\mathbb{E}\left[(\mathbb{Z} - \mu)(\mathbb{Z} - \mu)^T\right] = \mathbf{P}$, with the sum of probabilities as unity.

Lemma 1: The mean and covariance of the discrete random vector $\mathbb{Z}$ with weight probabilities w_i are $\mathbb{E}[\mathbb{Z}] = \sum_i w_i \mathbb{Z}_i = \mu$ and $\mathbb{E}\left[(\mathbb{Z} - \mu)(\mathbb{Z} - \mu)^T\right] = \mathbf{P}$.

Proof: For $4n + 1$ sigma points and taking $\Psi = \sum_{i=1}^n \alpha_i + \beta$, the expected value of $\mathbb{Z}$ can be calculated as

$$\mathbb{E}[\mathbb{Z}] = \sum_{i=1}^{4n+1} w_i \mathbb{Z}_i = \mu \left(1 - \frac{\sum_{i=1}^n \alpha_i}{2\Psi}\right) + \sum_{i=1}^n \left(\frac{m\alpha_i}{4\Psi}\right)$$

$$\times \left(\mu + \sqrt{\frac{\Psi}{m\alpha_i}} U_i\right) + \sum_{i=1}^n \left(\frac{m\alpha_i}{4\Psi}\right) \left(\mu - \sqrt{\frac{\Psi}{m\alpha_i}} U_i\right)$$

$$+ \sum_{i=1}^n \left(\frac{(1-m)\alpha_i}{4\Psi}\right) \times \left(\mu + \sqrt{\frac{\Psi}{(1-m)\alpha_i}} U_i\right)$$

$$+ \sum_{i=1}^n \left(\frac{(1-m)\alpha_i}{4\Psi}\right) \times \left(\mu - \sqrt{\frac{\Psi}{(1-m)\alpha_i}} U_i\right)$$

$$= \left(1 - \frac{\sum_{i=1}^{n} \alpha_i}{2\Psi}\right)\mu + 2\sum_{i=1}^{n}\left(\frac{m\alpha_i}{4\Psi}\right)\mu + 2\sum_{i=1}^{n}\left(\frac{\alpha_i\mu}{4\Psi}\right)$$

$$- 2\sum_{i=1}^{n}\left(\frac{m\alpha_i}{4\Psi}\right)\mu = \mu.$$

Similarly, for the covariance matrix, we have

$$\mathbb{E}\left[(\mathbb{Z}-\mu)(\mathbb{Z}-\mu)^T\right] = \sum_{i=1}^{4n+1} w_i(\mathbb{Z}_i-\mu)(\mathbb{Z}_i-\mu)^T$$

$$= \sum_{i=1}^{n}\left(\frac{m\alpha_i}{4\Psi}\right)\Theta\Theta^T + \sum_{i=1}^{n}\left(\frac{m\alpha_i}{4\Psi}\right)\tilde{\Theta}\tilde{\Theta}^T + \sum_{i=1}^{n}\left(\frac{(1-m)\alpha_i}{4\Psi}\right)$$

$$\times \Lambda\Lambda^T + \sum_{i=1}^{n}\left(\frac{(1-m)\alpha_i}{4\Psi}\right)\tilde{\Lambda}\tilde{\Lambda}^T$$

$$= \sum_{i=1}^{n}\left(\frac{m\alpha_i}{4\Psi}\right)\left(\frac{\Psi}{m\alpha_i}\right)U_iU_i^T + \sum_{i=1}^{n}\left(\frac{m\alpha_i}{4\Psi}\right)\left(\frac{\Psi}{m\alpha_i}\right)U_iU_i^T$$

$$+ \sum_{i=1}^{n}\left(\frac{(1-m)\alpha_i}{4\Psi}\right)\left(\frac{\Psi}{(1-m)\alpha_i}\right)U_iU_i^T + \sum_{i=1}^{n}\frac{(1-m)\alpha_i}{4\Psi}$$

$$\times \left(\frac{\Psi}{(1-m)\alpha_i}\right)U_iU_i^T = \sum_{i=1}^{n}U_iU_i^T = \mathbf{P}. \qquad \blacksquare$$

where $\Theta = \sqrt{\frac{\Psi}{m\alpha_i}}U_i$, $\tilde{\Theta} = -\sqrt{\frac{\Psi}{m\alpha_i}}U_i$, $\Lambda = \sqrt{\frac{\Psi}{(1-m)\alpha_i}}U_i$ and $\tilde{\Lambda} = -\sqrt{\frac{\Psi}{(1-m)\alpha_i}}U_i$.

The next lemma establishes the condition on parameter β under which the maximum probability weight corresponds to the mean.

Lemma 2: $\mathbb{P}(\mathbb{Z}=\mu) > \mathbb{P}(\mathbb{Z}=\mathbb{Z}_j)$, for $j=1,\cdots,4n$, if the following condition holds: $\beta \geq \frac{1}{4}\max(m\alpha_i) - \frac{1}{2}\sum_{i=1}^{n}\alpha_i$.

Proof: From definition, $\mathbb{P}(\mathbb{Z}=\mu) > \mathbb{P}(\mathbb{Z}=\mathbb{Z}_j)$, $j=1,\cdots,4n$ if

$$1 - \frac{\sum_{i=1}^{n}\alpha_i}{2\Psi} > \max\left(\frac{m\alpha_i}{4\Psi}, \frac{(1-m)\alpha_i}{4\Psi}\right), \text{ i.e.}$$

$$\frac{\sum_{i=1}^{n}\alpha_i + 2\beta}{2\Psi} > \frac{1}{4\Psi}\max(m\alpha_i), \text{ i.e.}$$

$$\sum_{i=1}^{n}\alpha_i + 2\beta > \frac{1}{2}\max(m\alpha_i),$$

which leads to the required result. $\blacksquare$

The choice of β helps in establishing the main idea of the new sigma point transformation that the points nearer to the mean shall have more probability of occurrence, *i.e.* weights (w_i). From Lemma 2, we can write $\beta > \beta_{min}$ where $\beta_{min} = \frac{1}{4}\max(m\alpha_i) - \frac{1}{2}\sum_{i=1}^{n}\alpha_i$. If $\beta = \beta_{min}$, out of $4n+1$ weights, $w_1 = w_i$, where $i=2,\cdots,2n+1$, which is clearly not what we propose. Hence β should be greater than β_{min}.

Let us consider $\beta = \beta_{min} + b$. Allocating a large value for b tends to concentrate maximum probability mass around the mean and increases the spread of $4n$ points around the mean. This directly implies that other points are assigned weights which are negligible, as shown in Fig. 1. This condition is also not recommended as it may affect the accuracy in approximating the probability densities. Hence we recommend a moderate value for b such that all points are assigned significant weight values. To illustrate this, we provide the plot of points and weights for $\mu = [1\ 1]^T$, $\mathbf{P} = diag([1,1])$ and $m = 0.7$, for various values of b.
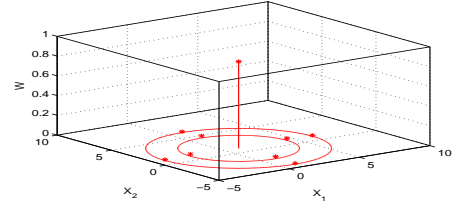

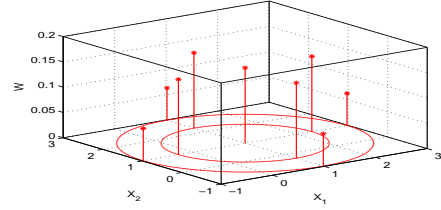Fig. 1. Points and weights of NSKF for $b = 5$

Fig. 2. Points and weights of NSKF for $b = 0$

Fig. 2 shows that when $b = 0$, $2n$ points which are far from the mean have the same probability weights as the mean. This is against the principle on which our work is based. Fig. 1 shows that taking $b = 5$ makes the probability of occurence of mean highly likely, assigning negligible weights for other points. This as mentioned above is not recommended.

Fig. 3 shows points and weights for $b = 1$ in comparison with the sigma points and weights of UKF. We recommend to represent the density function in such a way that mean has the highest probability weight, with points which are further away assigned with less probability weights, keeping in mind that they are not negligible. Choosing a larger value of b (and hence β) tends to concentrate more probability mass around the mean, *i.e.* it makes the distribution more 'peaky'. As the sigma points are symmetrically distributed across the mean, the third order moment, *i.e.* skewness will be zero.

The parameter m distributes the probability mass corresponding to U_i to four different support vectors, such that the number of sigma points around the mean is doubled. m being close to 1 makes one set of weights ($2n$) negligible while $m = 0.5$ makes the two sets ($4n$ points) to coincide. Then we will be left with only $2n+1$ points. For m values greater than 1, one set of points will become imaginary and their probability weights negative, which is not acceptable. Hence in practice, we suggest to take m in the range of $0.5 < m < 1$. It
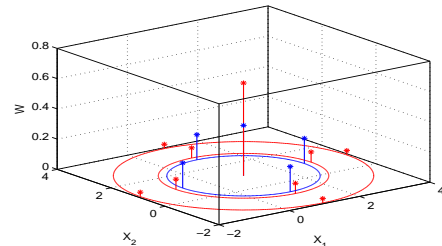

Fig. 3. Points and weights of a) UKF (blue) b) NSKF (red) for $b = 1$

is further worth noting that the definition of α_i does not affect the matching of the first two moments. Thus the algorithm can easily be modified to match a property other than adjusting the probabilities of U_i according to their inner product with the mean.

The proposed NSKF can be easily extended for formulating its square-root algorithm. Square-root algorithms are essential when the filter fails to keep or preserve the positive-definite nature of the prior and posterior covariance matrices, which can occur due to lack of arithmetic precision. In UKF, possibility of a negative weight may lead towards numerical instability as $n > 3$, which makes the square root version of the algorithm infeasible for higher order systems. But in case of NSKF, all the weights are positive (as defined in Lemma 2) irrespective of n. The square-root filtering algorithm for NSKF is the same as that given in [20].

IV. STABILITY OF THE UNSCENTED FILTERS

For the system described by Eq. (1) and (2), the prior and posterior estimation error can be defined as $\tilde{x}_{k+1|k} = x_{k+1} - \hat{x}_{k+1|k}$ and $\tilde{x}_{k+1|k+1} = x_{k+1} - \hat{x}_{k+1|k+1}$. Then $\tilde{x}_{k+1|k} = \phi(x_k) + \eta_k - \sum_{i=1}^{4n+1} w_i \phi(\mathbb{Z}_{i,k|k})$. Using Taylor series expansion around the posterior mean $\hat{x}_{k|k}$, nonlinear function $\phi(\cdot)$ can be linearized and $\tilde{x}_{k+1|k}$ is written as $\tilde{x}_{k+1|k} = A_k \tilde{x}_{k|k} + \eta_k + R(\tilde{x}_{k|k})$, where $A_k = \frac{\partial \phi(x)}{\partial x}\big|_{x=\hat{x}_{k|k}}$ and $R(\tilde{x}_{k|k})$ is the accumulated residue which arises due to ignoring the higher order terms of the Taylor series expansion and the error induced during calculation of prior mean with new sigma points and weights. It can be represented as $R(\tilde{x}_{k|k}) = B_k \Delta_{x,k} L_{x,k} \tilde{x}_{k|k} + M_k$ [21] such that B_k is a problem dependent scaling matrix, $\Delta_{x,k}$ is an unknown time varying matrix with the property $\Delta_{x,k}\Delta_{x,k}^T \leq I$, M_k represents the error induced due to new unscented transformation and matrix $L_{x,k}$ is used to provide an extra degree of freedom. Then the prior estimation error and covariance can be expressed as $\tilde{x}_{k+1|k} = (A_k + B_k \Delta_{x,k} L_{x,k})\tilde{x}_{k|k} + M_k + \eta_k$ and $P_{k+1|k} = \mathbb{E}[\tilde{x}_{k+1|k}\tilde{x}_{k+1|k}^T] = (A_k + B_k\Delta_{x,k}L_{x,k})P_{k|k}(A_k + B_k\Delta_{x,k}L_{x,k})^T + M_k M_k^T + Q_k$.

Similarly, $\gamma(x_{k+1})$ is linearized around $\hat{x}_{k+1|k}$ using Taylor series expansion. Then the innovation is represented as $\tilde{z}_{k+1|k} = z_{k+1} - \hat{z}_{k+1|k} = (C_{k+1} + E_{k+1}\Delta_{z,k+1}L_{z,k+1})\tilde{x}_{k+1|k} + N_{k+1} + v_{k+1}$, where $C_{k+1} = \frac{\partial \gamma(x)}{\partial x}\big|_{x=\hat{x}_{k+1|k}}$, E_{k+1}, $\Delta_{z,k+1}$, N_{k+1} and $L_{z,k+1}$ are similar to that defined earlier with $\Delta_{z,k+1}\Delta_{z,k+1}^T \leq I$. Now $P_{zz,k+1} = \mathbb{E}[\tilde{z}_{k+1|k}\tilde{z}_{k+1|k}^T] = (C_{k+1} + E_{k+1}\Delta_{z,k+1}L_{z,k+1})P_{k+1|k}(C_{k+1} + E_{k+1}\Delta_{z,k+1}L_{z,k+1})^T + N_{k+1}N_{k+1}^T + R_{k+1}$ and $P_{xz,k+1} = \mathbb{E}[\tilde{x}_{k+1|k}\tilde{z}_{k+1|k}^T] = P_{k+1}(C_{k+1} + E_{k+1}\Delta_{z,k+1}L_{z,k+1})^T$. Hence the Kalman gain is obtained as $K_{k+1} = P_{k+1}(C_{k+1} + E_{k+1}\Delta_{z,k+1}L_{z,k+1})^T[(C_{k+1} + E_{k+1}\Delta_{z,k+1}L_{z,k+1})P_{k+1|k}(C_{k+1} + E_{k+1}\Delta_{z,k+1}L_{z,k+1})^T + N_{k+1}N_{k+1}^T + R_{k+1}]^{-1}$. The posterior estimation error is expressed as $\tilde{x}_{k+1|k+1} = \tilde{x}_{k+1|k} - K_{k+1}\tilde{z}_{k+1|k} = [I - K_{k+1}(C_{k+1} + E_{k+1}\Delta_{z,k+1}L_{z,k+1})]\tilde{x}_{k+1|k} - K_{k+1}[N_{k+1} + v_{k+1}]$.

Lemma 3: Assuming that the linearized form of $\phi(x)$ and $\gamma(x)$ satisfy uniform observability condition, and there exist real constants $\underline{c}, \underline{l}, \bar{n}, \bar{r} > 0$ such that $\forall\ k \geq 0$; $\underline{c}^2 I \leq C_k C_k^T$, $\underline{l}^2 I \leq (E_k\Delta_{z,k}L_{z,k})(E_k\Delta_{z,k}L_{z,k})^T$, $N_k N_k^T \leq \bar{n}I$, $R_k \leq \bar{r}I$, then

$$P_{k+1|k+1} \leq \frac{\bar{n}+\bar{r}}{(\underline{c}+\underline{l})^2}I. \qquad (5)$$

Proof: From $P_{k+1|k+1} = P_{k+1|k} - K_{k+1}P_{zz,k+1}K_{k+1}^T$, and substituting the relations stated above,

$$P_{k+1|k+1} = (A_k + B_k\Delta_{x,k}L_{x,k})P_{k|k}(A_k + B_k\Delta_{x,k}L_{x,k})^T + M_k M_k^T + Q_k - \big\{\{(A_k + B_k\Delta_{x,k}L_{x,k})P_{k|k}(A_k + B_k\Delta_{x,k}L_{x,k})^T + M_k M_k^T + Q_k\}$$
$$\{C_{k+1} + E_{k+1}\Delta_{z,k+1}L_{z,k+1}\}^T \{\{C_{k+1} + E_{k+1}\Delta_{z,k+1}L_{z,k+1}\}\{(A_k + B_k\Delta_{x,k}L_{x,k})P_{k|k}(A_k + B_k\Delta_{x,k}L_{x,k})^T + M_k M_k^T + Q_k\}\{C_{k+1} + E_{k+1}\Delta_{z,k+1}L_{z,k+1}\}^T + N_{k+1}N_{k+1}^T + R_{k+1}\}^{-1}\{C_{k+1} + E_{k+1}\Delta_{z,k+1}L_{z,k+1}\}\{(A_k + B_k\Delta_{x,k}L_{x,k})P_{k|k}(A_k + B_k\Delta_{x,k}L_{x,k})^T + M_k M_k^T + Q_k\}\big\}.$$

Using the matrix inequality $(A+B+C)^{-1} > A^{-1} - A^{-1}BA^{-1} - A^{-1}CA^{-1}$ [22] for symmetric positive definite matrices A, B and C, expression for $P_{k+1|k+1}$ reduces to

$$P_{k+1|k+1} \leq [C_{k+1} + E_{k+1}\Delta_{z,k+1}L_{z,k+1}]^{-1}[N_{k+1}N_{k+1}^T + R_{k+1}][C_{k+1} + E_{k+1}\Delta_{z,k+1}L_{z,k+1}]^{-T},$$

from which using the above mentioned bounds, the result in Eq. (5) is obtained. ∎

This result proves the boundedness of estimation error covariance. Note that this result is generic and holds for unscented Kalman filters other than NSKF as well.

V. SIMULATION RESULTS

In this section, NSKF has been implemented on two well known nonlinear filtering problems and its performance is studied. For a fair comparison, popular nonlinear filters like the UKF, CKF and CQKF are also implemented. Here the accuracy level of CQKF is taken as 2 [8], such that it almost matches the number of points generated by the proposed NSKF. Further, to compare with an unscented filter with the same number of sigma points as that of NSKF, higher order unscented filter (HOUF) [15] is implemented.

Problem 1: Here we consider a single dimensional system with process model [8], [6] $x_{k+1} = \phi(x_k) + \eta_k$ and measurement model $z_k = \gamma(x_k) + v_k$, where $\phi(x_k) = x_k + T5x_k(1-x_k^2)$, $\gamma(x_k) = Tx_k(1-0.5x_k)$, $\eta \sim \aleph(0, \bar{b}^2 T)$, $v_k \sim \aleph(0, d^2 T)$ and $\aleph$ represents a normal distribution.

The sampling time $T = 0.01sec$ and the simulation is done for a total of 4 seconds. The value of simulation parameters $\bar{b}$, d and m are 0.5, 0.11 and 0.8 respectively. All the filters were fed with an initial estimate $\hat{x}_{0|0} = -0.8$ and initial error covariance $P_{0|0} = 2$, with the truth being initialized as $x_{0|0} = -0.2$. Filtering performance has been compared in terms of RMSE and the number of track-loss.

Fig. 4 shows the RMSE out of 1000 Monte Carlo runs. From this figure, we can see that the proposed filter performs with better accuracy in comparison to UKF, CKF, CQKF and HOUF. Filtering performance has also been studied in terms of diverged tracks. Track-loss is the number of cases when the estimation error goes beyond a pre-defined value, and here we took it as $|x_{kmax} - \hat{x}_{kmax}| < 1$, where $kmax$ is the final time step. Track-loss in percentage and number of points for different filters are summarized in Table I where we can observe the superior filtering accuracy of NSKF, with less number of diverged tracks.
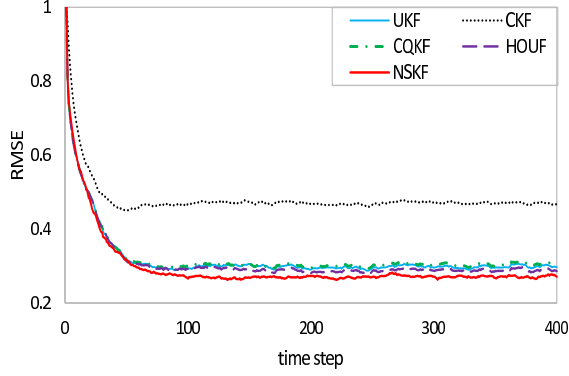
Fig. 4. RMSE for Problem 1

Parameters	Values
Initial range (r)	$5\ km$
Target speed (s)	$4\ knots$
Target course	-140°
Observer speed	$5\ knots$
Observer initial course	140°
Observer final course	20°
Observer manoeuvre	From 13th to 17th min
σ_θ	2°
σ_c	$\pi/\sqrt{12}$
$\tilde{q}$	$2.944 \times 10^{-6} km^2/\mathrm{min}^3$

TABLE I
TRACK-LOSS AND NUMBER OF POINTS FOR DIFFERENT FILTERS

Filter	Track-loss (%)	Number of points
CKF	5.5	2
UKF	2	3
CQKF	2	4
HOUF	1.8	5
NSKF	1.2	5

Problem 2: A real-life passive underwater bearings-only tracking (BOT) scenario is solved using the proposed new sigma point Kalman filter, where a constant velocity target is tracked using a manoeuvring observer. If the observer is following a non manoeuvring path, the dynamic model becomes unobservable and it is impossible to obtain target position accurately [5]. The target observer dynamics considered in this problem is shown in Fig 5. The state vector for the two vessels: namely target and observer containing states as position and velocity are defined as $\mathbf{x}_k^t = [x_k^t\ y_k^t\ \dot{x}_k^t\ \dot{y}_k^t]^T$ and $\mathbf{x}_k^o = [x_k^o\ y_k^o\ \dot{x}_k^o\ \dot{y}_k^o]^T$. Now, a relative state vector can be defined as $\mathbf{x}_k \triangleq \mathbf{x}_k^t - \mathbf{x}_k^o$. Hence the process model involving relative states is described as $\mathbf{x}_{k+1} = \mathbf{F}\mathbf{x}_k + \eta_k - \mathcal{U}_{k,k+1}$, where

$$\mathbf{F} = \begin{bmatrix} I_{2\times 2} & T I_{2\times 2} \\ O_{2\times 2} & I_{2\times 2} \end{bmatrix},$$

and $\mathcal{U}_{k,k+1}$ is a vector of observer inputs given by $\mathcal{U}_{k,k+1} = [x_{k+1}^o - x_k^o - T\dot{x}_k^o,\ y_{k+1}^o - y_k^o - T\dot{y}_k^o,\ \dot{x}_{k+1}^o - \dot{x}_k^o,\ \dot{y}_{k+1}^o - \dot{y}_k^o]^T$. Here $I_{2\times 2}$ is an identity matrix of order 2 and $O_{2\times 2}$ is a zero matrix of order 2. The process noise vector η_k is assumed to follow a normal distribution with zero mean and covariance Q, where Q is the same as mentioned in [23].

Now the nonlinear bearing measurement can be expressed as $z_k = \theta_k + v_k$, where v_k is modeled as a Gaussian noise with zero mean and standard deviation σ_θ. The true bearing measurement is expressed as $\theta_k = \tan^{-1}\left(\frac{x_k}{y_k}\right)$. These are obtained in a direction from the observer to the target with a reference clockwise positive to the y-axis. The parameters considered for simulating the tracking scenario is given in Table II, where the simulation period lasts for 30 min. The filters are all initialized based on the initial bearing measurement z_o as given in [23]. The initial state vector $\hat{\mathbf{x}}_{0|0}$ is assumed as

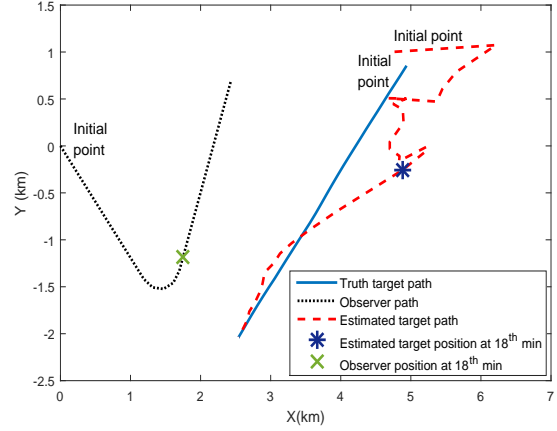

Fig. 5. Tracking scenario with estimated target path given by NSKF

$\hat{\mathbf{x}}_{0|0} = [r\sin(z_0)\ r\cos(z_0)\ s\sin(\bar{c}) - \dot{x}_0^o\ s\cos(\bar{c}) - \dot{y}_0^o]^T$ and $\mathbf{P}_{0|0}$ is taken according to [23].

Fig. 5 shows the estimated target path generated by the proposed NSKF. From this figure, it can be observed that only after observer manoeuvre the estimated target path starts to follow its truth path. To illustrate this, observer and estimated target position at 18^{th}min (after observer manoeuvre) is marked in Fig. 5. Performance comparison was done by calculating the RMSE in position and velocity, and the percentage of track-loss over 1000 Monte Carlo runs. Track-loss is calculated by counting the estimated track which move away from the truth track without converging by a predefined value. In this simulation, a track-loss is said to have occurred when the position error at the final time step is greater than or equal to 1 *km*. Percentage of track-loss incurred by all the filters is listed in Table III and it can be inferred that NSKF incurred far fewer instances of track-loss, at least 5 times fewer than the other filters being compared. Table III also shows the number of points and relative computational time for all the filters.

RMSE in position and velocity after observer manoeuvre for all the filters are plotted in Fig. 6 and 7. The plots were obtained by excluding the diverged tracks defined according to our predefined track-loss condition. From these figures, it is observed that while the position RMSEs are comparable for all filters, the RMSE in velocity is clearly and consistently lower for NSKF as compared to other filters. This shows the accurate tracking performance of NSKF.

TABLE III
TRACK-LOSS, NUMBER OF POINTS & RELATIVE COMPUTATIONAL TIME

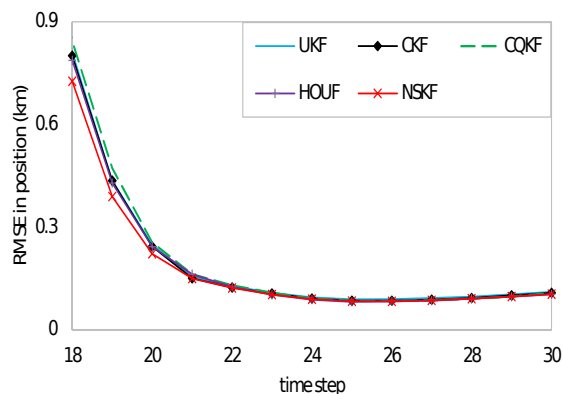| Filter	Track-loss (%)	No. of points	Relative compu. time
CKF	1	8	1
UKF	1.2	9	1.15
CQKF	1.5	16	1.55
HOUF	1	17	1.58
NSKF	0.2	17	1.58

Fig. 6. RMSE in position for Problem 2

We have also calculated the variance of squared error (VSE) in position and velocity for NSKF over 1000 ensemble. VSE at each step k represents the spread of squared errors over ensemble. It is found that VSEs obtained from different filters are comparable and decrease with time step.

As a further comparative study, we have also studied the performance of NSKF in comparison to all other filters mentioned here, for an another tracking problem (problem 2 in [24]). We have seen that velocity estimation with NSKF for that problem is more accurate compared to other filters reported here. Results of these studies are omitted for brevity and are available from the authors on request.

VI. CONCLUSIONS

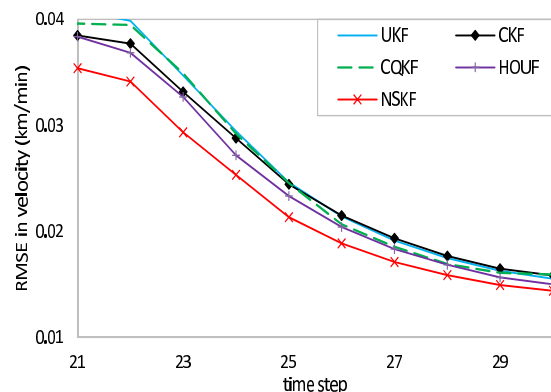In this work, a new sigma point Kalman filter is proposed which uses a new set of deterministic points and weights. The proposed transformation gives a unimodal distribution with a mode at its mean, and with sigma points which are nearer to the mean in inner product sense having higher probability of occurrence. Simulation results show the improved performance of the proposed filter in comparison to the existing UKF, CKF, CQKF and HOUF.

REFERENCES

[1] B. D. O. Anderson and J. B. Moore, Optimal filtering, Dover Publication, New York, 2005.
[2] Y. Bar-Shalom, X. R. Li and T. Kirubarajan, Estimation with application to tracking and navigation, Wiley-Interscience, New York, 2001.
[3] S. J. Julier and J. K. Uhlmann, "New extension of the Kalman filter to nonlinear systems," *Proc. SPIE, Signal Process., Sensor Fusion, and Target Recognition VI*, vol. 3068, pp. 182-193, July 1997.
[4] J. Simon, J. Uhlmann, F. Hugh and W. Durrant, "A new method for the nonlinear transformation of means and covariances in filters and estimators," *IEEE Trans. Autom. Control*, vol. 45, no. 3, pp. 477-482, March 2000.
[5] B. Ristic, S. Arulampalam, N. Gordon, Beyond the Kalman filter: Particle filters for tracking applications, Artech house Boston, 2004.
[6] K. Ito and K. Xiong, "Gaussian filters for nonlinear filtering problems," *IEEE Trans. Autom. Control.*, vol. 45, no. 5, pp. 910-927, May 2000.
[7] I. Arasaratnam and S. Haykin, "Cubature Kalman filter," *IEEE Trans. Autom. Control*, vol. 54, no. 6, pp. 1254-1269, June 2009.
[8] S. Bhaumik and Swati, "Cubature quadrature Kalman filter," *IET Signal Process.*, vol. 7, no. 7, pp. 533-541, April 2013.
[9] B. Jia, M. Xing and Y. Chin, "Sparse-grid quadrature nonlinear filtering," *Automatica*, vol. 48, no. 2, pp. 327-341, Feb. 2012.
[10] G. Yu. Kulikov and M. V. Kulikova, "Accurate continuous-discrete unscented Kalman filtering for estimation of nonlinear continuous-time stochastic models in radar tracking," *Signal Process.*, vol. 139, pp. 25-35, Oct. 2017.
[11] B. Ristic and M. S. Arulampalam, "Tracking a manoeuvring target using angle-only measurements: algorithms and performance," *Signal Process.*, vol. 83, no. 6, pp. 1223-1238, June 2003.
[12] S. J. Julier and J. K. Uhlmann, "Reduced sigma point filters for the propagation of means and covariances through nonlinear transformations," in *Proc. IEEE American Control Conf.*, 2002, pp. 887-892.
[13] S. J. Julier, "The spherical simplex unscented transformation," in *Proc. IEEE American Control Conf.*, 2003, pp. 2430-2434.
[14] H. Menegaz, J. Ishihara and G. Borges, "New minimum sigma set for unscented filtering," *Int. J. Robust Nonlinear Control*, vol. 25, no. 17, pp. 3286-3298, Nov. 2015.
[15] D. Tenne and T. Singh, "The higher order unscented filter," in *Proc. IEEE American Control Conf.*, 2003, pp. 2441-2446.
[16] K. Ponomareva and P. Date, "Higher order sigma point filter: a new heuristic for nonlinear time series filtering," *Applied Mathematics and Computation*, vol. 221, pp. 662-671, Sep. 2013.
[17] L. Chang, B. Hu, G. Chang and A. Li, "Marginalized iterated unscented Kalman filter," *IET Control Theory Appl.*, vol. 6, no. 6, pp. 847-854, April 2012.
[18] S. Bhaumik, S. Sadhu and T. Ghoshal, "Risk-sensitive formulation of unscented Kalman filter," *IET Control Theory Appl.*, vol. 3, no. 4, pp. 375-382, April 2009.
[19] H. Menegaz, J. Ishihara, G. Borges and A. Vargas, "A systematization of the unscented Kalman filter theory," *IEEE Trans. Autom. Control*, vol. 60, no. 10, pp. 2583-2598, Feb. 2015.
[20] S. Bhaumik and Swati, "Square-root cubature-quadrature Kalman filter," Asian J. of Control, vol. 16, no. 2, pp. 617-622, March 2014.
[21] L. Li, D. Yu, Y. Xia, and H. Yang, "Stochastic stability of a modified unscented Kalman filter with stochastic nonlinearities and multiple fading measurements," *J. Franklin Inst.*, vol. 354, no. 2, pp. 650-667, Jan. 2017.
[22] S. Kluge, K. Reif, and M. Brokate, "Stochastic stability of the extended Kalman filter with intermittent observation," *IEEE Trans. Autom. Control*, vol. 55, no. 2, pp. 514-518, Jan. 2010.
[23] R. Radhakrishnan, S. Bahumik and N. Tomar, "Gaussian sum shifted Rayleigh filter for underwater passive bearings-only target tracking problems," *IEEE J. Ocean. Eng.*, DOI: 10.1109/JOE.2018.2814218, April 2018.
[24] R. Radhakrishnan, A. K. Singh, S. Bahumik and N. Tomar, Multiple sparse-grid Gauss-Hermite filtering, *Appl. Math. Model.*, vol. 40, no. 7-8, pp. 4441-4450, April 2016.

Fig. 7. RMSE in velocity for Problem 2