Scaling parameters selection principle for the scaled unscented Kalman filter

NIE Yongfang[1,2] and ZHANG Tao[1,*]

1. Department of Automation, Tsinghua University, Beijing 100084, China;
2. Department of Strategic Missile and Underwater Weapon, Naval Submarine Academy, Qingdao 266071, China

Abstract: The paper deals with the state estimation of the widely used scaled unscented Kalman filter (UKF). In particular, the stress is laid on the scaling parameters selection principle for the scaled UKF. Several problems caused by recommended constant scaling parameters are highlighted. On the basis of the analyses, an effective scaled UKF is proposed with self-adaptive scaling parameters, which is easy to understand and implement in engineering. Two typical strong nonlinear examples are given and their simulation results show the effectiveness of the proposed principle and algorithm.

Keywords: nonlinear filtering, scaled unscented Kalman filter, scaling parameter selection principle.

1. Introduction

The unscented transformation (UT) was first proposed by Julier and Uhlman [1,2]. It can propagate mean and covariance through nonlinear transformation. On the basis of UT, a new filter named the unscented Kalman filter (UKF) is yielded. The UKF is widely used in many fields to solve nonlinear state estimators, such as control models [3], state estimation [4] and signal processing [5]. In the UT algorithm, the first two moments of the prior distribution are estimated by a set of deterministically selected sigma points with designated weights. The estimate performance of the UT depends on the number of sigma points and the corresponding weights. Therefore, the sigma point selection principle is important to ensure the good performance of UKF, which is always a challenge for some strong nonlinear systems, such as exponents and trigonometric functions.

Julier derived a symmetrically distributed set of $2n + 1$ sigma points whose mean and covariance match the prior

distribution [1]. Then, Julier discussed the minimal skew set of simplex points that minimize the magnitude of the third order moments [6] to reduce the computational costs, and proved that, for an n-dimensional state, only $n + 1$ points are required to represent the mean and covariance fully. Thereafter, Julier pioneered the spherical simplex UT [7] that requires $n + 2$ points for n dimensions to solve the weights numerical problem in the minimal skew.

However, the radius of the sphere that bounds the above sigma point sets will increase as the dimension increases. Julier developed the scaled UT [8] to solve this question.

The scaled sigma point selection principle is widely used in UKF applications. Although several basic recommendations on choice of the scaling parameters have been provided, they are independent of the nonlinear system or only depend on the dimension of the system. Thus, the parameters are chosen prior to the estimation nonlinear system and remain constant during the entire filtering. If the parameters of the sigma points are inappropriate to the given nonlinear system, the UKF may not achieve the ideal performance, and lack robustness [9].

Minimal attention has focus on the appropriate selection of the parameters. An offline approach is proposed using the discriminative learning method (maximizing the likelihood function) and the training set of data [10]. However, the offline approach is unsuitable for dynamical systems because the change of the working point requires a new training procedure. Adaptive algorithms are discussed based on maximizing the different performance criteria with respect to the scaling parameter at each time instant in [11 – 13], such as the mean squared error (MSE), the maximize likelihood (ML) and the maximize a posteriori probability (MAP). However, these algorithms all suffer from yielding extra computational cost, the amount of which depends on the number of possible values of scaling parameters. The algorithms in [11 – 13] were summarized and refined in [14] to reduce the computational cost, but the pre-

sented numerical results showed that this UKF is still sig-nificantly slower than the regular UKF. Recently, two com-plete off-line tuning algorithms for the UKF were proposed in [15], namely, a new stochastic tuning algorithm and a standard model-based optimizer. The offline processes do not impose extra computational cost to the filter at runtime. Whereas the model-based optimizer that needs to select the goal function and the good stochastic tuning algorithm re-lies on the assumption that the goal function shows some degree of smoothness. The performance of the tuned UKF is dependent on the goal function. Therefore, the objective of this paper is to highlight the scaling parameter selec-tion principle and proposes an effective algorithm for the scaled UKF, which is easy to understand and implement in engineering.

This paper is an extended version of the conference paper [16], where we only analyze the drawback of the constant scaling parameters for the scaled UKF with one example and propose an adaptive scaling parameter algo-rithm for α. In this paper, we discuss the problems caused by the constant scaling parameters in depth and propose a more perfect selection principle of the scaling parameters for the scaled UKF.

The rest of this paper is organized as follows. Section 2 reviews the main UKF variants and the basic recommen-dations of scaling parameters. Section 3 points out several problems in the literature caused by the constant scaling parameters. Section 4 proposes a new scaled UKF with self-adaptive scaling parameters selection principle. Sec-tion 5 presents two typical examples to demonstrate the ef-fectiveness of the proposed principle and algorithm. Con-clusions are drawn in Section 6.

2. Problem statements

2.1 Symmetrical UKF

Julier proposed the UKF based on the UT to improve the flaws of the extended Kalman filter (EKF) [1].

Consider the nonlinear filtering system:

$$\boldsymbol{\xi}_k = \psi(\boldsymbol{\xi}_{k-1}, \boldsymbol{u}_{k-1}) + \boldsymbol{\tau}_{k-1} \qquad (1)$$

$$\boldsymbol{\zeta}_k = \rho(\boldsymbol{\xi}_k, \boldsymbol{u}_k) + \boldsymbol{\mu}_k \qquad (2)$$

where $\boldsymbol{\xi}_k \in \mathbf{R}^n$ is the state vector, $\boldsymbol{u}_k \in \mathbf{R}^m$ is the control input vector and $\boldsymbol{\zeta}_k \in \mathbf{R}^p$ is the measurement output vec-tor. $\psi(.)$ and $\rho(.)$ are nonlinear known functions. $\boldsymbol{\tau}_{k-1}$ and $\boldsymbol{\mu}_k$ are independent process and measurement white Gaus-sian noise with zero means and covariance $\boldsymbol{\Gamma}_{k-1}$ and $\boldsymbol{R}_k$, respectively.

A set of N sigma points and weights $\{\widetilde{\boldsymbol{\chi}}_i, \widetilde{\omega}_i\}_{i=1}^N$ are

chosen to satisfy the following conditions [2]:

$$\begin{cases} \sum_{i=1}^N \widetilde{\omega}_i = 1 \\ \sum_{i=1}^N \widetilde{\omega}_i \widetilde{\boldsymbol{\chi}}_i = \overline{\boldsymbol{\xi}} \\ \sum_{i=1}^N \widetilde{\omega}_i (\widetilde{\boldsymbol{\chi}}_i - \overline{\boldsymbol{\xi}})(\widetilde{\boldsymbol{\chi}}_i - \overline{\boldsymbol{\xi}})^{\mathrm{T}} = \boldsymbol{P} \end{cases} . \qquad (3)$$

Many UKF variants, which differ from each other de-pending on how sigma point sets are selected, have been presented in the literature. (See details in Table I of [17]). Among them, the symmetrically distributed sigma point set $\{\widetilde{\boldsymbol{\chi}}_i, \widetilde{\omega}_i\}_{i=0}^{2n}$ is the most frequently used [1].

$$\begin{cases} \widetilde{\boldsymbol{\chi}}_0 = \overline{\boldsymbol{\xi}}, \quad \widetilde{\omega}_0 = \kappa/(n+\kappa) \\ \widetilde{\boldsymbol{\chi}}_i = \overline{\boldsymbol{\xi}} + (\sqrt{(n+\kappa)\boldsymbol{P}})_i, \quad i = 1, \ldots, n \\ \widetilde{\boldsymbol{\chi}}_{i+n} = \overline{\boldsymbol{\xi}} - (\sqrt{(n+\kappa)\boldsymbol{P}})_i, \quad i = 1, \ldots, n \\ \widetilde{\omega}_i = 1/2(n+\kappa), \quad i = 1, \ldots, 2n \end{cases} \qquad (4)$$

where κ is the scaling parameter for the UKF.

The symmetrical UKF keeps the structure of the Kalman filter for linear systems, which has two steps, namely, time update and measurement update.

Time update:

Step 1 Given $\boldsymbol{\xi}_0$ and $\boldsymbol{P}_0$, form the sigma points set $\{\widetilde{\boldsymbol{\chi}}_{k-1}^i, \widetilde{\omega}_i\}_{i=0}^{2n}$ according to (4).

$$\widetilde{\boldsymbol{\chi}}_{k-1}^i = [\boldsymbol{\xi}_{k-1}, \boldsymbol{\xi}_{k-1} \pm (\sqrt{(n+\kappa)\boldsymbol{P}_{k-1}})_i] \qquad (5)$$

Step 2 Propagate the sigma points through the pro-cess model (1), compute the predict mean and covariance as follows:

$$\boldsymbol{\xi}_{k|k-1} = \sum_{i=0}^{2n} \widetilde{\omega}_i \psi(\widetilde{\boldsymbol{\chi}}_{k-1}^i) \qquad (6)$$

$$\boldsymbol{P}_{k|k-1} = \sum_{i=0}^{2n} \widetilde{\omega}_i (\psi(\widetilde{\boldsymbol{\chi}}_{k-1}^i) - \boldsymbol{\xi}_{k|k-1}) \cdot$$
$$(\psi(\widetilde{\boldsymbol{\chi}}_{k-1}^i) - \boldsymbol{\xi}_{k|k-1})^{\mathrm{T}} + \boldsymbol{\Gamma}_{k-1}. \qquad (7)$$

Measurement update:

Step 3 Form the sigma points $\{\widetilde{\boldsymbol{\chi}}_{k|k-1}^i, \widetilde{\omega}_i\}_{i=0}^{2n}$ accord-ing to (4).

$$\widetilde{\boldsymbol{\chi}}_{k|k-1}^i = [\boldsymbol{\xi}_{k|k-1}, \boldsymbol{\xi}_{k|k-1} \pm (\sqrt{(n+\kappa)\boldsymbol{P}_{k|k-1}})_i] \quad (8)$$

Step 4 Propagate the sigma points through the mea-surement model (2), and compute the predicted mean, co-variance and the cross-covariance as follows:

$$\overline{\boldsymbol{\zeta}}_k = \sum_{i=0}^{2n} \widetilde{\omega}_i \rho(\widetilde{\boldsymbol{\chi}}_{k|k-1}^i) \qquad (9)$$

$$\boldsymbol{P}_{\zeta\zeta} = \sum_{i=0}^{2n} \widetilde{\omega}_i (\rho(\widetilde{\boldsymbol{\chi}}_{k|k-1}^i) - \overline{\boldsymbol{\zeta}}_k)(\rho(\widetilde{\boldsymbol{\chi}}_{k|k-1}^i) - \overline{\boldsymbol{\zeta}}_k)^{\mathrm{T}} \quad (10)$$

$$P_{\xi\zeta} = \sum_{i=0}^{2n} \widetilde{\omega}_i(\psi(\widetilde{\boldsymbol{\chi}}_{k-1}^i) - \boldsymbol{\xi}_{k|k-1})(\rho(\widetilde{\boldsymbol{\chi}}_{k|k-1}^i) - \overline{\boldsymbol{\zeta}}_k)^{\mathrm{T}}$$

(11)

Step 5 Compute the filter gain, the update state mean and covariance:

$$\boldsymbol{K}_k = \boldsymbol{P}_{\xi\zeta}(\boldsymbol{P}_{\zeta\zeta} + \boldsymbol{R}_k)^{-1} \tag{12}$$

$$\boldsymbol{\xi}_k = \boldsymbol{\xi}_{k|k-1} + \boldsymbol{K}_k(\boldsymbol{\zeta}_k - \overline{\boldsymbol{\zeta}}_k) \tag{13}$$

$$\boldsymbol{P}_k = \boldsymbol{P}_{k|k-1} - \boldsymbol{K}_k \boldsymbol{P}_{\zeta\zeta} \boldsymbol{K}_k^{\mathrm{T}}. \tag{14}$$

2.2 Scaled UT

When κ is fixed, as n increases, $\sqrt{(n+\kappa)}$ also increases. The distance of the sigma point from the center point may exceed the small neighborhood of the center point, which also results in the non-local effect. As the dimension increases, the non-local effect becomes seriously. Therefore, Julier proposed the scaled UT which scales the sigma points to an arbitrary dimension [8]. Then (4) can be rewritten as (15) and (16).

$$\widetilde{\boldsymbol{\chi}}'_i = \overline{\boldsymbol{\xi}} + \alpha(\widetilde{\boldsymbol{\chi}}_i - \overline{\boldsymbol{\xi}}) \tag{15}$$

$$\begin{cases} \lambda = \alpha^2(n+\kappa) - n \\ \widetilde{\boldsymbol{\chi}}'_0 = \overline{\boldsymbol{\xi}} \\ \widetilde{\omega}'^m_0 = \lambda/(n+\lambda) \\ \widetilde{\omega}'^c_0 = \widetilde{\omega}'^m_0 + (1 - \alpha^2 + \beta) \\ \widetilde{\boldsymbol{\chi}}'_i = \overline{\boldsymbol{\xi}} + (\sqrt{(n+\lambda)\boldsymbol{P}})_i, \quad i = 1, \ldots, n \\ \widetilde{\boldsymbol{\chi}}'_{i+n} = \overline{\boldsymbol{\xi}} - (\sqrt{(n+\lambda)\boldsymbol{P}})_i, \quad i = 1, \ldots, n \\ \widetilde{\omega}'^m_i = \widetilde{\omega}'^c_i = 1/2(n+\lambda) \end{cases} \tag{16}$$

where α, β, κ are the scaling parameters of the scaled UKF. Some basic recommendations on the selection of scaling parameters are provided. Generally, $\kappa \geqslant 0$ ($\kappa = 3 - n$ or $\kappa = 0$) is selected to guarantee positive semi-definiteness of the covariance matrix [8]. $\alpha \in [10^{-4}, 1]$ is a small number to avoid sampling non-local effects when the nonlinearities are strong. β is used to incorporate prior knowledge of the higher order moments of the distribution. For Gaussian distribution, $\beta = 2$ is optimal [8].

3. Several problems caused by the constant scaling parameters

In most UKF applications, scaling parameters are predetermined and remain constant during the entire experiment. Sometimes, the estimation performance of the UKF with constant scaling parameters is not as good as that of other filters, such as EKF [9], EKF2 [18], Gaussian Hermite filter (GHF) [19] or cubature Kalman filter (CKF) [20]. Examples are given below to point out several problems caused by the predefined constant scaling parameters for the UKF.

3.1 Problem 1

It is not the case that α should be selected arbitrarily small to minimize higher order effects.

Consider a nonlinear model called univariate nonstationary growth model (UNGM). It is particularly interesting and challenging for nonlinear filtering because of its highly nonlinear and bimodal. It is widely used in the literature to show the estimate performance of the proposed nonlinear filtering, such as [9,21,22].

The dynamic state space model for UNGM [9] can be written as follows:

$$\begin{cases} \xi_k = 0.5\xi_{k-1} + 25\xi_{k-1}/(1 + \xi_{k-1}^2) + \\ \qquad 8\cos(1.2(k-1)) + \tau_k \\ \zeta_k = \xi_k^2/20 + \mu_k, \quad k = 1, 2, \ldots, k \end{cases} \tag{17}$$

where $\tau_k \sim N(\tau_k; 0, 1)$ and $\mu_k \sim N(\mu_k; 0, 1)$, $\xi_0 = 0.1$, $P_0 = 1$.

In [9], using a set of constant parameters $\alpha = 0.5, \beta = 2, \kappa = 2$, the authors pointed out that the simulation result was very poor and even worse than that of the EKF. However, the authors only saw the superficial question and did not discuss the underlying reasons. Fig. 1 shows the poor result of [9].
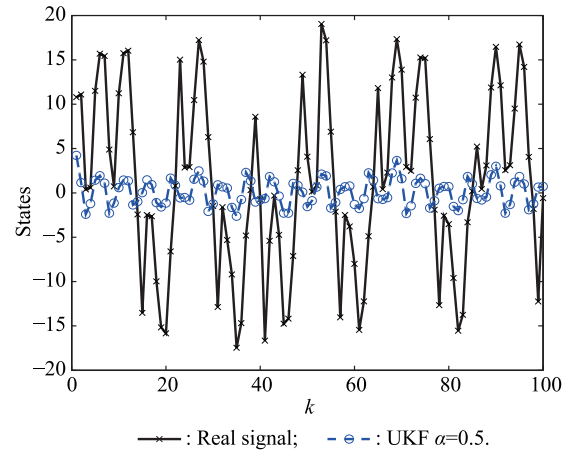

Fig. 1 UKF result with constant scaling parameters $\alpha = 0.5$, $\beta = 2, \kappa = 2$

The main reason for the poor result is that the scaling parameters of the sigma points are inappropriate, such that they do not actually capture the higher moments. In this model, when the other set of constant scaling parameters $\alpha = 1, \beta = 2, \kappa = 2$ are used, Fig. 2 illustrates a better performance than Fig. 1.

In this example, the authors select $\kappa = 3 - n = 2$ according to the recommendation. The poor result shown in Fig. 1 indicates that the ideal result cannot be obtained through the recommended scaling parameters when using the scaled UKF for a strong nonlinear system. Comparing

the results shown in Fig. 1 with Fig. 2, it can be seen that a smaller value of α does not mean an improved performance.
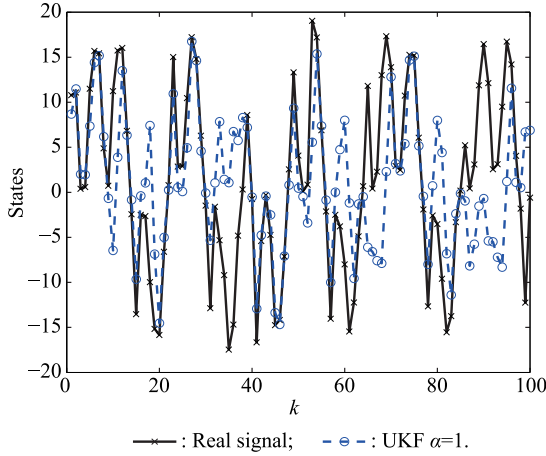

Fig. 2 UKF result with constant scaling parameters $\alpha = 1$, $\beta = 2, \kappa = 2$

3.2 Problem 2

κ is critical and should not be selected just according to recommendations.

Gustafsson demonstrated a counter-example to show that the scaled UKF cannot give the correct second order moments, even for the quadratic function that has an analytical solution [18], which was referenced in [17] and [19] as an example to show the same question of the scaled UKF for arbitrary nonlinearities.

$$\zeta = h(\boldsymbol{\xi}) = \boldsymbol{\xi}^{\mathrm{T}}\boldsymbol{\xi}, \quad \boldsymbol{\xi} \sim N(\boldsymbol{\xi}; \mathbf{0}_{n\times 1}, \boldsymbol{I}_n). \quad (18)$$

A typical chi-square distribution has the mean $\overline{\zeta} = n$ and covariance $P = 2n$. For this nonlinear system, the authors use EKF, EKF2, UKF1, UKF2 and CKF filters and compare their performances with the true mean and covariance [18]. For the completeness of the paper, some results of [18] are referenced and listed in Table 1.

Table 1   Calculation results with typical filters [18]

Filter	Mean	Covariance	Parameter
EKF	0	0	—
EKF2	n	$2n$	—
UKF1	n	$(3-n)n$	$\alpha = \sqrt{3/n}, \beta = 3/n - 1, \kappa = 0$
UKF2	n	$2n^2$	$\alpha = 10^{-3}, \beta = 2, \kappa = 0$
CKF	n	0	—

Obviously, only EKF2 works perfectly. The results show that UKF1 yields a negative covariance, whereas UKF2 overestimates the covariance. In fact, these incorrect results are due to the predefined constant scaling parameters.

Theorem 1   For the specific nonlinear system (18), its mean $\overline{\zeta} = n$ and covariance $P_{\zeta\zeta} = \kappa n$ with the symmetrical UKF.

Proof   Using (4) to obtain $\{\widetilde{\boldsymbol{\chi}}_i, \widetilde{\omega}_i\}_{i=0}^{2n}$, then substituting them into (6)–(8), we can calculate the mean $\overline{\zeta}$ and covariance $P_{\zeta\zeta}$ as follows:

$$\widetilde{\boldsymbol{\chi}}_0 = \mathbf{0}$$
$$\widetilde{\boldsymbol{\chi}}_i = \mathbf{0} + (\sqrt{(n+\kappa)\boldsymbol{I}_n})_i = \sqrt{n+\kappa}\boldsymbol{e}_i$$
$$\widetilde{\boldsymbol{\chi}}_{i+n} = -\widetilde{\boldsymbol{\chi}}_i, \quad i = 1, \ldots, n$$
$$h(\widetilde{\boldsymbol{\chi}}_0) = 0$$
$$h(\widetilde{\boldsymbol{\chi}}_i) = (\sqrt{n+\kappa}\boldsymbol{e}_i)^{\mathrm{T}}(\sqrt{n+\kappa}\boldsymbol{e}_i) = n+\kappa$$
$$h(\widetilde{\boldsymbol{\chi}}_{i+n}) = (-\sqrt{n+\kappa}\boldsymbol{e}_i)^{\mathrm{T}}(-\sqrt{n+\kappa}\boldsymbol{e}_i) = n+\kappa$$
$$\overline{\zeta} = \sum_{i=0}^{2n}\widetilde{\omega}_i h(\widetilde{\boldsymbol{\chi}}_i) = \widetilde{\omega}_0 h(\widetilde{\boldsymbol{\chi}}_0) +$$
$$\frac{1}{2(n+\kappa)}\sum_{i=1}^{2n} h(\widetilde{\boldsymbol{\chi}}_i) = \frac{2n(n+\kappa)}{2(n+\kappa)} = n$$
$$P_{\zeta\zeta} = \sum_{i=0}^{2n}\widetilde{\omega}_i(h(\widetilde{\boldsymbol{\chi}}_i) - \overline{\zeta})(h(\widetilde{\boldsymbol{\chi}}_i) - \overline{\zeta}) =$$
$$\widetilde{\omega}_0(h(\widetilde{\boldsymbol{\chi}}_0) - \overline{\zeta})(h(\widetilde{\boldsymbol{\chi}}_0) - \overline{\zeta}) +$$
$$\frac{1}{2(n+\kappa)}\sum_{i=1}^{2n}(h(\widetilde{\boldsymbol{\chi}}_i) - \overline{\zeta})(h(\widetilde{\boldsymbol{\chi}}_i) - \overline{\zeta}) =$$
$$\frac{\kappa n^2}{n+k} + \frac{2n\kappa^2}{2(n+\kappa)} = \kappa n$$

where $\boldsymbol{e}_i$ is the unit vector in $\mathbf{R}^n$ with the ith element being 1.   □

Therefore, when $\kappa = 2$, the symmetrical UKF can give the correct mean and covariance for the chi-square distribution. To satisfy the condition in (3), the sigma point set $\{\widetilde{\boldsymbol{\chi}}_i, \widetilde{\omega}_i\}_{i=0}^{2n}$ of the symmetrical UKF and the sigma point set $\{\boldsymbol{\chi}'_i, \omega'_i\}_{i=0}^{2n}$ of the scaled UKF should fulfill (15) and the following equation [8]:

$$\widetilde{\omega}'_i = \begin{cases} \widetilde{\omega}_0/\alpha^2 + (1 - 1/\alpha^2) \\ \widetilde{\omega}_i/\alpha^2, \quad i = 1, \ldots, N \end{cases}. \quad (19)$$

In fact, when the mean and covariance of a given nonlinear system are known, the optimal value of κ is a constant which cannot be randomly selected just according to recommendations. The value of κ is critical, which determines the weight of the center point. Therefore, κ should be properly selected first according to the properties of the specific nonlinear dynamical system, thereafter α can be arbitrarily selected to solve the non-local effect to satisfy (3), (15) and (19). A bearing-only tracking example was given in [11] to demonstrate that the recommendations of κ do not lead to the best results.

3.3 Problem 3

It should not be in doubt that the scaled UT (15) and (19) can preserve the first two moments for any set of sigma points.

The scaled UT (15) and (19) highlight that it allows any set of sigma points to be scaled by an arbitrary scaling factor and preserves the first two moments of the set [8]. The authors gave two examples to query this ability [17]. One example is an arbitrary set of sigma points as follows:

Suppose $\overline{\boldsymbol{\xi}} = [0]_{2\times 1}$ and $\boldsymbol{P} = \boldsymbol{I}_2$. The previous set of sigma points is

$$\widetilde{\boldsymbol{\chi}}_1 = [\sqrt{2}, 0]^{\mathrm{T}}, \quad \widetilde{\boldsymbol{\chi}}_2 = [0, \sqrt{2}]^{\mathrm{T}}, \quad \widetilde{\boldsymbol{\chi}}_3 = [-\sqrt{2}, 0]^{\mathrm{T}},$$

$\widetilde{\boldsymbol{\chi}}_4 = [0, -\sqrt{2}]^{\mathrm{T}}, \widetilde{\omega}_1 = \widetilde{\omega}_2 = \widetilde{\omega}_3 = \widetilde{\omega}_4 = 0.25.$

For $\alpha = 0.5$, the authors gave the mean $\overline{\boldsymbol{\xi}}' = [-\sqrt{2}, 0]^{\mathrm{T}}$ and covariance $\boldsymbol{P}' = \mathrm{diag}([-3, 1])^{\mathrm{T}}$ of the scaled sigma set directly without proof, which were not equal to the previous mean and covariance. The authors showed that the scaled UT cannot preserve the first two moments for an arbitrary sigma set according to the above results.

In fact, the authors obtained these incorrect results because they misunderstood the meaning of $\overline{\boldsymbol{\xi}}$ in (15). Julier pointed out that $\overline{\boldsymbol{\xi}}$ is the mean of the previous sigma point [8], whereas the authors thought $\overline{\boldsymbol{\xi}}$ was the first sigma point $\widetilde{\boldsymbol{\chi}}_1$ ($\widetilde{\boldsymbol{\chi}}'_i = \widetilde{\boldsymbol{\chi}}_1 + \alpha(\widetilde{\boldsymbol{\chi}}_i - \widetilde{\boldsymbol{\chi}}_1)$, see details in Table II of [17]). There we only give the mean calculation to show their mistake.

Proof The incorrect result in [17] is

$$\overline{\boldsymbol{\xi}}' = \sum_{i=1}^{4} \widetilde{\omega}'_i \widetilde{\boldsymbol{\chi}}'_i = \widetilde{\omega}'_1 \widetilde{\boldsymbol{\chi}}'_1 + \sum_{i=2}^{4} \widetilde{\omega}'_i \widetilde{\boldsymbol{\chi}}'_i =$$

$$(\widetilde{\omega}_1\alpha^{-2} + 1 - \alpha^{-2})\widetilde{\boldsymbol{\chi}}_1 + \sum_{i=2}^{4} \widetilde{\omega}_i \alpha^{-2}(\widetilde{\boldsymbol{\chi}}_1 + \alpha(\widetilde{\boldsymbol{\chi}}_i - \widetilde{\boldsymbol{\chi}}_1)) =$$

$$[(\widetilde{\omega}_1\alpha^{-2} + 1 - \alpha^{-2}) + 3(1-\alpha)\widetilde{\omega}_1\alpha^{-2}]\widetilde{\boldsymbol{\chi}}_1 +$$

$$\sum_{i=2}^{4} \widetilde{\omega}_i \alpha^{-1} \widetilde{\boldsymbol{\chi}}_i = \frac{\alpha - 3\widetilde{\omega}_1}{\alpha}\widetilde{\boldsymbol{\chi}}_1 + \frac{\widetilde{\omega}_1}{\alpha}(\widetilde{\boldsymbol{\chi}}_2 + \widetilde{\boldsymbol{\chi}}_3 + \widetilde{\boldsymbol{\chi}}_4) =$$

$$\frac{\alpha - 4\widetilde{\omega}_1}{\alpha}\widetilde{\boldsymbol{\chi}}_1 = \frac{\alpha - 1}{\alpha}\widetilde{\boldsymbol{\chi}}_1$$

substituting $\alpha = 0.5$, then $\overline{\boldsymbol{\xi}}' = [-\sqrt{2}, 0]^{\mathrm{T}}$.

In fact, the mean of the true scaled sigma set is $[0]_{2\times 1}$, which can hold for arbitrary α:

$$\widetilde{\boldsymbol{\chi}}'_i = \overline{\boldsymbol{\xi}} + \alpha(\widetilde{\boldsymbol{\chi}}_i - \overline{\boldsymbol{\xi}}) = \alpha\widetilde{\boldsymbol{\chi}}_i$$

$$\overline{\boldsymbol{\xi}}' = \sum_{i=0}^{4} \widetilde{\omega}'_i \widetilde{\boldsymbol{\chi}}'_i = \sum_{i=0}^{4} \widetilde{\omega}'_i \alpha\widetilde{\boldsymbol{\chi}}_i =$$

$$(\widetilde{\omega}_0\alpha^{-2} + 1 - \alpha^{-2})\overline{\boldsymbol{\xi}} + \sum_{i=1}^{4} \widetilde{\omega}_i \alpha^{-1} \widetilde{\boldsymbol{\chi}}_i = [0]_{2\times 1}$$

$$\square$$

The other example was given in [17] to show the scaled UT cannot preserve the covariance of the nonlinear system

(18) by using the UT algorithm proposed in [23]:

$$\begin{cases} \widetilde{\boldsymbol{\chi}}_0 = \overline{\boldsymbol{\xi}} \\ \widetilde{\boldsymbol{\chi}}_i = \overline{\boldsymbol{\xi}} + (\sqrt{n/(1-\widetilde{\omega}_0)\boldsymbol{P}})_i, \quad i = 1, \ldots, n \\ \widetilde{\boldsymbol{\chi}}_{i+n} = \overline{\boldsymbol{\xi}} - (\sqrt{n/(1-\widetilde{\omega}_0)\boldsymbol{P}})_i, \quad i = 1, \ldots, n \\ \widetilde{\omega}_i = (1-\widetilde{\omega}_0)/2n, \quad i = 1, \ldots, 2n \end{cases} \quad (20)$$

where $\widetilde{\omega}_0$ ($\widetilde{\omega}_0 < 1$) is the weight of the center point, which controls how the other sigma points will be repositioned. The symmetrical UKF is a special case of the UT (20), where $\widetilde{\omega}_0 = \kappa/(n+\kappa)$.

The authors give the transformed results $\overline{\zeta} = 3$, $P_{\zeta\zeta} = 3\alpha^2 - 8$ directly without proof (we cannot know how this result was derived). $P_{\zeta\zeta}$ is not equal to the true covariance ($P = 2n = 6$) of 3D chi-square distribution. The authors pointed out that this result showed two problems. First, the transformed covariance for this scaled UT is not matched up to the second order, but only to the first order. Second, the scaling factor modifies the covariance even for second order polynomial approximation. We will give the detail proof to show that $P_{\zeta\zeta}$ can obtain the true covariance with the scaled UKF if $\widetilde{\omega}_0$ and α are properly selected.

Theorem 2 For the specific nonlinear system (18), its mean $\overline{\zeta} = n$ and covariance $P_{\zeta\zeta} = n^2\widetilde{\omega}_0/(1-\widetilde{\omega}_0)$ with the UT (20), while the covariance is $P'_{\zeta\zeta} = n^2(\alpha^2 - 1 + \widetilde{\omega}_0)/(1-\widetilde{\omega}_0)$ with the UT (20) transformed by the scaled UT.

Proof For any given $\widetilde{\omega}_0 < 1$, using (20) to obtain $\{\widetilde{\boldsymbol{\chi}}_i, \widetilde{\omega}_i\}_{i=0}^{2n}$, then substituting them into (6)−(8), we can calculate mean $\overline{\zeta}$ and covariance $P_{\zeta\zeta}$ as follows:

$$\widetilde{\omega}_0 = \widetilde{\omega}_0$$

$$\widetilde{\omega}_i = (1-\widetilde{\omega}_0)/2n$$

$$\widetilde{\boldsymbol{\chi}}_0 = \boldsymbol{0}$$

$$\widetilde{\boldsymbol{\chi}}_i = \boldsymbol{0} + (\sqrt{(n/(1-\widetilde{\omega}_0))\boldsymbol{I}_n})_i = \sqrt{(n/(1-\widetilde{\omega}_0))}\boldsymbol{e}_i$$

$$\widetilde{\boldsymbol{\chi}}_{i+n} = -\widetilde{\boldsymbol{\chi}}_i, \quad i = 1, \ldots, n$$

$$h(\widetilde{\boldsymbol{\chi}}_0) = 0$$

$$h(\widetilde{\boldsymbol{\chi}}_i) = (\sqrt{(n/(1-\widetilde{\omega}_0))}\boldsymbol{e}_i)^{\mathrm{T}}(\cdot) = (n/(1-\widetilde{\omega}_0))$$

$$h(\widetilde{\boldsymbol{\chi}}_{i+n}) = (-\sqrt{(n/(1-\widetilde{\omega}_0))}\boldsymbol{e}_i)^{\mathrm{T}}(\cdot) = (n/(1-\widetilde{\omega}_0))$$

$$\overline{\zeta} = \sum_{i=0}^{2n} \widetilde{\omega}_i h(\widetilde{\boldsymbol{\chi}}_i) =$$

$$\widetilde{\omega}_0 h(\widetilde{\boldsymbol{\chi}}_0) + \frac{1-\widetilde{\omega}_0}{2n}\sum_{i=1}^{2n} h(\widetilde{\boldsymbol{\chi}}_i) =$$

$$\frac{1-\widetilde{\omega}_0}{2n} \cdot 2n \cdot \frac{n}{1-\widetilde{\omega}_0} = n$$

$$P_{\zeta\zeta} = \sum_{i=0}^{2n} \widetilde{\omega}_i (h(\widetilde{\boldsymbol{\chi}}_i) - \overline{\zeta})(\cdot)^{\mathrm{T}} =$$

$$\widetilde{\omega}_0 (h(\widetilde{\boldsymbol{\chi}}_0) - \overline{\zeta})(\cdot)^{\mathrm{T}} + \frac{1 - \widetilde{\omega}_0}{2n} \sum_{i=1}^{2n} (h(\widetilde{\boldsymbol{\chi}}_i) - \overline{\zeta})(\cdot)^{\mathrm{T}} =$$

$$\widetilde{\omega}_0 n^2 + \frac{1 - \widetilde{\omega}_0}{2n} \cdot 2n \cdot \left(\frac{n}{1 - \widetilde{\omega}_0} - n \right)^2 = \frac{n^2 \widetilde{\omega}_0}{1 - \widetilde{\omega}_0}.$$

$$\square$$

Likely, for any given α, using (15) and (19) to obtain the transformed $\{\widetilde{\boldsymbol{\chi}}'_i, \widetilde{\omega}'_i\}_{i=0}^{2n}$, then substituting them into (6)–(8), we can obtain $\overline{\zeta}' = n$ and $P'_{\zeta\zeta} = n^2(\alpha^2 - 1 + \widetilde{\omega}_0)/(1 - \widetilde{\omega}_0)$.

Therefore, when $n = 3$, we can get the right covariance of the nonlinear system (18) if we set $\widetilde{\omega}_0 = 0.4$. Then $\alpha = 1$ can also give the right covariance of the transformed sigma points.

3.4 Conclusions of the problems

From the above problems, we can draw two conclusions. First, the scaled UKF can preserve the first two moments for any set of sigma points as Julier declared. However, the "arbitrariness" of α must be conditioned on the appropriate selection of κ or $\widetilde{\omega}_0$. Second, for a given nonlinear dynamical system, it is not reasonable to predefine a set of constant scaling parameters during the entire UKF filtering process without considering particular system description, especially for the strong nonlinear systems.

4. Scaling parameters selection principle for the scaled UKF

The results shown in Table 1 indicate that, if κ is inappropriately selected first, then regardless of how small the value of α is, the correct covariance cannot be obtained. Therefore, deriving a self-adaptive scaling parameter α for a specific nonlinear system conditioned on the appropriate κ is better. The optimal value of α should be time-varying according to the time-varying covariance. β has only two choices, namely, $\beta = 0$ or $\beta = 2$, depending on whether the distribution is Gaussian or not. We propose a scaling parameter selection principle by the following procedures.

4.1 Estimating κ and β in the MMSE and the minimum distribution interval senses with Monte Carlo simulation

The state is estimated using the scaled UKF with $\alpha = 1$, $\beta = 0$, $\kappa = 0 : 1 : 3n$ and $\alpha = 1$, $\beta = 2$, $\kappa = 0 : 1 : 3n$. The maximum value of κ is set to $3n$ because κ determines the weight of the center point. When $\kappa = 3n$, $\widetilde{\omega}_0$ is 0.75, and the sum of the weights of the other $2n$ sigma points is only 0.25. The filter performance is measured in the mean

square error (MSE) sense with $M = 10^6$ Monte Carlo simulations [24], i.e.,

$$\mathrm{MSE} = \sum_{i=1}^{M} \sum_{k=0}^{N_k} (\boldsymbol{\xi}_k^i - \widehat{\boldsymbol{\xi}}_k^i)^2 / M N_k \tag{21}$$

where $\boldsymbol{\xi}_k^i$, $\widehat{\boldsymbol{\xi}}_k^i$ are the true and the estimated state vectors at time k in the ith Monte Carlo simulation. N_k is the number of sampling. In the case of two or more equal (or similar) MSEs, how do we select the appropriate κ? To solve this problem, suppose an MSE from one Monte Carlo simulation is a random variable. We define the covariance of the MSE to depict its distribution interval. From this point of view, MMSE is the mean of MSE variables. Therefore, we can select a set of scaling parameters $\alpha = 1$, $\beta = \beta_c$ and $\kappa = \kappa_c$ (where β_c and κ_c are known constants), which yield the MMSE. If there are two or more equal (or similar) MSEs, then we select κ_c which yields the minimum distribution interval.

4.2 Estimating the time-varying parameter α_k

The distance between the mean $\overline{\boldsymbol{\xi}}$ and the transformed sigma point $\widetilde{\boldsymbol{\chi}}_i$ is calculated as follows:

$$\delta(n, \alpha, \boldsymbol{P}) = [\alpha \sqrt{(n + \kappa_c)\boldsymbol{P}}]_i. \tag{22}$$

For a given nonlinear system, its dimension n is fixed, and the value of δ is determined by the values of α and $\boldsymbol{P}$. At each prediction and measurement update step, $\boldsymbol{P}$ is time-varying. Therefore, if α keeps constant, δ gets larger as $\boldsymbol{P}$ gets larger. Sometimes the larger δ does not satisfy the necessary condition of the Taylor series expansion, which results in the sampling non-local effects. Therefore, the scaling parameter α should be changed with the time-varying $\boldsymbol{P}$ at each step to keep the distance δ in a small neighborhood of the center point. Otherwise, the chosen sample points may not capture the posterior higher moments and introduce large errors.

Jiang proposed a self-adaptive scaled UT and applied it to the minimal skew simplex sampling [25]. The basic idea is to approximate the neighborhood of the true state with the estimated state $o(\widehat{\boldsymbol{\xi}}_k, \sqrt{\mathrm{trace}(\boldsymbol{P}_k)})$, which is the sampling range at step $k + 1$. The maximum distance between the scaled sigma point and the center point does not exceed $\sqrt{\mathrm{trace}(\boldsymbol{P}_k)}$.

At each $k - 1$ step, the sigma points are calculated with the current scaling parameters $\alpha = \alpha_{k-1}$, $\beta = \beta_c$, and $\kappa = \kappa_c$. When the updated covariance P_k is obtained, by means of Jiang's idea, we can calculate δ_k^{max} and α_k for the scaled UKF as follows:

$$\delta_k = [\sqrt{(n + \kappa_c)\boldsymbol{P}_k}]_i \tag{23}$$

$$\delta_k^{\max} = \max\{\delta_k(i,i), i \in \Omega\} \qquad (24)$$

$$\alpha_k = (\text{sqrt}(\text{tr}(\boldsymbol{P}_k)))/\delta_k^{\max} \qquad (25)$$

where Ω refers to all the sigma points.

4.3 Scaled UKF algorithm with the proposed scaling parameters selection principle

We name the scaled UKF algorithm with the proposed scaling parameters selection principle as UKF_G, and that with constant scaling parameters $\alpha = 1$, $\beta = \beta_c$ and $\kappa = \kappa_c$ as the default UKF. UKF_G can be computed as follows:

Step 1   Initialize the mean $\overline{\boldsymbol{\xi}} = \boldsymbol{\xi}_0$ and the covariance $\boldsymbol{P} = \boldsymbol{P}_0$.

Step 2   Estimate a set of scaling parameters $\alpha = 1$, $\beta = \beta_c$ and $\kappa = \kappa_c$ according to (16) and (21).

Step 3   Calculate the sigma points $\{\widetilde{\chi}_{ki}, \widetilde{\omega}_{ki}\}_{i=0}^{2n}$ at step k.

$$\begin{cases} \lambda_{k-1} = \alpha_{k-1}^2(n + \kappa_c) - n \\ \widetilde{\omega}_{k0}^m = \lambda_{k-1}/(n + \lambda_{k-1}) \\ \widetilde{\omega}_{k0}^c = \widetilde{\omega}_{k0}^m + (1 - \alpha_{k-1}^2 + \beta_c) \\ \widetilde{\omega}_{ki}^m = \widetilde{\omega}_{ki}^c = 1/2(n + \lambda_{k-1}) \end{cases} \qquad (26)$$

$$\widetilde{\chi}_{ki} = [\overline{\boldsymbol{\xi}}_{k-1}, \overline{\boldsymbol{\xi}}_{k-1} \pm (\alpha_{k-1}\sqrt{(n + \kappa_c)\boldsymbol{P}_{k-1}})_i] \qquad (27)$$

Step 4   Calculate the updated state mean $\boldsymbol{\xi}_k$ and covariance $\boldsymbol{P}_k$ according to (6)−(14).

Step 5   Calculate α_k according to (23)−(25).

Step 6   Let $k = k + 1$, repeat Step 3−Step 6.

Step 7   Compare the estimated covariance with that of the default UKF, select the smaller covariance and keep the associated mean as the final result.

5. Example and simulation results

5.1 UNGM example

Reconsider the UNGM example. Using Step 2, we can obtain the corresponding results from 10^6 Monte Carlo simulations, which are given in Table 2.

Table 2   MSE of different κ and β

κ	MSE ($\beta = 0$)	MSE ($\beta = 2$)
$\kappa = 0$	69.905 0	78.462 3
$\kappa = 1$	35.363 5	56.641 5
$\kappa = 2$	41.835 9	65.735 2
$\kappa = 3$	42.032 8	66.355 7

Obviously, $\alpha = 1$, $\beta = 0$ and $\kappa = 1$ yield the MMSE. Therefore, we select the set of scaling parameters $\alpha = 1$, $\beta_c = 0$ and $\kappa_c = 1$, and use them to estimate the state with UKF_G.

The default UKF simulation result is given in Fig. 3. Although Fig. 3 illustrates a better performance than Fig. 2, at steps $67-74$, $80-83$, the default UKF still has difficulties

in following the correct model, even giving the incorrect, reversed estimated state.

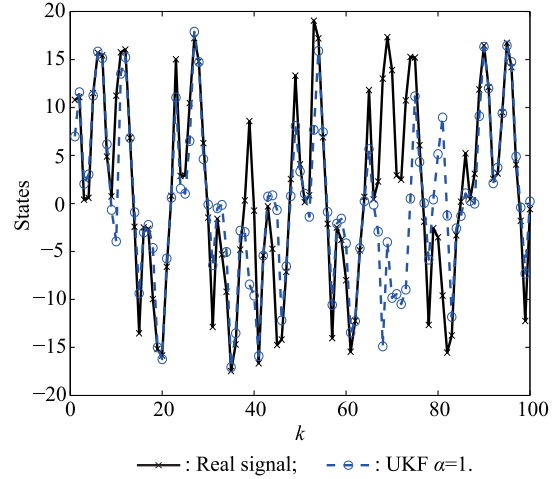—×—: Real signal;   – ⊖ –: UKF α=1.

Fig. 3   Default UKF result with constant scaling parameters $\alpha = 1, \beta_c = 0, \kappa_c = 1$

The simulation result of UKF_G is given in Fig. 4. UKF_G can track the correct state at almost every step and exhibits a better performance than the default UKF. The superiority of UKF_G is shown clearly in steps $67-74$ and $80-83$ which change the incorrect, reversed estimated state to the nearly right states.
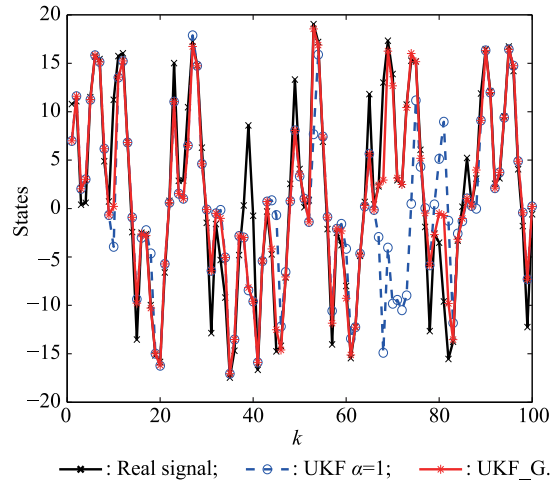

—×—: Real signal;   – ⊖ –: UKF α=1;   —✳—: UKF_G.

Fig. 4   UKF_G result

The time-varying parameter α_k demonstrates the good performance of UKF_G, as shown in Fig. 5. Because the performance of the UT is analyzed based on the multi-dimensional Taylor series, the transformed mean $\overline{\zeta'}$ and the non-transformed one $\overline{\zeta}$ only agree up to the second order if $\alpha_k^2 = \mu$. α_k has an effect on the rest high order terms as follows [8,16].

$$\overline{z}' = \psi(\overline{\xi}) + \frac{1}{2}\nabla^2\psi\frac{\alpha_k^2}{\mu}P_x + \frac{1}{3!}\nabla^3\psi\frac{\alpha_k^3}{\mu}E[\delta\xi^3] + \cdots \quad (27)$$
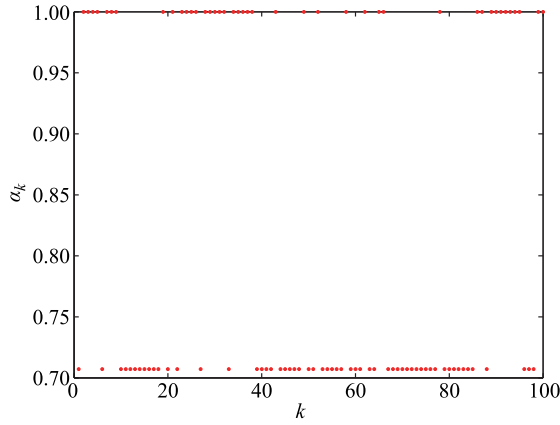
Fig. 5   Time-varying scaling parameter α_k

The results shown in Fig. 4 and Fig. 5 indicate that $\alpha_k < 1$ perfectly solves the non-local effects caused by the constant scaling parameters. In this example, α_k has only one time-varying value because the UNGM model is one-dimensional. Thus, its time-varying property is not evident in this example.

5.2   Multi-dimensional strong nonlinear example

Consider a 3D strong nonlinear model as follows:

$$\boldsymbol{\xi}_k = \begin{bmatrix} 3\sin(2\xi_{2,k-1}) \\ \xi_{1,k-1} + \exp^{-0.05\xi_{3,k-1}} + 10 \\ \dfrac{\xi_{1,k-1}(\xi_{2,k-1} + \xi_{3,k-1})}{5} \end{bmatrix} + \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}\boldsymbol{\tau}_{k-1}$$

$$(28)$$

$$\zeta_k = \cos(2\xi_{1,k}) + \xi_{2,k}\xi_{3,k} + \mu_k \qquad (29)$$

where $\tau = 0.3$, $\Gamma = 0.7$, $\mu = 0.5$, $R = 1$. The true and estimated initial states are $\boldsymbol{\xi}_0 = [-0.7, 1, 1]^{\mathrm{T}}$ and $\overline{\boldsymbol{\xi}}_0 = [0, 0, 0]^{\mathrm{T}}$ with $\boldsymbol{P}_0 = \boldsymbol{I}_3$. This example includes trigonometric and exponent functions, whose higher-order derivatives cannot be ignored. Thus, the poorest results are likely produced in the worst cases.

The initial scaling parameters are estimated by Step 2 with 10^6 Monte Carlo simulations ($\alpha = 1$), and the results are listed in Table 3.

Table 3   MSE of different κ and β

k	MSE ($\beta = 0$)	MSE ($\beta = 2$)
$\kappa = 0$	56.419 8	37.779 9
$\kappa = 1$	34.401 7	29.880 9
$\kappa = 2$	62.194 3	31.618 3
$\kappa = 3$	43.622 8	42.168 3
$\kappa = 4$	100.988 8	46.790 2
$\kappa = 5$	40.782 8	32.233 6
$\kappa = 6$	35.781 1	29.792 4
$\kappa = 7$	50.805 8	24.477 3
$\kappa = 8$	44.785 1	32.977 5
$\kappa = 9$	36.406 1	24.003 0

Two notable results are shown in Table 3. First, $\beta = 2$ is better for this system. Second, $\kappa = 7$ and $\kappa = 9$ yield

the similar MSE. We utilize 10^6 MSE variables from the Monte Carlo simulations of $\kappa = 0 : 9$. Their distributions are illustrated in Fig.6. As can be seen, $\kappa = 9$ yields the MMSE distribution interval. Therefore, we select $\beta_c = 2$ and $\kappa_c = 9$. The recommended $\kappa = 0$ does not give the MMSE with either $\beta = 0$ or $\beta = 2$.
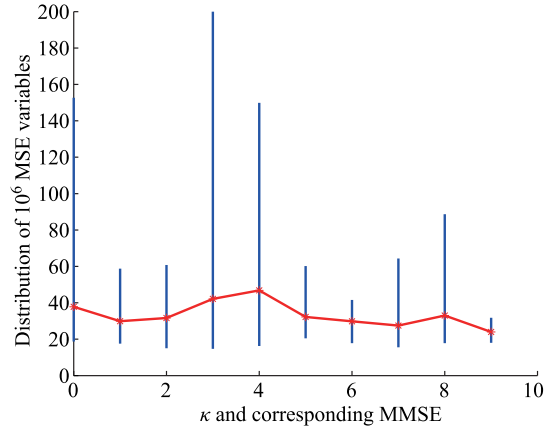

Fig. 6 κ and the corresponding MMSE

The estimation quality of $\boldsymbol{\xi}$ with the default UKF and UKF_G are compared in Fig. 7. Obviously, UKF_G exhibits a better performance than the default UKF. In particular, UKF_G can track the true states of $\xi_{2,k}$ and $\xi_{3,k}$ at almost every step. The estimated quality of $\xi_{1,k}$ with UKF_G is not as perfect as those of $\xi_{2,k}$ and $\xi_{3,k}$, but still produces a significantly improved performance than that with the default UKF. This example shows that UKF_G can effectively estimate the strong nonlinear systems.
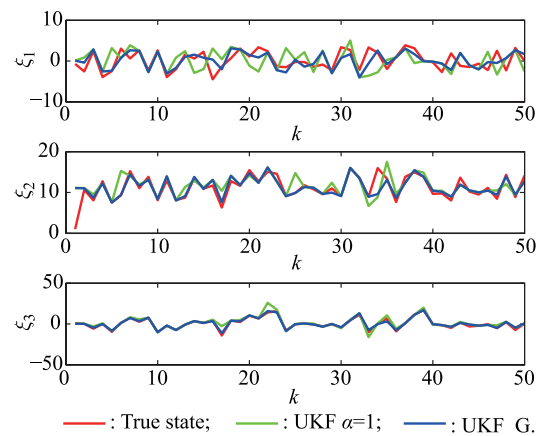

—— : True state;   —— : UKF $\alpha = 1$;   —— : UKF_G.

Fig. 7   Simulation comparison results of $\boldsymbol{\xi}$ with the default UKF and UKF_G

The time-varying parameter α_k is demonstrated in Fig. 8. Its time-varying property is more evident than in Fig. 5. The results shown in Fig. 7 and Fig. 8 indicate that, for $\xi_{2,k}$, the time-varying parameters $\alpha_6 = 0.361$,

$\alpha_{34} = 0.315\,6$ and $\alpha_{35} = 0.49$ correct the large errors of the default UKF at Steps 6, 34 and 35, which demonstrate the superior function of the time-varying parameter.
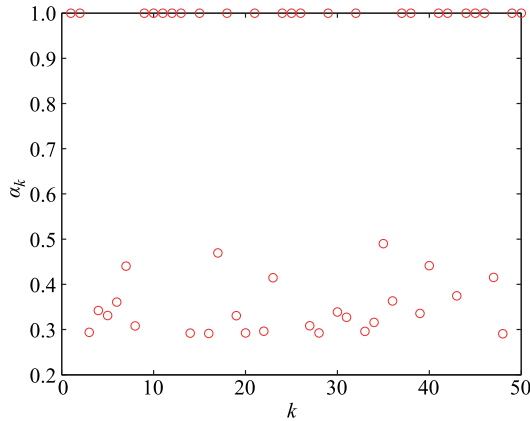

Fig. 8   Time-varying scaling parameter α_k

5.3   Data comparison of the computational cost and estimated performance

Table 4 lists the MSE and computational cost of the above-mentioned examples with 100 Monte Carlo simulations.

Table 4   Data comparison

Example	Parameter	Default UKF	UKF_G
Example1	MSE	38.951 8	8.473 2
	time/s	0.047	0.092
Example2	MSE	22.859 6	9.843 5
	time/s	0.072	0.106

Although the computational cost of UKF_G is approximately two times of the default UKF because it needs to compute two UKF filtering procedures, the computation is worth doing to significantly improve the estimation performance in the MSE sense. Moreover, we can compute the estimated state with the default UKF, keep the results, and use the known default results to compute UKF_G to reduce computational cost. Thus, UKF_G will only cost slightly more time than the default UKF to compute the time-varying parameter α_k at each step. From this point of view, UKF_G is an effective nonlinear filter than the default UKF, which can realize on-line adjustment.

6. Conclusions

The widely used scaling parameters recommendations are discussed and several problems are highlighted in this paper. On the basis of analyses, a scaling parameter selection principle is proposed. A self-adaptive scaling parameter selection algorithm for the scaled UKF is derived according to the above principle, which is named UKF_G. The performance of UKF_G is demonstrated by two typical strong nonlinear models. Simulation results show that UKF_G provides significantly higher quality in the MSE sense than UKF with constant parameters, which demonstrate its effectiveness. Experimental procedure also indicates that this algorithm is simple and understandable. The main objective of this paper is to propose an easily applicable and effective scaled UKF in engineering. Therefore, the proposed algorithm (UKF_G) is not compared in terms of performance with other works [10–15] which are based on different performance criteria.

References

[1] JULIER S J, UHLMANN J K, DURRANT-WHYTE H F. A new approach for filtering nonlinear systems. Proc. of the IEEE American Control Conference, 1995: 1628–1632.

[2] JULIER S J, UHLMANN J K, DURRANT-WHYTE H F. A new method for the nonlinear transformation of means and covariances in filters and estimators. IEEE Trans. on Automatic Control, 2000, 45(3): 477–482.

[3] JAYASIRI A, NANDAN A, IMTIAZ S, et al. Dynamic positioning of vessels using a UKF-based observer and an NMPC-based controller. IEEE Trans. on Automation Science and Engineering, 2017, 14(4): 1778–1785.

[4] KANDEPU R, FOSS B, IMSLAND L. Applying the unscented Kalman filter for nonlinear state estimation. Journal of Process Control, 2008, 18(7): 753–768.

[5] QIAN H, LIU K, LI L, et al. Widely linear UKF constant modulus algorithm for blind adaptive beam forming. Journal of Systems Engineering and Electronics, 2017, 28(3): 413–423.

[6] JULIER S J, UHLMANN J K. Reduced sigma point filters for propagation of means and covariances through nonlinear transformations. Proc. of the American Control Conference, 2002: 887–892.

[7] JULIER S J. The spherical simplex unscented transformation. Proc. of the American Control Conference, 2003: 2430–2434.

[8] JULIER S J, UHLMANN J K. The scaled unscented transformation (author's comments). IEEE Trans. on Automatic Control, 2002, 47: 1408–1409.

[9] HARTIKAINEN J, SOLIN A, SARKKA S. Optimal filtering with Kalman filters and smoothers – a manual for the matlab toolbox EKF/UKF. Greater Helsinki Finland: Aato University, 2011.

[10] SAKAI A, KURODA Y. Discriminative parameter training of unscented Kalman filter. IFAC Proceedings Volumes, 2010, 43(18): 677–682.

[11] DUNIK J, SIMANDL M, STRAKA O. Adaptive choice of scaling parameter in derivative-free local filters. IEEE Information Fusion, 2010: 1–8.

[12] DUNIK J, SIMANDL M, STRAKA O. Unscented Kalman filter: aspects and adaptive setting of scaling parameter. IEEE Trans. on Automatic Control, 2012, 57(9): 2411–2416.

[13] STRAKA O, DUNIK J, SIMANDL M. Scaling parameter in unscented transform: analysis and specification. Proc. of the American Control Conference, 2012: 5550–5555.

[14] STRAKA O, DUNIK J, SIMANDL M. Unscented Kalman filter with advanced adaptation of scaling parameter. Automatica, 2014, 50(10): 2657–2664.

[15] SCARDUA L A, CRUZ J J D. Complete offline tuning of the unscented Kalman filter. Automatica, 2017, 80: 54–61.

[16] NIE Y, ZHANG T. A self-adaptive scaling parameter selection algorithm for the unscented Kalman filter. Proc. of the Chinese Automation Congress, 2016: 86–90.

[17] MENEGAZ H M T, ISHIHARA J Y, BORGES G A, et al. A systematization of the unscented Kalman filter theory. IEEE Trans. on Automatic Control, 2015, 60(10): 2583 – 2598.

[18] GUSTAFSSON F, HENDEBY G. Some relations between extended and unscented Kalman filters. IEEE Trans. on Signal Processing, 2012, 60(2): 545 – 555.

[19] ROTH M, HENDEBY G, GUSTAFSSON F. Nonlinear Kalman filters explained: a tutorial on moment computations and sigma point methods. Journal of Advances in Information Fusion, 2016, 11(1): 47 – 70.

[20] JIA B, XIN M, CHENG Y. High-degree cubature Kalman filter. Automatica, 2013, 49(2): 510 – 518.

[21] SARKKA S, HARTIKAINEN J, SVENSSON L. On the relation between Gaussian process quadratures and sigma-point methods. Statistics, 2015: arXiv: 1504.05994.

[22] LIU Y, YU A X, ZHU J B, et al. Unscented Kalman filtering in the additive noise case. Science China Technological Sciences, 2010, 53(4): 929 – 941.

[23] JULIER S J, UHLMANN J K. Unscented filtering and nonlinear estimation. Proceeding of the IEEE, 2004, 92(3): 401 – 422.

[24] SIMANDL M, DUNÍK J. Derivative-free estimation methods: new results and performance analysis. Automatica, 2009, 45(7): 1749 – 1757.

[25] WANG J. Self-adaptive selecting method for scaling factor of scaled unscented transformation. Chinese Space Science & Technology, 2008, 28(3): 1 – 6, 28. (in Chinese)

Biographies
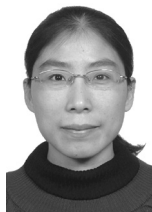
NIE Yongfang was born in 1976. She received her M.S degree in weapon system engineering from Naval Aeronautical Engineering Institute in 2002. She was a lecturer with Naval Submarine Academy from 2002 to 2014. Now, she is also a Ph.D. student in control science and engineering at Tsinghua University. Her current research activity focuses on simultaneous localization and mapping for autonomous underwater vehicles.
E-mail: nyf14@mails.tsinghua.edu.cn

ZHANG Tao was born in 1969. He received his B.S. degree, M.S. degree and Ph.D. degree from Tsinghua University, Beijing, China, in 1993, 1995 and 1999 respectively. He received his second Ph.D. degree from Saga University, Saga, Japan, in 2002. He is currently a professor and the deputy head of the Department of Automation, School of Information Science and Technology, Tsinghua University, Beijing, China. He is the author or coauthor of more than 200 papers and three books. His current research includes robotics, control theory, artificial intelligence, navigation and control of spacecraft, fault diagnosis and reliability analysis, body signal extraction and recognition.
E-mail: taozhang@mail.tsinghua.edu.cn