Study on initial value problem for fractional-order cubature Kalman filters of nonlinear continuous-time fractional-order systems

Chuang Yang · Zhe Gao · Yue Miao · Tao Kan

Received: 18 January 2021 / Accepted: 2 July 2021 / Published online: 25 July 2021
© The Author(s), under exclusive licence to Springer Nature B.V. 2021

Abstract To realize the state estimation of nonlinear continuous-time fractional-order systems, two types of fractional-order cubature Kalman filters are designed to solve problem on the initial value influence. For the first type of cubature Kalman filter (CKF), the initial value of the estimated system is also regarded as the augmented state, and the augmented state equation is constructed to obtain the CKF based on Grünwald–Letnikov difference. For the second type of CKF, the fractional-order hybrid extended-cubature Kalman filter is proposed to weaken the influence of initial value by the first-order Taylor expansion and the third-order spherical-radial rule. These two methods can reduce the influence of initial value on the state estimation effectively. Finally, the effectiveness of the proposed CKFs is verified by two simulation examples.

Keywords Nonlinear fractional-order systems · Extended Kalman filter · Cubature Kalman filter · State estimation · Initial value compensation

1 Introduction

In recent decades, fractional-order systems (FOSs) has been studied by many scholars, because fractional-order calculus can more accurately describe the phenomena in real-world systems [1–3]. The fractional-order calculus is applied to deal with the fractional-order integral and derivative, and it is widely used in the network system [4], the heat conduction [5], the material science [6] and so on. The fractional-order controller is a common application of fractional-order calculus [7]. Due to the simple structure of fractional-order PID controller, the fractional-order PID controller is used in most industrial control systems [8,9]. With the introduction of fractional-order operators, other types of fractional-order controllers appear, such as the fractional-order sliding mode controller [10], the fractional-order back-stepping controller [11], the fractional-order model predictive controller [12] and the fractional-order extremum seeking controller [13]. In practical control systems, the measurement signal inevitably contains certain noise disturbances, and some state information may not be obtained directly. Then robust state observers for FOSs need to be designed.

As one kind of robust state observers [14], the Bayesian filtering can achieve the state estimation of nonlinear systems. The key of solving the nonlinear filtering problem is to obtain the posterior probability distribution of the state information. Thus, the prob-
The problem of nonlinear filtering can be solved in the framework of Bayesian filtering [15]. As a special case of Bayesian filter, Kalman filter (KF) has become an increasing common tool to estimate the state information of integer-order systems (IOSs). The extended Kalman filter (EKF) is a very effective method to deal with nonlinear filtering problem based on the first-order Taylor expansion [16]. EKF has been widely used in the fields of the state-of-charge estimation of lithium-ion battery [17], the multi-sensor tracking of maneuvering targets [18] and the current transformer saturation [19]. However, the state estimation accuracy of EKF linearization method is not ideal, sometimes. If some nonlinear functions are not differentiable, EKF cannot effectively estimate the state information, which is also the limitation of EKF [20]. Hence, other types of KFs have been designed to deal with nondifferentiable nonlinear functions, such as unscented Kalman filter (UKF) [21–23] and cubature Kalman filter (CKF) [24–26]. The UKF uses the unscented transformation to cope with nonlinear functions in FOSs, then more accurate state information is obtained by comparing with EKF [27].

Meanwhile, fractional-order Kalman filters (FOKFs) are still appropriate to estimate state information for linear and nonlinear FOSs. The FOKFs can obtain more satisfactory state estimation than that of integer-order KFs. The fractional-order EKFs were proposed for linear FOSs with white Gaussian noises and colored noises in [28,29], respectively. In [30], KFs for linear and nonlinear discrete-time FOSs were extended to estimate the parameters or order of FOSs. Also, linear and nonlinear discrete-time FOSs were mentioned in [30]. In [31], the improved fractional-order UKF for nonlinear FOSs was designed to estimate more accurate state information. Besides, CKF can also estimate effectively state information for nonlinear FOSs. In [32], the CKF method was applied to obtain the higher estimation accuracy compared with EKF method for nonlinear continuous-time FOSs containing white Gaussian noises. Meanwhile, hybrid FOKFs for nonlinear FOSs were also offered to enhance the estimation accuracy in [33,34].

For a linear FOS, the initial value performs a significant impact on the state estimation and fractional-order identification if the fractional-order of the system is relatively small in the interval (0, 1), then an initial value compensation (IVC) is achieved by means of an augmented vector. The EKF with IVC for linear continuous-time FOSs with an unknown fractional-order was proposed in [35]. In this paper, the CKF for nonlinear FOSs with the IVC is investigated to obtain more effective state estimation. This paper introduces the Grünwald–Letnikov (G–L) difference method into two types of fractional-order cubature Kalman filters (FOCKFs) to obtain the effective state estimation of a nonlinear FOS with IVC.

The main contributions of this paper are generalized as follows: (1) The state estimation problem of nonlinear FOSs is investigated. (2) For a relatively small fractional-order in (0, 1), the augmented vector method is used to achieve the IVC. (3) The influence of initial value on state estimation is weakened by the model transformation method. (4) The FOCKF with the IVC and the fractional-order hybrid extended-cubature Kalman filter (HECKF) for nonlinear continuous-time FOSs is proposed to achieve the higher accuracy state estimation to compare with the CKF without the IVC in [32].

The remaining of this paper is arranged as follows. In Sect. 2, we introduce the definitions of fractional-order derivatives, the third-order spherical-radial rule and cubature points. We design the discretized equation of a nonlinear FOS, the FOCKF and HECKF to solve initial value problems in Sect. 3, respectively. In Sect. 4, the simulation examples are given to testify the validity of the proposed FOCKF and HECKF with the IVC. Section 5 summarizes this paper.

2 Preliminaries

2.1 Definitions of fractional-order derivatives

In the subsection, some definitions for fractional-order derivatives are introduced to describe dynamic characteristics of nonlinear FOSs. Considering the consistency of initial conditions of FOS, the Caputo definition has an increasingly wide application in real-world systems. Therefore, the establishment of nonlinear FOSs is achieved based on the Caputo definition and the IVC is considered, and the Riemann–Liouville (R–L) definition is also used in this study.

Definition 1 The R–L definition [36] with \( \alpha \)-order derivative for \( \varepsilon(t) \) is defined as

\[
\mathring{\mathring{D}_{t}^{\alpha}} \varepsilon(t) = \frac{1}{\Gamma(n-\alpha)} \frac{d^{n}}{dt^{n}} \left( \int_{0}^{t} (t-\mu)^{n-\alpha-1} \varepsilon(\mu) d\mu \right),
\]
where \( D_0^\alpha \varepsilon(t) \) represents \( \alpha \)-order derivative based on R–L definition, \( \alpha \in \mathbb{R}^+ \) is the fractional-order and satisfies \( n - 1 \leq \alpha < n \), \( n \in \mathbb{Z}^+ \), \( \mathbb{R}^+ \) and \( \mathbb{Z}^+ \) are the set of positive numbers and the set of positive integer numbers, respectively, and \( \Gamma(\cdot) \) is the Gamma function.

**Definition 2** The Caputo definition [36] with \( \alpha \)-order derivative for \( \varepsilon(t) \) is defined as

\[
C_0^\alpha D_t^\alpha \varepsilon(t) = \frac{1}{\Gamma(n-\alpha)} \int_0^t (t-\mu)^{n-\alpha-1} \varepsilon^{(n)}(\mu) d\mu,
\]

where \( C_0^\alpha D_t^\alpha \varepsilon(t) \) is \( \alpha \)-order derivative under Caputo definition, \( n - 1 < \alpha \leq n \) and \( n \in \mathbb{Z}^+ \).

It is obvious that the main difference between Definitions 1 and 2 is the order of differential operation and integral operation. In addition, the relationship between Definitions 1 and 2 is described by [37]

\[
C_0^\alpha D_t^\alpha \varepsilon(t) = R^\alpha D_t^\alpha \varepsilon(t) - \sum_{r=0}^{n-1} \frac{\varepsilon^{(r)}(0)}{\Gamma(r-\alpha+1)} t^{r-\alpha}.
\]

**Definition 3** The G–L difference [38] for \( \alpha \)-order is defined by

\[
\Delta^\alpha \varepsilon(r) = \frac{1}{T^\alpha} \sum_{m=0}^{r} (-1)^m b_m^\alpha \varepsilon(r - m),
\]

where \( \Delta^\alpha \) is the \( \alpha \)-order difference operation, \( T \) is the sampling period, and the coefficient \( b_m^\alpha \) is represented as follows

\[
b_m^\alpha = \begin{pmatrix} \alpha \\ m \end{pmatrix} = \begin{cases} 1, & m = 0 \\ \alpha(\alpha-1) \cdots (\alpha-m+1) / m!, & m \geq 1. \end{cases}
\]

Since \( \varepsilon(r - m) \approx 0 \) for \( r > m \), the R–L definition is approximately equivalent to G–L difference at \( t = rT \), then

\[
R^\alpha D_t^\alpha \varepsilon(t) \approx \Delta^\alpha \varepsilon(r) = \frac{1}{T^\alpha} \sum_{m=0}^{r} (-1)^m b_m^\alpha \varepsilon(r - m).
\]

For \( \alpha \in (0, 1) \), we take \( n = 1 \) in (1), then it follows that

\[
C_0^\alpha D_t^\alpha \varepsilon(t) \approx \frac{1}{T^\alpha} \sum_{m=0}^{r} (-1)^m b_m^\alpha \varepsilon(r - m) - \frac{A(r)}{T^\alpha} \varepsilon(0),
\]

where \( A(r) = r^{-\alpha} / \Gamma(1-\alpha) \).

To describe the impact of the initial value \( \varepsilon(0) \), we define \( \rho(r) = |A(r)| \). Figure 1 shows the relationship between \( \rho(r) \) and the iteration \( r \).

From Fig. 1, it is seen that \( \rho(r) \) decreases gradually with the increase of \( r \). If the fractional-order \( \alpha \) becomes closer to 1, \( \rho(r) \) decays faster with an increasing of \( r \). An decreasing fractional-order has a larger influence on the initial value. Therefore, the IVC can estimate more accurate state information for the FOS with small fractional-order in \((0, 1)\).

### 2.2 Third-order spherical-radial rule and construction of cubature points

The EKF algorithm uses Taylor series to expand the nonlinear function and retains the first-order term to realize the linearization of a nonlinear function. Other high-order terms are ignored, then the estimation accuracy becomes greatly reduced. Besides, the CKF algorithm is available if the nonlinear function is nondifferentiable. Due to the simple structure of EKF algorithm, it is suitable for systems with differentiable nonlinear functions. The CKF algorithm is performed based on the third-order spherical-radial rule via the cubature points to deal with the nonlinear function with a higher precision and suitable for the nondifferentiable nonlinear function.

The vector \( \varepsilon \in \mathbb{R}^n \) satisfies the Gaussian distribution, namely \( \varepsilon \sim \mathcal{N}(\bar{\varepsilon}; M) \), where \( \bar{\varepsilon} \) represents the mathematical expectation of \( \varepsilon \), the covariance matrix of \( \varepsilon \) is represented by \( M \) with the following transformation \( \varepsilon = \hat{\varepsilon} + \sqrt{M} \gamma \), where \( M = \sqrt{M} \) and \( \gamma \).

The nonlinear function \( l(\varepsilon) \) satisfies the following equation

![Fig. 1 Curves of \( \rho(r) \) for different fractional-orders \( \alpha \)](image-url)
\[ \int_{\mathbb{R}^n} l(\varepsilon)\mathcal{N}(\varepsilon; \hat{\varepsilon}, M)\,d\varepsilon = \int_{\mathbb{R}^n} l(\hat{\varepsilon} + \sqrt{M}\gamma)\mathcal{N}(\gamma; 0, I)\,d\gamma. \] (4)

The 2n cubature points are gained by performing the transformation \( \chi^{(i)} = \hat{\varepsilon} + \sqrt{M}\gamma^{(i)} \), where

\[ \gamma^{(i)} = \begin{cases} \sqrt{n}e_i, & i = 1, 2, \ldots, n \\ -\sqrt{n}e_{i-n}, & i = n + 1, n + 2, \ldots, 2n \end{cases} \] (5)

and \( e_i \in \mathbb{R}^n \) is the unit vector on the \( i \)th axis.

Based on the third-degree spherical-radial rule in [24], we use the 2n cubature points defined in (5), and the integral \( \int_{\mathbb{R}^n} l(\hat{\varepsilon} + \sqrt{M}\gamma)\mathcal{N}(\gamma; 0, I)\,d\gamma \) in (4) is approximately described by

\[ \int_{\mathbb{R}^n} l(\hat{\varepsilon} + \sqrt{M}\gamma)\mathcal{N}(\gamma; 0, I)\,d\gamma \approx \frac{1}{2n} \sum_{i=1}^{2n} l(\chi^{(i)}). \] (6)

Combining Eq. (6) with Eq. (4), the following equation holds

\[ \int_{\mathbb{R}^n} l(\varepsilon)\mathcal{N}(\varepsilon; \hat{\varepsilon}, M)\,d\varepsilon \approx \frac{1}{2n} \sum_{i=1}^{2n} l(\chi^{(i)}). \] (7)

3 Main results

3.1 Fractional-order cubature Kalman filter with IVC

by using augmented vector method

Consider the following nonlinear continuous-time FOS as

\[ C_0 D_0^\alpha \varepsilon(t) = \varphi(\varepsilon(t), u(t)) + Gw(t), \] (8)

\[ \eta(t) = y(\varepsilon(t)) + v(t), \] (9)

where \( \varphi(\varepsilon(t), u(t)) \) and \( y(\varepsilon(t)) \) represent the nonlinear functions, \( \alpha \in (0, 1) \) is fractional-order, \( \varepsilon(t) \in \mathbb{R}^n \), \( u(t) \in \mathbb{R}^p \) and \( \eta(t) \in \mathbb{R}^q \) represent the state vector, the input and the measurement output, respectively. \( w(t) \in \mathbb{R}^n \) and \( v(t) \in \mathbb{R}^q \) are process and measurement noises, and satisfy \( E[w(r)] = 0 \), \( E[v(r)] = 0 \), \( E[w(r)w^T(s)] = Q\delta(r-s) \) and \( E[v(r)v^T(s)] = R\delta(r-s) \), with \( E[w(r)v^T(s)] = 0 \), \( w(r) \) and \( v(r) \) are the sampling values at \( t = rT \) with the sampling period \( T \), \( E[\cdot] \) returns the mathematical expectation. \( \xi(r-s) \) fulfills \( \xi(r-s) = 0 \) for \( r \neq s \) and \( \xi(r-s) = 1 \) for \( r = s \). \( \varepsilon(r), u(r) \) and \( \eta(r) \) are denoted as the sampling values of \( \varepsilon(t), u(t) \) and \( \eta(t) \) at \( t = rT \), and \( 0 \) is a zero vector or a zero matrix with the appropriate dimension.

Combining Eqs. (3) with (8), we obtain

\[ \frac{1}{T^\alpha} \sum_{m=0}^{r} (-1)^m b_m^\alpha \varepsilon(r-m) - \frac{A(r)}{T^\alpha} \varepsilon(0) = \varphi(\varepsilon(r-1), u(r-1)) + Gw(r-1). \]

It follows that

\[ \varepsilon(r) = \sum_{m=1}^{r} B_m \varepsilon(r-m) \]

\[ + T^\alpha \varphi(\varepsilon(r-1), u(r-1)) \]

\[ + T^\alpha Gw(r-1) + A(r)\varepsilon(0), \] (10)

where \( B_m = (-1)^{m+1} b_m^\alpha \).

The vector \( \tilde{\varepsilon}(r) \) defined as the initial state \( \varepsilon(0) \) at the \( r \)th iteration satisfies

\[ \tilde{\varepsilon}(r) = \tilde{\varepsilon}(r-1) + \tilde{w}(r-1), \] (11)

where \( \tilde{w}(r) \) represents the white Gaussian noise and satisfies \( E[\tilde{w}(r)] = 0 \), \( E[\tilde{w}(r)\tilde{w}^T(s)] = Q_1\delta(r-s) \), \( Q_1 \) is a positive definite matrix with a very small norm.

In order to reduce the influence of initial value on a FOS, we define the augmented vector \( \zeta(r) = [\varepsilon^T(r), \tilde{\varepsilon}^T(r)]^T \), the augmented noise \( \tilde{w}(r) = [w^T(r), \tilde{w}^T(r)]^T \) with \( E[\tilde{w}(r)\tilde{w}^T(r)] = Q_2 = \begin{bmatrix} 0 & 0 \\ 0 & Q_1 \end{bmatrix} \) and \( E[\tilde{w}(r)v^T(s)] = 0 \).

Based on Eqs. (10) and (11), the corresponding augmented equation is given as follows

\[ \zeta(r) = \tilde{\varphi}(\zeta(r-1), u(r-1)) + \sum_{m=1}^{r} \tilde{B}_m(\zeta(r-m)) \]

\[ + G_\alpha \tilde{w}(r-1), \] (12)

where the nonlinear function \( \tilde{\varphi}(\zeta(r-1), u(r-1)) = \begin{bmatrix} T^\alpha \varphi(\varepsilon(r-1), u(r-1)) \\ 0 \end{bmatrix} \), \( \tilde{B}_m(\zeta(r-m)) = \begin{bmatrix} B_1 & A(r)I \\ 0 & I \end{bmatrix} \), \( m = 1 \), \( G_\alpha = \begin{bmatrix} T^\alpha G & 0 \\ 0 & I \end{bmatrix} \) and \( I \) is an identity matrix with the appropriate dimension.

A CKF algorithm with the IVC is proposed to achieve more accuracy estimation effect of nonlinear FOS (8). The prediction and estimation of \( \zeta(r) \) are defined as \( \hat{\zeta}(r|r-1) = E[\zeta(r)|\lambda(r-1)] \) and \( \hat{\zeta}(r|r) = E[\zeta(r)|\lambda(r)] \), where \( \lambda(r) \) covers \( \eta(0), \eta(1), \ldots, \eta(r) \), \( u(0), u(1), \ldots, u(r) \).

Assuming \( E[\zeta(r-m)|\lambda(r-1)] \cong E[\zeta(r-m)|\lambda(r-m)] \), the prediction \( \hat{\zeta}(r|r-1) \) is obtained on the basis of difference Eq. (12) as follows
\[
\hat{\zeta}(r| r-1) = E[\psi(\zeta(r-1), u(r-1)) + \sum_{m=1}^{r} B_m(r)\zeta(r-m) + G_\alpha\varpi(r-1)|\lambda(r-1)] = E[\psi(\zeta(r-1), u(r-1))|\lambda(r-1)] + \sum_{m=1}^{r} B_m(r)\hat{\zeta}(r-m|r-m).
\]  

(13)

The probability density \( p(\zeta(r-1)|\lambda(r-1)) \) is supposed to satisfy the Gaussian distribution, that is

\[
p(\zeta(r-1)|\lambda(r-1)) = \mathcal{N}(\zeta(r-1), \hat{\zeta}(r-1|r-1), M(r-1|r-1)).
\]

Therefore, the prediction \( \hat{\zeta}(r| r-1) \) is derived by

\[
\hat{\zeta}(r| r-1) = \int_{\mathbb{R}^n} \psi(\zeta(r-1), u(r-1))\mathcal{N}(\zeta(r-1); \hat{\zeta}(r-1|r-1), M(r-1|r-1))d\zeta(r-1) + \sum_{m=1}^{r} B_m(r)\hat{\zeta}(r-m|r-m).
\]

The nonlinear function \( \hat{\mathcal{Y}}(\hat{\zeta}(r-1|r-1), u(r-1)) \) is defined as

\[
\hat{\mathcal{Y}}(\hat{\zeta}(r-1|r-1), u(r-1)) = \int_{\mathbb{R}^n} \psi(\zeta(r-1), u(r-1))\mathcal{N}(\zeta(r-1); \hat{\zeta}(r-1|r-1), M(r-1|r-1))d\zeta(r-1).
\]

For \( i \neq m \), the condition \( E[(\hat{\zeta}(r-i| r-i) - \zeta(r-i))\zeta(r-m|r-m) - \zeta(r-m)] = 0 \) is assumed in [38], and the prediction error matrix \( M(r| r-1) \) is defined as

\[
M(r| r-1) = E[(\zeta(r) - \hat{\zeta}(r| r-1)) (\zeta(r) - \hat{\zeta}(r| r-1))^T].
\]

Then, the prediction error matrix \( P(r| r-1) \) is derived as follows

\[
P(r| r-1) = E[(\psi(\zeta(r-1), u(r-1)) - \hat{\mathcal{Y}}(\hat{\zeta}(r-1|r-1), u(r-1)) + \sum_{m=1}^{r} B_m(r)(\zeta(r-m) - \hat{\zeta}(r-m|r-m)) + G_\alpha\varpi(r-1) - \hat{\mathcal{Y}}(\hat{\zeta}(r-1|r-1), u(r-1)) + \sum_{m=1}^{r} B_m(r)(\zeta(r-m) - \hat{\zeta}(r-m|r-m)) + G_\alpha\varpi(r-1))^T] = \int_{\mathbb{R}^n} h(r-1|r-1)h^T(r-1|r-1) \times \mathcal{N}(\zeta(r-1); \hat{\zeta}(r-1|r-1), M(r-1|r-1)d\zeta(r-1) + \sum_{m=1}^{r} B_m(r)(M(r-m|r-m) - B_m(r)^T) + G_\alpha QG_\alpha^T.
\]

where \( h(r-1|r-1) = \psi(\zeta(r-1), u(r-1)) - \hat{\mathcal{Y}}(\hat{\zeta}(r-1|r-1), u(r-1)) \).

The probability density of \( \zeta(r) \) is supposed to satisfy the Gaussian distribution. Based on the prediction \( \hat{\zeta}(r| r-1) \) and the prediction error matrix \( M(r| r-1) \), hence

\[
\hat{\eta}(r| r-1) = \int_{\mathbb{R}^n} y(\zeta(r))\mathcal{N}(\zeta(r); \hat{\zeta}(r| r-1), M(r| r-1))d\zeta(r)
\]

holds.

Thus, the Kalman gain matrix \( K(r) \) is represented as follows

\[
K(r) = P(r)Z^{-1}(r),
\]

where

\[
P(r) = \int_{\mathbb{R}^n} (\zeta(r) - \hat{\zeta}(r| r-1))(y(\zeta(r)) - \hat{\eta}(r| r-1))^T \times \mathcal{N}(\zeta(r); \hat{\zeta}(r| r-1), M(r| r-1))d\zeta(r),
\]

\[
Z(r) = \int_{\mathbb{R}^n} (y(\zeta(r)) - \hat{\eta}(r| r-1))(y(\zeta(r)) - \hat{\eta}(r| r-1))^T \times \mathcal{N}(\zeta(r); \hat{\zeta}(r| r-1), M(r| r-1))d\zeta(r) + R.
\]

The updating formulas for \( \hat{\zeta}(r| r) \) and \( M(r| r) \) are shown, respectively, as follows

\[
\hat{\zeta}(r| r) = \hat{\zeta}(r| r-1) + K(r)(\eta(r) - \hat{\eta}(r| r-1)),
\]

\[
M(r| r) = M(r| r-1) - K(r)Z(r)K^T(r).
\]

Based on the approximate method for numerical integration (7), the cubature points are defined by

\[
\chi^{(i)}(r-1|r-1) = \hat{\zeta}(r-1|r-1) + \sqrt{M(r-1|r-1)}\gamma^{(i)},
\]

where \( \gamma^{(i)} \) is given in Eq. (5). Therefore, the prediction and the prediction error matrix are presented as follows, respectively.
\[ \hat{\zeta}(r| r-1) \approx \frac{1}{2n} \sum_{i=1}^{2n} \varphi(\chi^{(i)}(r-1|r-1), u(r-1)) \]

\[ + \sum_{m=1}^{r} \overline{B}_m(r) \hat{\zeta}(r-m|r-m), \quad (18) \]

\[ M(r| r-1) \]

\[ \approx \frac{1}{2n} \sum_{i=1}^{2n} \varphi(\chi^{(i)}(r-1|r-1), u(r-1)) \]

\[ \times \varphi^T(\chi^{(i)}(r-1|r-1), u(r-1)) \]

\[ - \left[ \frac{1}{2n} \sum_{i=1}^{2n} \varphi(\chi^{(i)}(r-1|r-1), u(r-1)) \right]^T \]

\[ \times \left[ \frac{1}{2n} \sum_{i=1}^{2n} \varphi(\chi^{(i)}(r-1|r-1), u(r-1)) \right] \]

\[ + \sum_{m=1}^{r} \overline{B}_m(r) M(r-m|r-m) \overline{B}_m^T(r) \]

\[ + G_a Q_2 G_a^T. \quad (19) \]

Similarly, we define \( \chi^{(i)}(r| r-1) = \hat{\zeta}(r| r-1) + \sqrt{M(r| r-1)} y^{(i)}. \)

(20)

Hence, \( \hat{\eta}(r| r-1), P(r) \) and \( Z(r) \) are approximately expressed by

\[ \hat{\eta}(r| r-1) \approx \frac{1}{2n} \sum_{i=1}^{2n} y(\chi^{(i)}(r| r-1)), \quad (21) \]

\[ P(r) \approx \frac{1}{2n} \sum_{i=1}^{2n} \chi^{(i)}(r| r-1) y^T(\chi^{(i)}(r| r-1)) \]

\[ - \hat{\zeta}(r| r-1) \times \left[ \frac{1}{2n} \sum_{i=1}^{2n} y(\chi^{(i)}(r| r-1)) \right]^T \]

(22)

and

\[ Z(r) \approx \frac{1}{2n} \sum_{i=1}^{2n} y(\chi^{(i)}(r| r-1)) y^T(\chi^{(i)}(r| r-1)) \]

\[ - \left[ \frac{1}{2n} \sum_{i=1}^{2n} y(\chi^{(i)}(r| r-1)) \right]^T \]

\[ \times \left[ \frac{1}{2n} \sum_{i=1}^{2n} y(\chi^{(i)}(r| r-1)) \right] + R. \quad (23) \]

**Algorithm 1** The FOCKF with the IVC for nonlinear FOS (8) with measurement (9) is implemented by the following steps.

Prediction:

Step 1. Compute the cubature points \( \chi^{(i)}(r-1|r-1) \) for \( i = 1, 2, \ldots, 2n \).

Step 2. Calculate approximations (18) and (19) for \( \hat{\zeta}(r| r-1) \) and \( M(r| r-1) \).

Step 3. Compute the cubature points \( \chi^{(i)}(r| r-1) \) for \( i = 1, 2, \ldots, 2n \).

Updating:

Step 4. Substitute these cubature points \( \chi^{(i)}(r| r-1) \) into \( \eta(r) \) to gain \( \hat{\eta}(r| r-1) \) by (21).

Step 5. Use \( \hat{\eta}(r| r-1) \) and \( \chi(r| r-1) \) to achieve \( K(r) \) by \( P(r) \) and \( Z(r) \) based on (22) and (23).

Step 6. Compute \( \hat{\zeta}(r| r) \) and \( M(r| r) \) by (15) and (16).

**Remark 1** Because the storage capacity is limited, we select a suitable truncation \( L \) in the practical engineering. Therefore, \( \hat{\zeta}(r| r-1) \) and \( M(r| r-1) \) with a truncation \( L \) are modified by

\[ \hat{\zeta}(r| r-1) \approx \frac{1}{2n} \sum_{i=1}^{2n} \varphi(\chi^{(i)}(r-1|r-1), u(r-1)) \]

\[ + \sum_{m=1}^{L} \overline{B}_m(r) \hat{\zeta}(r-m|r-m) \]

and

\[ M(r| r-1) \]

\[ \approx \frac{1}{2n} \sum_{i=1}^{2n} \varphi(\chi^{(i)}(r-1|r-1), u(r-1)) \]

\[ \times \varphi^T(\chi^{(i)}(r-1|r-1), u(r-1)) \]

\[ - \left[ \frac{1}{2n} \sum_{i=1}^{2n} \varphi(\chi^{(i)}(r-1|r-1), u(r-1)) \right]^T \]

\[ \times \left[ \frac{1}{2n} \sum_{i=1}^{2n} \varphi(\chi^{(i)}(r-1|r-1), u(r-1)) \right] \]

\[ + \sum_{m=1}^{L} \overline{B}_m(r) M(r-m|r-m) \overline{B}_m^T(r) \]

\[ + G_a Q_2 G_a^T. \]

3.2 Hybrid extended-cubature Kalman filter to weaken impact of initial value

The advantages of EKF algorithm and CKF algorithm are concerned to weaken the influence of initial value in this subsection. Based on the definition of Caputo derivative, the following equation for \( \alpha \in (0, 1) \) is taken as

\[ _{D_0}^\alpha D_t^\alpha e(t) = _{D_0}^R D_t^{(1-\alpha)} e(t). \quad (24) \]
Therefore, it follows that
\[ \varepsilon^{(1)}(t) = \left. \frac{\partial}{\partial \varepsilon(t)} \left( \varepsilon(t), u(t) \right) \right|_{t = \varepsilon(r-m|r-m)}, \]  
(25)

Using the G–L difference to discretize Eq. (25) at \( t = rt \), the following equation is gained as
\[ \varepsilon(r) - \varepsilon(r-1) \]
\[ = \frac{1}{r^{1-\alpha}} \sum_{m=0}^{r-1} (-1)^{m} b_{m}^{1-\alpha} \phi(\varepsilon(r-m), u(r-m)) \]
\[ + \frac{1}{r^{1-\alpha}} \sum_{m=0}^{r-1} (-1)^{m} b_{m}^{1-\alpha} G w(r-m), \]
(26)

where the absolute value of the coefficient for \( Gw(r-m) \) is related to \( b_{m}^{1-\alpha} \). It is difficult to design fractional-order Kalman filter since the noise term is related to historical information, hence the noise term needs to be simplified. The coefficient \( b_{m}^{1-\alpha} \) decreases rapidly with the increases of \( m \), while \( b_{m}^{1-\alpha} \) tends to zero if \( m \) becomes very large. We retain the two terms \( w(r) \) and \( w(r-1) \), then Eq. (26) is approximated by
\[ \varepsilon(r) = T^{\alpha} \sum_{m=0}^{r} (-1)^{m} b_{m}^{1-\alpha} \phi(\varepsilon(r-m), u(r-m)) \]
\[ + T^{\alpha} \sum_{m=0}^{r} (-1)^{m} b_{m}^{1-\alpha} G w(r-m) + \varepsilon(r-1) \]
\[ = T^{\alpha} \phi(\varepsilon(r), u(r)) + T^{\alpha}(\alpha-1) \phi(\varepsilon(r-1), u(r-1)) \]
\[ + b_{m}^{1-\alpha} \phi(\varepsilon(r-m), u(r-m)) \]
\[ + T^{\alpha} G w(r) + T^{\alpha}(\alpha-1) G w(r-1). \]
(27)

The estimation of \( \varepsilon(r) \) is denoted by \( \hat{\varepsilon}(r|\varepsilon) \) = \( E[\varepsilon(r)|\lambda(r)] \), where \( \lambda(r) \) covers \( \eta(0), \eta(1), \ldots, \eta(r), u(0), \ldots, u(r), \) hence the first-order Taylor expansion of the nonlinear function \( \phi(\varepsilon(r), u(r)) \) at \( \varepsilon(r) = \varepsilon(r-1|r-1) \) is given as follows
\[ \phi(\varepsilon(r), u(r)) \approx \phi(\varepsilon(r-1|r-1), u(r)) + C(r-1)(\varepsilon(r) - \varepsilon(r-1|r-1)), \]
(28)

where \( C(r-1) = \frac{\partial \phi(\varepsilon(u(r)))}{\partial \varepsilon} \big|_{\varepsilon=r-1|r-1}. \)

For \( m \geq 2 \), the first-order Taylor expansion of the nonlinear function \( \phi(\varepsilon(r-m), u(r-m)) \) at \( \varepsilon(r-m) = \hat{\varepsilon}(r-m|r-m) \) is represented by
\[ \phi(\varepsilon(r-m), u(r-m)) \]
\[ \approx \phi(\hat{\varepsilon}(r-m|r-m), u(r-m)) + N(r-m) \phi(\varepsilon(r-m) - \hat{\varepsilon}(r-m|r-m)), \]
(29)

where \( N(r-m) = \frac{\partial \phi(\varepsilon(u(r-m)))}{\partial \varepsilon} \big|_{\varepsilon=r-1|r-1}. \)

The initial values \( C(-2) \) and \( C(-1) \) are expressed by
\[ C(-2) = C(-1) = \frac{\partial \phi(\varepsilon(u(0)))}{\partial \varepsilon} \big|_{\varepsilon=0}. \]

Substituting Eqs. (28) and (29) into Eq. (27), then the following equation holds
\[ \varepsilon(r) = T^{\alpha} \phi(\hat{\varepsilon}(r-1|r-1), u(r)) \]
\[ + T^{\alpha} C(r-1)(\varepsilon(r) - \hat{\varepsilon}(r-1|r-1)) \]
\[ + T^{\alpha}(\alpha-1) \phi(\varepsilon(r-1), u(r-1)) \]
\[ + \varepsilon(r-1) + T^{\alpha} \sum_{m=2}^{r} (-1)^{m} b_{m}^{1-\alpha} \]
\[ \times \phi(\hat{\varepsilon}(r-m|r-m), u(r-m)) \]
\[ + T^{\alpha} \sum_{m=2}^{r} (-1)^{m} b_{m}^{1-\alpha} N(r-m) \]
\[ \times (\varepsilon(r-m) - \hat{\varepsilon}(r-m|r-m)) + T^{\alpha} G w(r) \]
\[ + T^{\alpha}(\alpha-1) G w(r-1). \]

It follows that
\[ \varepsilon(r) = D(\varepsilon(r-1), u(r)) \]
\[ + \tilde{C}(r-1) \hat{\varepsilon}(r-1|r-1) \]
\[ + \Omega(\varepsilon(r-1), u(r-1)) \]
\[ + \sum_{m=2}^{r} D(r)(-1)^{m} b_{m}^{1-\alpha} \]
\[ \times \hat{\varepsilon}(r-m|r-m), u(r-m)) \]
\[ + \sum_{m=2}^{r} \mathcal{L}_{m}(r)(\varepsilon(r-m) - \hat{\varepsilon}(r-m|r-m)) \]
\[ + D(r) G w(r) \]
\[ + J(r) G w(r-1), \]
(30)

where \( D(r) = L(r) T^{\alpha}, \tilde{C}(r-1) = -D(r) C(r-1), \) the nonlinear function \( \Omega(\varepsilon(r-1), u(r-1)) = J(r) \phi(\varepsilon(r-1), u(r-1)) + L(r) \phi(\varepsilon(r-1), u(r-1)), \mathcal{L}_{m}(r) = D(r)(-1)^{m} b_{m}^{1-\alpha} N(r-m), J(r) = D(r)(\alpha-1) \) and 
\[ L(r) = (I - T^{\alpha} C(r-1))^{-1}. \]

Therefore, \( L(0) \) and \( L(-1) \) can be represented by \( L(0) = (I - C(-1))^{-1} \) and \( L(-1) = (I - C(-2))^{-1}. \)

In terms of Eq. (30) and the assumption in Eq. (13), we obtain
\[ \hat{\varepsilon}(r|\varepsilon) = D(\varepsilon(r-1), u(r)) \]
\[ + \tilde{C}(r-1) \hat{\varepsilon}(r-1|r-1) \]
\[ + \Omega(\varepsilon(r-1), u(r-1)) \]
\[ + \sum_{m=2}^{r} D(r)(-1)^{m} b_{m}^{1-\alpha} \]
\[ \times \hat{\varepsilon}(r-m|r-m), u(r-m)) \]
\[ + \sum_{m=2}^{r} \mathcal{L}_{m}(r)(\varepsilon(r-m) - \hat{\varepsilon}(r-m|r-m)) \]
\[ + D(r) G w(r) \]
\[ + J(r) G w(r-1), \]
\[
+ \sum_{m=2}^{r} \mathcal{L}_m(r)(\varepsilon(r - m)
- \hat{\varepsilon}(r - m|r - m)) + D(r)Gw(r)
+ J(r)Gw(r - 1)\lambda(r - 1)).
\]

Similarly, supposing that the probability density function \(p(\varepsilon(r - 1)\lambda(r - 1))\) for the \((r - 1)\)th iteration satisfies the Gaussian distribution, \(p(\varepsilon(r - 1)|\lambda(r - 1)) = \mathcal{N}(\varepsilon(r - 1); \hat{\varepsilon}(r - 1|r - 1), M(r - 1|r - 1))\) holds. Assuming \(w(r)\) is the white Gaussian noise, the prediction \(\hat{\varepsilon}(r|1)\) is rewritten by

\[
\hat{\varepsilon}(r|1) = \int_{\mathbb{R}^a} \mathcal{D}(\varepsilon(r - 1), u(r - 1), \mathcal{N}(\varepsilon(r - 1);
- \hat{\varepsilon}(r - 1|1), M(r - 1|1))de(r - 1),
M(r - 1|1))de(r - 1).
\]

Supposing \(\hat{\varepsilon}(r|r) = \hat{\varepsilon}(r|r - 1) + K(r)(\eta(r) - y(\hat{\varepsilon}(r|r - 1)))\), where \(K(r)\) is the Kalman gain matrix. The first-order Taylor expansion of the nonlinear function \(y(\varepsilon(r))\) in (9) at \(\varepsilon(r) = \hat{\varepsilon}(r|1)\) is represented by

\[
y(\varepsilon(r)) \approx y(\hat{\varepsilon}(r|1)) + Y(r)(\varepsilon(r) - \hat{\varepsilon}(r|1)),
\]

where \(Y(r) = -\frac{\partial y(\varepsilon(r))}{\partial \varepsilon(r)}|_{\varepsilon(\hat{\varepsilon}(r|1))}\).

Hence, the estimation error \(\varepsilon(r) - \hat{\varepsilon}(r|r)\) is expressed by

\[
\varepsilon(r) - \hat{\varepsilon}(r|r) = \varepsilon(r) - \hat{\varepsilon}(r|1) - K(r)(\eta(r) - y(\hat{\varepsilon}(r|1)))
\approx [I - K(r)Y(r)][\mathcal{D}(\varepsilon(r - 1), u(r - 1)) - \hat{\mathcal{D}}(\hat{\varepsilon}(r - 1|1), u(r - 1)) - \hat{\mathcal{D}}(\hat{\varepsilon}(r - 1|1), u(r - 1)) + \sum_{m=2}^{r} \mathcal{L}_m(r)(\varepsilon(r - m) + D(r)Gw(r)
+ J(r)Gw(r - 1)) - K(r)v(r).
\]

Supposing that the conditions \(\hat{\varepsilon}(r - m|r - m) \approx \varepsilon(r - m)\) for \(m \geq 2\) and \(\mathcal{D}(\hat{\varepsilon}(r - 1|1), u(r - 1)) \approx \mathcal{D}(\varepsilon(r - 1), u(r - 1))\), we have

\[
\varepsilon(r) - \hat{\varepsilon}(r|r) \approx (I - K(r)Y(r))(D(r)Gw(r)
+ J(r)Gw(r - 1)) - K(r)v(r).
\]

(31)

For \(i \neq m\), the following equation holds

\[
E\left[\left(\sum_{i=2}^{r} \mathcal{L}_i(r)(\varepsilon(r - i) - \hat{\varepsilon}(r - i|r - i))\right) \times \left(\sum_{m=2}^{r} \mathcal{L}_m(r)(\varepsilon(r - m) - \hat{\varepsilon}(r - m|r - m))\right)^T\right]
\approx E\left[\left(\sum_{i=2}^{r} \mathcal{L}_i(r)(I - K(r - i)Y(r - i)) \times (D(r - i)Gw(r - i) + J(r - i)Gw(r - i - 1)) - K(r - i)v(r - i))\right) \times (I - K(r - m)Y(r - m)) \times (D(r - m)Gw(r - m) + J(r - m)Gw(r - m - 1)) - K(r - m)v(r - m))\right)^T\left(\sum_{m=2}^{r} \mathcal{L}_m(r)\right)
\approx \sum_{m=2}^{r} \sum_{i=2}^{r} \mathcal{L}_m(r)J(r - m)GQG^TD^T(r - m - 1)\mathcal{L}_m^T(r) + \sum_{m=2}^{r} \sum_{i=2}^{r} \mathcal{L}_m(r)J(r - m)GQG^TD^T(r - m)\mathcal{L}^T_{m+1}(r),
\]

where \(\mathcal{L}_m(r) = \mathcal{L}_m(r)(I - K(r - m)Y(r - m))\).

In terms of Eq. (31), we gain

\[
E\left[\left(\sum_{i=2}^{r} \mathcal{L}_i(r)(\varepsilon(r - i) - \hat{\varepsilon}(r - i|r - i))\right) \times (D(r)Gw(r) + J(r)Gw(r - 1))^T\right] = 0.
\]

For \(i = m\), we gain

\[
E\left[\left(\sum_{m=2}^{r} \mathcal{L}_m(r)(\varepsilon(r - m) - \hat{\varepsilon}(r - m|r - m))\right) \times (\sum_{m=2}^{r} \mathcal{L}_m(r)(\varepsilon(r - m) - \hat{\varepsilon}(r - m|r - m)))^T\right]
\approx \sum_{m=2}^{r} \sum_{i=2}^{r} \mathcal{L}_m(r)M(r - m|r - m)\mathcal{L}_m^T(r)
\]

and

\[
E[(D(r)Gw(r) + J(r)Gw(r - 1)) \times (D(r)Gw(r) + J(r)Gw(r - 1))^T] = E[(D(r)Gw(r) + J(r)Gw(r - 1)) \times (D(r)Gw(r) + J(r)Gw(r - 1))^T] = (D(r)GQG^TD^T(r) + J(r)GQG^TJ^T(r).
Meanwhile, we also obtain
\[
E[(\Omega(\varepsilon(r-1), u(r-1)) - \hat{\Omega}(\hat{\varepsilon}(r-1)|r-1), u(r-1)))(\Omega(\varepsilon(r-1), u(r-1)) - \hat{\Omega}(\hat{\varepsilon}(r-1)|r-1), u(r-1))] = \int_{\mathbb{R}^n} \Omega(\varepsilon(r-1), u(r-1))\Omega^T(\varepsilon(r-1), u(r-1)) \, d\varepsilon(r-1) - \hat{\Omega}(\hat{\varepsilon}(r-1)|r-1), u(r-1)) \times \hat{\Omega}^T(\hat{\varepsilon}(r-1)|r-1), u(r-1)).
\]

The first-order Taylor expansion of the nonlinear function \( \Omega(\varepsilon(r-1), u(r-1)) \) at \( \varepsilon(r-1) = \hat{\varepsilon}(r-1)|r-1) \) is given as follows
\[
\Omega(\varepsilon(r-1), u(r-1)) \approx \Omega(\hat{\varepsilon}(r-1)|r-1), u(r-1)) + F(r-1)(\varepsilon(r-1) - \hat{\varepsilon}(r-1)|r-1)) + \Omega(\varepsilon(r-1), u(r-1)) - \hat{\Omega}(\hat{\varepsilon}(r-1)|r-1), u(r-1)) \times \hat{\Omega}^T(\hat{\varepsilon}(r-1)|r-1), u(r-1)).
\]

Assuming \( \Omega(\hat{\varepsilon}(r-1)|r-1), u(r-1)) \approx \hat{\Omega}(\hat{\varepsilon}(r-1)|r-1), u(r-1)) \) for \( i \neq m \), we obtain
\[
E[(\Omega(\varepsilon(r-1), u(r-1)) - \hat{\Omega}(\hat{\varepsilon}(r-1)|r-1), u(r-1))) \times (D(r)GW(r) + J(r)GW(r-1)) = E[(\Omega(\varepsilon(r-1), u(r-1)) + F(r-1)(\varepsilon(r-1) - \hat{\varepsilon}(r-1)|r-1)) + \Omega(\varepsilon(r-1), u(r-1)) - \hat{\Omega}(\hat{\varepsilon}(r-1)|r-1), u(r-1)) \times \hat{\Omega}^T(\hat{\varepsilon}(r-1)|r-1), u(r-1)),
\]

and
\[
E[(\Omega(\varepsilon(r-1), u(r-1)) - \hat{\Omega}(\hat{\varepsilon}(r-1)|r-1), u(r-1))] \times (\sum_{m=2}^{r} \mathcal{L}_m(r)(\varepsilon(r-m) - \hat{\varepsilon}(r-m)|r-m)),
\]

where
\[
\begin{align*}
H(r-1) &= F(r-1)(I - K(r-1)Y(r-1)) \\
&\times (D(r)GW(r) + J(r)GW(r-1)) \times \hat{\Omega}^T(\hat{\varepsilon}(r-1)|r-1), u(r-1)) \times (\sum_{m=2}^{r} \mathcal{L}_m(r)(\varepsilon(r-m) - \hat{\varepsilon}(r-m)|r-m))
\end{align*}
\]

Similar to Eqs. (15) and (16), the updated formulas for \( \hat{\varepsilon}(r|r) \) and \( M(r|r) \) are showed by
\[
\hat{\varepsilon}(r|r) = \hat{\varepsilon}(r-1) + K(r)\eta(r)|r-1)
\]

and
\[
M(r|r) = M(r-1) - K(r)Z(r)K^T(r).
\]
Being similar to the cubature points in (17) and (20), we define the cubature points as follows
\[
\chi^{(i)}(r-1|r-1) = \hat{\epsilon}(r-1|r-1) + \sqrt{M(r-1|r-1)} y^{(i)}
\]
and
\[
\chi^{(i)}(r|r-1) = \hat{\epsilon}(r|r-1) + \sqrt{M(r|r-1)} y^{(i)}.
\]

Then, the prediction \(\hat{\epsilon}(r|r-1)\), the prediction error matrix \(M(r|r-1)\), \(\hat{\eta}(r|r-1)\), \(P(r)\) and \(Z(r)\) are approximately represented by
\[
\hat{\epsilon}(r|r-1) \approx \sum_{m=2}^{r} D(r)(-1)^{m} b_{m}^{1-\alpha} \times \varphi(\hat{\epsilon}(r-m|r-m), u(r-m)),
\]
(34)
\[
M(r|r-1) = \sum_{m=1}^{2n} \Omega^{T}(\chi^{(i)}(r-1|r-1), u(r-1)) \\
\times \Omega(\chi^{(i)}(r-1|r-1), u(r-1)) \\
- \left[ \sum_{m=1}^{2n} \Omega^{T}(\chi^{(i)}(r-1|r-1), u(r-1)) \right]^{T} \\
+ \sum_{m=2}^{r} \mathcal{L}_{m}(r) M(r-m|r-m) \mathcal{L}_{m}^{T}(r) \\
+ D(r) G Q G^{T} D^{T}(r) + J(r) G Q G^{T} J^{T}(r) \\
+ \sum_{m=2}^{r-1} \tilde{\mathcal{L}}_{m}(r) \\
\times J(r-m) G Q G^{T} D^{T}(r-m-1) \tilde{\mathcal{L}}_{m+1}^{T}(r) \\
+ \sum_{m=2}^{r-1} \tilde{\mathcal{L}}_{m+1}(r) \\
\times D(r-m-1) G Q G^{T} J^{T}(r-m) \tilde{\mathcal{L}}_{m}(r) \\
+ H(r-1) D(r-1) G Q G^{T} J^{T}(r) \\
+ J(r) G Q G^{T} D^{T}(r-1) H^{T}(r-1) \\
+ H(r-1) J(r-1) G Q G^{T} D^{T}(r-2) \tilde{\mathcal{L}}_{2}^{T}(r) \\
+ \tilde{\mathcal{L}}_{2}(r) D(r-2) G Q G^{T} J^{T}(r-1) H^{T}(r-1),
\]
(35)
\[
\hat{\eta}(r|r-1) \approx \frac{1}{2n} \sum_{i=1}^{2n} y(\chi^{(i)}(r|r-1)),
\]
(36)
\[
P(r) = \frac{1}{2n} \sum_{i=1}^{2n} \chi^{(i)}(r|r-1) y^{T}(\chi^{(i)}(r|r-1)) \\
- \hat{\epsilon}(r|r-1) \times \left[ \frac{1}{2n} \sum_{i=1}^{2n} y(\chi^{(i)}(r|r-1)) \right]^{T}
\]
(37)
\[
Z(r) = \frac{1}{2n} \sum_{i=1}^{2n} y(\chi^{(i)}(r|r-1)) y^{T}(\chi^{(i)}(r|r-1)) \\
- \left[ \frac{1}{2n} \sum_{i=1}^{2n} y(\chi^{(i)}(r|r-1)) \right]^{T} \times \left[ \frac{1}{2n} \sum_{i=1}^{2n} y(\chi^{(i)}(r|r-1)) \right]^{T} + R.
\]

Algorithm 2 The FOCKF for nonlinear FOS (8) with measurement (9) is implemented by the following steps.

Prediction:

Step 1. Calculate the cubature points \(\chi^{(i)}(r-1|r-1)\) for \(i = 1, 2, \ldots, 2n\).

Step 2. Compute approximations (34) and (35) for \(\hat{\epsilon}(r|r-1)\) and \(M(r|r-1)\).

Step 3. Construct the cubature points \(\chi^{(i)}(r|r-1)\) for \(i = 1, 2, \ldots, 2n\).

Updating:

Step 4. Replace \(\chi^{(i)}(r|r-1)\) into \(\eta(r)\) to gain the approximation \(\hat{\eta}(r|r-1)\) by (36).

Step 5. Utilize \(\eta(r|r-1)\) and \(\hat{\epsilon}(r|r-1)\) to achieve \(K(r)\) by \(P(r)\) and \(Z(r)\) based on (37) and (38).

Step 6. Compute \(\hat{\epsilon}(r|r)\) and \(M(r|r)\) by (32) and (33).

Remark 2 Because the storage capacity is limited, we select a suitable truncation \(L\) in the practical engineering. Therefore, \(\hat{\epsilon}(r|R)\) and \(M(r|R)\) with a truncation \(L\) are modified by
\[
\hat{\epsilon}(r|R) \approx \frac{1}{2n} \sum_{i=1}^{2n} \Omega^{T}(\chi^{(i)}(r-1|r-1), u(r-1)) \\
+ D(r) \varphi(\hat{\epsilon}(r-1|r-1), u(r)) \\
+ \tilde{\mathcal{C}}(r-1) \hat{\epsilon}(r-1|r-1) \\
+ \sum_{m=2}^{L} D(r)(-1)^{m} b_{m}^{1-\alpha}
\]
(35)
4 Numerical examples

Example 1 is given to study the state estimation for nonlinear FOS (39) and verify the validity of the proposed Algorithm 1 based on the CKF. Meanwhile, the calculation time is compared with Algorithm 1 based on the CKF in [32]. Example 2 is used to investigate the state estimation for nonlinear FOS (40) and check the validity of the proposed Algorithm 2 based on the HECKF. Moreover, the calculation time for state estimation is compared with Algorithm 1 based on the CKF in [32].

Example 1 Considering the following nonlinear FOS

\[
\begin{cases}
C D^2_t \varepsilon_1(t) = \varepsilon_2(t) + u_1(t) + u_1(t) \\
0 D^2_t \varepsilon_2(t) = -2\varepsilon_1(t) - 2\varepsilon_2(t) + 0.5\sin(\varepsilon_2(t))\varepsilon_1(t) + w_2(t) + u_2(t)
\end{cases}
\]  

(39)

where \( u_1(t) = 4\sin(0.9t) \), \( u_2(t) = 5\sin(0.9t + \pi/3) \), and \( \eta(t) = \cos(\varepsilon_1(t)) + \varepsilon_2(t) + v(t) \).

The sampling period \( T \) is 0.1 s, and the running time is 40 s. The initial value \( \varepsilon(0) \) in FOS (39) is set as \( \varepsilon(0) = [3, -3]^T \). The process noise \( w(r) = [w_1(r), w_2(r)]^T \) and measurement noise \( v(r) \) are uncorrelated white Gaussian noises and satisfy \( E[w(r)] = 0, E[v(r)] = 0, E[w(r)w^T(s)] = Q\xi(r - s), E[v(r)v^T(s)] = R\xi(r - s), Q = \text{diag}(0.001, 0.001) \) and \( R = 1 \).

The nonlinear functions \( \varphi_1(\varepsilon(t), u(t)) \) and \( \varphi_2(\varepsilon(t), u(t)) \) are \( \varphi_1(\varepsilon(t), u(t)) = \varepsilon_2(t) + u_1(t), \varphi_2(\varepsilon(t), u(t)) = -2\varepsilon_1(t) - 2\varepsilon_2(t) + 0.5\sin(\varepsilon_2(t))\varepsilon_1(t) + u_2(t) \). In order to reduce the initial value influence, we construct the augmented vector as \( \zeta(r) = [\varepsilon_1(r), \varepsilon_2(r), \tilde{\varepsilon}_1(r), \tilde{\varepsilon}_2(r)]^T \) with \( \tilde{\varepsilon}(r) = [\tilde{\varepsilon}_1(r), \tilde{\varepsilon}_2(r)]^T \) and \( Q_1 = \text{diag}(10^{-5}, 10^{-5}) \). Then, the nonlinear function in the corresponding augmented equation is \( \overline{\varphi}(\zeta(r), u(r)) = [T^a \varphi_1(\varepsilon(r), u(r)), T^a \varphi_2(\varepsilon(r), u(r)), 0, 0]^T \), and the augmented noise is \( \overline{w}(r) = [w_1(r), w_2(r), \tilde{\varepsilon}_1(r), \tilde{\varepsilon}_2(r)]^T \) with \( \tilde{w}(r) = [\tilde{w}_1(r), \tilde{w}_2(r)]^T \). The initial estimation \( \tilde{\zeta}(0) \) and the initial error matrix \( M(0|0) \) of \( \zeta(r) \) are set as \( \tilde{\zeta}(0|0) = [0, 0, 0, 0]^T \) and \( M(0|0) = I \).

To evaluate the estimation impact by the proposed Algorithm 1 based on CKF with the IVC, the following indicator is set by

\[
E = \frac{1}{S + 1} \sum_{r=0}^{S} \frac{\sqrt{\sum_{i=1}^{n}(\varepsilon_i(r) - \tilde{\varepsilon}_i(r))^2}}{\sqrt{\sum_{i=1}^{n}\varepsilon_i^2(r)}},
\]

where \( S + 1 \) is the number of the sampling values of \( \varepsilon(r) \).
Let the truncation $L$ be $L = 25, \ldots, 50$ with step 5, the fractional-order $\alpha$ is set from 0.2 to 0.8 with the step 0.1, then the estimation error $E$ for Algorithm 1 is shown in Table 1.

Hence, the estimation error $E$ of system (39) has a tendency to increase for the fractional-order $\alpha$ setting from 0.2 to 0.8 with the step 0.1, with the increase of truncation $L$. However, the calculation burden is increased by a larger truncation $L$. To balance the estimation accuracy and the calculation complexity, an appropriate truncation $L$ is needed to be selected in a practical engineering. To verify the effectiveness of the proposed Algorithm 1 to estimate the state information of system (39), we select $L = 35$ and $\alpha = 0.3$. Then, Fig. 2 draws the measurement $\eta(r)$. Figures 3 and 4 offer the real values and the estimations of $\varepsilon_1(r)$ and $\varepsilon_2(r)$, respectively.

The red and blue curves in Figs. 3 and 4 represent the curves of the state estimations via Algorithm 1 considering the influence of initial value and Algorithm 1 without the IVC in [32], respectively. If the number of iteration $r$ is relatively small, the estimation of initial value is also performed in the dynamic adjustment. Therefore, the responses of the state estimations may fluctuate before the initial value reaches the real value. Figures 3 and 4 show that the proposed Algorithm 1 based on the CKF can more effectively estimate the
state information compared with Algorithm 1 based on the CKF in [32] of FOS (39). It is verified that the IVC method can effectively reduce the influence of initial value on the state estimation.

To compare with the estimation effect of the CKF without the IVC in [32] and the proposed Algorithm 1 with the IVC in this paper, we set the fractional-order as 0.2, 0.3, ..., 0.8 and the truncation as \( L = 35 \), then the error indexes under the index \( E \) are shown in Table 2, where \( E_1 \) and \( T_1 \) represent estimation error and calculation time of Algorithm 1 in [32], respectively. \( E_2 \) and \( T_2 \) are estimation error and calculation time of the proposed Algorithm 1 in this paper, respectively. The simulation software is MATLAB 2014a, hardware configuration is AMD A10-5750M APU with Radeon(tm) HD Graphics, the processing speed of CPU and the RAM are 2.50 GHz and 4.0 GB, respectively.

From Table 2, it is obvious that the state estimation based on CKF proposed by Algorithm 1 is more accurate for different fractional-orders \( \alpha = 0.2, 0.3, \ldots, 0.8 \), but a longer time is taken for calculating. Besides, the difference between \( E_1 \) and \( E_2 \) via Algorithm 1 and Algorithm 1 in [32] decreases with increasing of fractional-order \( \alpha \), and it indicates that the effect of state estimation based on CKF via Algorithm 1 with IVC is more obvious if the fractional-order \( \alpha \in (0, 1) \) is relatively small. Thus, the CKF based on Algorithm 1 can be applied to estimate the state information for a nonlinear continuous-time FOS with a relatively small fractional-order \( \alpha \in (0, 1) \), and the CKF based on Algorithm 1 in [32] is more suitable for the system with a relatively large fractional-order \( \alpha \in (0, 1) \) if the calculation time of state estimation is concerned.

The sampling period of model discretization also determines the dynamic characteristics of a nonlinear fractional-order system. We set the sampling period \( T \) from 0.002 s to 0.25 s with the step 0.008 s.

Then, the curves of Algorithm 1 and Algorithm 1 in [32] under different sampling periods \( T \) are shown in Fig. 5 with \( L = 35 \).

At \( T = 0.002 \) s, the estimation error of Algorithm 1 is slightly larger than that of Algorithm 1 in [32], but the estimation errors via two algorithms are both relatively large. After \( T = 0.02 \) s, the estimation errors present by Algorithm 1 are clearly smaller than that given by Algorithm 1 in [32]. The sampling period should be selected reasonably in terms of an actual model. If the sampling period is set too large or too small, the estimation error may increase. If the estimation errors of the two algorithms are both large, it means that the corresponding sampling period \( T \) is not appropriate.

**Example 2** Considering the following nonlinear FOS

\[
\begin{align*}
\dot{C_0} D_0^\alpha \varepsilon_1(t) &= \varepsilon_2(t) + w_1(t) + u_1(t) \\
\dot{C_0} D_0^\alpha \varepsilon_2(t) &= -1.5 \varepsilon_1(t) - 1.5 \varepsilon_2(t) + 0.5 \sin(\varepsilon_2(t)) \varepsilon_1(t) + w_2(t) + u_2(t)
\end{align*}
\] (40)

where \( u_1(t) = 5 \sin(1.1t) \), \( u_2(t) = 3 \sin(1.1t + \pi/3) \), and \( \eta(t) = \cos(\varepsilon_1(t)) + \varepsilon_2(t) + v(t) \).

The initial value \( \varepsilon(0) \) in FOS (40) is set as \( \varepsilon(0) = [3, -3]^T \). The sampling period \( T \) is 0.1 s, and the running time is 30 s. The process noise \( w(r) = [w_1(r), w_2(r)]^T \) and measurement noise \( v(r) \) are uncorrelated white Gaussian noises and satisfy \( E[w(r)] = 0 \), \( E[v(r)] = 0 \), \( E[w(r) w^T(s)] = Q \xi(r-s) \), \( E[v(r) v^T(s)] = R \xi(r-s) \). \( Q = \text{diag}(0.001, 0.001) \) and \( R = 0.5 \). The initial estimation \( \hat{\varepsilon}(0|0) \) and the esti-
The red and blue curves in Figs. 7 and 8 represent the estimation error curves of Algorithm 2 considering the influence of initial value and Algorithm 1 without the IVC in [32], respectively. It can be seen that the proposed Algorithm 2 based on the HECKF can usually obtain more accurate state estimation, compared with Algorithm 1 based on CKF in [32] from the curves of Figs. 7 and 8. It verifies that the model transformation method can weaken the influence of initial value on state estimation.

The estimation effect of Algorithm 1 in [32] and the proposed Algorithm 2 based on the HECKF in this paper is evaluated by the error index $E$. The fractional-order $\alpha$ is set as 0.2, 0.3, . . . , 0.8, the truncation $L = 35$, and the error indexes under the index $E$ are shown in Table 4, where $E_3$ and $T_3$ represent the estimation error and the calculation time of Algorithm 1 in [32], respectively. The $E_4$ and $T_4$ are estimation error and calculation time of proposed Algorithm 2 in this paper, respectively.

Table 3 Relationship between $E$ and $L$ with different fractional-orders $\alpha$ via Algorithm 2

| $L$ | $\alpha$ | 0.2 | 0.3 | 0.4 | 0.5 | 0.6 | 0.7 | 0.8 |
|-----|----------|-----|-----|-----|-----|-----|-----|-----|
| 25  | 0.2638   | 0.2335 | 0.1904 | 0.1408 | 0.1025 | 0.0946 | 0.0990 |
| 30  | 0.2211   | 0.2046 | 0.1738 | 0.1324 | 0.0989 | 0.0907 | 0.0950 |
| 35  | 0.1863   | 0.1793 | 0.1577 | 0.1229 | 0.0929 | 0.0860 | 0.0900 |
| 40  | 0.1578   | 0.1571 | 0.1425 | 0.1133 | 0.0852 | 0.0775 | 0.0818 |
| 45  | 0.1341   | 0.1376 | 0.1287 | 0.1044 | 0.0778 | 0.0661 | 0.0707 |
| 50  | 0.1145   | 0.1204 | 0.1164 | 0.0964 | 0.0720 | 0.0557 | 0.0591 |
From Table 4, it is obvious that the state estimation based on CKF by the proposed Algorithm 2 is more accurate for different fractional-orders $\alpha = 0.2, 0.3, \ldots, 0.8$, but it needs to take a longer time for calculation. In addition, the difference of state estimation errors between Algorithm 2 and Algorithm 1 in [32] gradually decreases as the fractional-order $\alpha$ increases, and a more accurate estimation effect of Algorithm 2 based on CKF for a relatively small fractional-order $\alpha \in (0, 1)$ is obtained to compare with Algorithm 1 in [32]. Moreover, the CKF based on Algorithm 1 in [32] is more applicable for FOS (40) with a relatively large fractional-order $\alpha \in (0, 1)$ if the calculation time of state estimation is taken into account.

Similarly, the sampling period $T$ is set from 0.003 s to 0.3 s with the step 0.009 s, and the estimation error curves of Algorithm 2 and Algorithm 1 in [32] with the truncation $L = 35$ are drawn in Fig. 9.

Figure 9 presents the estimation errors via Algorithm 2 and Algorithm 1 in [32], respectively. It is similar to the conclusion in Example 1, and the sampling period also affects the accuracy of state estimation. However, the algorithm proposed in this paper is the better for the estimation accuracy with the appropriate sampling period.

| $\alpha$ | $E_3$  | $E_4$  | $E_3 - E_4$ | $T_3$ (s) | $T_4$ (s) |
|---------|--------|--------|-------------|-----------|-----------|
| 0.2     | 0.9792 | 0.1863 | 0.7929      | 2.7007    | 3.6005    |
| 0.3     | 0.5225 | 0.1793 | 0.3432      | 2.6071    | 3.5816    |
| 0.4     | 0.3493 | 0.1577 | 0.1916      | 2.3306    | 3.1347    |
| 0.5     | 0.2349 | 0.1229 | 0.1120      | 2.2731    | 3.3592    |
| 0.6     | 0.1717 | 0.0929 | 0.0788      | 2.0796    | 2.9506    |
| 0.7     | 0.1405 | 0.0860 | 0.0545      | 2.2288    | 3.0905    |
| 0.8     | 0.1110 | 0.0900 | 0.0210      | 2.2285    | 3.1283    |
5 Conclusion

This paper presents two types of FOCKFs to solve problem on the initial value influence. The proposed Algorithms 1 and 2 based on CKF and HECKF can both effectively reduce the influence of initial value on the state estimation, respectively. For a relatively small fractional-order $\alpha \in (0, 1)$, the more accurate state estimation is obtained by Algorithms 1 and 2 compared to Algorithm 1 in [32] based on CKF without the initial value compensation. But the CKF proposed in Algorithm 1 in [32] may be more suitable for larger fractional-orders $\alpha \in (0, 1)$ if the calculation time is taken into account. The future work is to consider the initial value problem in the adaptive fractional-order CKF with the unknown covariance matrices $Q$ and $R$ for nonlinear fractional-order systems.

Funding This work was supported by Liaoning Revitalization Talents Program under Grant XLYC1807229, China Postdoctoral Science Foundation Funded Project under Grant 2019M651206, Scientific Research Fund of Liaoning Provincial Education Department, China, under Grant LJC202010 and Liaoning University Science Research Fund LDGY2019020.

Declarations

Conflict of interest The authors declare that they have no conflict of interest.

References

1. Sun, H.G., Zhang, Y., Baleanu, D., Chen, W., Chen, Y.Q.: A new collection of real world applications of fractional calculus in science and engineering. Commun. Nonlinear Sci. Numer. Simul. 64, 213–231 (2018)
2. Tarasov, V.E., Tarasova, V.V.: Macroeconomic models with long dynamic memory: fractional calculus approach. Appl. Math. Comput. 338, 466–486 (2018)
3. Liang, B.Y., Zheng, S.Q., Ahn, C.K., Liu, F.: Adaptive fuzzy control for fractional-order interconnected systems with unknown control directions. IEEE Trans. Fuzzy Syst. (2020). https://doi.org/10.1109/TFUZZ.2020.3031694
4. West, B.J., Turalska, M., Grigolini, P.: Fractional calculus ties the microscopic and macroscopic scales of complex network dynamics. Physics 17(4), ID. 045009 (2015)
5. Zhao, D.Z., Luo, M.K.: Representations of acting processes and memory effects: general fractional derivative and its application to theory of heat conduction with finite wave speeds. Appl. Math. Comput. 346, 531–544 (2019)
6. Failla, G., Zingales, M.: Advanced materials modelling via fractional calculus: challenges and perspectives. Philos. Trans. R. Soc. A Math. Phys. Eng. Sci. 378(2172), ID. 20200050 (2020)
7. Boudjehem, B., Boudjehem, D.: Fractional order controller design for desired response. Proc. Inst. Mech. Eng. Part I J. Syst. Control Eng. 227(2), 243–251 (2013)
8. Zhao, H.M., Deng, W., Yang, X.H., Li, X.M., Dong, C.: An optimized fractional order PID controller for suppressing vibration of AC motor. J. Vibroeng. 18(4), 2205–2220 (2016)
9. Mohan, V., Chhabra, H., Rani, A., Singh, V.: Robust self-tuning fractional order PID controller dedicated to nonlinear dynamic system. J. Intell. Fuzzy Syst. 34(3), 1467–1478 (2018)
10. Pashaie, S., Badamchizadeh, M.: A new fractional-order sliding mode controller via a nonlinear disturbance observer for a class of dynamical systems with mismatched disturbances. ISA Trans. 63, 39–48 (2016)
11. Mozei, S.A., Zakeri, E., Eghtesad, M.: Optimal adaptive interval type-2 fuzzy fractional-order backstepping sliding mode control method for some classes of nonlinear systems. ISA Trans. 93, 23–39 (2019)
12. Ntouskas, S., Sarimveis, H., Sopasakis, P.: Model predictive control for offset-free reference tracking of fractional order systems. Control Eng. Pract. 71, 26–33 (2018)
13. Necabia, A., Ladaci, S., Charef, A., Loiseau, J.J.: Fractional order extremum seeking approach for maximum power point tracking of photovoltaic panels. Front. Energy 9(1), 43–53 (2015)
14. Pasha, S.A., Tuan, H.D., Vo, B.N.: Nonlinear Bayesian filtering using the uncentred linear fractional transformation model. IEEE Trans. Signal Process. 58(2), 477–489 (2010)
15. Liu, T.Y., Wei, Y.H., Yin, W.D., Wang, Y., Liang, Q.: State estimation for nonlinear discrete-time fractional systems a Bayesian perspective. Signal Process. 165, 250–261 (2019)
16. Ljung, L.: Asymptotic-behavior of the extended Kalman filter as a parameter estimator for linear systems. Autom. Control IEEE Trans. 24(1), 36–50 (2013)
17. Charkhgard, M., Farrokhi, M.: State-of-charge estimation for lithium-ion batteries using neural networks and EKF. IEEE Trans. Ind. Electron. 57(12), 4178–4187 (2010)
18. Wang, X.W., Xu, Z.S., Gou, X.J., Trajkovic, L.: Tracking a maneuvering target by multiple sensors using extended Kalman filter with nested probabilistic-numerical linguistic information. IEEE Trans. Fuzzy Syst. 28(2), 346–360 (2020)
19. Naseri, F., Kazemi, Z., Farjah, E., Ghanbari, T.: Fast detection and compensation of current transformer saturation using extended Kalman filter. IEEE Trans. Power Deliv. 34(3), 91, 1087–1097 (2019)
20. Zeng, N.Y., Wang, Z.D., Zhang, H.: Inferring nonlinear lateral flow immunoassay state-space models via an uncentred Kalman filter. Sci. China Inf. Sci. 59(11), ID. 112204 (2016)
21. Wan, E., Van Der Merwe, R.: The uncentred Kalman filter. In: Kalman Filtering and Neural Networks. Wiley, pp 123–174 (2001)
22. Xiong, K., Zhang, H.Y., Chan, C.W.: Performance evaluation of UKF-based nonlinear filtering. Automatica 42(2), 261–270 (2006)
23. Sarkka, S.: Uncentred Kalman filtering for state estimation of continuous-time nonlinear systems. IEEE Trans. Autom. Control 52(9), 1631–1641 (2007)
24. Arasaratnam, I., Haykin, S.: Cubature Kalman filters. IEEE Trans. Autom. Control 54(6), 1254–1269 (2009)
25. Arasaratnam, I., Haykin, S., Hurd, T.R.: Cubature Kalman filtering for continuous-discrete systems: theory and simulations. IEEE Trans. Signal Process. 58(10), 4977–4993 (2010)

26. Cui, B.B., Wei, X.H., Chen, X.Y., Li, J.Y., Wang, A.C.: Robust cubature Kalman filter based on variational Bayesian and transformed posterior sigma points error. ISA Trans. 86, 18–28 (2018)

27. Kolås, S., Foss, B.A., Schei, T.S.: Constrained nonlinear state estimation based on the UKF approach. Comput. Chem. Eng. 33(8), 1386–1401 (2009)

28. Liu, F.H., Gao, Z., Yang, C., Ma, R.C.: Fractional-order Kalman filters for continuous-time fractional-order systems involving correlated and uncorrelated process and measurement noises. Trans. Inst. Meas. Control 41(7), 1933–1947 (2019)

29. Yang, C., Gao, Z., Liu, F.H.: Kalman filters for linear continuous-time fractional-order systems involving colored noises using fractional-order average derivative. IET Control Theory Appl. 12(4), 456–465 (2018)

30. Sierociuk, D., Dzielinski, A.D.: Fractional Kalman filter algorithm for the states, parameters and order of fractional system estimation. Int. J. Appl. Math. Comput. Sci. 16(1), 129–140 (2006)

31. Ramezani, A., Safarinejad, B.: A modified fractional-order unscented Kalman filter for nonlinear fractional-order systems. Circuits Syst. Signal Process. 37(9), 3756–3784 (2018)

32. Gao, Z.: Cubature Kalman filters for nonlinear continuous-time fractional-order systems with uncorrelated and correlated noises. Nonlinear Dyn. 96, 1805–1817 (2019)

33. Ramezani, A., Safarinejad, B., Zarei, J.: Novel hybrid robust fractional interpolatory cubature Kalman filters. J. Frankl. Inst. 357(1), 704–725 (2020)

34. Yang, C., Gao, Z., Huang, X.M., Kan, T.: Hybrid extended-cubature Kalman filters for nonlinear continuous-time fractional-order systems involving uncorrelated and correlated noises using fractional-order average derivative. IET Control Theory Appl. 14(11), 1424–1437 (2020)

35. Huang, X.M., Gao, Z., Chen, X.J.: Extended Kalman filter for linear fractional-order systems with unknown fractional-order. ICIC Express Lett. 14(5), 431–441 (2020)

36. Prodlubny, I.: Fractional Differential Equation. Academic Press, New York (1999)

37. Caputo, M.C., Torres, D.F.M.: Duality for the left and right fractional derivatives. Signal Process. 107, 265–271 (2015)

38. Sierociuk, D., Dzielinski, A.D.: Fractional Kalman filter algorithm for the states, parameters and order of fractional system estimation. Int. J. Appl. Math. Comput. Sci. 1(16), 129–140 (2006)

Publisher’s Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.