

Study on the Initial Value Problem for Fractional-order Cubature Kalman Filters of Nonlinear Continuous-time Fractional-order Systems

Chuang Yang

Liaoning University

Zhe Gao (✉ gaozhe@lnu.edu.cn)

Liaoning University <https://orcid.org/0000-0002-7084-5426>

Yue Miao

Liaoning University

Tao Kan

Liaoning University

Research Article

Keywords: Nonlinear fractional-order systems, extended Kalman filter, cubature Kalman filter, state estimation, initial value compensation

Posted Date: February 24th, 2021

DOI: <https://doi.org/10.21203/rs.3.rs-207512/v1>

License:   This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Version of Record: A version of this preprint was published at Nonlinear Dynamics on July 25th, 2021. See the published version at <https://doi.org/10.1007/s11071-021-06726-3>.

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

Chuang Yang¹, Zhe Gao^{1,2,3*}, Yue Miao¹,
Tao Kan¹

Received: date / Accepted: date

Abstract To realize the state estimation of a nonlinear continuous-time fractional-order system, two types of fractional-order cubature Kalman filters (FOCKFs) 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 are also regarded as the augmented state, 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 (HECKF) 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 effectively reduce the influence of initial value on the state estimation. 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, the investigation of fractional-order systems (FOSs) has been studied by many scholars, because fractional-order calculus can more accurately describe the phenomena in real worlds [1], [2], [3]. The fractional-order calculus is applied to deal with fractional-order integral and derivative, and 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

Z. Gao is the corresponding author

1. School of Mathematics, Liaoning University, Shenyang, P R China, 110036

2. College of Light Industry, Liaoning University, Shenyang, P R China, 110036

3. Department of Control Science and Engineering, Jilin University, Changchun, P R China, 130025

Tel.: +86-24-62202139

E-mail: gaozhe@lnu.edu.cn, ORCID: 0000-0002-7084-5426

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 the practical control systems, the measurement signal inevitably contains certain noise disturbances and some states information may not be obtained directly, then robust state observers for FOSs need to be designed.

As one 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 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 been more and more commonly used 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 have 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, this method can not effectively estimate the state information, which is also the limitation of EKF [20]. Hence, other types of KFs have been designed to deal with non-differentiable nonlinear functions, such as unscented Kalman filter (UKF) [21], [22], [23] and cubature Kalman filter (CKF) [24], [25], [26]. The UKF uses unscented transformation to cope with nonlinear functions in FOSs, then a more proper 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 integer-order KFs. The fractional-order EKFs were proposed for linear FOSs with white Gaussian noise and coloured noise respectively in [28], [29]. 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 [31], the improved fractional-order UKF to nonlinear FOSs was designed to estimate more accurately state information. Besides, CKF can also estimate effectively state information for the nonlinear FOSs. In [32], the CKF method was applied to obtain the higher estimation accuracy compared with EKF method for nonlinear continuous-time FOSs contain white Gaussian noise. Meanwhile, hybrid FOKFs for nonlinear FOSs were also offered to enhance the estimation accuracy in [33] and [34].

For the 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 the means of augmented vector. The EKF with IVC for continuous-time linear FOSs with unknown fractional-order was proposed in [35]. In this paper, the CKF for nonlinear FOSs with the IVC is investigated to obtain effective state estimations. This paper introduces the G-L difference method into two types of fractional-order cubature Kalman filters (FOCKF) to obtain the effective state estimation of nonlinear FOSs with IVC.

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

The remaining of this paper is arranged as follows. In Section 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 and the FOCKF and HECKF to solve initial value problems in Section 3 respectively. In Section 4, the simulation examples are gave to testify the validity of the proposed fractional-order CKF 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 conformity 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 Caputo definition and the IVC problem is considered, the Riemman-Liouville (R-L) definition is also used in this study.

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

$${}^R_0D_t^\alpha \varepsilon(t) = \frac{1}{\Gamma(n-\alpha)} \frac{d^n}{dt^n} \left(\int_0^t (t-\mu)^{n-\alpha-1} \varepsilon(\mu) d\mu \right),$$

where ${}^R_0D_t^\alpha \varepsilon(t)$ represents α -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 α -order derivative for $\varepsilon(t)$ is defined as

$${}_0^C D_t^\alpha \varepsilon(t) = \frac{1}{\Gamma(n-\alpha)} \int_0^t (t-\mu)^{n-\alpha-1} \varepsilon^{(n)}(\mu) d\mu,$$

where ${}_0^C D_t^\alpha \varepsilon(t)$ is α -order derivative under Caputo definition, and $n-1 < \alpha \leq n$.

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

$${}_0^C D_t^\alpha \varepsilon(t) = {}_0^R D_t^\alpha \varepsilon(t) - \sum_{r=0}^{n-1} \frac{\varepsilon^{(r)}(0)}{\Gamma(r-\alpha+1)} t^{r-\alpha}. \quad (1)$$

Definition 3 The Grünwald-Letnikov (G-L) difference [38] for α -order is defined by

$$\Delta^\alpha \varepsilon(r) = \frac{1}{T^\alpha} \sum_{m=0}^r (-1)^m b_m^\alpha \varepsilon(r-m), \quad (2)$$

where Δ^α is the α -order difference operation, T is the sampling period and the coefficient b_m^α is represented as follows

$$b_m^\alpha = \binom{\alpha}{m} = \begin{cases} 1, & m=0 \\ \frac{\alpha(\alpha-1)\cdots(\alpha-m+1)}{m!}, & m \geq 1 \end{cases}.$$

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

$${}_0^R D_t^\alpha \varepsilon(t) \approx \Delta^\alpha \varepsilon(r) = \frac{1}{T^\alpha} \sum_{m=0}^r (-1)^m b_m^\alpha \varepsilon(r-m).$$

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

$${}_0^C D_t^\alpha \varepsilon(t) \approx \frac{1}{T^\alpha} \sum_{m=0}^r (-1)^m b_m^\alpha \varepsilon(r-m) - \frac{A(r)}{T^\alpha} \varepsilon(0), \quad (3)$$

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

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

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

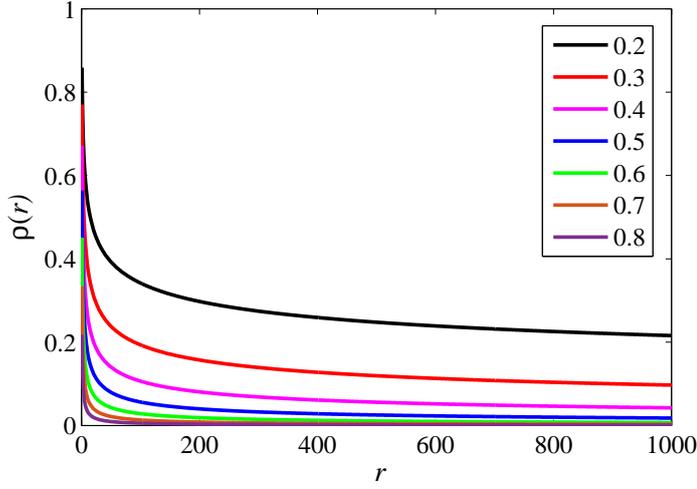


Fig. 1 Curves of $\rho(r)$ for different fractional-order α .

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 the 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 non-differentiable. 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 non-differentiable nonlinear function.

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

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

$$\int_{\mathbb{R}^n} l(\varepsilon) \mathcal{N}(\varepsilon; \hat{\varepsilon}, M) d\varepsilon = \int_{\mathbb{R}^n} l(\hat{\varepsilon} + \sqrt{M}\gamma) \mathcal{N}(\gamma; \mathbf{0}, I) d\gamma. \quad (4)$$

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

$$\gamma^{(i)} = \begin{cases} \sqrt{n}e_i, & i = 1, 2, \dots, n \\ -\sqrt{n}e_{i-n}, & i = n+1, n+2, \dots, 2n \end{cases} \quad (5)$$

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), the integral $\int_{\mathbb{R}^n} l(\hat{\varepsilon} + \sqrt{M}\gamma)\mathcal{N}(\gamma; \mathbf{0}, I)d\gamma$ in (4) is approximately described by

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

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

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

3 Main results

3.1 Fractional-order cubature Kalman filter with initial value compensation by using augmented vector method

Consider the following nonlinear continuous-time FOS as

$${}^C_0 D_t^\alpha \varepsilon(t) = \varphi(\varepsilon(t), u(t)) + Gw(t), \quad (8)$$

$$\eta(t) = y(\varepsilon(t)) + v(t), \quad (9)$$

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 $E[w(r)] = \mathbf{0}$, $E[v(r)] = \mathbf{0}$, $E[w(r)w^T(s)] = Q\xi(r-s)$, $E[v(r)v^T(s)] = R\xi(r-s)$ and $E[w(r)v^T(s)] = \mathbf{0}$, $w(r)$ and $v(r)$ are the sampling values at $t = rT$ with the sampling period T , $E[\cdot]$ returns the mathematical expectation. $\xi(r-s)$ fulfils $\xi(r-s) = \mathbf{0}$ 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 $\mathbf{0}$ is a zero vector or a zero matrix with an appropriate dimension.

Combining equations (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

$$\begin{aligned} \varepsilon(r) = & \sum_{m=1}^r B_m \varepsilon(r-m) + T^\alpha \varphi(\varepsilon(r-1), u(r-1)) + \\ & T^\alpha Gw(r-1) + A(r)\varepsilon(0), \end{aligned} \quad (10)$$

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

$$\tilde{\varepsilon}(r) = \tilde{\varepsilon}(r-1) + \tilde{w}(r-1), \quad (11)$$

where $\tilde{w}(r)$ represents the white Gaussian noise, and satisfies $E[\tilde{w}(r)] = \mathbf{0}$, $E[\tilde{w}(r)\tilde{w}^T(s)] = \xi(r-s)$, 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 $\bar{w}(r) = [w^T(r), \tilde{w}^T(r)]^T$ with $E[\bar{w}(r)\bar{w}^T(r)] = Q_2 = \begin{bmatrix} Q & \mathbf{0} \\ \mathbf{0} & Q_1 \end{bmatrix}$ and $E[\bar{w}(r)v^T(s)] = \mathbf{0}$.

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

$$\zeta(r) = \bar{\varphi}(\zeta(r-1), u(r-1)) + \sum_{m=1}^r \bar{B}_m(r)\zeta(r-m) + G_\alpha \bar{w}(r-1), \quad (12)$$

where the nonlinear function

$$\bar{\varphi}(\zeta(r-1), u(r-1)) = \begin{bmatrix} T^\alpha \varphi(\varepsilon(r-1), u(r-1)) \\ \mathbf{0} \end{bmatrix},$$

the matrix $\bar{B}_m(r) = \begin{cases} \begin{bmatrix} B_1 & A(r)I \\ \mathbf{0} & I \end{bmatrix}, & m=1 \\ \begin{bmatrix} B_m & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix}, & m>1 \end{cases}$, and $G_\alpha = \begin{bmatrix} T^\alpha G & \mathbf{0} \\ \mathbf{0} & I \end{bmatrix}$, I is an

identity matrix with the appropriate dimension.

A CKF algorithm with the initial value compensation is proposed to achieve more accuracy estimation effect of the 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 $\hat{\zeta}(r) = E[\zeta(r)|\lambda(r)]$, where $\lambda(r)$ covers $\eta(0), \eta(1), \dots, \eta(r), u(0), u(1), \dots, u(r)$.

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 the difference equation (12) as follows

$$\begin{aligned} & \hat{\zeta}(r|r-1) \\ &= E[\bar{\varphi}(\zeta(r-1), u(r-1)) + \sum_{m=1}^r \bar{B}_m(r)\zeta(r-m) + G_\alpha \bar{w}(r-1)|\lambda(r-1)] \quad (13) \\ &= E[\bar{\varphi}(\zeta(r-1), u(r-1))|\lambda(r-1)] + \sum_{m=1}^r \bar{B}_m(r)\hat{\zeta}(r-m|r-m). \end{aligned}$$

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

$$\begin{aligned} \hat{\zeta}(r|r-1) &= \int_{\mathbb{R}^n} \bar{\varphi}(\zeta(r-1), u(r-1)) \mathcal{N}(\zeta(r-1); \hat{\zeta}(r-1|r-1), \\ & \quad M(r-1|r-1)) d\zeta(r-1) + \sum_{m=1}^r \bar{B}_m(r)\hat{\zeta}(r-m|r-m). \end{aligned}$$

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

$$\begin{aligned} & \widehat{Y}(\widehat{\zeta}(r-1|r-1), u(r-1)) \\ &= \int_{\mathbb{R}^n} \overline{\varphi}(\zeta(r-1), u(r-1)) \mathcal{N}(\zeta(r-1); \widehat{\zeta}(r-1|r-1), M(r-1|r-1)) d\zeta(r-1). \end{aligned}$$

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

$$M(r|r-1) = E[(\zeta(r) - \widehat{\zeta}(r|r-1))(\zeta(r) - \widehat{\zeta}(r|r-1))^T],$$

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

$$\begin{aligned} M(r|r-1) &= E[(\overline{\varphi}(\zeta(r-1), u(r-1)) - \widehat{Y}(\widehat{\zeta}(r-1|r-1), u(r-1)) + \\ & \quad \sum_{m=1}^r \overline{B}_m(r)(\zeta(r-m) - \widehat{\zeta}(r-m|r-m)) + G_\alpha \overline{w}(r-1)) \times \\ & \quad (\overline{\varphi}(\zeta(r-1), u(r-1)) - \widehat{Y}(\widehat{\zeta}(r-1|r-1), u(r-1)) + \\ & \quad \sum_{m=1}^r \overline{B}_m(r)(\zeta(r-m) - \widehat{\zeta}(r-m|r-m)) + G_\alpha \overline{w}(r-1))^T] \\ &= \int_{\mathbb{R}^n} h(r-1|r-1) h^T(r-1|r-1) \times \\ & \quad \mathcal{N}(\zeta(r-1); \widehat{\zeta}(r-1|r-1), M(r-1|r-1)) d\zeta(r-1) + \\ & \quad \sum_{m=1}^r \overline{B}_m(r) M(r-m|r-m) \overline{B}_m^T(r) + G_\alpha Q_2 G_\alpha^T, \end{aligned}$$

where $h(r-1|r-1) = \overline{\varphi}(\zeta(r-1), u(r-1)) - \widehat{Y}(\widehat{\zeta}(r-1|r-1), u(r-1))$.

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

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

holds.

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

$$K(r) = P(r) Z^{-1}(r), \quad (14)$$

where

$$\begin{aligned} P(r) &= \int_{\mathbb{R}^n} (\zeta(r) - \widehat{\zeta}(r|r-1))(y(\zeta(r)) - \hat{\eta}(r|r-1))^T \times \\ & \quad \mathcal{N}(\zeta(r); \widehat{\zeta}(r|r-1), M(r|r-1)) d\zeta(r), \\ Z(r) &= \int_{\mathbb{R}^n} (y(\zeta(r)) - \hat{\eta}(r|r-1))(y(\zeta(r)) - \hat{\eta}(r|r-1))^T \times \\ & \quad \mathcal{N}(\zeta(r); \widehat{\zeta}(r|r-1), M(r|r-1)) d\zeta(r) + R. \end{aligned}$$

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

$$\hat{\zeta}(r|r) = \hat{\zeta}(r|r-1) + K(r)(\eta(r) - \hat{\eta}(r|r-1)), \quad (15)$$

$$M(r|r) = M(r|r-1) - K(r)S(r)K^T(r). \quad (16)$$

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

$$\chi^{(i)}(r-1|r-1) = \hat{\zeta}(r-1|r-1) + \sqrt{M(r-1|r-1)}\gamma^{(i)}, \quad (17)$$

where $\gamma^{(i)}$ is given in the equation (5). Therefore, the prediction and prediction error matrix are presented as follows, respectively.

$$\hat{\zeta}(r|r-1) \approx \frac{1}{2n} \sum_{i=1}^{2n} \bar{\varphi}(\chi^{(i)}(r-1|r-1), u(r-1)) + \sum_{m=1}^r \bar{B}_m(r) \hat{\zeta}(r-m|r-m), \quad (18)$$

$$\begin{aligned} M(r|r-1) \approx & \frac{1}{2n} \sum_{i=1}^{2n} \bar{\varphi}(\chi^{(i)}(r-1|r-1), u(r-1)) \times \\ & \bar{\varphi}^T(\chi^{(i)}(r-1|r-1), u(r-1)) - \\ & \left[\frac{1}{2n} \sum_{i=1}^{2n} \bar{\varphi}(\chi^{(i)}(r-1|r-1), u(r-1)) \right] \times \\ & \left[\frac{1}{2n} \sum_{i=1}^{2n} \bar{\varphi}(\chi^{(i)}(r-1|r-1), u(r-1)) \right]^T + \\ & \sum_{m=1}^r \bar{B}_m(r) M(r-m|r-m) \bar{B}_m^T(r) + G_\alpha Q_2 G_\alpha^T. \end{aligned} \quad (19)$$

Similarly, based on the definition

$$\chi^{(i)}(r|r-1) = \hat{\zeta}(r|r-1) + \sqrt{M(r|r-1)}\gamma^{(i)}. \quad (20)$$

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

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

$$\begin{aligned} P(r) \approx & \frac{1}{2n} \sum_{i=1}^{2n} \chi^{(i)}(r|r-1) y^T(\chi^{(i)}(r|r-1)) - \hat{\zeta}(r|r-1) \times \\ & \left[\frac{1}{2n} \sum_{i=1}^{2n} y(\chi^{(i)}(r|r-1)) \right]^T, \end{aligned} \quad (22)$$

and

$$\begin{aligned} Z(r) \approx & \frac{1}{2n} \sum_{i=1}^{2n} y(\chi^{(i)}(r|r-1)) y^T(\chi^{(i)}(r|r-1)) - \\ & \left[\frac{1}{2n} \sum_{i=1}^{2n} y(\chi^{(i)}(r|r-1)) \right] \left[\frac{1}{2n} \sum_{i=1}^{2n} y(\chi^{(i)}(r|r-1)) \right]^T + R. \end{aligned} \quad (23)$$

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

Prediction:

Step one. Compute the cubature point $\chi^{(i)}(r-1|r-1)$ for $i = 1, 2, \dots, 2n$.

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

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

Updating:

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

Step five. Use $\hat{\eta}(r|r-1)$ and $\hat{\zeta}(r|r-1)$ to ensure $K(r)$ by $P(r)$ and $Z(r)$ based on (22) and (23).

Step six. 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

$$\hat{\zeta}(r|r-1) \approx \frac{1}{2n} \sum_{i=1}^{2n} \bar{\varphi}(\chi^{(i)}(r-1|r-1), u(r-1)) + \sum_{m=1}^L \bar{B}_m(r) \hat{\zeta}(r-m|r-m),$$

and

$$\begin{aligned} M(r|r-1) &\approx \frac{1}{2n} \sum_{i=1}^{2n} \bar{\varphi}(\chi^{(i)}(r-1|r-1), u(r-1)) \times \\ &\quad \bar{\varphi}^T(\chi^{(i)}(r-1|r-1), u(r-1)) - \\ &\quad \left[\frac{1}{2n} \sum_{i=1}^{2n} \bar{\varphi}(\chi^{(i)}(r-1|r-1), u(r-1)) \right] \times \\ &\quad \left[\frac{1}{2n} \sum_{i=1}^{2n} \bar{\varphi}(\chi^{(i)}(r-1|r-1), u(r-1)) \right]^T + \\ &\quad \sum_{m=1}^L \bar{B}_m(r) M(r-m|r-m) \bar{B}_m^T(r) + G_\alpha Q_2 G_\alpha^T. \end{aligned}$$

3.2 Hybrid extended-cubature Kalman filter to solve the 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 is taken as

$${}_0^C D_t^\alpha \varepsilon(t) = {}_0^R D_t^{-(1-\alpha)} \varepsilon^{(1)}(t). \quad (24)$$

Therefore, it follows that

$$\varepsilon^{(1)}(t) = {}_0^R D_t^{(1-\alpha)} (\varphi(\varepsilon(t), u(t)) + Gw(t)). \quad (25)$$

Using G-L difference to discretize the equation (25) at $t = rT$, the following equation is gained as

$$\begin{aligned} \frac{\varepsilon(r) - \varepsilon(r-1)}{T} &= \frac{1}{T^{1-\alpha}} \sum_{m=0}^r (-1)^m b_m^{1-\alpha} \varphi(\varepsilon(r-m), u(r-m)) + \\ &\quad \frac{1}{T^{1-\alpha}} \sum_{m=0}^r (-1)^m b_m^{1-\alpha} Gw(r-m), \end{aligned} \quad (26)$$

where the absolute value of coefficient for the term $Gw(r-m)$ is related to $b_m^{1-\alpha}$. Since the noise term is also related to the historical information, it is very difficult to design the Kalman filters and need to be simplified. The coefficient $b_m^{1-\alpha}$ decreases rapidly with the increases of m , $b_m^{1-\alpha}$ tends to zero if m is very large. We retain the two terms $w(r)$ and $w(r-1)$, then the equation (26) is approximately by

$$\begin{aligned} \varepsilon(r) &= T^\alpha \sum_{m=0}^r (-1)^m b_m^{1-\alpha} \varphi(\varepsilon(r-m), u(r-m)) + \\ &\quad T^\alpha \sum_{m=0}^1 (-1)^m b_m^{1-\alpha} Gw(r-m) + \varepsilon(r-1) \\ &= T^\alpha \varphi(\varepsilon(r), u(r)) + T^\alpha (\alpha - 1) \varphi(\varepsilon(r-1), u(r-1)) + \\ &\quad \varepsilon(r-1) + T^\alpha \sum_{m=2}^r (-1)^m b_m^{1-\alpha} \varphi(\varepsilon(r-m), u(r-m)) + \\ &\quad T^\alpha Gw(r) + T^\alpha (\alpha - 1) Gw(r-1). \end{aligned} \quad (27)$$

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

$$\varphi(\varepsilon(r), u(r)) \approx \varphi(\hat{\varepsilon}(r-1|r-1), u(r)) + C(r-1)(\varepsilon(r) - \hat{\varepsilon}(r-1|r-1)), \quad (28)$$

where $C(r-1) = \frac{\partial \varphi(\varepsilon, u(r))}{\partial \varepsilon} \Big|_{\varepsilon = \hat{\varepsilon}(r-1|r-1)}$.

For $m \geq 2$, the first-order Taylor expansion of the nonlinear function $\varphi(\varepsilon(r-m), u(r-m))$ at $\varepsilon(r-m) = \hat{\varepsilon}(r-m|r-m)$ is represented by

$$\begin{aligned} \varphi(\varepsilon(r-m), u(r-m)) &\approx \varphi(\hat{\varepsilon}(r-m|r-m), u(r-m)) + \\ &\quad N(r-m)(\varepsilon(r-m) - \hat{\varepsilon}(r-m|r-m)), \end{aligned} \quad (29)$$

where $N(r-m) = \frac{\partial \varphi(\varepsilon, u(r-m))}{\partial \varepsilon} \Big|_{\varepsilon = \hat{\varepsilon}(r-m|r-m)}$.

The initial values $C(-2)$ and $C(-1)$ are expressed by $C(-2) = C(-1) = \frac{\partial \varphi(\varepsilon, u(0))}{\partial \varepsilon} \Big|_{\varepsilon=0}$.

Substituting equations (28) and (29) into equation (27), then the following equation holds

$$\begin{aligned} \varepsilon(r) = & T^\alpha \varphi(\hat{\varepsilon}(r-1|r-1), u(r)) + T^\alpha C(r-1)(\varepsilon(r) - \hat{\varepsilon}(r-1|r-1)) + \\ & T^\alpha (\alpha - 1) \varphi(\varepsilon(r-1), u(r-1)) + \varepsilon(r-1) + T^\alpha \sum_{m=2}^r (-1)^m b_m^{1-\alpha} \times \\ & \varphi(\hat{\varepsilon}(r-m|r-m), u(r-m)) + T^\alpha \sum_{m=2}^r (-1)^m b_m^{1-\alpha} N(r-m) \times \\ & (\varepsilon(r-m) - \hat{\varepsilon}(r-m|r-m)) + T^\alpha Gw(r) + T^\alpha (\alpha - 1) Gw(r-1). \end{aligned}$$

It follows that

$$\begin{aligned} \varepsilon(r) = & D(r) \varphi(\hat{\varepsilon}(r-1|r-1), u(r)) + \tilde{C}(r-1) \hat{\varepsilon}(r-1|r-1) + \\ & \Omega(\varepsilon(r-1), u(r-1)) + \sum_{m=2}^r D(r) (-1)^m b_m^{1-\alpha} \times \\ & \varphi(\hat{\varepsilon}(r-m|r-m), u(r-m)) + \sum_{m=2}^r \mathcal{L}_m(r) (\varepsilon(r-m) - \\ & \hat{\varepsilon}(r-m|r-m)) + D(r) Gw(r) + J(r) Gw(r-1), \end{aligned} \quad (30)$$

where $D(r) = L(r)T^\alpha$, $\tilde{C}(r-1) = -D(r)C(r-1)$, the nonlinear function $\Omega(\varepsilon(r-1), u(r-1)) = J(r)\varphi(\varepsilon(r-1), u(r-1)) + L(r)\varepsilon(r-1)$, $\mathcal{L}_m(r) = D(r)(-1)^m b_m^{1-\alpha} N(r-m)$, $J(r) = D(r)(\alpha - 1)$ 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 term of the equation (30) and the assumption in the equation (13), we obtain

$$\begin{aligned} \hat{\varepsilon}(r|r-1) = & E[D(r) \varphi(\hat{\varepsilon}(r-1|r-1), u(r)) + \tilde{C}(r-1) \hat{\varepsilon}(r-1|r-1) + \\ & \Omega(\varepsilon(r-1), u(r-1)) + \sum_{m=2}^r D(r) (-1)^m b_m^{1-\alpha} \times \\ & \varphi(\hat{\varepsilon}(r-m|r-m), u(r-m)) + \sum_{m=2}^r \mathcal{L}_m(r) (\varepsilon(r-m) - \\ & \hat{\varepsilon}(r-m|r-m)) + D(r) Gw(r) + J(r) Gw(r-1) | \lambda(r-1)]. \end{aligned}$$

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

$$\begin{aligned} \hat{\varepsilon}(r|r-1) = & \int_{\mathbb{R}^n} \Omega(\varphi(r-1), u(r-1)) \mathcal{N}(\varepsilon(r-1); \hat{\varepsilon}(r-1|r-1), \\ & M(r-1|r-1)) d\varepsilon(r-1) + D(r) \varphi(\hat{\varepsilon}(r-1|r-1), u(r)) + \\ & \tilde{C}(r-1) \hat{\varepsilon}(r-1|r-1) + \sum_{m=2}^r D(r) (-1)^m b_m^{1-\alpha} \times \\ & \varphi(\hat{\varepsilon}(r-m|r-m), u(r-m)). \end{aligned}$$

The nonlinear function $\hat{\Omega}(\hat{\varepsilon}(r-1|r-1), u(r-1))$ is defined as

$$\begin{aligned} & \hat{\Omega}(\hat{\varepsilon}(r-1|r-1), u(r-1)) \\ &= \int_{\mathbb{R}^n} \Omega(\varepsilon(r-1), u(r-1)) \mathcal{N}(\varepsilon(r-1); \hat{\varepsilon}(r-1|r-1), M(r-1|r-1)) d\varepsilon(r-1). \end{aligned}$$

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 $\varepsilon(r) = \hat{\varepsilon}(r|r-1)$ is represented by

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

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

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

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

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

$$\varepsilon(r) - \hat{\varepsilon}(r|r) \approx (I - K(r)Y(r))(D(r)Gw(r) + J(r)Gw(r-1)) - K(r)v(r). \quad (31)$$

If $i \neq m$, the following equation holds

$$\begin{aligned} & E[(\sum_{i=2}^r \mathcal{L}_i(r)(\varepsilon(r-i) - \hat{\varepsilon}(r-i|r-i))) \times \\ & \quad (\sum_{m=2}^r \mathcal{L}_m(r)(\varepsilon(r-m) - \hat{\varepsilon}(r-m|r-m)))^T] \\ & \approx E[(\sum_{i=2}^r \mathcal{L}_i(r)((I - K(r-i)Y(r-i))(D(r-i)Gw(r-i) + \\ & \quad J(r-i)Gw(r-i-1)) - K(r-i)v(r-i)))(\sum_{m=2}^r \mathcal{L}_m(r) \times \\ & \quad ((I - K(r-m)Y(r-m))(D(r-m)Gw(r-m) + \\ & \quad J(r-m)Gw(r-m-1)) - K(r-m)v(r-m)))^T] \\ & = \sum_{m=2}^{r-1} \tilde{\mathcal{L}}_m(r) J(r-m) G Q G^T D^T(r-m-1) \tilde{\mathcal{L}}_{m+1}^T(r) + \\ & \quad \sum_{m=2}^{r-1} \tilde{\mathcal{L}}_{m+1}(r) D(r-m-1) G Q G^T J^T(r-m) \tilde{\mathcal{L}}_m^T(r), \end{aligned}$$

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

In term of the equation (31), we gain

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

If $i = m$, we gain

$$\begin{aligned} & E[(\sum_{m=2}^r \mathcal{L}_m(r)(\varepsilon(r-m) - \hat{\varepsilon}(r-m|r-m))) \times \\ & (\sum_{m=2}^r \mathcal{L}_m(r)(\varepsilon(r-m) - \hat{\varepsilon}(r-m|r-m)))^T] \\ & = \sum_{m=2}^r \mathcal{L}_m(r)M(r-m|r-m)\mathcal{L}_m^T(r), \end{aligned}$$

and

$$\begin{aligned} & E[(D(r)Gw(r) + J(r)Gw(r-1))(D(r)Gw(r) + J(r)Gw(r-1))^T] \\ & = D(r)GQG^T D^T(r) + J(r)GQG^T J^T(r). \end{aligned}$$

Meanwhile, we also obtain

$$\begin{aligned} & E[(\Omega(\varepsilon(r-1), u(r-1)) - \hat{\Omega}(\hat{\varepsilon}(r-1|r-1), u(r-1))) \\ & (\Omega(\varepsilon(r-1), u(r-1)) - \hat{\Omega}(\hat{\varepsilon}(r-1|r-1), u(r-1)))^T] \\ & = \int_{\mathbb{R}^n} \Omega(\varepsilon(r-1), u(r-1))\Omega^T(\varepsilon(r-1), u(r-1)) \times \\ & \mathcal{N}(\varepsilon(r-1); \hat{\varepsilon}(r-1|r-1), M(r|r-1)) d\varepsilon(r-1) - \\ & \hat{\Omega}(\hat{\varepsilon}(r-1|r-1), u(r-1))\hat{\Omega}^T(\hat{\varepsilon}(r-1|r-1), u(r-1)). \end{aligned}$$

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

$$\begin{aligned} \Omega(\varepsilon(r-1), u(r-1)) & \approx \Omega(\hat{\varepsilon}(r-1|r-1), u(r-1)) + \\ & F(r-1)(\varepsilon(r-1) - \hat{\varepsilon}(r-1|r-1)), \end{aligned}$$

where $F(r-1) = \frac{\partial \Omega(\varepsilon, u(r-1))}{\partial \varepsilon} \Big|_{\varepsilon = \hat{\varepsilon}(r-1|r-1)}$.

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

$$\begin{aligned} & E[(\Omega(\varepsilon(r-1), u(r-1)) - \hat{\Omega}(\hat{\varepsilon}(r-1|r-1), u(r-1))) \times \\ & (D(r)Gw(r) + J(r)Gw(r-1))^T] \\ & \approx E[(\Omega(\hat{\varepsilon}(r-1|r-1), u(r-1)) + F(r-1)(\varepsilon(r-1) - \hat{\varepsilon}(r-1|r-1)) - \\ & \hat{\Omega}(\hat{\varepsilon}(r-1|r-1), u(r-1)))(D(r)Gw(r) + J(r)Gw(r-1))^T] \\ & \approx E[(F(r-1)((I - K(r-1))Y(r-1))(D(r-1)Gw(r-1) + J(r-1)G \times \\ & w(r-2) - K(r-1)v(r-1)))(D(r)Gw(r) + J(r)Gw(r-1))^T] \\ & = H(r-1)D(r-1)GQG^T J^T(r), \end{aligned}$$

and

$$\begin{aligned}
& E[(\Omega(\varepsilon(r-1), u(r-1)) - \hat{\Omega}(\hat{\varepsilon}(r-1|r-1), u(r-1))) \times \\
& \quad (\sum_{m=2}^r \mathcal{L}_m(r)(\varepsilon(r-m) - \hat{\varepsilon}(r-m|r-m)))^T] \\
& \approx E[(F(r-1)((I - K(r-1)Y(r-1))(D(r-1)Gw(r-1) + J(r-1)G \times \\
& \quad w(r-2)) - K(r-1)v(r-1)))(\sum_{m=2}^r \mathcal{L}_m(r)((I - K(r-m)Y(r-m)) \times \\
& \quad (D(r-m)Gw(r-m) + J(r-m)Gw(r-m-1)) - K(r-m)v(r-m)))^T] \\
& = H(r-1)J(r-1)GQG^T D^T(r-2)\tilde{\mathcal{L}}_2^T(r),
\end{aligned}$$

where $H(r-1) = F(r-1)(I - K(r-1)Y(r-1))$.

The prediction error matrix $M(r|r-1)$ is approximated by

$$\begin{aligned}
& M(r|r-1) \\
& = \int_{\mathbb{R}^n} \Omega(\varepsilon(r-1), u(r-1))\Omega^T(\varepsilon(r-1), u(r-1)) \times \\
& \quad \mathcal{N}(\varepsilon(r-1); \hat{\varepsilon}(r-1|r-1), M(r-1|r-1))d\varepsilon(r-1) - \\
& \quad \hat{\Omega}(\hat{\varepsilon}(r-1|r-1), u(r-1))\hat{\Omega}^T(\hat{\varepsilon}(r-1|r-1), u(r-1)) + \\
& \quad \sum_{m=2}^r \mathcal{L}_m(r)M(r-m|r-m)\mathcal{L}_m^T(r) + D(r)GQG^T D^T(r) + \\
& \quad J(r)GQG^T J^T(r) + \sum_{m=2}^{r-1} \tilde{\mathcal{L}}_m(r)J(r-m)GQG^T D^T(r-m-1)\tilde{\mathcal{L}}_{m+1}^T(r) + \\
& \quad \sum_{m=2}^{r-1} \tilde{\mathcal{L}}_{m+1}(r)D(r-m-1)GQG^T J^T(r-m)\tilde{\mathcal{L}}_m^T(r) + H(r-1)D(r-1) \times \\
& \quad GQG^T J^T(r) + J(r)GQG^T D^T(r-1)H^T(r-1) + H(r-1)J(r-1) \times \\
& \quad GQG^T D^T(r-2)\tilde{\mathcal{L}}_2^T(r) + \tilde{\mathcal{L}}_2(r)D(r-2)GQG^T J^T(r-1)H^T(r-1).
\end{aligned}$$

Similar to equations (15) and (16), the updated formulas for $\hat{\varepsilon}(r|r)$ and $M(r|r)$ are showed by

$$\hat{\varepsilon}(r|r) = \hat{\varepsilon}(r|r-1) + K(r)(\eta(r) - \hat{\eta}(r|r-1)), \quad (32)$$

and

$$M(r|r) = M(r|r-1) - K(r)Z(r)K^T(r). \quad (33)$$

Being similar to the cubature points in (17) and (20), we define the cubature points as follows

$$\chi^{(i)}(r-1|r-1) = \hat{\varepsilon}(r-1|r-1) + \sqrt{M(r-1|r-1)}\gamma^{(i)},$$

and

$$\chi^{(i)}(r|r-1) = \hat{\varepsilon}(r|r-1) + \sqrt{M(r|r-1)}\gamma^{(i)}.$$

Then, the prediction $\hat{\varepsilon}(r|r-1)$, the prediction error matrix $M(r|r-1)$, $\hat{\eta}(r|r-1)$, $P(r)$ and $Z(r)$ are approximately represented by

$$\begin{aligned} \hat{\varepsilon}(r|r-1) \approx & \frac{1}{2n} \sum_{i=1}^{2n} \Omega(\chi^{(i)}(r-1|r-1), u(r-1)) + \\ & D(r)\varphi(\hat{\varepsilon}(r-1|r-1), u(r)) + \tilde{C}(r-1)\hat{\varepsilon}(r-1|r-1) + \\ & \sum_{m=2}^r D(r)(-1)^m b_m^{1-\alpha} \varphi(\hat{\varepsilon}(r-m|r-m), u(r-m)), \end{aligned} \quad (34)$$

and

$$\begin{aligned} M(r|r-1) \approx & \frac{1}{2n} \sum_{i=1}^{2n} \Omega(\chi^{(i)}(r-1|r-1), u(r-1)) \times \\ & \Omega^T(\chi^{(i)}(r-1|r-1), u(r-1)) - \\ & \left[\frac{1}{2n} \sum_{i=1}^{2n} \Omega(\chi^{(i)}(r-1|r-1), u(r-1)) \right] \times \\ & \left[\frac{1}{2n} \sum_{i=1}^{2n} \Omega(\chi^{(i)}(r-1|r-1), u(r-1)) \right]^T + \\ & \sum_{m=2}^r \mathcal{L}_m(r) M(r-m|r-m) \mathcal{L}_m^T(r) + \\ & D(r) G Q G^T D^T(r) + J(r) G Q G^T J^T(r) + \\ & \sum_{m=2}^{r-1} \tilde{\mathcal{L}}_m(r) J(r-m) G Q G^T D^T(r-m-1) \tilde{\mathcal{L}}_{m+1}^T(r) + \\ & \sum_{m=2}^{r-1} \tilde{\mathcal{L}}_{m+1}(r) D(r-m-1) G Q G^T J^T(r-m) \tilde{\mathcal{L}}_m^T(r) + \\ & H(r-1) D(r-1) G Q G^T J^T(r) + \\ & J(r) G Q G^T D^T(r-1) H^T(r-1) + \\ & H(r-1) J(r-1) G Q G^T D^T(r-2) \tilde{\mathcal{L}}_2^T(r) + \\ & \tilde{\mathcal{L}}_2(r) D(r-2) G Q G^T J^T(r-1) H^T(r-1). \end{aligned} \quad (35)$$

$$\hat{\eta}(r|r-1) \approx \frac{1}{2n} \sum_{i=1}^{2n} y(\chi^{(i)}(r|r-1)), \quad (36)$$

$$\begin{aligned} P(r) \approx & \frac{1}{2n} \sum_{i=1}^{2n} \chi^{(i)}(r|r-1) y^T(\chi^{(i)}(r|r-1)) - \hat{\varepsilon}(r|r-1) \times \\ & \left[\frac{1}{2n} \sum_{i=1}^{2n} y(\chi^{(i)}(r|r-1)) \right]^T, \end{aligned} \quad (37)$$

and

$$\begin{aligned} Z(r) \approx & \frac{1}{2n} \sum_{i=1}^{2n} y(\chi^{(i)}(r|r-1)) y^T(\chi^{(i)}(r|r-1)) - \\ & \left[\frac{1}{2n} \sum_{i=1}^{2n} y(\chi^{(i)}(r|r-1)) \right] \left[\frac{1}{2n} \sum_{i=1}^{2n} y(\chi^{(i)}(r|r-1)) \right]^T + R. \end{aligned} \quad (38)$$

Algorithm 2. The fractional-order CKF for the nonlinear FOS (8) with the measurement output (9) is implemented by the following steps.

Prediction:

Step one. Calculate the cubature point $\chi^{(i)}(r-1|r-1)$ for $i = 1, 2, \dots, 2n$.

Step two. Compute the approximations (34) and (35) for $\hat{\varepsilon}(r|r-1)$ and $M(r|r-1)$.

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

Updating:

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

Step five. Utilize $\hat{\eta}(r|r-1)$ and $\hat{\varepsilon}(r|r-1)$ to ensure $K(r)$ by $P(r)$ and $Z(r)$ based on (37) and (38).

Step six. Compute $\hat{\varepsilon}(r|r)$ 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, $\hat{\varepsilon}(r|r-1)$ and $M(r|r-1)$ with a truncation L are modified by

$$\begin{aligned} \hat{\varepsilon}(r|r-1) &\approx \frac{1}{2n} \sum_{i=1}^{2n} \Omega(\chi^{(i)}(r-1|r-1), u(r-1)) + \\ &D(r)\varphi(\hat{\varepsilon}(r-1|r-1), u(r)) + \tilde{C}(r-1)\hat{\varepsilon}(r-1|r-1) + \\ &\sum_{m=2}^L D(r)(-1)^m b_m^{1-\alpha} \varphi(\hat{\varepsilon}(r-m|r-m), u(r-m)), \end{aligned}$$

$$\begin{aligned} M(r|r-1) &\approx \frac{1}{2n} \sum_{i=1}^{2n} \Omega(\chi^{(i)}(r-1|r-1), u(r-1)) \times \\ &\Omega^T(\chi^{(i)}(r-1|r-1), u(r-1)) - \\ &[\frac{1}{2n} \sum_{i=1}^{2n} \Omega(\chi^{(i)}(r-1|r-1), u(r-1))] \times \\ &[\frac{1}{2n} \sum_{i=1}^{2n} \Omega(\chi^{(i)}(r-1|r-1), u(r-1))]^T + \\ &\sum_{m=2}^L \mathcal{L}_m(r)M(r-m|r-m)\mathcal{L}_m^T(r) + \\ &D(r)GQG^T D^T(r) + J(r)GQG^T J^T(r) + \\ &\sum_{m=2}^{L-1} \tilde{\mathcal{L}}_m(r)J(r-m)GQG^T D^T(r-m-1)\tilde{\mathcal{L}}_{m+1}^T(r) + \\ &\sum_{m=2}^{L-1} \tilde{\mathcal{L}}_{m+1}(r)D(r-m-1)GQG^T J^T(r-m)\tilde{\mathcal{L}}_m^T(r) + \\ &H(r-1)D(r-1)GQG^T J^T(r) + \\ &J(r)GQG^T D^T(r-1)H^T(r-1) + \\ &H(r-1)J(r-1)GQG^T D^T(r-2)\tilde{\mathcal{L}}_2^T(r) + \\ &\tilde{\mathcal{L}}_2(r)D(r-2)GQG^T J^T(r-1)H^T(r-1). \end{aligned}$$

The CKF and HECKF are proposed to weaken the initial value influence on state estimation for the nonlinear FOS (8) compared with Algorithm 1 based on the CKF in [32].

4 Numerical examples

Example 1. Considering the following nonlinear FOS

$$\begin{cases} {}_0^C D_t^\alpha \varepsilon_1(t) = \varepsilon_2(t) + w_1(t) + u_1(t) \\ {}_0^C D_t^\alpha \varepsilon_2(t) = -2\varepsilon_1(t) - 2\varepsilon_2(t) + 0.5\sin(\varepsilon_2(t))\varepsilon_1(t) + w_2(t) + u_2(t) \end{cases}, \quad (39)$$

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 0.1s, and the running time is 40s. The initial value $\varepsilon(0)$ in the 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 Gaussian noises, and satisfy $E[w(r)] = \mathbf{0}$, $E[v(r)] = 0$, $E[w(r)w^T(s)] = Q\xi(r-s)$, $E[v(r)v^T(s)] = R\xi(r-s)$, $Q = \text{diag}(0.001, 0.001)$ and $R = 1$.

The nonlinear functions $\varphi_1(\varepsilon(t), u(t))$ and $\varphi_2(\varepsilon(t), u(t))$ are $\varphi_1(\varepsilon(t), u(t)) = \varphi_2(t) + u_1(t)$, $\varphi_2(\varepsilon(t), u(t)) = -2\varepsilon_1(t) - 2\varepsilon_2(t) + 0.5\sin(\varepsilon_2(t))\varepsilon_1(t) + u_2(t)$. 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 $\bar{\varphi}(\zeta(r-1), u(r-1)) = [T^\alpha \varphi_1(\varepsilon(r-1), u(r-1)), T^\alpha \varphi_2(\varepsilon(r-1), u(r-1)), 0, 0]^T$, and the augmented noise is $\bar{w}(r-1) = [w_1(r-1), w_2(r-1), \tilde{w}_1(r-1), \tilde{w}_2(r-1)]^T$ with $\tilde{w}(r-1) = [\tilde{w}_1(r-1), \tilde{w}_2(r-1)]^T$. The initial estimate vector $\hat{\zeta}(0|0)$ and error matrix $M(0|0)$ of $\zeta(r)$ are set as $\hat{\zeta}(0|0) = [0, 0, 0, 0]^T$ and $M(0|0) = I$.

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

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

where $S+1$ is the number of the sampling values of $\varepsilon(r)$.

Let the truncation L be $L = 25, \dots, 50$ with the step value 5, the fractional-order α is set from 0.2 to 0.8 with the step value 0.1, hence the estimation error E for Algorithm 1 is shown in Table 1.

Table 1 Relationship between E and L with different fractional-orders α via Algorithm 1

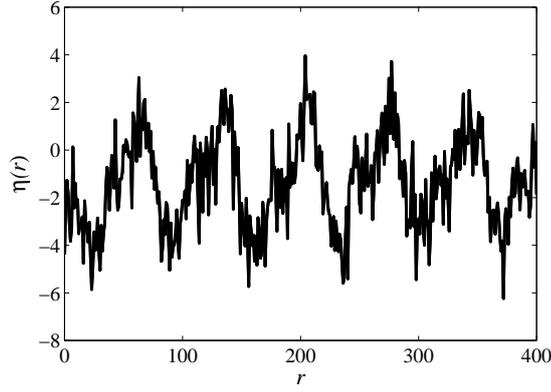


Fig. 2 Sampling value of measurement $\eta(r)$ in Example 1

L	α						
	0.2	0.3	0.4	0.5	0.6	0.7	0.8
25	0.2096	0.1431	0.1276	0.1185	0.1138	0.1081	0.0934
30	0.2047	0.1408	0.1248	0.1148	0.1081	0.1006	0.0880
35	0.1959	0.1337	0.1174	0.1069	0.0996	0.0920	0.0814
40	0.1857	0.1239	0.1073	0.0967	0.0896	0.0829	0.0745
45	0.1756	0.1135	0.0965	0.0858	0.0794	0.0743	0.0684
50	0.1672	0.1037	0.0863	0.0758	0.0704	0.0671	0.0639

Hence, with the the truncation L increasing, the estimation error E of the system (39) has a tendency to increase for the fractional-order α set from 0.2 to 0.8 with the step value 0.1. However, the calculation burden will be increased by a larger truncation L , to balance between the estimation accuracy and the calculation complexity, hence an appropriate truncation L is selected in practical engineering. Therefore, we select $L = 35$ and $\alpha=0.3$, Fig.2 draws the sampling value of $\eta(r)$, Fig.3 and Fig.4 respectively offer the real values and the estimations of $\varepsilon_1(r)$ and $\varepsilon_2(r)$.

Fig.3 and Fig.4 show that the proposed Algorithm 1 based on the CKF can more effectively estimate the state information compared with the Algorithm 1 based on the CKF without the initial value compensation in [32] of the FOS (39).

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, \dots , 0.8, and the truncation $L=35$, then the error indexes under the indexes 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

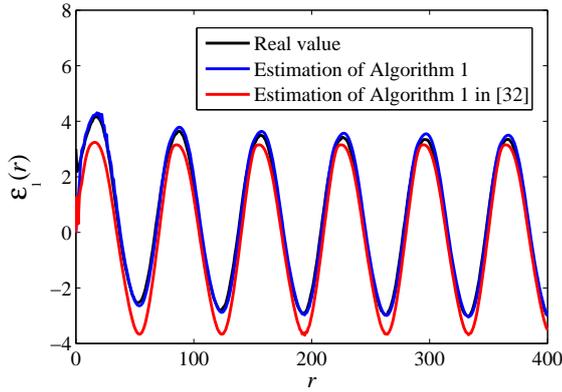


Fig. 3 Estimation and real value of $\varepsilon_1(r)$ in Example 1

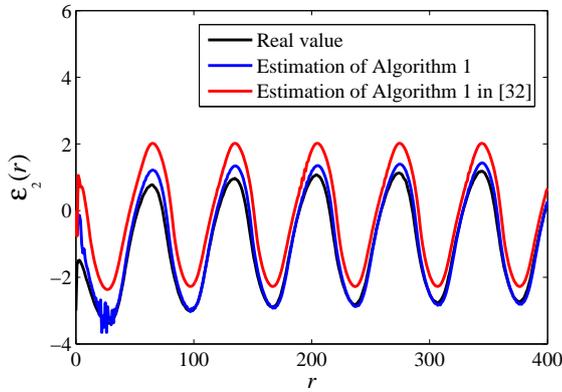


Fig. 4 Estimation and real value of $\varepsilon_2(r)$ in Example 1

and T_2 are estimation error and calculation time of the proposed Algorithm 1 in this paper, 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 $\alpha=0.2, 0.3, \dots, 0.8$, but the longer time is taken for calculating. Besides, the difference between E_1 and E_2 of state estimation error Algorithm 1 and the Algorithm 1 in [32] decreases with increasing of fractional-order α , 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 continuous-time nonlinear FOSs 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 we consider calculation time of state estimation.

Table 2 Comparison results of state estimation error for different fractional-orders α

α	E_1	E_2	$E_1 - E_2$	$T_1(/s)$	$T_2(/s)$
0.2	0.6980	0.1959	0.5021	4.2052	4.5786
0.3	0.4627	0.1337	0.3290	3.3881	4.1696
0.4	0.3052	0.1174	0.1878	3.0818	3.7685
0.5	0.2003	0.1069	0.0934	2.9652	3.7757
0.6	0.1426	0.0996	0.0430	2.5884	3.5994
0.7	0.1114	0.0920	0.0194	2.7013	3.6801
0.8	0.0872	0.0814	0.0058	2.6979	3.2418

Example 1 is given to study the state estimation for nonlinear FOS (39) to verify the validity of the proposed Algorithm 1 based on the CKF. Meanwhile, the state estimation of 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) to testify 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 2 Considering the following nonlinear FOS

$$\begin{cases} {}_0^C D_t^\alpha \varepsilon_1(t) = \varepsilon_2(t) + w_1(t) + u_1(t) \\ {}_0^C D_t^\alpha \varepsilon_2(t) = -1.5\varepsilon_1(t) - 1.5\varepsilon_2(t) + 0.5\sin(\varepsilon_2(t))\varepsilon_1(t) + w_2(t) + u_2(t) \end{cases}, \quad (40)$$

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 the FOS (40) is set as $\varepsilon(0) = [3, -3]^T$. The sampling period T is 0.1s, and the running time is 30s. The process noises $w(r) = [w_1(r), w_2(r)]^T$ and measurement noise $v(r)$ are uncorrelated Gaussian noises, and satisfy $E[w(r)] = \mathbf{0}$, $E[v(r)] = 0$, $E[w(r)w^T(s)] = Q\xi(r-s)$, $E[v(r)v^T(s)] = R\xi(r-s)$, $Q = \text{diag}(0.001, 0.001)$ and $R = 0.5$. The initial estimate vector $\hat{\varepsilon}(0|0)$ and error matrix $M(0|0)$ of $\varepsilon(r)$ are set as $\hat{\varepsilon}(0|0) = [0, 0]^T$ and $M(0|0) = I$.

The truncation L is set as 25, 30, \dots , 50, the fractional-order α is set from 0.2 to 0.8 with the step value 0.1, then the error indexes E for Algorithm 2 is shown in Table 3.

Table 3 Relationship between E and L with different fractional-orders α via Algorithm 2

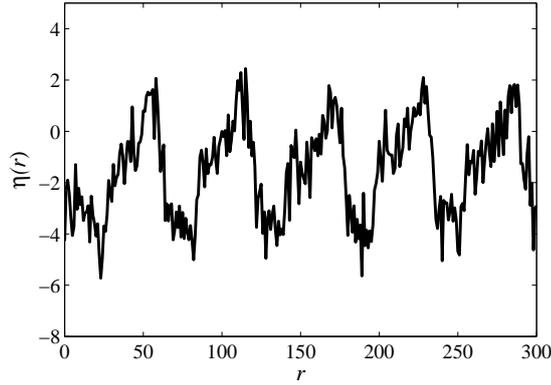


Fig. 5 Sampling value of measurement $\eta(r)$ in Example 2

L	α						
	0.2	0.3	0.4	0.5	0.6	0.7	0.8
25	0.2638	0.2335	0.1904	0.1407	0.1025	0.0946	0.0990
30	0.2211	0.2046	0.1738	0.1324	0.0989	0.0907	0.0949
35	0.1863	0.1793	0.1577	0.1229	0.0929	0.0860	0.0900
40	0.1577	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 3, the index E becomes smaller with increasing of the truncation L for the fractional-order α set from 0.2 to 0.8 with the step value 0.1. A small estimation error of the FOS (40) with a relatively small fractional-order α in $(0, 1)$ is still obtained by the proposed Algorithm 2 based on the HECKF. We select $L = 35$ and $\alpha = 0.3$, then Fig.5 offers the sampling value of $\eta(r)$ in Example 2. Meanwhile, Fig.6 and Fig.7 respectively draw the real values and the estimations of $\varepsilon_1(r)$ and $\varepsilon_2(r)$.

It can be seen that the proposed Algorithm 2 based on the HECKF can usually obtain more accurate state estimation to compare with Algorithm 1 in [32] from the curves of Fig.6 and Fig.7.

The estimation effect of the Algorithm in [32] and the proposed Algorithm 2 based on the HECKF in this paper is evaluated by the error indexes E . The fractional-order α is set as 0.2, 0.3, \dots , 0.8, the truncation $L=35$ and the error indexes under the indexes 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 the proposed Algorithm 2 in this paper, respectively.

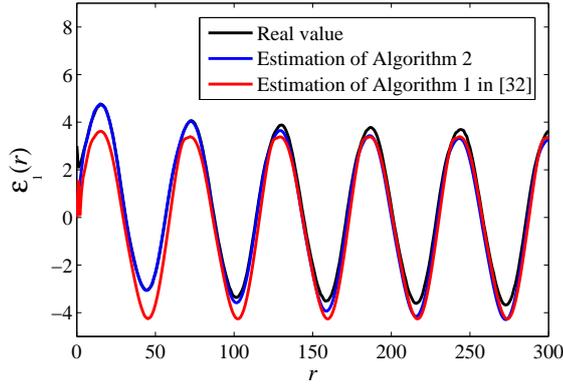


Fig. 6 Estimation and real value of $\varepsilon_1(r)$ in Example 2

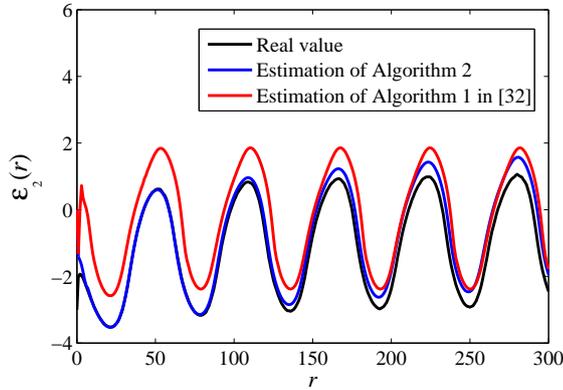


Fig. 7 Estimation and real value of $\varepsilon_2(r)$ in Example 2

From Table 4, it is obvious that the state estimation based on CKF by proposed Algorithm 2 is more accurate for different fractional-orders $\alpha = 0.2, 0.3, \dots, 0.8$, but it needs to take a longer time for calculation. In addition, the difference of state estimation error between Algorithm 2 and Algorithm 1 in [32] gradually decreases as the fractional-order α increases, a more accurate estimation effect of the 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 the FOS (40) with a relatively large fractional-order $\alpha \in (0, 1)$ if the calculation time of state estimation is taken into account.

Table 4 Comparison results of state estimation error for different fractional-orders α

α	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 Algorithm 1 and Algorithm 2 based on EKF 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 Algorithm 1 and Algorithm 2 compared to Algorithm 1 in [32] based on CKF without 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 running time is taken into account. The future work is to consider the initial value problem in the adaptive fractional-order CKF with unknown covariance matrices Q and R for nonlinear fractional-order systems.

Conflict of Interest

The authors declare that they have no conflict of interest.

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.

References

References

1. H. Sun, Y. Zhang, D. Baleanu, W. Chen, Y. Chen, "A new collection of real world applications of fractional calculus in science and engineering," *Communications in Nonlinear Science and Numerical Simulation*, 64(2018), pp. 213–231.
2. V. E. Tarasov, and V. V. Tarasova, "Macroeconomic models with long dynamic memory: Fractional calculus approach," *Applied Mathematics and Computation*, 338(2018), pp. 466–486.
3. B. Liang, S. Zheng, C. K. Ahn, and F. Liu. "Adaptive fuzzy control for fractional-order interconnected systems with unknown control directions," *IEEE Transactions on Fuzzy Systems*, DOI: 10.1109/TFUZZ.2020.3031694, 2020.
4. B. J. West, M. Turala, and P. Grigolini, "Fractional calculus ties the microscopic and macroscopic scales of complex network dynamics," *Physics*, 17(4)(2015), ID. 045009.
5. D. Zhao, and M. Luo, "Representations of acting processes and memory effects: General fractional derivative and its application to theory of heat conduction with finite wave speeds," *Applied Mathematics and Computation*, 346(2019), pp. 531–544.
6. G. Failla, and M. Zingales, "Advanced materials modelling via fractional calculus: challenges and perspectives," *Philosophical Transactions of The Royal Society A Mathematical Physical and Engineering Sciences*, 378(2172)(2020), ID. 20200050.
7. B. Boudjehem, and D. Boudjehem, "Fractional order controller design for desired response," *Proceedings of the Institution of Mechanical Engineers Part I Journal of Systems and Control Engineering*, 227(12)(2013), pp. 243–251.
8. H. Zhao, W. Deng, X. Yang, X. Li, and C. Dong, "An optimized fractional order PID controller for suppressing vibration of AC motor," *Journal of Vibroengineering*, 18(4)(2016), pp. 2205–2220.
9. V. Mohan, H. Chhabra, A. Rani, V. Singh. Robust self-tuning fractional order PID controller dedicated to non-linear dynamic system. *Journal of Intelligent & Fuzzy Systems*, 34(3)(2018), pp. 1467–1478.
10. S Pashaei, M Badamchizadeh. A new fractional-order sliding mode controller via a nonlinear disturbance observer for a class of dynamical systems with mismatched disturbances. *ISA Transactions*, 63(2016), pp. 39–48.
11. S A Moezi, E Zakeri, M Eghtesad. Optimal adaptive interval type-2 fuzzy fractional-order backstepping sliding mode control method for some classes of nonlinear systems. *ISA Transactions*, 93(2019), pp. 23–39.
12. S. Ntouskas, H. Sarimveis, P. Sopasakis. Model predictive control for offset-free reference tracking of fractional order systems. *Control Engineering Practice*, 71(2018), pp. 26–33.
13. A. Necaibia, S. Ladaci, A. Charef, J. J. Loiseau. Fractional order extremum seeking approach for maximum power point tracking of photovoltaic panels. *Frontiers in Energy*, 9(1)(2015), pp. 43–53.
14. S. A. Pasha, H. D. Tuan, B. N. Vo. Nonlinear Bayesian filtering using the unscented linear fractional transformation model. *IEEE Transactions on Signal Processing*, 58(2)(2010), pp. 477–489.
15. T. Liu, Y. Wei, W. Yin, Y. Wang, Q. Liang. State estimation for nonlinear discrete-time fractional systems a Bayesian perspective. *Signal Processing*, 165(2019), pp. 250–261.
16. L. Ljung Asymptotic behavior of the extended Kalman filter as a parameter estimator for linear systems. *Automatic Control IEEE Transactions on*, 24(1)(2013), pp. 36–50.
17. M. Charkhgard, M. Farrokhi. State-of-charge estimation for lithium-ion batteries using neural networks and EKF. *IEEE Transactions on Industrial Electronics*, 57(12)(2010), pp. 4178–4187.
18. X. Wang, Z. Xu, X. Gou, L. Trajkovic. Tracking a maneuvering target by multiple sensors using extended Kalman filter with nested probabilistic-numerical linguistic information. *IEEE Transactions on Fuzzy Systems*, 28(2)(2020), pp. 346–360.
19. F. Naseri, Z. Kazemi, E. Farjah, T. Ghanbari. Fast detection and compensation of current transformer saturation using extended Kalman filter. *IEEE Transactions on Power Delivery*, 34(3)(2019), pp. 1087–1097.

20. N. Zeng, Z. Wang, H. Zhang. Inferring nonlinear lateral flow immunoassay state-space models via an unscented Kalman filter. *SCIENCE CHINA Information Sciences*, 59(11)(2016), ID: 112204.
21. E. Wan, R. Van Der Merwe. The unscented Kalman filter. *Kalman Filtering and Neural Networks*, John Wiley & Sons, Inc. (2002).
22. K. Xiong, H. Zhang, C. Chan. Performance evaluation of UKF-based nonlinear filtering. *Automatica*, 42(2)(2006), pp. 261–270.
23. S. Sarkka. On. Unscented Kalman Filtering for State Estimation of Continuous-Time Nonlinear Systems. *IEEE Transactions on Automatic Control*, 52(9)(2007), pp. 1631–1641.
24. I. Arasaratnam, S. Haykin. Cubature Kalman filters. *IEEE Transactions on Automatic Control*, 54(6)(2009), pp. 1254–1269.
25. I. Arasaratnam, S. Haykin, T. R. Hurd. Cubature Kalman filtering for continuous-discrete systems: theory and simulations. *IEEE Transactions on Signal Processing*, 58(10)(2010), pp. 4977–4993.
26. B. Cui, X. Wei, X. Chen, J. Li, A. Wang. Robust cubature Kalman filter based on variational Bayesian and transformed posterior sigma points error. *ISA Transactions*, 86(2018), pp. 18–28.
27. S. Kolás, B. A. Foss, T. S. Schei. Constrained nonlinear state estimation based on the UKF approach. *Computers & Chemical Engineering*, 33(8)(2009), pp. 1386–1401.
28. F. Liu, Z. Gao, C. Yang, R. Ma. Fractional-order Kalman filters for continuous-time fractional-order systems involving correlated and uncorrelated process and measurement noises. *Transactions of the Institute of Measurement and Control*, 41(7)(2019), pp. 1933–1947.
29. C. Yang, Z. Gao, F. Liu. Kalman filters for linear continuous-time fractional-order systems involving colored noises using fractional-order average derivative. *IET Control Theory and Applications*, 12(4)(2018), pp. 456–465.
30. D. Sierociuk, A. D. Dzieliński. Fractional Kalman filter algorithm for the states, parameters and order of fractional system estimation. *International Journal of Applied Mathematics & Computer Science*, 16(1)(2006), pp. 129–140.
31. A. Ramezani, B. Safarinejadian. A modified fractional-order unscented Kalman filter for nonlinear fractional-order systems. *Circuits, Systems, and Signal Processing*, 37(9)(2018), pp. 3756–3784.
32. Z. Gao. Cubature Kalman filters for nonlinear continuous-time fractional-order systems with uncorrelated and correlated noises. *Nonlinear Dynamics*, 96(2019), pp. 1805–1817.
33. A. Ramezani, B. Safarinejadian, J. Zarei. Novel hybrid robust fractional interpolatory cubature Kalman filters. *Journal of the Franklin Institute*, 357(1)(2020), pp. 704–725.
34. C. Yang, Z. Gao, X. Huang, T. Kan. 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 & Applications*, 14(11)(2020), pp. 1424–1437.
35. X. Huang, Z. Gao, X. Chen. Extended Kalman filter for linear fractional-order systems with unknown fractional-order. *ICIC Express Letters*, 14(5)(2020), pp. 431–441.
36. I. Prodlubny. Fractional differential equation. *Academic Press*, New York, (1999).
37. M. C. Caputo, D. F. M. Torres. Duality for the left and right fractional derivatives. *Signal Processing*, 107(2015), pp. 265–271.
38. D. Sierociuk, A. D. Dzieliński. Fractional Kalman filter algorithm for the states, parameters and order of fractional system estimation. *International Journal of Applied Mathematics & Computer Science*, 1(16)(2006), pp. 129–140.

Figures

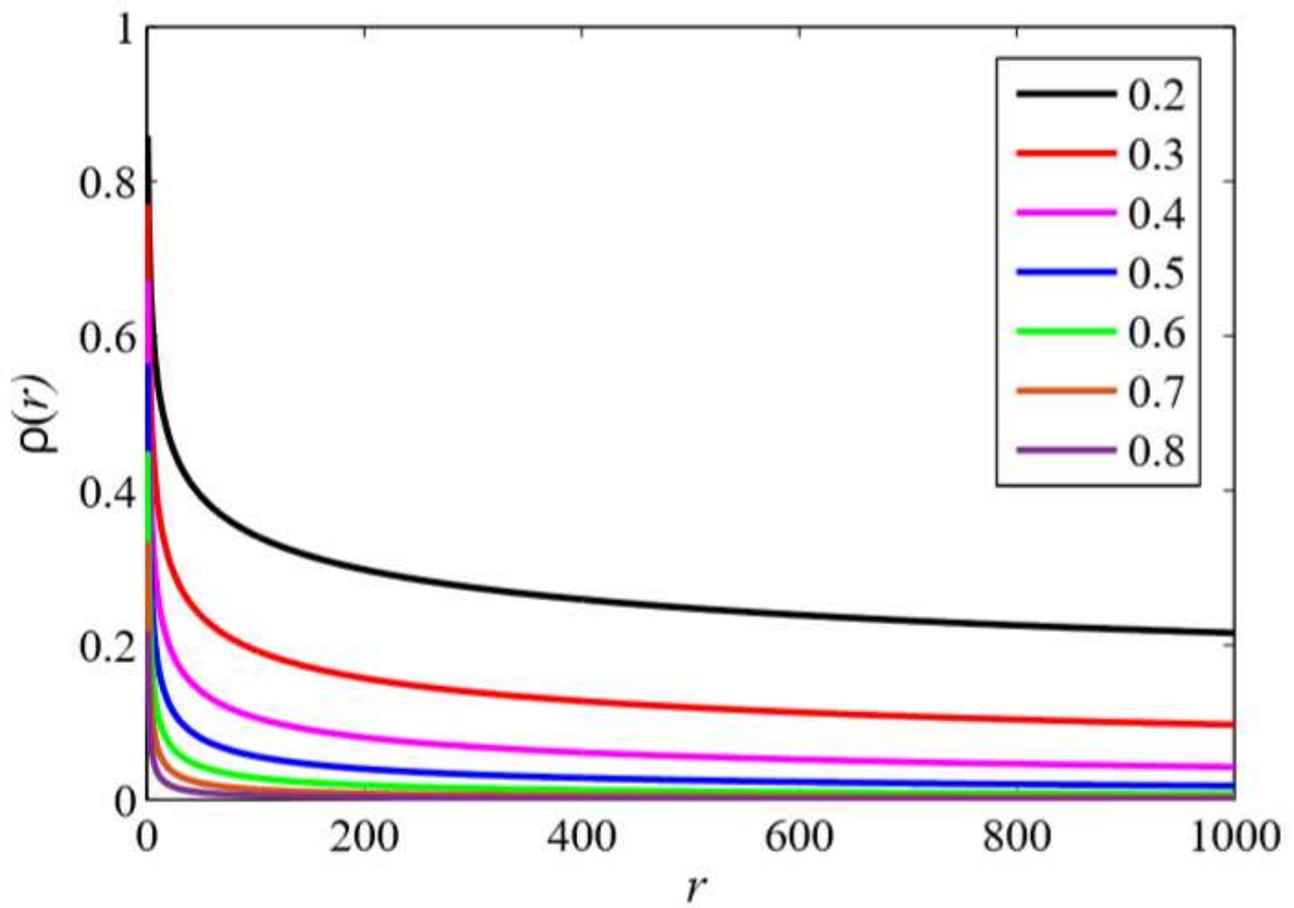


Figure 1

Curves of $\rho(r)$ for different fractional-order α .

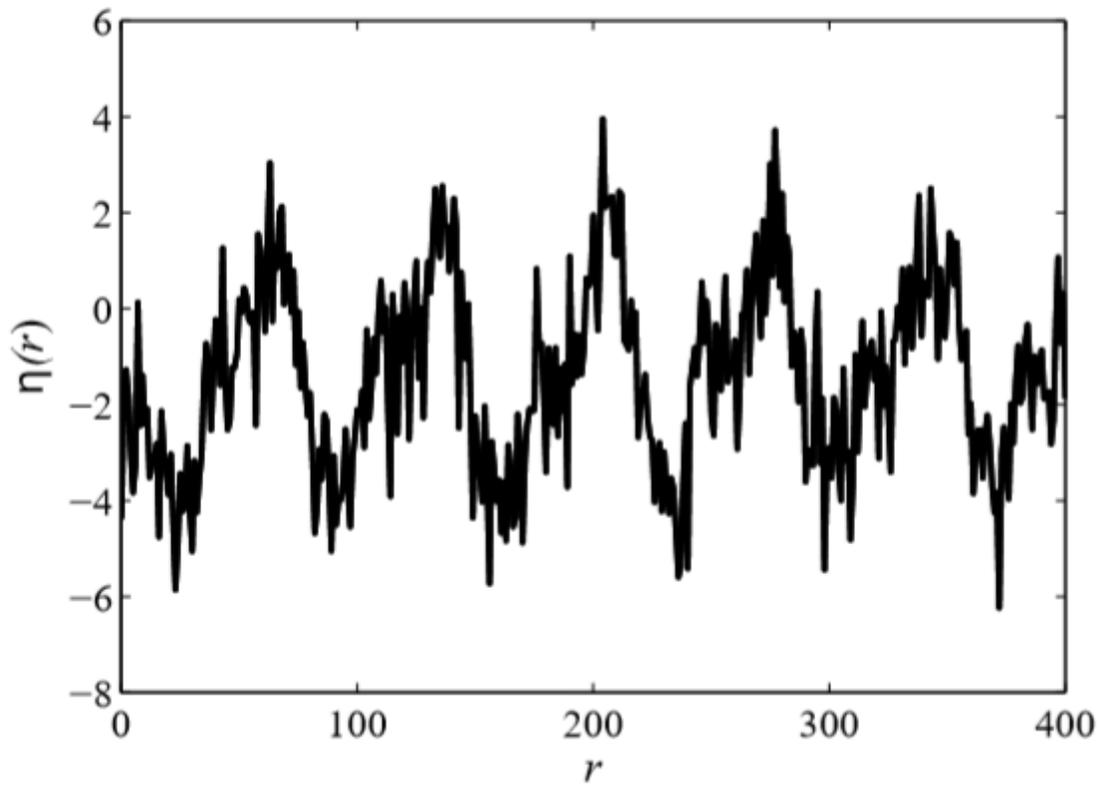


Figure 2

Sampling value of measurement $\eta(r)$ in Example 1

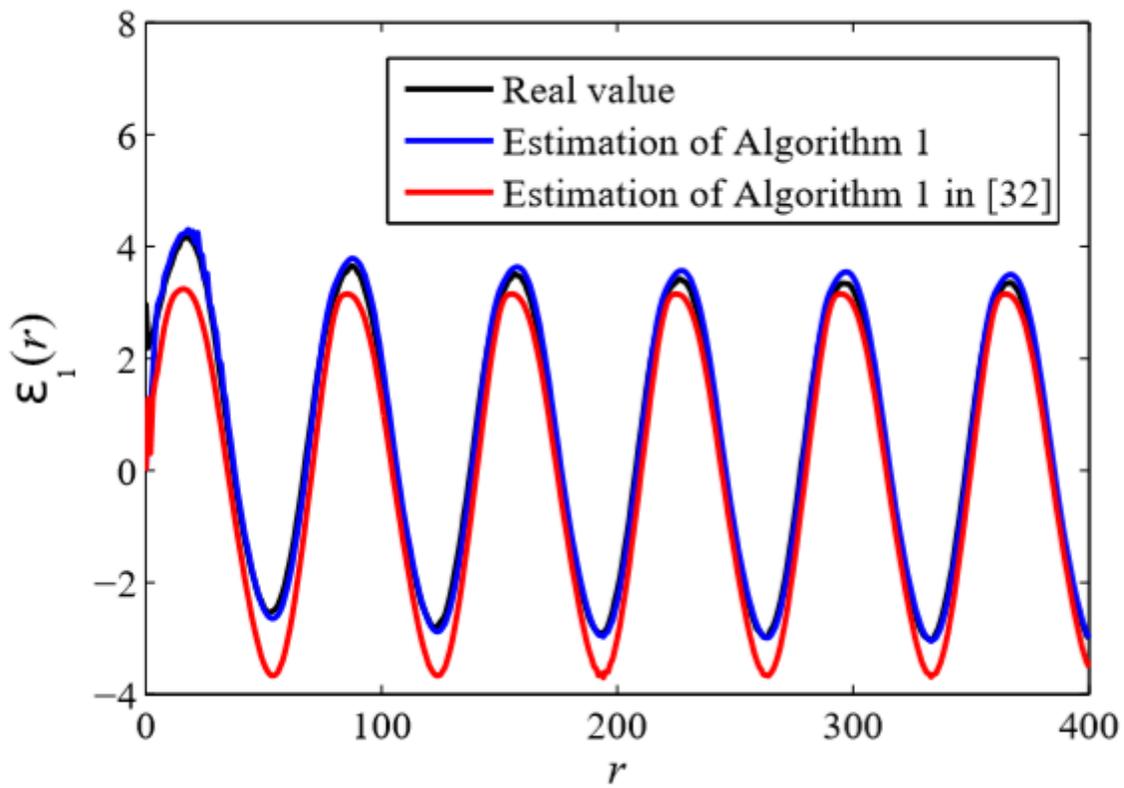


Figure 3

Estimation and real value of $\epsilon_1(r)$ in Example 1

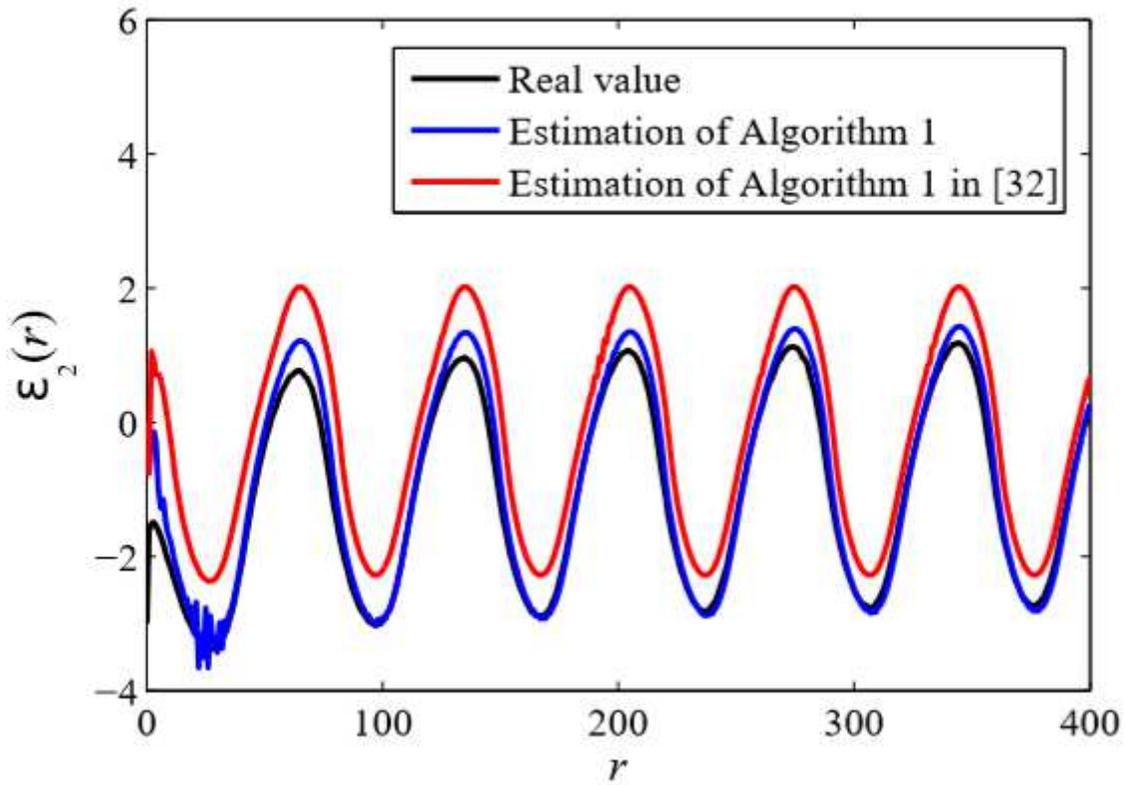


Figure 4

Estimation and real value of $\epsilon_2(r)$ in Example 1

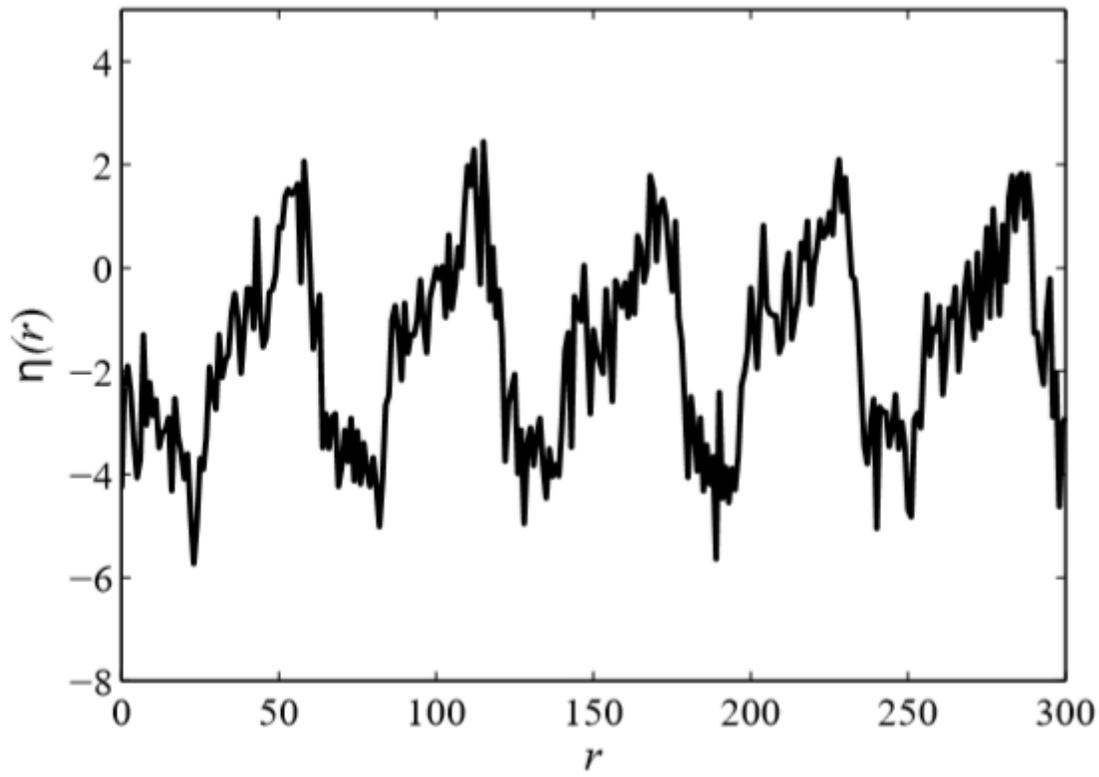


Figure 5

Sampling value of measurement $\eta(r)$ in Example 2

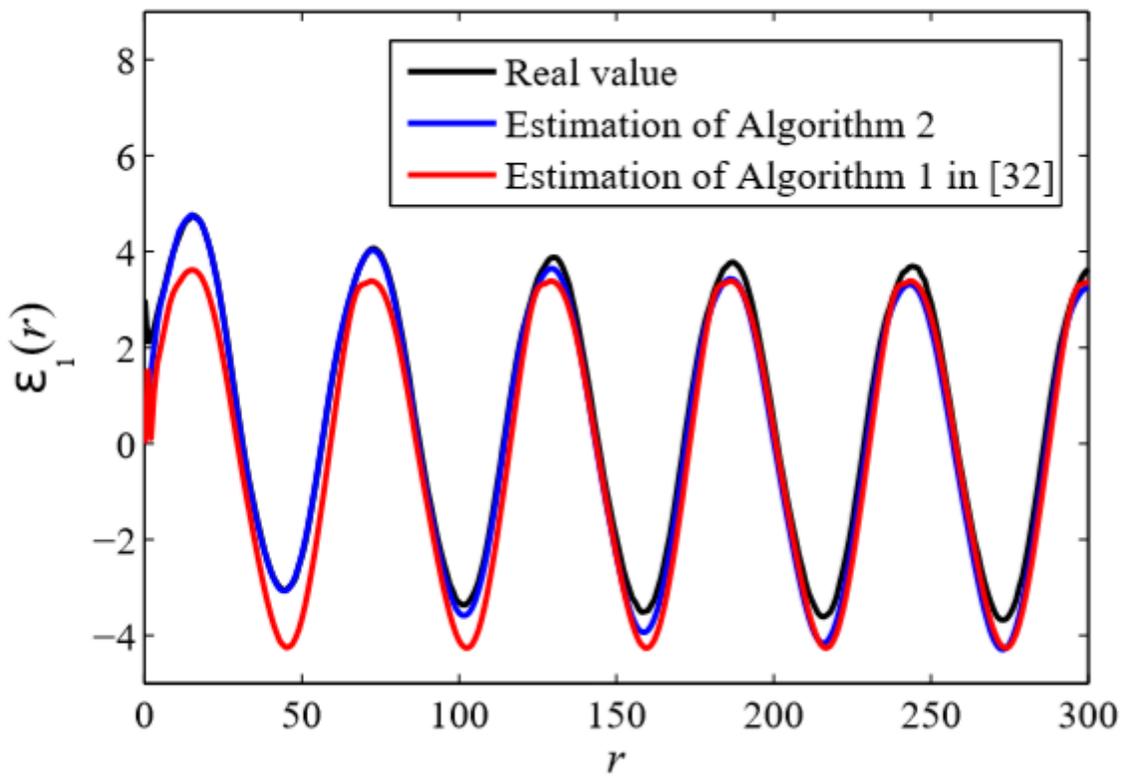


Figure 6

Estimation and real value of $\epsilon_1(r)$ in Example 2

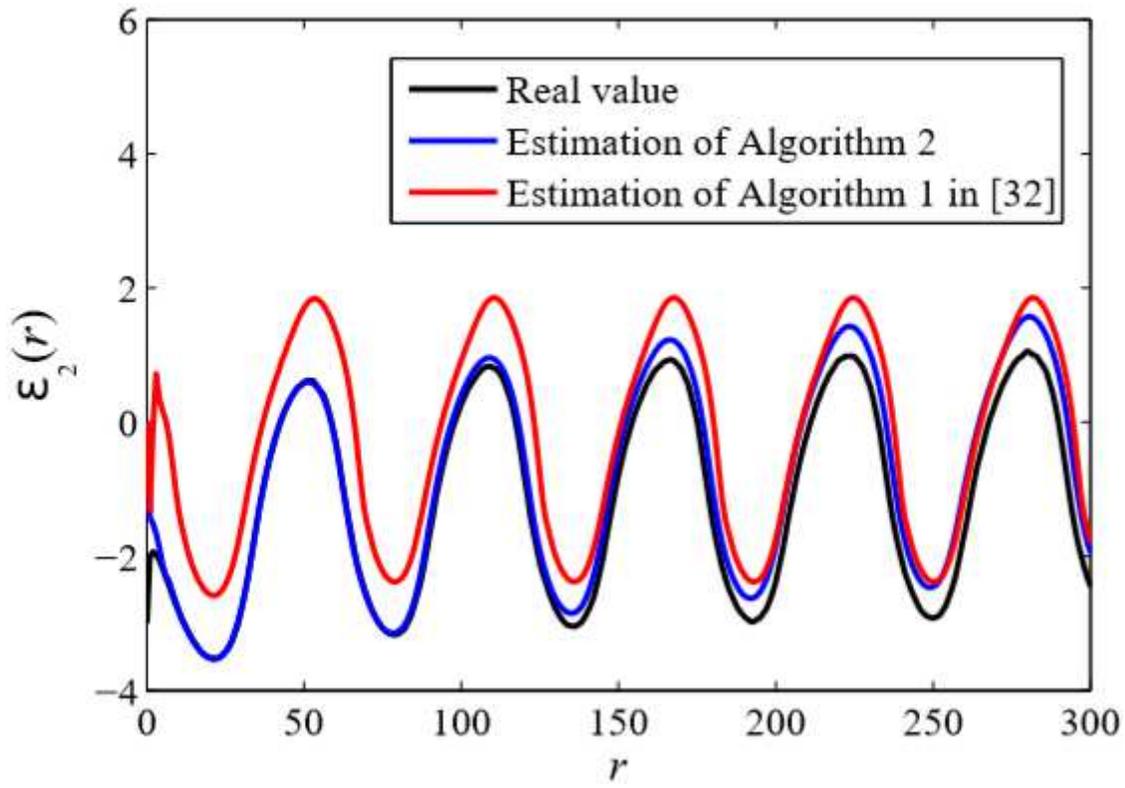


Figure 7

Estimation and real value of $\epsilon_2(r)$ in Example 2