Extended Fractional Singular Kalman Filter

Komeil Nosrati (✉ komeil.nosrati@taltech.ee)

Tallinn University of Technology: Tallinna Tehnikaulikool <https://orcid.org/0000-0001-9507-961X>

Juri Belikov

Tallinn University of Technology: Tallinna Tehnikaulikool

Aleksei Teplyakov

Tallinn University of Technology: Tallinna Tehnikaulikool

Eduard Petlenkov

Tallinn University of Technology: Tallinna Tehnikaulikool

Research Article

Keywords: Fractional calculus, Singular systems, Kalman filter, Nonlinear dynamics, Chaos synchronization

Posted Date: December 28th, 2021

DOI: <https://doi.org/10.21203/rs.3.rs-1148628/v1>

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

[Read Full License](#)

Extended fractional singular Kalman filter

Komeil Nosrati^{1*†}, Juri Belikov^{2†}, Aleksei Tepljakov^{1†} and Eduard Petlenkov^{1†}

^{1*}Department of Computer Systems, Tallinn University of Technology, 12618, Estonia.

²Department of Software Science, Tallinn University of Technology, 12618, Estonia.

*Corresponding author(s). E-mail(s): komeil.nosrati@taltech.ee;

Contributing authors: juri.belikov@taltech.ee; aleksei.tepljakov@taltech.ee;
eduard.petlenkov@taltech.ee;

[†]These authors contributed equally to this work.

Abstract

Effective and accurate state estimation is a staple of modern modeling. On the other hand, non-linear fractional-order singular (FOS) systems are an attractive modeling tool as well since they can provide accurate descriptions of systems with complex dynamics. Consequently, developing accurate state estimation methods for such systems is highly relevant since it provides vital information about the system including related memory effects and long interconnection properties with constraint elements. However, missing features in transforming structures such as violation of constraints in non-singular versions of such systems may affect the performance of the estimation result. This paper proposes the state estimation algorithm design for the original and non-transformed stochastic nonlinear FOS system. We introduce a deterministic data-fitting based framework which helps us to take steps directly towards Kalman filter (KF) derivation of the system, called extended fractional singular KF (EFSKF). Using stochastic reasoning, we demonstrate how to construct recursive form of the filter. Analysis of the filter shows how the proposed algorithm reduces to the nominal nonlinear filters when the system is in its usual state-space form making said algorithm highly flexible. Finally, simulation results verify that the estimation of nonlinear states can be accomplished with the proposed EFSKF algorithm with a reasonable performance.

Keywords: Fractional calculus, Singular systems, Kalman filter, Nonlinear dynamics, Chaos synchronization

1 Introduction

Fractional calculus—also referred to as *non-integer calculus*—appears in modeling of various dynamical systems in engineering applications such as power networks and biological systems [1–4]. Corresponding fractional-order models can describe complex phenomena and as such allow to model said systems more accurately [2]. Meanwhile, singular models or *differential algebraic equations* (DAEs) are a general form of the standard state-space systems, which can be formulated

with both differential and algebraic equations [5]. This description of a system has many extensive applications in a variety of sciences such as cyber-physical systems, biology and robotics [5, 6].

Fractional-order singular (FOS) models, which consist of both fractional and singular features, can be also considered for modeling of more complicated dynamical processes [7–9]. Some primary research were directed towards solvability [10], stability [11], and estimability [1, 12–14] of linear FOS models in the both time invariant (TI)

and time varying (TV) forms; however, there have been few reports connected to the nonlinear cases.

The estimation problem and extension of the Kalman filter (KF) to the fractional order models was considered in [15], which has been termed as fractional KF (FKF). In the case of singular models (in both TI and TV cases) and related KF estimation methods, one solution is the so-called indirect approach. In this methodology, the existing KF algorithms were applied to estimate the states of a transformed system of original one [16]. However, the updated estimates may lead to an inconsistent initialization for the next step because they will not necessarily satisfy the algebraic constraints, and makes those measurements, which are affected by the algebraic states, redundant for state estimation. So, viewed from a different perspective, some optimization approaches such as the moving horizon estimation [17], the maximum likelihood (ML) criterion [18], and the deterministic method [19] were manipulated to solve the estimation problem of these systems directly to introduce the singular KF (SKF).

Due to the nonlinear behavior of most practical dynamical systems, many attempts were made to apply the KF approach to nonlinear cases. Towards this, many works were done at NASA Ames [20] where the extended KF (EKF) was introduced as a nonlinear version of the KF. Adaptive KF (AKF), unscented KF (UKF), cubature KF (CKF), particle filter, and other variants [21] represent some other KF derivatives; each of them offers unique features of accuracy and speed of nonlinear state estimation. Sierociuk and Dzielinski discussed a time domain approach of FKF for nonlinear state estimation problems in detail [15], and extended FKF (EFKF) was proposed according to the Taylor series expansion. In [22], the UKFs for nonlinear fractional-order systems including both uncorrelated and correlated process and measurement noises have been introduced. Besides, the application of Kalman filtering for nonlinear singular systems was described in [23, 24], where a modified EKF was proposed. However, this technique only treats measurements as functions of differential state variables. This technique was therefore extended in [25] to accommodate measurements including algebraic state variables according to EKF and UKF filtering techniques.

The filtering problem of the both TI and TV linear FOS systems has been considered in literature based on direct [1, 13, 14] and indirect methods [26]. Using the deterministic filtering and ML techniques, the authors in [1] and [14] derived a kalman-type algorithm to estimate the states of the TI cases, and introduced the fractional singular KF (FSKF) in a direct methodology. In the following and according to a same formulation, the designed FSKF algorithm was extended to a TV FOS system based on the given measurement including the correlated noises [13]. However, the problem of state estimation for nonlinear FOS (NFOS) cases has not yet been considered.

The research motivation here addresses the filtering problem of stochastic NFOS models in its discrete TI linear form including uncorrelated process and measurement noises for the first time, and addresses the following objectives:

- Formulate a discrete TI stochastic NFOS system and study the index concept and causality condition.
- Introduce the notion of *estimability* for stochastic NFOS systems in the general case.
- Design a new and general form of KF for state estimation based on deterministic arguments in a completely self-contained way besides the stochastic reasoning.
- Analyze the established algorithm by comparing it with the nominal nonlinear KFs and demonstrate its general features.
- Introduce an extended form of the SKF to estimate the states of nonlinear singular systems which has not been considered before.
- Apply the presented nonlinear estimation technique towards the potential applications of chaos control for synchronization of FOS chaotic systems.

Our computationally simple algorithm covers the traditional nonlinear filtering algorithms EKF and EKF. Also, Also, our methods can be easily reformulated to a symmetric structure in three-block forms which provide some features over nonlinear filtering techniques with the standard dynamics including, but with non-invertible noise. Furthermore, we discovered how known derived recursive expressions can be achieved without stochastic assumptions on the system noises which is more effective when the system is uncertain,

because the deterministic approach can be manipulated towards the derivation of recursive robust estimations.

The rest of this workflow is organized as follows. In Section 2, we focus on some preliminaries related to the stochastic TI NFOS models. Then, after illustrating the deterministic data-fitting approach in Section 3, we address the estimation problem for the discrete TI NFOS system and then formulate the estimation algorithm in Section 4 according to said approach. Finally, numerical examples are presented in Section 5 to verify our algorithm. Finally, we draw conclusions based on this work in the last section.

2 Problem Formulation

An ideal continuous-time nonlinear stochastic fractional system in its more general form, can be presented by

$$\begin{aligned} ED^\alpha \mathbf{x}(t) &= \mathbf{H}(\mathbf{x}(t), \mathbf{u}(t)) \\ \mathbf{y}(t) &= \mathbf{J}(\mathbf{x}(t), \mathbf{u}(t)), \end{aligned} \quad (1)$$

where $t \in \mathbb{R} \geq 0$, $\mathbf{H} : \mathbb{R}^n \rightarrow \mathbb{R}^n$ and $\mathbf{J} : \mathbb{R}^n \rightarrow \mathbb{R}^p$ are nonlinear continuous functions, $\mathbf{x} \in \mathbb{R}^n$, $\mathbf{u} \in \mathbb{R}^m$, and $\mathbf{y} \in \mathbb{R}^p$ are state variables, input and output vectors, respectively, and $\alpha \in \mathbb{R}^n$. Also, $\mathbf{E} \in \mathbb{R}^{n \times n}$ is a singular matrix, that is rank $\mathbf{E} < n$, and D^α is the fractional derivative operator [11].

Definition 1 ([2]) The Grunwald-Letnikov (GL) operator for a continuous function $f : \mathbb{R}^+ \rightarrow \mathbb{R}$ can be defined as

$${}_a^{GL} D_t^\alpha f(t) = \lim_{h \rightarrow 0} \frac{1}{h^\alpha} \sum_{j=0}^{\lfloor \frac{t-a}{h} \rfloor} (-1)^j \Upsilon_j^\alpha f(t-jh), \quad (2)$$

where a and t are the lower and upper terminals, respectively, h is the sampling interval, and Υ_j^α equals 1 if $j = 0$, or $\Gamma(\alpha+1)[\Gamma(\alpha+1-j)\Gamma(j+1)]^{-1}$, otherwise, where Γ is the gamma function.

Definition 2 ([2]) Let $t = a + kh$, where k is the number of samples for which the derivative is calculated. Then, (2) can lead to the following approximation

$${}_0^{GL} \Delta_k^\alpha f_k = \frac{1}{h^\alpha} \sum_{j=0}^k (-1)^j \Upsilon_j^\alpha f_{k-j}, \quad (3)$$

where ${}_0^{GL} \Delta_k^\alpha$ denotes the GL difference operator.

According to Definition 1, one can construct the following equivalent discrete system of (1)

$$\begin{aligned} \mathbf{E}_0^{GL} \Delta_k^\alpha \mathbf{x}_{k+1} &= \mathbf{H}(\mathbf{x}_k, \mathbf{u}_k) \\ \mathbf{y}_k &= \mathbf{J}(\mathbf{x}_k, \mathbf{u}_k), \end{aligned} \quad (4)$$

where $k \in \mathbb{Z} \geq 0$ and the functions and variables are defined same as before.

Lemma 1 ([7]) Consider the nonlinear system (4) without inputs. The equilibrium points ($\mathbf{x}^{eq}$ and $\mathbf{u}^{eq}$) of the given system can be given by the following equality:

$$\mathbf{E} \mathbf{x}^{eq} = \mathbf{H}(\mathbf{x}^{eq}, \mathbf{u}^{eq}) - \mathbf{E} \mathbf{x}^{eq} \sum_{m=1}^{k+1} (-1)^m \Upsilon_m^\alpha, \quad (5)$$

where

$$\Upsilon_m^\alpha = \text{diag} \left[\binom{\alpha_1}{m}, \dots, \binom{\alpha_n}{m} \right],$$

and $\alpha_1, \dots, \alpha_n$ are the unequal orders of the system equations.

Without loss of generality, we can assume that $\mathbf{E} = \{\mathbf{I}_{n_1}, \mathbf{0}_{n_2}\}$, where $n_1 + n_2 = n$, $\mathbf{I}$ and $\mathbf{0}$ are identity and null square matrices, respectively. Consider linearized FOS systems given as

$$\begin{aligned} \mathbf{E}_0^{GL} \Delta_k^\alpha \mathbf{x}_{k+1} &= \mathbf{A} \mathbf{x}_k + \mathbf{B} \mathbf{u}_k \\ \mathbf{y}_k &= \mathbf{C} \mathbf{x}_k + \mathbf{D} \mathbf{u}_k, \end{aligned} \quad (6)$$

where $\mathbf{A}$, $\mathbf{B}$, $\mathbf{C}$, and $\mathbf{D}$ are constant Jacobian matrices of nonlinear functions $\mathbf{H}$ and $\mathbf{J}$ in their equilibria. Assume that $\{\mathbf{E}, \mathbf{A}\}$ is regular, that is $\det(\tau_0 \mathbf{E} - \mathbf{A}) \neq 0$, where $\tau_0 = z_0(1 - z_0^{-1})^\alpha$, for a constant scalar $z_0 \in \mathbb{C}$. According to Kronecker's theorem [13, 27], one can transform (6) into the following form

$$\begin{aligned} {}_0^{GL} \Delta_k^\alpha \mathbf{x}_{1,k+1} &= \mathbf{A}_1 \mathbf{x}_{1,k} + \mathbf{B}_1 \mathbf{u}_k \\ \mathbf{N}_0^{GL} \Delta_k^\alpha \mathbf{x}_{2,k+1} &= \mathbf{x}_{2,k} + \mathbf{B}_2 \mathbf{u}_k \\ \mathbf{y}_k &= \mathbf{C}_1 \mathbf{x}_{1,k} + \mathbf{C}_2 \mathbf{x}_{2,k} + \mathbf{D} \mathbf{u}_k, \end{aligned} \quad (7)$$

where $\mathbf{x}_1 \in \mathbb{R}^{n_1}$, $\mathbf{x}_2 \in \mathbb{R}^{n_2}$, and $\mathbf{A}_1$, $\mathbf{B}_1$, $\mathbf{B}_2$, $\mathbf{C}_1$, and $\mathbf{C}_2$ are constant matrices having compatible dimensions. Also, $\mathbf{N}$ is a nilpotent matrix with the index μ . For the nilpotent index $\mu > 1$, the transfer function, $\mathbf{G}(z) = \mathbf{C}(\tau \mathbf{E} - \mathbf{A})^{-1} \mathbf{B} + \mathbf{D}$ with $\tau = z(1 - z^{-1})^\alpha$, of the system (6) might be improper.

Definition 3 The regular FOS systems (6) and (7) are said to be of index 1 if the matrix $\mathbf{N}$ has the nilpotency index equal to 1.

This definition can be extended to the nonlinear FOS system (4) as follows.

Definition 4 The linearized FOS of (4) is said to be of index 1 at origin if the constant coefficient system $\mathbf{E}_0^{GL} \Delta_k^\alpha \mathbf{x}_{k+1} = \mathbf{A} \mathbf{x}_k$ has index 1, where $\mathbf{A} = \left. \frac{\partial \mathbf{H}}{\partial \mathbf{x}} \right|_{(0,0)}$.

Remark 1 A necessary and sufficient condition for causality of a linear discrete-time FOS system (6) and (7) is $\mathbf{N}^i \mathbf{B}_2 = 0$ for $i \geq 1$, since the first term of the transfer function can be rewritten as $\mathbf{C}_1(\tau \mathbf{I}_{n_1} - \mathbf{A}_1)^{-1} \mathbf{B}_1 + \mathbf{C}_2(\tau \mathbf{N} - \mathbf{I}_{n_2})^{-1} \mathbf{B}_2$ where $(\tau \mathbf{N} - \mathbf{I}_{n_2})^{-1} = \mathbf{I}_{n_2} + \tau \mathbf{N} + \dots + \tau^{\mu-1} \mathbf{N}^{\mu-1}$. On the other hand, being index 1 means $\mathbf{N} = 0$ which guarantees that the FOS systems (6) and (7) are causal.

Recently, the filtering problem of linear FOS system (6) has been considered in the literature [1, 13]. Using the data-fitting problem and maximum likelihood approach, a filtering algorithm was derived directly from the original FOS model. Here, we are interested in filtering problem of the NFOS system (4) in its stochastic form as

$$\mathbf{E}_0^{GL} \Delta_{k+1}^\alpha \mathbf{x}_{k+1} = \mathbf{H}(\mathbf{x}_k, \mathbf{u}_k) + \mathbf{w}_k, \quad (8a)$$

$$\begin{aligned} \mathbf{E} \mathbf{x}_{k+1} &= \mathbf{E}_0^{GL} \Delta_{k+1}^\alpha \mathbf{x}_{k+1} \\ &\quad - \sum_{j=1}^{k+1} \mathbf{E}(-1)^j \Upsilon_j^\alpha \mathbf{x}_{k+1-j}, \end{aligned} \quad (8b)$$

$$\mathbf{y}_k = \mathbf{J}(\mathbf{x}_k, \mathbf{u}_k) + \mathbf{v}_k \quad (8c)$$

with the following assumptions.

Assumption 1 The discrete stochastic FOS system (8) is causal, that is the linearized system of (8) is of index 1.

Assumption 2 The initial state $\mathbf{x}_0$ is a random variable of normally distributed with mean $\bar{\mathbf{x}}_0$ and positive definite (PD) covariance of prior information $\bar{\mathbf{P}}_0$.

Assumption 3 The independent vector noises $\mathbf{w}_k \in \mathbb{R}^m$ and $\mathbf{v}_k \in \mathbb{R}^p$ are zero mean white sequences with PD covariances $\mathbf{Q}_k$ and $\mathbf{R}_k$, respectively.

Assumption 4 The nonlinear functions $\mathbf{H}(\cdot, \cdot)$ and $\mathbf{J}(\cdot, \cdot)$ are assumed to be smooth or a member of class C^∞ , meaning that they have derivatives of all order. The Taylor series for these infinitely differentiable can be given as

$$\begin{aligned} \mathbf{H}(\mathbf{x}_k, \mathbf{u}_k) &= \mathbf{H}(\mathbf{x}^{eq}, \mathbf{u}^{eq}) \\ &\quad + \left. \frac{\partial \mathbf{H}}{\partial \mathbf{x}_k} \right|_{(\mathbf{x}^{eq}, \mathbf{u}^{eq})} (\mathbf{x}_k - \mathbf{x}^{eq}) + \text{H.O.T}, \end{aligned}$$

where the last term is higher order terms.

Assumption 5 The vector of observed signals from the stochastic FOS model (8) are given as $\mathbf{Y}_l = [y_0, \dots, y_l]$.

According to the observed measurements $\mathbf{Y}_l$, we are interested in seeking a suitable dynamical model which explains these observations. The following section is devoted to the KF algorithm derivation for the causal system (8) based on deterministic data-fitting approach. Relying on this technique, we will see how, while providing an optimal recursive algorithm, it overcomes the following constraints that result from the coexistence of fractional and singular features and the use of usual approaches:

1. *Singular filter form*: Instead of $\mathbf{x}_k$, the term $\mathbf{E} \mathbf{x}_k$ is estimated resulting in a singular filter. In such filters, in addition to the input and output of a target system, which are contaminated with noise and disturbance, their derivatives also appear. Then, the sensitivity to noise will be high due to the increase in the noise power caused by the derivative. In discrete case, the main problem is that to obtain a filter response, it is necessary to know the properties of noise in the future.
2. *Regularity conditions*: This condition is intended to consider a FOS system indirectly and in its decomposed form. Despite this, not only will a particular class of the systems be considered, but it will result in the loss of some information. For example, when considering the possibility of faults in a power grid leading to sudden structural or topological changes the underlying constraints change, then the resulting decomposed system is not compatible anymore with that of original form.

3. *Inflexibility*: Previous approaches have been limited to time-invariant, independent measurement noise with noninvertible covariance matrices. Also, all systems are considered in causal and linear mode.
4. *Infinite terms*: In KF design using the initial definition of the covariance matrix in the form of $\mathbf{P}_k = \varepsilon[\mathbf{x}_k \mathbf{x}_k^T]$, applying the fractional derivative to both sides of this equation causes to create infinite terms as $D^\alpha(\mathbf{x}_k \mathbf{x}_k^T) = \sum_{k=0}^{\infty} \Upsilon_k^\alpha D^{\alpha-k} \mathbf{x}_k D^k \mathbf{x}_k$ resulting in fundamental limitation.

3 Deterministic Data-Fitting Method

For the given observations $\mathbf{Y}_l$, the filtering problem comprises of computing an estimate of $\mathbf{x}_k$ based on these observations for $k = l$, otherwise we find the case as smoothing ($k < l$) or prediction ($k > l$) problems. The KF as an optimal linear estimator, minimizes the squared estimation error in an expected value configuration. In Gaussian processes, this linear filter coincides to those of the optimal nonlinear, maximum likelihood and conditional mean estimators [28]. In this study, a purely deterministic data-fitting approach is applied to derive a nonlinear filtering algorithm for the causal nonlinear system (8). Instead of considering the vector noises $\mathbf{w}_k$ and $\mathbf{v}_k$ as well-defined random input variables, the basic idea towards the filter derivation is taking these vectors as the following unknown character (fitting) errors

$$\mathbf{w}_i = \mathbf{E}_0^{GL} \Delta_{i+1}^\alpha \mathbf{x}_{i+1} - \mathbf{H}(\mathbf{x}_i, \mathbf{u}_i), \quad (9a)$$

$$\mathbf{v}_i = \mathbf{y}_i - \mathbf{J}(\mathbf{x}_i, \mathbf{u}_i), \quad (9b)$$

$$\bar{\mathbf{P}}_0 = \mathbf{x}_0 - \bar{\mathbf{x}}_0 \quad (9c)$$

for $0 \leq i \leq k$. According to the given observations, we seek a sequence of states $\mathbf{X}_l = [\mathbf{x}_0, \dots, \mathbf{x}_l]$ so that the fitting errors are small at the same time. In other words, we are interested in some solutions of (8a) which pass through the given observations as closely possible. Inspired by the classical least square method, the following cost function

$$L_k = \|\mathbf{x}_0 - \bar{\mathbf{x}}_0\|_{\bar{\mathbf{P}}_0^{-1}}^2 + \sum_{i=0}^{k-1} \|\mathbf{w}_i\|_{\mathbf{Q}_i^{-1}}^2 + \sum_{i=0}^k \|\mathbf{v}_i\|_{\mathbf{R}_i^{-1}}^2 \quad (10)$$

can be minimized with respect to the sequence $\mathbf{X}_l$ and subject to the constraint (9). The first term presents the prior estimate which present the initial uncertainty, and the two last ones are the process and measurement residuals. The weighting PD matrices $\mathbf{Q}_j$ and $\mathbf{R}_j$ can be interpreted as ‘fiddling parameters’ which can be manipulated to obtain a good least-squares fit. When there is no error in the process equation, that is $\mathbf{Q}_j = \mathbf{0}$, then the minimization problem in (10) collapses to a least square error problem [29].

In this approach, the minimization problem is set out in deterministic concepts without stochastic reasoning. The result here are based on the deterministic (not probabilistic) meaning which is difficult to interpret. For example, the weighting matrices are arbitrary without any statistical basis, and meaning of solutions are not clear. By reinterpreting the errors $\mathbf{w}_k$ and $\mathbf{v}_k$ as in Assumption 3, one can see that the deterministic data-fitting problem is equivalent to that of estimation approaches in the stochastic sense.

4 Filter Design

Using the deterministic data-fitting method, in this section we will try to achieve a recursive state estimation algorithm for the NFOS system (8) and develop the EFSKF as a new and general form of filtering algorithm. The results here, which is based on the linearization technique in each step using the best state vector estimates as reference values to reduce approximation errors and increase estimation accuracy, would be a basis for practical applications in the future.

4.1 Estimability conditions

To pave the way of KF derivation, it is necessary to define the cost function (10) under adaptive errors (9) in first step as

$$\begin{aligned} L_k(\mathbf{X}_k) = & \|\mathbf{x}_0 - \bar{\mathbf{x}}_0\|_{\bar{\mathbf{P}}_0^{-1}}^2 \\ & + \sum_{i=0}^k \|\mathbf{y}_i - \mathbf{J}(\mathbf{x}_i, \mathbf{u}_i)\|_{\mathbf{R}_i^{-1}}^2 \\ & + \sum_{i=0}^{k-1} \|\mathbf{E}_0^{GL} \Delta_{i+1}^\alpha \mathbf{x}_{i+1} - \mathbf{H}(\mathbf{x}_i, \mathbf{u}_i)\|_{\mathbf{Q}_i^{-1}}^2. \end{aligned} \quad (11)$$

The problem of matching the data-fitting based filter is to find the sequence of $\mathbf{X}_k$ so that $\mathbf{L}_k$ is minimized. The notion of estimability is determining a unique solution of the problem (11) for the given information, which can be given based on the following theorem.

Theorem 1 *The optimization problem (11) has a unique solution if and only if the following matrix*

$$[\mathbf{E} \ \mathbf{J}_i]^T, \quad 1 \leq i \leq k, \quad (12)$$

where $\mathbf{J}_i = \frac{\partial \mathbf{J}(\mathbf{x}_i, \mathbf{u}_i)}{\partial \mathbf{x}_i} \Big|_{\mathbf{x}_i^{eq}}$, has a full column rank, i.e.,

$$\text{rank} \left(\begin{bmatrix} \mathbf{E} \\ \mathbf{J}_i \end{bmatrix} \right) = n. \quad (13)$$

Proof It is not difficult to rewrite the problem (11) as

$$\min_{\mathbf{X}_k} (\mathbb{A}_k \mathbf{X}_k - \mathbb{Y}_k)^T \mathbb{R}_k (\mathbb{A}_k \mathbf{X}_k - \mathbb{Y}_k) \quad (14)$$

for $k \geq 0$, where

$$\mathbb{A}_k = \begin{bmatrix} C_k & \mathcal{A}_1^{k-1} & \cdots & \mathcal{A}_{k-1}^1 & \mathcal{A}_k^0 \\ 0 & C_{k-1} & \cdots & \mathcal{A}_{k-2}^1 & \mathcal{A}_{k-1}^0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & C_1 & \mathcal{A}_1^0 \\ 0 & 0 & \cdots & 0 & C_0 \end{bmatrix}, \quad \mathbf{X}_k = \begin{bmatrix} \mathbf{x}_k \\ \mathbf{x}_{k-1} \\ \vdots \\ \mathbf{x}_1 \\ \mathbf{x}_0 \end{bmatrix},$$

$$\mathbb{R}_k = \text{diag}(\mathcal{R}_k, \mathcal{R}_{k-1}, \dots, \mathcal{R}_0), \quad \mathbb{Y}_k = \begin{bmatrix} \mathcal{Y}_k \\ \mathcal{Y}_{k-1} \\ \vdots \\ \mathcal{Y}_1 \\ \mathcal{Y}_0 \end{bmatrix}$$

in which

$$C_i = \begin{bmatrix} \mathbf{E}_i \\ \mathbf{J}_i \end{bmatrix}, \quad \mathcal{Y}_i = \begin{bmatrix} \bar{\mathbf{x}}_i - \mathbf{H}_{i-1} \mathbf{x}_{i-1} + \mathbf{H}(\mathbf{x}_{i-1}, \mathbf{u}_{i-1}) \\ \mathbf{y}_i^T + (\mathbf{J}_i \mathbf{x}_i) - \mathbf{J}(\mathbf{x}_{i-1}, \mathbf{u}_{i-1}) \end{bmatrix},$$

$$\mathcal{R}_i = \begin{bmatrix} \mathbf{Q}_{i-1}^{-1} & \mathbf{0} \\ \mathbf{0} & \mathbf{R}_i^{-1} \end{bmatrix}, \quad \mathcal{A}_j^i = \begin{bmatrix} (-1)^j (\mathbf{E} \mathbf{Y}_j^\alpha) - \mathbf{H}_j^i \\ \mathbf{0} \end{bmatrix}$$

for the parameter values $\mathbf{E}_0 = \mathbf{I}_n$, $\mathbf{Q}_{-1} = \bar{\mathbf{P}}_0^{-1}$, $\mathbf{E}_i = \mathbf{E}$, $\bar{\mathbf{x}}_i = \mathbf{0}$, for $1 \leq i \leq k$, and $\mathbf{H}_1^i = \mathbf{H}_i$, $\mathbf{H}_j^i = \mathbf{0}$, for $0 \leq i \leq k-1$ and $2 \leq j \leq k$. According to the Assumption 4, the matrices $\mathbf{H}_i$ and $\mathbf{J}_i$ can be defined as $\mathbf{H}_i = \frac{\partial \mathbf{H}(\mathbf{x}_i, \mathbf{u}_i)}{\partial \mathbf{x}_i} \Big|_{\mathbf{x}_i^{eq}}$ and $\mathbf{J}_i = \frac{\partial \mathbf{J}(\mathbf{x}_i, \mathbf{u}_i)}{\partial \mathbf{x}_i} \Big|_{\mathbf{x}_i^{eq}}$, where $\mathbf{x}_i^{eq}$ is the best state vector estimate in each step.

Now, the minimizing solutions of (11) are the values of $\mathbf{x}_k$ for which the derivative of (14) is zero, that is $\mathbf{X}_k = (\mathbb{A}_k^T \mathbb{R}_k \mathbb{A}_k)^{-1} \mathbb{A}_k^T \mathbb{R}_k \mathbb{Y}_k$. This solution is unique if and only if $\mathbb{A}_k^T \mathbb{R}_k \mathbb{A}_k$ is an invertible matrix.

Necessity: Assume that the solution of (11), or equivalently (14), is unique, that is the matrix

$\mathbb{A}_k^T \mathbb{R}_k \mathbb{A}_k$ is invertible. Therefore, thanks to the positive definiteness of $\mathbb{R}_k$, the triangular matrix $\mathbb{A}_k$ in (15), or equivalently its entries on its main diagonal, that is C_i is a full column rank. This implies that the relation (13) holds for $1 \leq i \leq k$, since C_0 has an identity matrix in its first element.

Sufficiency: Assume that the relation (13) holds. Therefore, all entries on main diagonal of matrix $\mathbb{A}_k$ are from full column rank which together with the positive definiteness of the matrix. $\mathbb{R}_k$ implies that the matrix $\mathbb{A}_k^T \mathbb{R}_k \mathbb{A}_k$ is invertible, meaning that the solution $\mathbf{X}_k = (\mathbb{A}_k^T \mathbb{R}_k \mathbb{A}_k)^{-1} \mathbb{A}_k^T \mathbb{R}_k \mathbb{Y}_k$ is unique. $\square$

Definition 5 The NFOS system (8) is said to be estimable if for all $0 \leq k \leq M$, it is sufficient to determine uniquely the state $0 \leq k \leq M$ by knowing (8a) and the vector of observed signals $\mathbf{Y}_M$ from (8c).

Remark 2 According to Theorem 1 and Definition 5, the NFOS system (8) is said to be estimable if the relation (13) holds.

In what follows and with respect to the conditions of uniqueness, we seek a unique solution for the minimization problem (11), which steer us towards a filtering algorithm derivation.

4.2 Filter derivation

Theorem 1 together with Remark 2 and associated with the recursive concept of (11) lead to a recursive solution for the optimization problem. To extract the recursive concept of the cost function (11), it is not difficult to rewrite it recursively as

$$\begin{aligned} \mathbf{L}_k(\mathbf{X}_k) &= \mathbf{L}_{k-1}(\mathbf{X}_{k-1}) + \|\mathbf{y}_k - \mathbf{J}(\mathbf{x}_k, \mathbf{u}_k)\|_{\mathbf{R}_k}^2 \\ &\quad + \|\mathbf{E}_0^{GL} \Delta_k^\alpha \mathbf{x}_k - \mathbf{H}(\mathbf{x}_{k-1}, \mathbf{u}_{k-1})\|_{\mathbf{Q}_{k-1}^{-1}}^2 \end{aligned} \quad (16)$$

starting from

$$\mathbf{L}_0(\mathbf{X}_0) = \|\mathbf{x}_0 - \bar{\mathbf{x}}_0\|_{\bar{\mathbf{P}}_0^{-1}}^2 + \|\mathbf{y}_0 - \mathbf{J}(\mathbf{x}_0, \mathbf{u}_0)\|_{\mathbf{R}_0^{-1}}^2.$$

This recursive concept of (11) and its solution according to the stochastic NFOS system (8) introduce a new recursive algorithm called EFSKF. Therefore, the first step towards derivation of this algorithm is finding a unique solution for the optimization problem (11).

$$\mathbb{A}_k = \begin{bmatrix} \begin{bmatrix} \mathbf{E} \\ \mathbf{J}_k \end{bmatrix} & \begin{bmatrix} -\mathbf{E}\Upsilon_1^\alpha - \mathbf{H}_{k-1} \\ \mathbf{0} \end{bmatrix} & \dots & \begin{bmatrix} (-1)^{k-1}\mathbf{E}\Upsilon_{k-1}^\alpha \\ \mathbf{0} \end{bmatrix} & \begin{bmatrix} (-1)^k\mathbf{E}\Upsilon_k^\alpha \\ \mathbf{0} \end{bmatrix} \\ \mathbf{0} & \begin{bmatrix} \mathbf{E} \\ \mathbf{J}_{k-1} \end{bmatrix} & \dots & \begin{bmatrix} (-1)^{k-2}\mathbf{E}\Upsilon_{k-2}^\alpha \\ \mathbf{0} \end{bmatrix} & \begin{bmatrix} (-1)^{k-1}\mathbf{E}\Upsilon_{k-1}^\alpha \\ \mathbf{0} \end{bmatrix} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ \mathbf{0} & \mathbf{0} & \dots & \begin{bmatrix} \mathbf{E} \\ \mathbf{J}_1 \end{bmatrix} & \begin{bmatrix} -\mathbf{E}\Upsilon_1^\alpha - \mathbf{H}_0 \\ \mathbf{0} \end{bmatrix} \\ \mathbf{0} & \mathbf{0} & \dots & \mathbf{0} & \begin{bmatrix} \mathbf{I}_n \\ \mathbf{J}_0 \end{bmatrix} \end{bmatrix} \quad (15)$$

Theorem 2 Let matrix (13) holds for $1 \leq i \leq k$. The problem (11) has a unique recursive and optimal solution $\hat{\mathbf{x}}_k$, $k > 0$ as follows

$$\hat{\mathbf{x}}_k = \mathbf{P}_k \mathbf{E}^\top \tilde{\mathbf{P}}_k^{-1} \tilde{\mathbf{x}}_k + \mathbf{P}_k \mathbf{J}_k^\top \mathbf{R}_k^{-1} \tilde{\mathbf{y}}_k \quad (17)$$

starting from $\hat{\mathbf{x}}_0 = \mathbf{P}_0(\tilde{\mathbf{P}}_0^{-1}\tilde{\mathbf{x}}_0 + \mathbf{J}_0^\top \mathbf{R}_0^{-1}\tilde{\mathbf{y}}_0)$, where $\tilde{\mathbf{y}}_0 = (\mathbf{y}_0 + \mathbf{J}_0\tilde{\mathbf{x}}_0 - \mathbf{J}(\tilde{\mathbf{x}}_0, \mathbf{u}_0))$, $\tilde{\mathbf{y}}_k = \mathbf{y}_k + \mathbf{J}_k\tilde{\mathbf{x}}_k - \mathbf{J}(\tilde{\mathbf{x}}_k, \mathbf{u}_k)$, $\mathbf{P}_0 = (\tilde{\mathbf{P}}_0^{-1} + \mathbf{J}_0^\top \mathbf{R}_0^{-1} \mathbf{J}_0)^{-1}$, and $\mathbf{P}_k = (\mathbf{E}^\top(\tilde{\mathbf{P}}_k)^{-1}\mathbf{E} + \mathbf{J}_k^\top \mathbf{R}_k^{-1} \mathbf{J}_k)^{-1}$. Also, $\tilde{\mathbf{P}}_k = \mathbf{Q}_{k-1} + \Omega_{k-1}^1$ and $\tilde{\mathbf{x}}_k = -\mathbf{H}_{k-1}\hat{\mathbf{x}}_{k-1} + \mathbf{H}(\hat{\mathbf{x}}_{k-1}, \mathbf{u}_{k-1}) - \Omega_{k-1}^2$ in which

$$\begin{aligned} \Omega_{k-1}^1 &= (\mathbf{E}\Upsilon_1^\alpha + \mathbf{H}_{k-1})\mathbf{P}_{k-1}(\mathbf{E}\Upsilon_1^\alpha + \mathbf{H}_{k-1})^\top \\ &+ \sum_{j=2}^k ((-1)^j \mathbf{E}\Upsilon_j^\alpha) \mathbf{P}_{k-j} ((-1)^j \mathbf{E}\Upsilon_j^\alpha)^\top \end{aligned}$$

and

$$\Omega_{k-1}^2 = (-\mathbf{E}\Upsilon_1 - \mathbf{H}_{k-1})\hat{\mathbf{x}}_{k-1} + \sum_{j=2}^k (-1)^j \mathbf{E}\Upsilon_j^\alpha \hat{\mathbf{x}}_{k-j}.$$

Proof For $k = 0$, and according to the minimization solution of (14), one has

$$\begin{aligned} \hat{\mathbf{x}}_0 &= (\mathbb{A}_0^\top \mathbb{R}_0 \mathbb{A}_0)^{-1} \mathbb{A}_0^\top \mathbb{R}_0 \mathbb{A}_0 = (\mathcal{C}_0^\top \mathcal{R}_0 \mathcal{C}_0)^{-1} \mathcal{C}_0^\top \mathcal{R}_0 \mathcal{Y}_0 \\ &= \left(\begin{bmatrix} \mathbf{I}_n \\ \mathbf{J}_0 \end{bmatrix}^\top \begin{bmatrix} \tilde{\mathbf{P}}_0^{-1} & \mathbf{0} \\ \mathbf{0} & \mathbf{R}_0^{-1} \end{bmatrix} \begin{bmatrix} \mathbf{I}_n \\ \mathbf{J}_0 \end{bmatrix} \right)^{-1} \\ &\times \begin{bmatrix} \mathbf{I}_n \\ \mathbf{J}_0 \end{bmatrix}^\top \begin{bmatrix} \tilde{\mathbf{P}}_0^{-1} & \mathbf{0} \\ \mathbf{0} & \mathbf{R}_0^{-1} \end{bmatrix} \begin{bmatrix} \tilde{\mathbf{x}}_0 \\ \tilde{\mathbf{y}}_0 \end{bmatrix} \\ &= (\tilde{\mathbf{P}}_0^{-1} + \mathbf{J}_0^\top \mathbf{R}_0^{-1} \mathbf{J}_0)^{-1} (\tilde{\mathbf{P}}_0^{-1} \tilde{\mathbf{x}}_0 + \mathbf{J}_0^\top \mathbf{R}_0^{-1} \tilde{\mathbf{y}}_0), \end{aligned}$$

where $\tilde{\mathbf{y}}_0 = \mathbf{y}_0 + \mathbf{J}_0\tilde{\mathbf{x}}_0^h - \mathbf{J}(\tilde{\mathbf{x}}_0^h, \mathbf{u}_0)$. By defining $\mathbf{P}_0 = \tilde{\mathbf{P}}_0^{-1} + \mathbf{J}_0^\top \mathbf{R}_0^{-1} \mathbf{J}_0)^{-1}$ and $\mathbf{x}_0^h = \tilde{\mathbf{x}}_0$ the solution in the first step is proven.

For $k > 0$, one can use the following partitioned matrices

$$\mathbb{A}_k = \begin{bmatrix} \mathcal{C}_k & \mathfrak{S}_k \\ \mathbf{0} & \mathbb{A}_{k-1} \end{bmatrix}, \quad \mathbb{Y}_k = \begin{bmatrix} \mathcal{Y}_k \\ \mathbb{Y}_{k-1} \end{bmatrix},$$

$$\mathbb{R}_k = \begin{bmatrix} \mathcal{R}_k & \mathbf{0} \\ \mathbf{0} & \mathbb{R}_{k-1} \end{bmatrix}, \quad \mathbb{X}_k = \begin{bmatrix} \mathbf{x}_k \\ \mathbb{X}_{k-1} \end{bmatrix},$$

where $\mathfrak{S}_k = [\mathcal{A}_1^{k-1} \mathcal{A}_2^{k-2} \dots \mathcal{A}_k^0]$, to have a recursive solution. Therefore, the solution of the optimization problem (11) using the equivalent recursive problem (15) can be given by the following relation

$$\begin{aligned} \hat{\mathbb{X}}_k &= \begin{bmatrix} \hat{\mathbf{x}}_k \\ \hat{\mathbb{X}}_{k-1} \end{bmatrix} \\ &= \left(\begin{bmatrix} \mathcal{C}_k & \mathfrak{S}_k \end{bmatrix}^\top \begin{bmatrix} \mathcal{R}_k & \mathbf{0} \\ \mathbf{0} & \mathbb{R}_{k-1} \end{bmatrix} \begin{bmatrix} \mathcal{C}_k & \mathfrak{S}_k \end{bmatrix} \right)^{-1} \\ &\times \begin{bmatrix} \mathcal{C}_k & \mathfrak{S}_k \end{bmatrix}^\top \begin{bmatrix} \mathcal{R}_k & \mathbf{0} \\ \mathbf{0} & \mathbb{R}_{k-1} \end{bmatrix} \begin{bmatrix} \mathcal{Y}_k \\ \mathbb{Y}_{k-1} \end{bmatrix}. \end{aligned}$$

Using partitioned matrix inversion lemma [30], the Woodbury formula, and some algebraic manipulations, it is not difficult to see that this solution can be simplified as

$$\hat{\mathbf{x}}_k = \mathbf{P}_{11,k} \mathcal{C}_k^\top [(\mathcal{R}_k^{-1} + \mathfrak{S}_k \eta_{k-1} \mathfrak{S}_k^\top)^{-1} (\mathcal{Y}_k - \mathfrak{S}_k \hat{\mathbb{X}}_{k-1})], \quad (18)$$

where,

$$\begin{aligned} &\begin{bmatrix} \mathbf{P}_{11,k} & \mathbf{P}_{12,k} \\ \mathbf{P}_{21,k} & \mathbf{P}_{22,k} \end{bmatrix} \\ &= \begin{bmatrix} (\mathcal{C}_k^\top \Sigma_k^1 \mathcal{C}_k)^{-1} & -\mathbf{P}_{11,k} \mathcal{C}_k^\top \Sigma_k^1 \mathfrak{S}_k \eta_{k-1} \\ -\Sigma_k^2 \mathfrak{S}_k^\top \mathcal{R}_k \mathcal{C}_k \mathbf{P}_{11,k} & \Sigma_k^2 (I - \mathfrak{S}_k^\top \mathcal{R}_k \mathcal{C}_k \mathbf{P}_{12,k}) \end{bmatrix} \\ &\triangleq \eta_k \end{aligned}$$

in which $\Sigma_k^1 = (\mathcal{R}_k^{-1} + \mathfrak{S}_k \eta_{k-1} \mathfrak{S}_k^\top)^{-1}$ and $\Sigma_k^2 = (\mathfrak{S}_k^\top \mathcal{R}_k \mathfrak{S}_k + \eta_{k-1}^{-1})^{-1}$. Now, with defining an auxiliary variable as $\mathbf{P}_k = \mathbf{P}_{11,k}$, one has

$$\begin{aligned} \mathbf{P}_k^{-1} &= \mathcal{C}_k^\top \left(\mathcal{R}_k^{-1} + \mathcal{A}_1^{k-1} \mathbf{P}_{k-1} \left(\mathcal{A}_1^{k-1} \right)^\top \right. \\ &+ \left[\mathcal{A}_2^{k-2} \dots \mathcal{A}_k^0 \right] \mathbf{P}_{12,k-1} \left(\mathcal{A}_1^{k-1} \right)^\top \\ &+ \mathcal{A}_1^{k-1} \mathbf{P}_{21,k-1} \left[\mathcal{A}_2^{k-2} \dots \mathcal{A}_k^0 \right]^\top \\ &+ \left[\mathcal{A}_2^{k-2} \mathcal{A}_3^{k-3} \dots \mathcal{A}_k^0 \right]^\top \mathbf{P}_{22,k-1} \\ &\left. \times \left[\mathcal{A}_2^{k-2} \mathcal{A}_3^{k-3} \dots \mathcal{A}_k^0 \right] \right)^\top \mathcal{C}_k. \end{aligned}$$

The auxiliary variable $\mathbf{P}_k$ is defined with no intuitive concept, but – following from the given PD matrices $\mathbf{Q}_k$ and $\mathbf{R}_k$ as degrees of fitting uncertainties or fiddling parameters – it can be considered as the uncertainty degree of the estimated error $\mathbf{e}_k = \mathbf{x}_k - \hat{\mathbf{x}}_k$. This deterministic interpretation parallels to the stochastic meaning as the error covariance, meaning that $\mathbf{P}_k = \varepsilon[\mathbf{e}_k \mathbf{e}_k^\top]$. This concept can be applied for the other three variables as $\mathbf{P}_{12,k} = \varepsilon[\mathbf{e}_k \nu_{k-1}^\top]$, $\mathbf{P}_{21,k} = \varepsilon[\nu_{k-1} \mathbf{e}_k^\top]$ and $\mathbf{P}_{22,k} = \varepsilon[\nu_{k-1} \nu_{k-1}^\top]$, where $\nu_k = \mathbb{X}_k - \hat{\mathbb{X}}_k$. In the following the two cross-correlations $\mathbf{P}_{12,k}$ and $\mathbf{P}_{21,k}$ will be neglected since they have no significant impact on the derived results. Therefore, using this simplification, rewriting $\mathbf{P}_{22,k-1}$ as

$$\begin{aligned} \mathbf{P}_{22,k-1} &= \Sigma_{k-1}^2 \left(\mathbf{I} + \gamma_{k-1}^\top \mathcal{C}_{k-1} \mathbf{P}_{11,k-1} \left(\mathcal{C}_{k-1}^\top \gamma_{k-1} \right) \right. \\ &\quad \left. \times \Sigma_{k-1}^2 \right) \\ &= \Sigma_{k-1}^2 + \zeta_{k-1}^\top \Sigma_{k-1}^1 \left(\mathcal{C}_{k-1} \mathbf{P}_{11,k-1} \mathcal{C}_{k-1}^\top \right) \\ &\quad \times \Sigma_{k-1}^1 \zeta_{k-1} \\ &= \Sigma_{k-1}^2 + \zeta_{k-1}^\top \Sigma_{k-1}^1 \zeta_{k-1} \\ &= \eta_{k-2} \left(\mathbf{I} - \gamma_{k-1}^\top \zeta_{k-1} \left(\mathbf{I} + \gamma_{k-1}^\top \zeta_{k-1} \right)^{-1} \right. \\ &\quad \left. + \gamma_{k-1}^\top \zeta_{k-1} \left(\mathbf{I} + \gamma_{k-1}^\top \zeta_{k-1} \right)^{-1} \right) \\ &= \eta_{k-2}, \end{aligned}$$

where $\zeta_{k-1} = \mathbb{S}_{k-1} \eta_{k-2}$ and $\gamma_{k-1} = \mathcal{R}_{k-1} \mathbb{S}_{k-1}$, and continuing this process recursively, and also substituting the auxiliary variables $\mathbf{P}_{k-j} = \mathbf{P}_{11,k-j}$ for $j = 1, 2, \dots, k$, the covariance matrix $\mathbf{P}_k$ can be reduced to

$$\begin{aligned} \mathbf{P}_k^{-1} &= \mathcal{C}_k^\top \left(\mathcal{R}_k^{-1} + \mathcal{A}_1^{k-1} \mathbb{P}_{k-1} (\mathcal{A}_1^{k-1})^\top \right. \\ &\quad \left. + \left[\mathcal{A}_2^{k-2} \ \mathcal{A}_3^{k-3} \ \dots \ \mathcal{A}_k^0 \right]^\top \eta_{k-2} \right. \\ &\quad \left. \times \left[\mathcal{A}_2^{k-2} \ \mathcal{A}_3^{k-3} \ \dots \ \mathcal{A}_k^0 \right]^\top \right)^{-1} \mathcal{C}_k \\ &= \mathcal{C}_k^\top \left(\mathcal{R}_k^{-1} + (\mathcal{A}_1^{k-1}) \mathbf{P}_{k-1} (\mathcal{A}_1^{k-1})^\top \right. \\ &\quad \left. + (\mathcal{A}_2^{k-2}) \mathbf{P}_{k-2} (\mathcal{A}_2^{k-2})^\top + (\mathcal{A}_3^{k-3}) \mathbf{P}_{k-3} (\mathcal{A}_3^{k-3})^\top \right. \\ &\quad \left. + \dots + (\mathcal{A}_k^0) \mathbf{P}_0 (\mathcal{A}_k^0)^\top \right)^{-1} \mathcal{C}_k. \end{aligned}$$

Now by substituting the variables $\mathcal{A}_j^i$ for $i = 0, 2, \dots, k-1$ and $j = 1, 2, \dots, k$, $\mathcal{C}_k$ and $\mathcal{R}_k$ in the recent equality, one has

$$\begin{aligned} \mathbf{P}_k^{-1} &= \begin{bmatrix} \mathbf{E} \\ \mathbf{J}_k \end{bmatrix}^\top \left(\begin{bmatrix} \mathbf{Q}_{k-1}^{-1} & \mathbf{0} \\ \mathbf{0} & \mathbf{R}_k^{-1} \end{bmatrix}^{-1} \right. \\ &\quad \left. + \begin{bmatrix} -\mathbf{E} \Upsilon_1^\alpha - \mathbf{H}_{k-1} \\ \mathbf{0} \end{bmatrix} \mathbf{P}_{k-1} \begin{bmatrix} -\mathbf{E} \Upsilon_1^\alpha - \mathbf{H}_{k-1} \\ \mathbf{0} \end{bmatrix}^\top \right) \end{aligned}$$

$$\begin{aligned} &+ \begin{bmatrix} (-1)^2 \mathbf{E} \Upsilon_2^\alpha \\ \mathbf{0} \end{bmatrix} \mathbf{P}_{k-2} \begin{bmatrix} (-1)^2 \mathbf{E} \Upsilon_2^\alpha \\ \mathbf{0} \end{bmatrix}^\top + \dots \\ &+ \begin{bmatrix} (-1)^k \mathbf{E} \Upsilon_k^\alpha \\ \mathbf{0} \end{bmatrix} \mathbf{P}_0 \begin{bmatrix} (-1)^k \mathbf{E} \Upsilon_k^\alpha \\ \mathbf{0} \end{bmatrix}^\top \Big)^{-1} \begin{bmatrix} \mathbf{E} \\ \mathbf{J}_k \end{bmatrix} \\ &= \mathbf{E}^\top (\tilde{\mathbf{P}}_k)^{-1} \mathbf{E} + \mathbf{J}_k^\top \mathbf{R}_k^{-1} \mathbf{J}_k, \end{aligned}$$

where $\tilde{\mathbf{P}}_k = \mathbf{Q}_{k-1} + \Omega_{k-1}^1$ in which

$$\begin{aligned} \Omega_{k-1}^1 &= (\mathbf{E} \Upsilon_1^\alpha + \mathbf{H}_{k-1}) \mathbf{P}_{k-1} (\mathbf{E} \Upsilon_1^\alpha + \mathbf{H}_{k-1})^\top \\ &+ \sum_{j=2}^k ((-1)^j \mathbf{E} \Upsilon_j^\alpha) \mathbf{P}_{k-j} ((-1)^j \mathbf{E} \Upsilon_j^\alpha)^\top. \end{aligned}$$

Also, from (18) one has

$$\begin{aligned} \hat{\mathbf{x}}_k &= \mathbf{P}_k \mathcal{C}_k^\top \left(\mathcal{R}_k^{-1} + (\mathcal{A}_1^{k-1}) \mathbf{P}_{k-1} (\mathcal{A}_1^{k-1})^\top \right. \\ &\quad \left. + (\mathcal{A}_2^{k-2}) \mathbf{P}_{k-2} (\mathcal{A}_2^{k-2})^\top + \dots \right. \\ &\quad \left. + (\mathcal{A}_{k-1}^1) \mathbf{P}_1 (\mathcal{A}_{k-1}^1)^\top + (\mathcal{A}_k^0) \mathbf{P}_0 (\mathcal{A}_k^0)^\top \right)^{-1} \\ &\quad \times \left(\mathcal{Y}_k - \begin{bmatrix} \mathcal{A}_1^{k-1} & \mathcal{A}_2^{k-2} & \dots & \mathcal{A}_k^0 \end{bmatrix} \begin{bmatrix} \hat{\mathbf{x}}_{k-1} \\ \hat{\mathbf{x}}_{k-2} \\ \vdots \\ \hat{\mathbf{x}}_0 \end{bmatrix} \right). \end{aligned}$$

By substituting the variables $\mathcal{A}_j^i$ for $i = 0, 2, \dots, k-1$ and $j = 1, 2, \dots, k$, $\mathcal{C}_k$, $\mathcal{R}_k$ and $\mathcal{Y}_k$ in the recent equality, one has $\hat{\mathbf{x}}_k = \mathbf{P}_k (\mathbf{E}^\top (\tilde{\mathbf{P}}_k)^{-1} \tilde{\mathbf{x}}_k + \mathbf{J}_k^\top \mathbf{R}_k^{-1} \tilde{\mathbf{y}}_k)$, where $\tilde{\mathbf{x}}_k = -\mathbf{H}_{k-1} \mathbf{x}_{k-1}^f + \mathbf{H}(\mathbf{x}_{k-1}^f, \mathbf{u}_{k-1}) - \Omega_{k-1}^2$ and $\tilde{\mathbf{y}}_k = \mathbf{y}_k + \mathbf{J}_k \mathbf{x}_k^h - \mathbf{J}(\mathbf{x}_k^h, \mathbf{u}_{k-1})$ in which

$$\Omega_{k-1}^2 = (-\mathbf{E} \Upsilon_1 - \mathbf{H}_{k-1}) \hat{\mathbf{x}}_{k-1} + \sum_{j=2}^k (-1)^j \mathbf{E} \Upsilon_j^\alpha \hat{\mathbf{x}}_{k-j}.$$

By defining $\mathbf{x}_{k-1}^f = \hat{\mathbf{x}}_{k-1}$ and $\mathbf{x}_k^h = \tilde{\mathbf{x}}_k$, the proof is completed. $\square$

We are now at a point where we can derive the estimation algorithm. According to the minimization problem (11), the obtained recursive sequence in Theorem 2 is our solution for the deterministic filtered least-squares fitting problem. Inspired by the classic KF, we organize this recursive algorithm by incorporating the prediction and the update concepts to each step and summarize the overall procedure in the form of a pseudo-code shown in Algorithm 1. We call this recursive algorithm as EFSKF for the state estimation of the NFOS system (8) summarized in the following remark.

Remark 3 Suppose that the stochastic NFOS system (8) is estimable that means the condition (13)

holds. For the given sequence $\mathbf{Y}_k$, the state estimation (EFSKF) algorithm of this nonlinear system is given by Algorithm 1.

Remark 4 At each time step, the Jacobian matrices of $\mathbf{H}_i$ and $\mathbf{J}_i$ are evaluated with current estimated states and transformed estimated states through the NFOS function (8a), respectively. This process at projection step essentially linearizes the nonlinear functions around the current estimate which is the best state vector estimate.

Algorithm 1 EFSKF algorithm

- 1: Initialization:
 - 2: Input: $\bar{\mathbf{x}}_0, \bar{\mathbf{P}}_0, \mathbf{R}_0, \mathbf{y}_0, \mathbf{u}_0$
 - 3: Compute the Jacobian matrix:
 $\mathbf{J}_0 \leftarrow \frac{\partial \mathbf{J}(\mathbf{x}_0, \mathbf{u}_0)}{\partial \mathbf{x}_0} \Big|_{\bar{\mathbf{x}}_0}$
 - 4: Output:
 - 5: Initiate the state estimation:
 $\hat{\mathbf{x}}_0 \leftarrow \mathbf{P}_0(\bar{\mathbf{P}}_0^{-1}\bar{\mathbf{x}}_0 + \mathbf{J}_0^T \mathbf{R}_0^{-1} \bar{\mathbf{y}}_0)$
 - 6: Initiate the error covariance:
 $\mathbf{P}_0 \leftarrow (\bar{\mathbf{P}}_0^{-1} + \mathbf{J}_0^T \mathbf{R}_0^{-1} \mathbf{J}_0)^{-1}$
 - 7: Return: $\hat{\mathbf{x}}_0, \mathbf{P}_0$
 - 8: **for** $i = 0$ to k **do**
 - 9: Input: $\mathbf{x}_i, \mathbf{y}_i, \mathbf{Q}_i, \mathbf{R}_i, \mathbf{u}_i$
 - 10: Projection:
 - 11: Compute the Jacobian matrix:
 $\mathbf{H}_i \leftarrow \frac{\partial \mathbf{H}(\mathbf{x}_i, \mathbf{u}_i)}{\partial \mathbf{x}_i} \Big|_{\hat{\mathbf{x}}_{i-1}}$
 - 12: Project the a priori prediction:
 $\tilde{\mathbf{x}}_i \leftarrow -\mathbf{H}_{i-1} \hat{\mathbf{x}}_{i-1} + \mathbf{H}(\hat{\mathbf{x}}_{i-1}, \mathbf{u}_{i-1}) - \Omega_{i-1}^2$
 - 13: Compute the Jacobian matrix:
 $\mathbf{J}_i \leftarrow \frac{\partial \mathbf{J}(\mathbf{x}_i, \mathbf{u}_i)}{\partial \mathbf{x}_i} \Big|_{\tilde{\mathbf{x}}_i}$
 - 14: Project the a priori error covariance:
 $\tilde{\mathbf{P}}_i \leftarrow \mathbf{Q}_{i-1} + \Omega_{i-1}^1$
 - 15: Output (Updating):
 - 16: update the error covariance:
 $\mathbf{P}_i \leftarrow (\mathbf{E}^T(\tilde{\mathbf{P}}_i)^{-1} \mathbf{E} + \mathbf{J}_i^T \mathbf{R}_i^{-1} \mathbf{J}_i)^{-1}$
 - 17: update the estimation states:
 $\hat{\mathbf{x}}_i \leftarrow \mathbf{P}_i \mathbf{E}^T \tilde{\mathbf{P}}_i^{-1} \tilde{\mathbf{x}}_i + \mathbf{P}_i \mathbf{J}_i^T \mathbf{R}_i^{-1} \tilde{\mathbf{y}}_i$
 - 18: **end for**
 - 19: Return $\hat{\mathbf{x}}_k, \mathbf{P}_k$
-

4.3 Filter Analysis

Algorithm 1 induces the most general kalman filtering among various filters presented in the literature. The proposed NFSKF covers the traditional EKF and also the linear filtering algorithms such as SKF, FKF and KF.

Corollary 1 Consider the stochastic NFOS system (8) with causes the system collapses to a stochastic nonlinear singular model. In this case, the EFSKF algorithm collapses to the extended SKF (ESKF) algorithm to an updated Algorithm 1 in projection step as $\tilde{\mathbf{x}}_i = \mathbf{H}(\hat{\mathbf{x}}_{i-1}, \mathbf{u}_{i-1}) + \mathbf{E}\hat{\mathbf{x}}_{i-1}$ and

$$\tilde{\mathbf{P}}_i = \mathbf{Q}_{i-1} + (\mathbf{H}_{i-1} + \mathbf{E})\mathbf{P}_{i-1}(\mathbf{H}_{i-1} + \mathbf{E})^T.$$

Proof In this case, we need to just consider the derived results with $\Upsilon_1^\alpha = \mathbf{I}_n$ and Υ_j^α for $2 \leq j \leq k$ equal to zero matrix since

$$\binom{1}{j} = 0, \quad \text{for } j > 1.$$

It is worth to mention that the additional terms $\mathbf{E}\hat{\mathbf{x}}_{i-1}$ and two matrices $\mathbf{E}$, respectively, in $\tilde{\mathbf{x}}_i$ and $\tilde{\mathbf{P}}_i$, are because of the additional term $\mathbf{E}\mathbf{x}_k$ in (8a) after converting to the integer order case as

$$\mathbf{E}\mathbf{x}_{k+1} - \mathbf{E}\mathbf{x}_k = \mathbf{H}(\mathbf{x}_k, \mathbf{u}_k) + \mathbf{w}_k.$$

□

Corollary 2 Consider the stochastic NFOS system (8) with $\mathbf{E} = \mathbf{I}_n$ causes the system collapses to a stochastic nonlinear fractional-order standard model. In this case, the EFSKF algorithm collapses to the EFKF reported in [15].

Proof Let's $\mathbf{E} = \mathbf{I}_n$ in the derived EFSKF algorithm shown in Algorithm 1. Then, the a priori prediction $\tilde{\mathbf{x}}_i$ and error covariance $\tilde{\mathbf{P}}_i$ can be rewritten as

$$\tilde{\mathbf{x}}_i = \mathbf{H}(\hat{\mathbf{x}}_{i-1}, \mathbf{u}_{i-1}) - \sum_{j=1}^i (-1)^j \Upsilon_j^\alpha \hat{\mathbf{x}}_{i-j},$$

$$\begin{aligned} \tilde{\mathbf{P}}_i &= \mathbf{Q}_{i-1} + (\mathbf{H}_{i-1} + \Upsilon_1^\alpha) \mathbf{P}_{i-1} (\mathbf{H}_{i-1} + \Upsilon_1^\alpha)^T \\ &\quad + \sum_{j=2}^i (\Upsilon_j^\alpha) \mathbf{P}_{i-j} (\Upsilon_j^\alpha)^T. \end{aligned}$$

Also, applying the Woodbury formula on the updated error covariance $\hat{\mathbf{P}}_i$, one has

$$\mathbf{P}_i = (\tilde{\mathbf{P}}_i^{-1} + \mathbf{J}_i^T \mathbf{R}_i^{-1} \mathbf{J}_i)^{-1}$$

$$\begin{aligned}
&= \tilde{\mathbf{P}}_i - \tilde{\mathbf{P}}_i \mathbf{J}_i^\top \mathbf{R}_i^{-1} \mathbf{J}_i \tilde{\mathbf{P}}_i (\mathbf{I} + \mathbf{J}_i^\top \mathbf{R}_i^{-1} \mathbf{J}_i \tilde{\mathbf{P}}_i)^{-1} \\
&= \tilde{\mathbf{P}}_i - \tilde{\mathbf{P}}_i \mathbf{J}_i^\top (\mathbf{R}_i + \mathbf{J}_i \tilde{\mathbf{P}}_i \mathbf{J}_i^\top)^{-1} \mathbf{J}_i \tilde{\mathbf{P}}_i.
\end{aligned}$$

By defining $\mathbf{K}_i := \tilde{\mathbf{P}}_i \mathbf{J}_i^\top (\mathbf{R}_i + \mathbf{J}_i \tilde{\mathbf{P}}_i \mathbf{J}_i^\top)^{-1}$ as KF gain matrix, one can obtain the Riccati equation as $\mathbf{P}_i = (\mathbf{I} - \mathbf{K}_i \mathbf{J}_i) \tilde{\mathbf{P}}_i$. Also,

$$\begin{aligned}
\hat{\mathbf{x}}_i &= (\tilde{\mathbf{P}}_i - \mathbf{K}_i \mathbf{J}_i \tilde{\mathbf{P}}_i) (\tilde{\mathbf{P}}_i^{-1} \tilde{\mathbf{x}}_i + \mathbf{J}_i^\top \mathbf{R}_i^{-1} \tilde{\mathbf{y}}_i) \\
&= \Phi_i \tilde{\mathbf{x}}_i + (\tilde{\mathbf{P}}_i - \tilde{\mathbf{P}}_i \mathbf{J}_i^\top \Psi_i^{-1} \mathbf{R}_i^{-1} \mathbf{J}_i \tilde{\mathbf{P}}_i) \mathbf{J}_i^\top \mathbf{R}_i^{-1} \tilde{\mathbf{y}}_i \\
&= \Phi_i \tilde{\mathbf{x}}_i + \tilde{\mathbf{P}}_i \mathbf{J}_i^\top (\mathbf{R}_i^{-1} - \Psi_i^{-1} \mathbf{R}_i^{-1} \mathbf{J}_i \tilde{\mathbf{P}}_i \mathbf{J}_i^\top \mathbf{R}_i^{-1}) \tilde{\mathbf{y}}_i \\
&= \Phi_i \tilde{\mathbf{x}}_i + \tilde{\mathbf{P}}_i \mathbf{J}_i^\top (\mathbf{R}_i^{-1} - \mathbf{R}_i^{-1} \mathbf{J}_i \tilde{\mathbf{P}}_i \mathbf{J}_i^\top \Psi_i^{-1} \mathbf{R}_i^{-1}) \tilde{\mathbf{y}}_i \\
&= \Phi_i \tilde{\mathbf{x}}_i + \tilde{\mathbf{P}}_i \mathbf{J}_i^\top (\mathbf{R}_i + \mathbf{J}_i \tilde{\mathbf{P}}_i \mathbf{J}_i^\top)^{-1} \tilde{\mathbf{y}}_i,
\end{aligned}$$

where $\Phi_i = (\mathbf{I} - \mathbf{K}_i \mathbf{J}_i)$ and $\Psi_i = (\mathbf{I} + \mathbf{R}_i^{-1} \mathbf{J}_i \tilde{\mathbf{P}}_i \mathbf{J}_i^\top)$. By substituting the defined Kalman gain in the recent equality, one has the following updated estimation state equation in its standard form

$$\begin{aligned}
\hat{\mathbf{x}}_i &= (\mathbf{I} - \mathbf{K}_i \mathbf{J}_i) \tilde{\mathbf{x}}_i + \mathbf{K}_i (\mathbf{y}_i + \mathbf{J}_i \tilde{\mathbf{x}}_i - \mathbf{J}(\tilde{\mathbf{x}}_i, \mathbf{u}_i)) \\
&= \tilde{\mathbf{x}}_i + \mathbf{K}_i (\mathbf{y}_i - \mathbf{J}(\tilde{\mathbf{x}}_i, \mathbf{u}_i)).
\end{aligned}$$

□

Remark 5 Consider the stochastic NFOS system (8) with $\alpha = \mathbf{I}_n$ and $\mathbf{E} = \mathbf{I}_n$ causes the system collapses to a stochastic nonlinear model. In this case, it is not difficult to show that the EFSKF algorithm coincides to the traditional EKF algorithm reported in [20].

As seen, not only the EFSKF covers the three other nonlinear filtering ESKF, EFKF and EKF but also it covers FSKF, SKF, FKF and KF algorithms introduced in [1, 13], [17–19], [15] and [14], respectively, when the system (8) collapses to a linear FOS system. In a hierarchical structure, Figure 1 illustrates these connections in visual perspective. The resulting KF algorithm is a common type of filter that covers all Kalman filter algorithms among all types of KFs, which have been introduced in the literature.

Remark 6 With respect to the inversion lemma reported in [18], the Algorithm 1 can be reformulated in the form of a three-block framework as $[\hat{\mathbf{x}}_k \ \mathbf{P}_k] = \mathcal{U}_k^\top \Pi_k^{-1} [-\mathcal{U}_k \ \Omega_k]$, where $\mathcal{U}_k^\top = [\mathbf{0}_{n \times n} \ \mathbf{0}_{p \times n} \ \mathbf{I}_n]^\top$, and

$$\Pi_k = \begin{bmatrix} \tilde{\mathbf{P}}_k & \mathbf{0}_{n \times p} & \mathbf{E} \\ \mathbf{0}_{p \times n} & \mathbf{R}_k & \mathbf{J}_k \\ \mathbf{E}^\top & \mathbf{J}_k^\top & \mathbf{0}_n \end{bmatrix}, \Omega_k = \begin{bmatrix} \tilde{\mathbf{x}}_k \\ \tilde{\mathbf{y}}_k \\ \mathbf{0}_{n \times 1} \end{bmatrix},$$

with initial conditions

$$\Pi_0 = \begin{bmatrix} \tilde{\mathbf{P}}_0 & \mathbf{0}_{n \times p} & \mathbf{I}_n \\ \mathbf{0}_{p \times n} & \mathbf{R}_0 & \mathbf{J}_0 \\ \mathbf{I}_n & \mathbf{J}_0^\top & \mathbf{0}_n \end{bmatrix}, \Omega_0 = \begin{bmatrix} \tilde{\mathbf{x}}_0 \\ \tilde{\mathbf{y}}_0 \\ \mathbf{0}_{n \times 1} \end{bmatrix},$$

Fig. 1 Hierarchical structure on KF algorithms connections

and given $\tilde{\mathbf{y}}_k$ and $\tilde{\mathbf{y}}_0$. This alternative representation brings a number of advantages such as consistency with some difficulties including singular measurement noise and missing redundant perfect information. In this matter, this solution not only can use the pseudo-inverses instead of standard matrix inverses, it consists less complexity in computational processes. Moreover, this framework can drive us towards derivation of the Algorithm 1 based on the ML approach.

For the future work, the derived EFSKF algorithm here can be extended to a more general form of NFOS, where it is rectangular and time variant (TV) accompanied with the correlated noises. Some similar works were done in [1, 13] for the linear cases called general KF (GKF). Also, it would be interesting to [extend the proposed results to a NFOS system subject to uncertainties and derive a robust KF for such systems.](#)

5 Simulation Results

In this section, we try to implement some numerical simulations to illustrate the efficiency of the presented algorithm. First, we consider a mathematical NFOS model and try to estimate its states, and then, considering a physical chaotic system, we synchronize its trajectories via EFSKF.

Example 1 Consider the NFOS system (8) with the following parameters:

$$\mathbf{H}(\mathbf{x}_k, \mathbf{u}_k) = \begin{bmatrix} 5x_{2,k} - 0.8x_{1,k}x_{2,k} \\ 4x_{1,k}x_{2,k} - x_{2,k} - 20 \end{bmatrix}, \mathbf{E} = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix},$$

$$\mathbf{J}(\mathbf{x}_k, \mathbf{u}_k) = 0.1x_{1,k} + 0.3x_{2,k}^2 + x_{2,k}, \alpha = \text{diag}\{0.5\}.$$

From the given nonlinear functions, one can obtain the following Jacobian matrices $\mathbf{J}_i = [0.1 \ 0.6x_{2,i} + 1]$ and

$$\mathbf{H}_i = \begin{bmatrix} -0.8x_{2,i} & 5 - 0.8x_{1,i} \\ 4x_{2,i} & 4x_{1,i} - 1 \end{bmatrix}.$$

With respect to the positive initial conditions $x_{1,0} = 2$ and $x_{2,0} = 1$, the aforementioned system is estimable because

$$\text{rank} \begin{bmatrix} \mathbf{E} \\ \mathbf{J}_i \end{bmatrix} = \text{rank} \begin{bmatrix} 1 & 0 \\ 0 & 0 \\ 0.1 & 0.6x_{2,i} + 1 \end{bmatrix} = 2, \quad \text{for } i > 0,$$

meaning that the estimability condition meets, since the last entry of the matrix will not equal to zero through the always positive values of the second state. As seen from Fig. 2 the two states are estimated with high accuracy.

Fig. 2 State estimation of the NFOS system in Example 1 using the EFSKF algorithm

Example 2 (Chaos Synchronization) In this example, we try to consider the FOS logistic model reported in [7] in its chaotic mode and analyze the synchronization performance via EFSKF algorithm. Towards

this, consider the NFOS system (8) with the following parameters:

$$\mathbf{H}(\mathbf{x}_k, \mathbf{u}_k) = \begin{bmatrix} rx_k \left(1 - \frac{x_{1,k}}{K}\right) - h(\mathbf{x}_k) \\ \rho h(\mathbf{x}_k) - cx_{2,k} - m \end{bmatrix}, \quad \mathbf{E} = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$$

and $\mathbf{J}(\mathbf{x}_k, \mathbf{u}_k) = 0.1x_{1,k} + 0.3x_{2,k}$ in which the biological parameters are $K = 8$, $\rho = 5$, $c = 1$, the harvest function is considered as $h(\mathbf{x}_k) = x_{1,k}x_{2,k}$ and $0 < \alpha < 1$. Besides the fractional order α , the variation of the economic profit m and growth rate r cause the dynamics of system changes and undergoes to a period-doubling bifurcation rout to chaos and also a reverse behavior. For example, when the order α increases, a period becomes chaotic through the period doubling. Figure 3 shows this behavior in detail for the first state with two nearby initial conditions. Also, according to the bifurcation diagrams against the variable α depicted in Figure 4 (4a and 4b), when two parameters m and r increases and decreases, respectively, the distance between the symmetric transitions in two opposite (period-doubling and reverse period-doubling) treatments decreases until the band of chaotic behavior vanishes for $m = 5$ and $r = 1.8$.

Fig. 3 The graphs of the first state with two nearby initial conditions $m = 2$ and $r = 2$

Fig. 4 Bifurcation diagrams of the first state against variation of α

Now, let's assume consider the FOS logistic map in its chaotic phase by considering $\alpha = \text{diag}\{0.85\}$, $m = 1$ and $r = 2.8$ with initial conditions $x_{1,0} = 2$ and $x_{2,0} = 1$, and apply the EFSKF presented in Algorithm 1 to estimate the states of the system towards synchronization process of the trajectories. The considered system is estimable since (13) holds according to the following equality.

$$\text{rank} \begin{bmatrix} \mathbf{E} \\ \mathbf{J}_i \end{bmatrix} = \text{rank} \begin{bmatrix} 1 & 0 \\ 0 & 0 \\ 0.1 & 0.3 \end{bmatrix} = 2, \quad \text{for } i > 0.$$

Figure 5 shows the two chaotic trajectories of the FOS logistic map system and their estimates obtained by the EFSKF algorithm developed in this paper. As can be seen, after a short time, the estimates are synchronized with the original states. Figure 6 illustrates the error vectors $e_{1,k} = x_{1,k} - \hat{x}_{1,k}$ and $e_{2,k} = x_{2,k} - \hat{x}_{2,k}$ in which both converge to zero. To illustrate the efficiency of the algorithm more, the mean square error (MSE) for the two states has been calculated according to the following equation

$$\text{MSE}_j = \frac{1}{N} \sum_{i=0}^N (x_{j,i} - \hat{x}_{j,i})^2, \quad \text{for } j = 1, 2,$$

where $x_{j,i}$ and $\hat{x}_{j,i}$ are j th state variable and its estimation at step i , respectively, and N is the number of samples. The analysis shows that the MSE value for two states of the systems are as $\text{MSE}_1 = 0.034$ and $\text{MSE}_2 = 0.038$ which reflect a satisfactory result

of the algorithm performance. It is worth mentioning that the chaotic system here and numerical simulation towards its synchronization can be considered as a future set of stage to implement secure communication and encryption technique.

Fig. 5 Two states of FOS logistic map and their estimations using the EFSKF algorithm

Fig. 6 Synchronization errors

6 Conclusion

In this paper, we have proposed a state estimation algorithm design for stochastic NFOS systems for the first time. Besides formulating and investigating the index and causality concepts of these systems, we have introduced the notion of estimability for such systems and designed a new general form of the KF called EFSKF for state estimation of said systems on deterministic arguments in a

completely self-contained way besides the stochastic reasoning. We have also demonstrated how the features and capabilities of the derived algorithm are superior to the existing KF technologies. The newly designed algorithm overcomes some constraints such as the singular filter form, regularity conditions, inflexibility and infinite terms that result from the coexistence of fractional and singular features when using usual approaches. Furthermore, we have illustrated the efficiency of the presented algorithm and have shown how it can be manipulated for a physical chaotic system towards synchronization of its trajectories. Derivation of the estimation algorithm for more general NFOS cases including rectangular and TV systems with the correlated process and measurement noises, and also, based on pure stochastic reasoning such as ML approach can be seen as material for future investigation. The authors also are working on robust KF problems of the FOS systems in both linear and nonlinear cases.

Acknowledgments. The work of K. Nosrati in the project “ICT programme” was supported by the European Union through European Social Fund. The work of K. Nosrati, A. Tepljakov, and E. Petlenkov was also partly supported by the Estonian Research Council grant PRG658.

Declarations

- **Conflict of interest** The authors declare that they have no conflict of interest.
- **Data availability** All data generated or analysed during this study are included in this article.

References

- [1] Nosrati, K., Shafiee, M.: Kalman filtering for discrete-time linear fractional-order singular systems. *IET Control Theory Appl.* 12(2), 1254–1266 (2018)
- [2] Podlubny, I., Fractional Differential Equations. Academic Press, New York (1999)
- [3] Tepljakov, A., Baykant Alagoz, B., Yeroglu, C., Gonzalez, E.A., Hosseinnia, S.H., Petlenkov, E., Ates, A., Cech, M.: Towards Industrialization of FOPID Controllers: A Survey on Milestones of Fractional-Order Control and Pathways for Future Developments. *IEEE Access* 9, 21016–21042 (2021)
- [4] Chen, X., Xi, L., Zhang, Y., Ma, H., Huang, Y., Chen, Y.: Fractional techniques to characterize non-solid aluminum electrolytic capacitors for power electronic applications. *Nonlinear Dyn.* 98(4), 3125–3141 (2019)
- [5] Mu, Y., Zhang, H., Su, H., Ren, H.: Unknown input observer synthesis for discrete-time T–S fuzzy singular systems with application to actuator fault estimation. *Nonlinear Dyn.* 100(4), 3399–3412 (2020)
- [6] Chowdhury, N., Belikov, J., Baimel, D., Levron, Y.: Observer-based detection and identification of sensor attacks in networked CPSs. *Automatica* 121, 109166 (2020)
- [7] Nosrati, K., Shafiee, M.: Fractional-order singular logistic map: Stability, bifurcation and chaos analysis. *Chaos Solitons Fract.* 115, 224–238 (2018)
- [8] Kaczorek, T., Rogowski, K.: Fractional Linear Systems and Electrical Circuits. Springer, Switzerland (2015)
- [9] Nosrati, K., Shafiee, M.: Dynamic analysis of fractional-order singular Holling type-II predator–prey system. *Appl. Math. Comput.* 313, 159–179 (2017)
- [10] Sajewski, L.: Descriptor fractional discrete-time linear system with two different fractional orders and its solution. *Bull. Pol. Acad. Sci.: Tech. Sci.* 64(1), 15–20 (2016)
- [11] Liu, S., Zhou, X.F., Li, X., Jiang, W.: Stability of fractional nonlinear singular systems and its applications in synchronization of complex dynamical networks. *Nonlinear Dyn.* 84(4), 2377–2385 (2016)
- [12] Komachali, F.P., Shafiee, M.: Sensor fault diagnosis in fractional-order singular systems using unknown input observer. *Int. J. Syst. Sci.* 51(1), 116–132 (2019)

- [13] Nosrati, K., Shafiee, M.: State estimation of time varying fractional singular systems. In: 2020 28th Iranian Conference on Electrical Engineering (ICEE), pp. 1-5. IEEE (2020)
- [14] Nosrati, K., Abbott, D., Shafiee, M.: Maximum likelihood estimation of stochastic fractional singular models. *IEEE Access* 9, 128276-128287 (2021)
- [15] Sierociuk, D., Dzieliński, A.: Fractional Kalman filter algorithm for states, parameters and order of fractional system estimation. *Int. J. Appl. Math. Comput. Sci.* 16(1), 129-140 (2006)
- [16] Hsieh, C.S.: State estimation for descriptor systems via the unknown input filtering method. *Automatica* 49(5), 1281–1286 (2013)
- [17] Boukroune, B., Darouach, M., Zasadzinski, M.: Moving horizon state estimation for linear discrete-time singular systems. *IET Control Theory Appl.* 4(3), 339–350 (2010)
- [18] Nikoukhah, R., Willsky, A.L., Levy, B.C.: Kalman filtering and Riccati equations for descriptor systems. *IEEE Trans. Autom. Control* 37(9), 1325–1342 (1992)
- [19] Ishihara, J.Y., Terra, M.H., Campos, J.C.: Optimal recursive estimation for discrete-time descriptor systems. *Int. J. Syst. Sci.* 36(10), 605–615 (2005)
- [20] McElhoe, B.A.: An Assessment of the Navigation and Course Corrections for a Manned Flyby of Mars or Venus, *IEEE Trans. Aerosp. Electron. Syst.* 2(4), 613–623 (1966)
- [21] Simon, D.: *Optimal State Estimation: Kalman, H Infinity, and Nonlinear Approaches*. Wiley, Hoboken (2006)
- [22] Gao, Z.: Cubature Kalman filters for nonlinear continuous-time fractional-order systems with uncorrelated and correlated noises. *Nonlinear Dyn.* 96(3), 1805-1817 (2019)
- [23] Aliyu, M., Boukas, E.: Kalman filtering for affine nonlinear descriptor systems. *Circuits Syst. Signal Process.* 30(1), 125-142 (2011)
- [24] Becerra, V., Roberts, P., Griffiths, G.: Applying the extended Kalman filter to systems described by nonlinear differential-algebraic equations. *Control Eng. Pract.* 9(3), 267–281 (2001)
- [25] Mandela, R.K., Rengaswamy, R., Narasimhan, S., Sridhar, L.N.: Recursive state estimation techniques for nonlinear differential algebraic systems. *Chem. Eng. Sci.* 65(16), 4548–4556 (2010)
- [26] Ashayeri, L., Shafiee, M., Menhaj, M.: Kalman filter for fractional order singular systems. *Am. J. Sci.* 9(1), 209–216 (2013)
- [27] Gantmacher, F.R.: *The Theory of Matrices II*. Chelsea Publishing Company, New York (1960)
- [28] Willems, J.C.: Deterministic least squares filtering. *J. Econom.*, 118(1), 341–373 (2004)
- [29] Sorenson, H.W.: *Least-squares estimation: from Gauss to Kalman*. *IEEE Spectr.* 7(7), 63-68 (1970)
- [30] Henderson, H.V., Searle, S.R.: On deriving the inverse of a sum of matrices. *SIAM Rev.* 23(1), 53–60 (1981)