Tobit Kalman Filtering for Fractional-Order Systems with Stochastic Nonlinearities under Round-Robin Protocol

Hang Geng, Zidong Wang, Xiaojian Yi, Fuad E. Alsaadi and Yuhua Cheng*

Abstract

In this paper, the Tobit Kalman filtering problem is investigated for a class of discrete time-varying fractional-order systems in the presence of measurement censoring and stochastic nonlinearities under the Round-Robin protocol (RRP). The fractional-order dynamic model is described by the Grunwald-Letnikov difference equation, and the statistical means are utilized to characterize the stochastic nonlinearities that include state-dependent stochastic disturbances as a special case. The RRP is employed to decide the transmission sequence of sensors so as to alleviate undesirable data collisions. Under the RRP scheduling, only one sensor is permitted to transmit its measurement over the network at each time instant. In light of the renowned orthogonality projection principle, a protocol-based fractional Tobit Kalman filter (TKF) is devised with the fractional dynamics and stochastic nonlinearities elaborately addressed. In the pursuit of the filter design, a couple of new terms appear which are in relation to the RRP, fractional dynamics and stochastic nonlinearities arise, and these terms are adequately handled recursively or off-line. Simulation results are provided to demonstrate the usefulness of the proposed method.

Index Terms

Tobit Kalman filtering, fractional-order system, measurement censoring, Round-Robin protocol, stochastic nonlinearities.

I. INTRODUCTION

The last decade has seen ever-increasing research enthusiasm in fractional calculus due mainly to its strong capability in characterizing phenomena exhibiting fractional dynamics. For example, it has been revealed that, dynamics of rubber isolators, traffic situations within information networks, relaxation...
processes of organic dielectric materials, quantum evolutions of complex systems, etc. can all be described by fractional differential/difference equations (FDDEs) [31], [39], [49]. To date, FDDEs have been successfully applied to the modeling, estimation and control problems in a great variety of practical scenarios such as biological systems, networked systems, immune systems and complex networks, see [1], [9], [41] and the references therein.

The introduction of FDDEs in system modeling gives rise to the so-called fractional-order systems that offer adaptability and flexibility to the controller/estimator design, see e.g. [9], [32], [37], [41]. Taking advantage of the memory property of the fractional-order system, a fractional-order human immunodeficiency virus (HIV) model has been put forward in [9] to account for the memory function of immune responses, and the associate control problem has been resolved via properly optimizing a prescribed objective function. In [41], the pinning control problem for fractional-order complex networks has been researched, and it has been uncovered that the network can manage to stabilize itself below a certain order without pinning a single node.

Regarding the state estimation problems of fractional-order systems, two different fractional Kalman filters (FKFs) have been established in [32] and [37], respectively, on the basis of different system representations. Due to its recursive structure and concise implementation, the FKF has been quickly empowered to deal with correlated noises [36], colored noises [38] and Lévy noises [46]. To further extend its application scope, much research effort has been devoted to the development of FKFs suitable for various systems, for instance, multi-sensor systems, singular systems and nonlinear systems [12], [30], [47].

It has been well recognized that nonlinearities are ubiquitous that occur in almost all sorts of physical systems, and the investigation on nonlinearities has played an important role in the context of state estimation. In case of noisy circumstances (e.g. sensor networks), the nonlinearities may appear in the form of random disturbances stemming from randomly perturbed network environments and/or transmission limitations. In this regard, stochastic nonlinearities would become unavoidable and, if not properly settled, they might give rise to severe deterioration of system performance. Accordingly, much research attention has been paid to the filtering/control problems for systems with stochastic nonlinearities.

Aiming at the state estimation problem for a general class of uncertain nonlinear stochastic systems, a linear matrix inequality approach has been developed in [48] where stochastic nonlinearities have been characterized via statistical means. Such kind of characterization has then been extensively exploited in the design of filters for the state estimation problems in the concurrence of stochastic nonlinearities and networked-induced phenomena (e.g. packet dropouts, time delays and fading measurements) [6]–[8], [15], [16]. Additionally, the distributed filtering issue has been addressed in [25] for systems contaminated by stochastic nonlinearities and sensor degradation, where a sufficient condition has been acquired to ensure the mean-square boundedness of the associated estimation errors. The resilient $H_\infty$ filtering problem has been addressed in [18] for a class of stochastically uncertain discrete-time systems with randomly occurring gain variations, nonlinearities and channel fadings. The distributed variance-constrained robust filtering problem has been studied in [45] for a class of time-varying stochastic systems subject to ran-
domly occurring nonlinearities and missing measurements, where a sufficient condition has been provided to guarantee the boundedness of the filtering error covariance. Recently, the resilient state estimation problem has been resolved in [11] for uncertain time-varying recurrent neural networks with randomly varying nonlinearities and missing measurements, where the variance constraint and the prescribed $H_\infty$ performance have been ensured.

Apart from the aforementioned stochastic nonlinearities, another category of measurement nonlinearity, namely censored observation, has become a hotspot of research in recent years owing to its prevalent existence in engineering domains where the low-cost commercial off-the-shelf sensors are deployed [2], [4]. In the event of censored observations, the classic Kalman filter (KF) becomes inapplicable because measurement noises nearing the censored region turn out to be non-Gaussian with unknown statistics [5]. As such, the Tobit Kalman filter (TKF) has been proposed in [3] to provide a fully recursive state estimation paradigm for handling censored observations. Due to its succinct structure and recursive form, the TKF has immediately been employed in a broad range of application scenarios, for example, the cooperative localization, fault detection and target tracking, see [10], [13], [14], [17], [22].

In the context of TKF, the state estimation problem has been settled in [22] with both censored observations and multiplicative noises. Later on, TKFs under Lévy noises [13] and redundant channel transmission [14] have been devised, and examples concerning maneuvering target tracking have been presented to elucidate the efficiency of the proposed algorithms. To mitigate impacts from modeling uncertainties, a modified TKF has been designed in [10] with illustrative examples on unmanned aerial vehicle systems. In addition, the fault detection problem has also been tackled in [17] for systems with dead-zone-like censoring.

In the past decade, there has been a growing demand on the installation of networked systems due largely to their broad applications in practical areas such as the communication, patient monitoring and target localization [19], [21], [24], [25], [29], [40], [50]. In networked systems, a typical assumption is that all system components (e.g. actuators, controllers, filters and sensors) have the permission to broadcast their information via the shared medium. This assumption, however, is not really reasonable as limited-bandwidth-induced data collisions are likely to happen which would impair the communication efficiency especially when data transmission is launched simultaneously by more than one component [23], [27], [28], [33], [34]. To mitigate unnecessary data collisions, communication protocols have been utilized to orchestrate the transmission sequences by only permitting a single component to enter the network at each time instant [35], [42]–[44]. Among various communication protocols, the Round-Robin protocol (RRP) has gained a particular popularity because of its succinct execution manner, where the data transmission among system components is implemented in a fixed circular order, see e.g. [26], [51]–[53].

The primary objective of this paper is to fill in such a gap by designing a protocol-based fractional TKF robust to stochastic nonlinearities. This appears to be a non-trivial task for the following difficulties: 1) it is indistinct what fractional difference equation should be selected to describe the fractional dynamics of the system; 2) it is pretty challenging to build a protocol-based fractional Tobit regression model where effects of stochastic nonlinearities are also accommodated; and 3) it is fairly challenging to exam-
ine the performance of the resultant filtering algorithm with influences from fractional-order dynamics, measurement censoring and stochastic nonlinearities also considered. This motivates us to carry out the present study. In response to these difficulties, the main contributions are highlighted as follows. 1) To the best of the authors’ knowledge, this paper represents one of the first few attempts at the protocol-based Tobit Kalman filtering problem for fractional-order systems corrupted by stochastic nonlinearities, where the formulated model is holistic in catering for the general practice; 2) a protocol-based fractional TKF is developed to account for the fractional dynamics and RRP, and also mitigate the adverse impacts caused by censored observations and stochastic nonlinearities; 3) the performance of the developed filtering algorithm is ensured through examining the exponential boundedness of the filtering error dynamics in the mean square sense.

The remainder of the paper is organized as follows. In Section II, the problem under consideration is formulated. In Section III, a modified Tobit regression model is built, based on which a protocol-based fractional TKF is designed. In Section IV, a numerical example is provided to show the usefulness of the filter, and some conclusions are drawn in Section V.

**Notation** The notation used here is fairly standard except where otherwise stated. $\mathbb{R}^n$ denotes the $n$-dimensional Euclidean space. $I$ and 0 represent identity and zero matrices with compatible dimensions, respectively. The subscripts “-1” and “$T$” represent inverse and transpose operations, respectively. $A_{mn}$ and $A_m$ represent the $(m, n)$th sub-block and the $m$th row of matrix or variable $A$, respectively. $y_{1:k}$ stands for all the measurements up to time instant $k$. $\mathbb{E}\{x\}$ and $\mathbb{E}\{x|y\}$ will, respectively, mean the expectation of $x$ and the expectation of $x$ conditional on $y$. diag${X_m}$ ($m=1, 2, \ldots, p$) stands for a block-diagonal matrix with matrices $X_m$ on the diagonal. vec${x_m} = [x_1 \ x_2 \ \cdots \ x_p]^T$. var${x}$ denotes the variance of $x$. $\|x\|$ is the Euclidean norm of vector $x$. sign($x$) stands for the sign function where $\text{sign}(x) = -1$ if $x < 0$, $\text{sign}(x) = 0$ if $x = 0$, and $\text{sign}(x) = 1$ if $x > 0$. $\delta(\cdot) \in \{0, 1\}$ is the Dirac delta function.

**II. Problem Formulation**

Consider the following discrete-time fractional-order system in the presence of stochastic nonlinearities (see [36]):

$$\Delta^q x_{k+1} = A_k x_k + f(x_k, \eta_k) + \omega_k,$$  \hspace{1cm} (1)

$$x_{k+1} = \Delta^q x_{k+1} - \sum_{j=1}^{k+1} (-1)^j \Upsilon_j x_{k+1-j},$$  \hspace{1cm} (2)

$$z_k = C_k x_k + g(x_k, \zeta_k) + \nu_k,$$  \hspace{1cm} (3)

where

$$\Delta^q x_{k+1} = \left[ \begin{array}{cccc} \Delta^{q_1} x_{1,k+1} & \Delta^{q_2} x_{2,k+1} & \cdots & \Delta^{q_n} x_{n,k+1} \end{array} \right]^T,$$

$$\Upsilon_j = \text{diag} \left\{ \left( \frac{q_1}{j} \right), \left( \frac{q_2}{j} \right), \cdots, \left( \frac{q_n}{j} \right) \right\},$$
where \( \bar{y}_m \) is updated as follows (see [26], [51]):

\[
\bar{y}_{m,k} = \begin{cases} 
    z_{m,k}, & \text{if } \text{mod}(k - m, p) = 0, \\
    \bar{y}_{m,k-1}, & \text{otherwise},
\end{cases}
\]

(4)

Taking advantage of the update coefficient \( \Gamma_{m,h_k} \), (4) is transformed into

\[
\bar{y}_{m,k} = \sum_{l=0}^{p-1} \Gamma_{m,h_k-l} z_{m,k-l},
\]

(5)

where \( h_k \) is the unique non-negative remainder on division of \( k - m \) by \( p \), and \( z_{m,k} \) is the measurement update coefficient where \( \Gamma_{m,h_k} \) is the measurement update coefficient where \( \Gamma_{m,h_k} \) is the measurement update coefficient where \( \bar{y}_m \) is the actual measurement arriving at the estimator after network transmission with the zero-order holder strategy.

Denote \( \text{mod}(k - m, p) \) as the unique non-negative remainder on division of \( k - m \) by \( p \), and \( \bar{y}_m \) is the actual measurement arriving at the estimator after network transmission with the zero-order holder strategy.

Denote \( \text{mod}(k - m, p) \) as the unique non-negative remainder on division of \( k - m \) by \( p \), and \( \bar{y}_m \) is the actual measurement arriving at the estimator after network transmission with the zero-order holder strategy.

Denote \( \text{mod}(k - m, p) \) as the unique non-negative remainder on division of \( k - m \) by \( p \), and \( \bar{y}_m \) is the actual measurement arriving at the estimator after network transmission with the zero-order holder strategy.

Denote \( \text{mod}(k - m, p) \) as the unique non-negative remainder on division of \( k - m \) by \( p \), and \( \bar{y}_m \) is the actual measurement arriving at the estimator after network transmission with the zero-order holder strategy.

Denote \( \text{mod}(k - m, p) \) as the unique non-negative remainder on division of \( k - m \) by \( p \), and \( \bar{y}_m \) is the actual measurement arriving at the estimator after network transmission with the zero-order holder strategy.
where $\gamma_{m,k}$ $(m = 1, 2, \ldots, p)$ are Bernoulli random variables governing the censoring phenomena of $\bar{y}_{m,k}$ with following probability distributions:

$$\text{Prob}\{\gamma_{m,k} = 1\} = \bar{\gamma}_{m,k}, \text{Prob}\{\gamma_{m,k} = 0\} = 1 - \bar{\gamma}_{m,k}.$$ (8)

Here, $\bar{\gamma}_{m,k}$ are known non-negative constants. It is supposed that $\gamma_{m,k}$ are uncorrelated with $\eta_k$, $\zeta_k$ and other noise signals.

Let $y_{1:k}$ and $y_{m,1:k}$ be the measurement collections from, respectively, all sensors and the $m$th sensor up till time $k$, and

$$y_k \triangleq \text{vec}\{y_{m,k}\}, \quad \hat{x}_k \triangleq \mathbb{E}\{x_k|y_{1:k-1}\},$$

$$\hat{x}_k^\perp \triangleq \mathbb{E}\{x_k|y_{1:k-1}\}, \quad \hat{y}_k \triangleq \mathbb{E}\{y_k|y_{1:k-1}\},$$

$$\bar{x}_k \triangleq x_k - \hat{x}_k, \quad \bar{y}_k \triangleq y_k - \hat{y}_k,$$

$$P_{\bar{x}_k} \triangleq \mathbb{E}\{\bar{x}_k\bar{x}_k^T\}, \quad P_{\bar{y}_k} \triangleq \mathbb{E}\{\bar{y}_k\bar{y}_k^T\}.$$  

**Assumption 1:** 1) The initial state $x_0$ has the mean $\bar{x}_0$ and covariance $P_0$. 2) $x_0$, $\eta_k$, $\zeta_k$, $\omega_k$ and $\upsilon_k$ are mutually independent.

**Assumption 2:** The nonlinearity functions $f(x_k, \eta_k)$ and $g(x_k, \zeta_k)$ satisfy $f(0, \eta_k) = 0$, $g(0, \zeta_k) = 0$ with the following first-order moment

$$\mathbb{E}\left\{ \begin{bmatrix} f(x_k, \eta_k) \\ g(x_k, \zeta_k) \end{bmatrix} \left| x_k \right. \right\} = 0,$$

and covariance

$$\mathbb{E}\left\{ \begin{bmatrix} f(x_k, \eta_k) \\ g(x_k, \zeta_k) \end{bmatrix} \begin{bmatrix} f(x_t, \eta_t) \\ g(x_t, \zeta_t) \end{bmatrix}^T \left| x_k \right. \right\} = 0, k \neq t,$$

$$\mathbb{E}\left\{ \begin{bmatrix} f(x_k, \eta_k) \\ g(x_k, \zeta_k) \end{bmatrix} \begin{bmatrix} f(x_t, \eta_t) \\ g(x_t, \zeta_t) \end{bmatrix}^T \left| x_k \right. \right\} = \sum_{i=1}^{r} \Pi_i x_k^T \Gamma_i x_k, k = t,$$

where $\Pi_i = \text{diag}\{\Pi_{1i}, \Pi_{2i}\}$ $(i = 1, 2, \ldots, r)$, $r$ is a known positive integer and $\Gamma_i$ are known matrices with appropriate dimensions.

**Remark 1:** In contrast with its integer-order counterpart, the fractional-order dynamic model (which is based on the fractional calculus) is capable of providing more appropriate characterizations of certain nonlinear dynamics ranging from friction to slipping. Among the various fractional calculus definitions (e.g. the Grunwald-Letnikov definition, Riemann-Liouville definition and Caputo definition), the Grunwald-Letnikov definition has been recognized as particularly suitable for modeling discrete-time systems with fractional orders because of its backward difference form. As such, in this paper, the fractional-order dynamic model (1)–(2) (originated from the Grunwald-Letnikov definition) is adopted for the subsequent filter design.
Remark 2: As is shown in Assumption 2, the concerned stochastic nonlinearities are fairly generic due to its coverage of some well-studied nonlinearities including 1) state-dependent noises, i.e. \( f(x_k, \eta_k) = B_k x_k \eta_k \) and \( g(x_k, \zeta_k) = D_k x_k \zeta_k \) where \( B_k \) and \( D_k \) are known compatible matrices; 2) random sequences whose powers rely on sector-bounded nonlinear state functions, i.e. \( f(x_k, \eta_k) = \tilde{f}(x_k) \eta_k \) and \( g(x_k, \zeta_k) = \tilde{g}(x_k) \zeta_k \) where \( ||\tilde{f}(x_k)|| \leq \alpha ||x_k||, ||\tilde{g}(x_k)|| \leq \beta ||x_k|| \) and \( \alpha \) and \( \beta \) are known scalars; and 3) random sequences whose powers rely on the signs of nonlinear state functions, i.e. \( f(x_k, \eta_k) = B_k \text{sign}[f(x_k)] x_k \eta_k \) and \( g(x_k, \zeta_k) = D_k \text{sign}[g(x_k)] x_k \zeta_k \). In this sense, the nonlinear stochastic model considered in this paper is with respect to the nonlinearities in our paper is considerably general to reflect the engineering practice.

Remark 3: Under the scheduling of the RRP, equal priority is assigned to each sensor and the measurement from individual sensor is admitted to enter the network in a fixed circular manner. In the event that \( z_{m,k} \) from sensor \( m \) has no entry into the network, the zero-order holder strategy is utilized for the generation of \( \bar{y}_{m,k} \) for the purpose of offsetting \( \bar{y}_{m,k} \). Consequently, at time instant \( k - l \), only the component \( \bar{y}_{h_{k-l},k-l} \) of the actually arrived measurement \( \bar{y}_{k-l} \) is updated, whilst the rest components \( \bar{y}_{m,k-l} \) \((m = 1, 2, \ldots, p, m \neq h_{k-l})\) remain the same as their counterparts in \( \bar{y}_{k-l-1} \). Obeying the fixed circular order for information propagation, \( \bar{y}_{m,k} \) \((m = 1, 2, \ldots, p)\) can be represented by the sum of \( \Gamma_{m,h_{k-l}} z_{m,k-l} \) \((l = 0, 1, \ldots, p - 1)\) as shown in (5).

Remark 4: It is observed from (7) that random variables \( \gamma_{m,k} \) \((m = 1, 2, \ldots, p)\) are used to describe the censoring phenomena of \( \bar{y}_{m,k} \). In accordance with (7), if no censoring occurs for \( \bar{y}_{m,k} \), i.e. \( \gamma_{m,k} = 1 \), we have \( y_{m,k} = \bar{y}_{m,k} \), which means that the output observation is equivalent to the latent one. If the censoring occurs for \( \bar{y}_{m,k} \), i.e. \( \gamma_{m,k} = 0 \), we have \( y_{m,k} = \mathcal{I}_m \), which means that the censoring threshold is allocated to the output observation. Here, the censoring probabilities \( \bar{\gamma}_{m,k} \) are known a priori via some statistical experiments. Alternatively, as with [3], \( \bar{\gamma}_{m,k} \) can also be approximated by

\[
\bar{\gamma}_{m,k} \approx \Phi \left( \frac{\sum_{l=0}^{p-1} \Gamma_{m,h_{k-l}} C_{m,k-l} s_{k-l} - \mathcal{I}_m}{\sqrt{\sum_{l=0}^{p-1} \Gamma_{m,h_{k-l}}^2 R_{m,k-l}}} \right),
\]

where \( s_{k-l} \triangleq \hat{x}_{k-l} \) for \( l = 0 \), \( s_{k-l} \triangleq \hat{x}_{k-l} \) for \( l = 1, 2, \ldots, p - 1 \), and \( \Phi(\cdot) \) is the cumulative distribution function (CDF) of the random variable “·” which obeys the standard normal distribution.

The objective of this paper is to design the TKF for the discrete-time fractional-order system (1)–(8) under the RRP and Assumptions 1–2.

III. MAIN RESULTS

This section aims at formalizing a specifically tailored protocol-based fractional Tobit Kalman filtering paradigm to surmount the identified challenges brought by the fractional dynamics, stochastic nonlinearities and RRP. The formulation procedure differentiates itself from the conventional TKF through the following perspectives: (1) an ameliorated Tobit regression model is presented that stems from the involvement of the fractional dynamics, stochastic nonlinearities and RRP; and (2) extra computations with respect to the state prediction, measurement prediction, filter gain as well as associate covariances are encompassed.
Let
\[
\tilde{y}_{m,k} \triangleq \xi_{m,k} + \nu_{m,k}, \quad R_{m,k} \triangleq \sum_{l=0}^{p-1} \Gamma_{m,h_{k-l}}^2 R_{m,k-l},
\]
\[
\varphi_m \triangleq \frac{I_m - \xi_{m,k}}{R_{m,k}}, \quad \nu_{m,k} \triangleq \sum_{l=0}^{p-1} \Gamma_{m,h_{k-l}}\nu_{m,k-l}
\]
\[
\xi_{m,k} \triangleq \sum_{l=0}^{p-1} \Gamma_{m,h_{k-l}} \left( C_{m,k-l}x_{k-l} + g_m(x_{k-l}, \zeta_{k-l}) \right).
\]

The Tobit regression model for system (1)–(8) is obtained below.

**Lemma 1:** The mean and variance of \( y_{m,k} \) (\( m = 1, 2, \ldots, p \)) are given by
\[
\mathbb{E}\{y_{m,k}|x_k\} = [1 - \Phi(\varphi_{m,k})] \left[ \xi_{m,k} + \sqrt{R_{m,k}}\lambda(\varphi_{m,k}) \right] + \Phi(\varphi_{m,k}) I_m,
\]
\[
\text{var}\{y_{m,k}|x_k\} = R_{m,k}[1 - \varphi(\varphi_{m,k})],
\]
where
\[
\varphi(\varphi_{m,k}) = \lambda(\varphi_{m,k}) \left[ \lambda(\varphi_{m,k}) - \varphi_{m,k} \right],
\]
\[
\lambda(\varphi_{m,k}) = \frac{\phi(\varphi_{m,k})}{1 - \Phi(\varphi_{m,k})},
\]
and \( \phi(\varphi_{m,k}), \Phi(\varphi_{m,k}) \) are, respectively, the probability density function (PDF) and CDF of the Gaussian random variable \( \varphi_{m,k} \) with \( \phi(\varphi_{m,k}) \) and \( \Phi(\varphi_{m,k}) \) defined as
\[
\phi(\varphi_{m,k}) = \frac{1}{\sqrt{2\pi}} e^{-\frac{(y_{m,k}-\zeta_{m,k})^2}{2R_{m,k}}},
\]
\[
\Phi(\varphi_{m,k}) = \int_{-\infty}^{\varphi_{m,k}} \frac{1}{\sqrt{2\pi R_{m,k}}} e^{-\frac{(y_{m,k}-\zeta_{m,k})^2}{2R_{m,k}}} dy_{m,k}.
\]

**Proof:** See Appendix–A.

**Remark 5:** The Tobit regression model manifested in Lemma 1 embodies the mean and variance of \( y_{m,k} \). In contrast with its counterpart in [3] where merely the censoring phenomenon is concerned, it is seen that the term \( C_{m,k}x_k \) is now substituted by the term
\[
\xi_{m,k} \triangleq \sum_{l=0}^{p-1} \Gamma_{m,h_{k-l}} \left( C_{m,k-l}x_{k-l} + g_m(x_{k-l}, \zeta_{k-l}) \right).
\]

What is more, the original measurement noise covariance \( R_{m,k} \) is replaced by the sum of \( p \) noise covariances \( R_{m,k} \triangleq \sum_{l=0}^{p-1} \Gamma_{m,h_{k-l}}^2 R_{m,k-l} \). These two replacements sketch clearly the influences from stochastic nonlinearities and RRP on the desired regression model. Also, the impact from the fractional dynamics is reflected from the state \( x_{k-l} \) integrated in the term \( \xi_{m,k} \) by noticing the correlation between \( x_{k-l} \) and \( \Delta^\mu x_{k-l} \) as expressed in (1)–(2). It is noteworthy that, the impact from the fractional dynamics embedded in \( x_{s,k-l} \) gives rise to the emergence of a few new terms in the subsequent filter design.

For \( m = 1, 2, \ldots, p \), we let
\[
\mathcal{I} \triangleq \text{vec}\{\mathcal{I}_m\}, \quad \bar{\gamma}_k \triangleq \text{diag}\{\bar{\gamma}_{m,k}\}, \quad \bar{\xi}_k \triangleq \text{vec}\{\bar{\xi}_{m,k}\}, \quad \bar{\varphi}_k \triangleq \text{vec}\{\bar{\varphi}_{m,k}\},
\]
\[ R_k \triangleq \text{diag} \{ R_{m,k} \}, \quad \lambda (\bar{\vartheta}_k) \triangleq \text{vec} \{ \lambda (\bar{\vartheta}_{m,k}) \}, \quad \varphi (\bar{\vartheta}_k) \triangleq \text{diag} \{ \varphi (\bar{\vartheta}_{m,k}) \} \]

where

\[ \hat{\xi}_{m,k} \triangleq \sum_{l=0}^{p-1} \Gamma_{m,h_{k-1}} C_{m,k-l} s_{k-l}, \quad \bar{\vartheta}_{m,k} \triangleq \frac{T_m - \hat{\xi}_{m,k}}{R_{m,k}} \]

and \( \varphi (\bar{\vartheta}_{m,k}) \) and \( \lambda (\bar{\vartheta}_{m,k}) \) can be, respectively, computed via (12) and (13) by replacing \( \xi_{m,k} \) with \( \hat{\xi}_{m,k} \).

On the basis of Lemma 1, the protocol-based TKF for system (1)–(8) is derived in the following theorem.

**Theorem 1:** For system (1)–(8), its protocol-based fractional TKF is established as follows:

\[ \hat{x}_{k} = (A_{k-1} + \mathcal{Y}_1) \hat{x}_{k-1} - \sum_{j=2}^{k} (-1)^j \mathcal{Y}_j \hat{x}_{k-j}, \tag{16} \]

\[ \hat{y}_k = \tilde{\gamma}_k \left[ \hat{\xi}_k - \sqrt{R_k} \lambda (\bar{\vartheta}_k) \right] + (I - \tilde{\gamma}_k) \mathcal{Z}, \tag{20} \]

\[ K_k = P_{\hat{x}_k\hat{y}_k}^{-1} \tag{21} \]

where

\[ P_{\hat{x}_k\hat{y}_k} = P_{\hat{x}_k\hat{y}_k} - \sum_{i=1}^{r} \Pi_{1i} \text{tr} \left( P_{\hat{x}_{k-1}} \mathcal{Y}_i \right) \sum_{j=2}^{k} (-1)^j \mathcal{Y}_j P_{\hat{x}_{k-j}} \mathcal{Y}_j^T. \tag{24} \]

**Proof:** See Appendix–B.

**Remark 6:** Making comparison between the desired protocol-based fractional TKF in Theorem 1 and the traditional TKF in [3], two remarkable distinctions can be encapsulated. One is the replacement of the term \( C_k \hat{x}_k \) (which is the product of the measurement matrix \( C_k \) and the one-step state prediction \( \hat{x}_k \) by the term \( \hat{\xi}_k = \Gamma_{h_k} C_k \hat{x}_k + \sum_{l=1}^{p-1} \Gamma_{h_{k-1}} C_{k-l} \hat{x}_{k-l} \) which is the sum of \( p \) products riding on the measurement update coefficient \( \Gamma_{h_{k-1}} \), measurement matrix \( C_{k-l} \), one-step state prediction \( \hat{x}_{k-l} \) and past state estimate \( \hat{x}_{k-l} \). The other is the emergence of a suite of new terms \( \Delta \hat{q}_{\hat{x}_k}, P_{\hat{x}_{k-1}} \sum_{i=1}^{r} \Pi_{1i} \text{tr} \left( P_{\hat{x}_{k-1}} \mathcal{Y}_i \right) \) and \( \sum_{j=1}^{k} (-1)^j \mathcal{Y}_j P_{\hat{x}_{k-j}} \mathcal{Y}_j^T \) in computing the state prediction and filter gain. The first distinction results from the RRP whilst the second one roots in the fractional dynamics and stochastic nonlinearities.
TABLE I: The Pseudocode of the Protocol-Based Fractional TKF

<table>
<thead>
<tr>
<th>Algorithm: Protocol-Based TKF</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Input:</strong> $\bar{x}<em>0, P_0, y</em>{1:k}$</td>
</tr>
<tr>
<td><strong>Output:</strong> $\hat{x}<em>k, P</em>{\hat{x}_k}$</td>
</tr>
<tr>
<td>1: let $\hat{x}_0 = \bar{x}<em>0, P</em>{\hat{x}_0} = P_0$</td>
</tr>
<tr>
<td>2: for $k = 1 : N$ do</td>
</tr>
<tr>
<td>3: calculate the predicted estimate $\hat{x}<em>k$ by (16) and associated covariance $P</em>{\hat{x}_k}$ by (18);</td>
</tr>
<tr>
<td>4: calculate the gain matrix $K_k$ by (21)–(24);</td>
</tr>
<tr>
<td>5: calculate the updated estimate $\tilde{x}<em>k$ by (17) and associated covariance $P</em>{\tilde{x}_k}$ by (19);</td>
</tr>
<tr>
<td>6: end for</td>
</tr>
</tbody>
</table>

Lemma 1, together with Theorem 1, gives rise to the protocol-based fractional Tobit Kalman filtering algorithm with its pseudocode outlined in Table I.

Remark 7: Thanks to Lemma 1 and Theorem 1, a reinforced Tobit Kalman filtering framework is devised with a view to solving the novel filtering problem for fractional-order systems susceptible to stochastic nonlinearities under the RRP. System (1)–(8) under investigation is holistic that not only incorporates fractional behaviors of the system but also accounts for important randomly occurring phenomena (i.e. censored observations and stochastic nonlinearities) which are prevalently confronted in application scenarios ranging from networked control to object tracking. These disparate phenomena are tackled within a unified yet effective scheme. In addition, the RRP is adopted to avoid data collisions and boost transmission reliability.

Next, we move forward to discussion the performance of the designed fractional TKF. Due to the time-varying nature of the protocol-induced measurement update coefficient $\Gamma_{\bar{h}_k}$, the convergence of the developed fractional TKF cannot be guaranteed in general. Thus, we turn to pursue the boundedness of the developed algorithm where the exponential boundedness of the filtering error dynamics is analyzed in the mean square sense.

To start with, the following definition is first introduced for discussion convenience.

**Definition 1:** For real numbers $\epsilon > 0$, $\varepsilon > 0$ and $0 < \chi < 1$, if for all $k \geq 0$,

$$
\mathbb{E} \left\{ \| \tilde{x}_k \|^2 \right\} \leq c \mathbb{E} \left\{ \| \tilde{x}_0 \|^2 \right\} \chi^k + \varepsilon,
$$

holds, then the stochastic process $\tilde{x}_k$ is exponentially bounded in the mean square sense.

The following assumption is essential in deriving our results.

**Assumption 3:** The following conditions are satisfied for all $k \geq 0$:

$$
\| \Pi_{2i} \| \leq \bar{\pi}, \| \gamma_k \| \leq \bar{\gamma}, \| C_k \| \leq \bar{c},
$$

$$
\| \Upsilon_j \| \leq \bar{r}, \| \Pi_{1i} \| \leq \bar{\pi}, \| \hat{x}_k \| \leq \bar{\chi}, q \leq \bar{q} I \leq Q_k \leq \bar{q} I
$$

$$
\| A_k \| \leq \bar{a}, \| \Gamma_i \| \leq \bar{\tau}, \| \text{tr}(P_{x_{k-1}}) \| \leq \bar{s}, \| P_{y_{k}}^{-1} \| \leq \bar{\eta}.
$$

**Theorem 2:** Consider the discrete-time fractional-order system (1)–(8) with its filter in the form of (16)–(24). Under Assumption 3, the resulting filtering error dynamics is exponentially bounded in mean square sense.
Proof: See Appendix–C.

Remark 8: It is noteworthy that the developed filter in Theorem 1 is a generalization of those in [3] and [37]. In the case that model (1)–(2) is an integer-order one without the nonlinearity $f(x_k, \eta_k)$, the nonlinearity $g(x_k, \zeta_k)$ is omitted in (3) and the RRP is abandoned, the modified Tobit regression model in Lemma 1 and the protocol-based fractional TKF in Theorem 1 will, respectively, degenerate to the traditional Tobit regression model and TKF in [3]. In the case that the aforementioned randomly occurring phenomena are absent and the RRP is discarded, the Tobit regression model in Lemma 1 will disappear and the protocol-based fractional TKF in Theorem 1 will reduce to the renowned FKF in [37].

IV. ILLUSTRATIVE EXAMPLE

In this section, a numerical example (modified from [37]), a spring-mass-damper example (modified from [20]) and an oscillator example (modified from [3]) are leveraged to elucidate the usefulness of the presented filter design approach and associate filtering performance. Denote the root mean-squared errors (RMSEs) of $x_{1,k}$ and $x_{2,k}$, respectively, as

$$\text{RMSE}_1 \triangleq \sqrt{(1/M) \sum_{i=1}^{M} (\hat{x}_{1,k}^{(i)} - x_{1,k})^2},$$

$$\text{RMSE}_2 \triangleq \sqrt{(1/M) \sum_{i=1}^{M} (\hat{x}_{2,k}^{(i)} - x_{2,k})^2},$$

where $M$ is the number of Monte Carlo trials.

A. Numerical Example

Consider system (1)–(8) with parameters:

$$A_k = \begin{bmatrix} 0 & 1 \\ -0.1 & -0.2 \end{bmatrix}, C_k = \begin{bmatrix} 0.1 & 0.3 \\ 0.5 & 1.5 \end{bmatrix},$$

$$Q_k = R_k = 0.3I_2, P_{x_0} = P_{\tilde{x}_0} = 100I_2,$$

$q_1 = 0.7, q_2 = 1.2, x_0 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}^T, \mathcal{I} = \begin{bmatrix} -1 & -1 \end{bmatrix}^T.$

The stochastic nonlinearities $f(x_k, \eta_k)$ and $g(x_k, \zeta_k)$ are

$$f(x_k, \eta_k) = \begin{bmatrix} 0.2 \\ 0.3 \end{bmatrix}^T 0.3 \sin (x_{1,k}) x_{1,k} \eta_{1,k} + \begin{bmatrix} 0.2 \\ 0.3 \end{bmatrix}^T 0.4 \sin (x_{2,k}) x_{2,k} \eta_{2,k},$$

$$g(x_k, \zeta_k) = \begin{bmatrix} 0.1 \\ 0.1 \end{bmatrix}^T 0.3 \sin (x_{1,k}) x_{1,k} \zeta_{1,k} + \begin{bmatrix} 0.1 \\ 0.1 \end{bmatrix}^T 0.4 \sin (x_{2,k}) x_{2,k} \zeta_{2,k},$$

where $\eta_{1,k}, \eta_{2,k}, \zeta_{1,k}$ and $\zeta_{2,k}$ are zero-mean uncorrelated Gaussian white noises with unity covariances. Apparently, the above stochastic nonlinearities satisfy Assumption 2 with

$$r = 1, \quad \Pi_{11} = \begin{bmatrix} 0.04 & 0.06 \\ 0.06 & 0.09 \end{bmatrix},$$

$$\Pi_{21} = \begin{bmatrix} 0.01 & 0 \\ 0 & 0.01 \end{bmatrix}, \quad \Gamma_1 = \begin{bmatrix} 0.09 & 0 \\ 0 & 0.16 \end{bmatrix}.$$
Fig. 1: State and its estimate.

Fig. 2: Performance comparison in RMSE1 and RMSE2.

Fig. 1 depicts the true state values and associate estimates provided by the protocol-based TKF (which is named as PBTKF and is barely capable of tackling censored observations under the RRP in case of \( q_1 = q_2 = 1 \)) and the protocol-based fractional TKF with stochastic nonlinearities (which is named as PBFTKF-SN and is capable of addressing both censoring and nonlinearities under the RRP in case of \( q_1 = 0.7 \) and \( q_2 = 1.2 \)). Fig. 2 plots the comparison result in RMSE between the PBTKF and PBFTKF-SN after 1000 independent Mote Carlo trials.

It is confirmed from Fig. 1 that, our PBFTKF-SN manages to track the true state values precisely,
whilst the PBTKF has considerable deviations from true state values. Additionally, it is observed from Fig. 2 that the RMSE curve of our PBFTKF-SN always resides lower than that of the PBTKF, implying that issues of the fractional dynamics and stochastic nonlinearities are well settled in our PBFTKF-SN whereas they are not disposed of in the PBTKF.

Besides, to better illustrate the relationship between the censoring threshold and the filtering performance, simulations of our PBFTKF-SN based on different censoring thresholds have been conducted where corresponding state estimates and RMSEs are demonstrated in Figs. 3–4. It is apparently observed from Figs. 3–4 that, as censoring thresholds \([I_1, I_2]\) increase, the filtering accuracy with respect to the
system state deteriorates. This well coincides with the fact that, in case of larger censoring thresholds, measurements are more likely to be censored and less information can be used for state estimation, leading to performance degeneracy of our PBFTKF-SN.

B. Spring-Mass-Damper Example

Consider a simple mass-spring-damper system [20] as illustrated in Fig. 5. The equation of motion for such a model is given by

$$m \frac{d^2 x_t}{dt^2} + c \frac{dx_t}{dt} + k x_t = f_t,$$

where $m$ is the mass, $k$ is the spring constant and $c$ is the damping coefficient. $x_t$ is the system state and $f_t$ is the control term. Let $f_t = \Delta^{0.5}x_t + u_t$ where $\Delta$ is the the fractional differential/difference operator and $u_t$ is the known input. Referring to [20], the above model can be transformed into the following fractional-order differential equation to describe the viscoelastically damped structure:

$$m\Delta^2 x_t + c\Delta^1 x_t + k x_t - \Delta^{0.5}x_t = u_t.$$

Letting $m = 0.1$, $c = 0.4$, $m = 0.1$, the fractional-order difference counterpart of the above model is obtained as follows:

$$\Delta^{0.5}x_{k+1} = A_k x_k + B_k u_k,$$

$$x_{k+1} = \Delta^{0.5}x_{k+1} - \sum_{j=1}^{k+1} (-1)^j \Upsilon_j x_{k+1-j},$$

where

$$A_k = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ -0.1 & -0.1 & -0.4 & 0 \end{bmatrix}, B_k = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \end{bmatrix}.$$

Taking into account the possible noise and the stochastic nonlinearity, the above model is further converted into

$$\Delta^q x_{k+1} = A_k x_k + B_k u_k + f(x_k, \eta_k) + \omega_k,$$
\( x_{k+1} = \Delta^q x_{k+1} - \sum_{j=1}^{k+1} (-1)^j \Upsilon_j x_{k+1-j} \),

with its measurement model being

\( z_k = C_k x_k + g(x_k, \zeta_k) + \nu_k \),

which are exactly the same system and measurement models considered in this paper.
Consider the above system with parameters:

\[
C_k = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 \end{bmatrix}, \quad T = \begin{bmatrix} -2 \\ -2 \end{bmatrix}, \quad R_k = 0.0025I_2, \quad P_{x_0} = P_{z_0} = I_4, \\
Q_k = 10^{-4} \begin{bmatrix} 0.1665 & 0.0704 & -0.0198 & -0.1206 \\ 0.0704 & 0.1444 & 0.0257 & -0.1127 \\ -0.0198 & 0.0257 & 0.0991 & -0.0067 \\ -0.1206 & -0.1127 & -0.0067 & 0.2439 \end{bmatrix}, \quad x_0 = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \quad u_k = \sin\left(\frac{\pi}{12} k\right).
\]

The stochastic nonlinearities \( f(x_k, \eta_k) \) and \( g(x_k, \zeta_k) \) are

\[
f(x_k, \eta_k) = \begin{bmatrix} 0.2 & 0.3 & 0.2 & 0.3 \end{bmatrix}^T \begin{bmatrix} 0.3 \sin(x_{1,k})x_{1,k}\eta_{1,k} + 0.4 \sin(x_{2,k})x_{2,k}\eta_{2,k} \\ + 0.3 \sin(x_{3,k})x_{3,k}\eta_{3,k} + 0.4 \sin(x_{4,k})x_{4,k}\eta_{4,k} \end{bmatrix},
\]

\[
g(x_k, \zeta_k) = 0.1 \begin{bmatrix} 0.3 \sin(x_{1,k})x_{1,k}\zeta_{1,k} + 0.4 \sin(x_{2,k})x_{2,k}\zeta_{2,k} \\ + 0.3 \sin(x_{3,k})x_{3,k}\zeta_{3,k} + 0.4 \sin(x_{4,k})x_{4,k}\zeta_{4,k} \end{bmatrix},
\]

where \( \eta_{1,k}, \eta_{2,k}, \zeta_{1,k} \) and \( \zeta_{2,k} \) are zero-mean uncorrelated Gaussian white noises with unity covariances.

Apparently, the above stochastic nonlinearities satisfy Assumption 2 with

\[
r = 1, \quad \Pi_{11} = \begin{bmatrix} 0.04 & 0.06 & 0.04 & 0.06 \\ 0.06 & 0.09 & 0.06 & 0.09 \\ 0.04 & 0.06 & 0.04 & 0.06 \\ 0.06 & 0.09 & 0.06 & 0.09 \end{bmatrix},
\]

\[
\Pi_{21} = 0.01, \quad \Gamma_1 = \text{diag}\{0.09, 0.16, 0.09, 0.16\}.
\]

Figs. 6–7 plot the true state values and associate estimates provided by the PBTKF and PBFTKF-SN, from which it is confirmed that, our PBFTKF-SN manages to track the true state values precisely, whilst the PBTKF has considerable deviations from true state values.

C. Oscillator Example

In the event that the fractional order \( q \) converges to 1, the considered system becomes a first-difference model and the corresponding filtering problem can be resolved based on the existing filtering algorithm for integer order systems. In case of \( q = 1 \), the original fraction-order difference system is reduced to the following first-order difference system:

\[
\Delta^1 x_{k+1} = A_k x_k + f(x_k, \eta_k) + \omega_k,
\]

\[
x_{k+1} = \Delta^1 x_{k+1} + x_k,
\]

\[
z_k = C_k x_k + g(x_k, \zeta_k) + \nu_k,
\]

which is further transformed into

\[
x_{k+1} = \tilde{A}_k x_k + f(x_k, \eta_k) + \omega_k,
\]
Fig. 8: True values of the first and second dimensions of the state and their estimates.

Fig. 9: Performance comparison in RMSE.

\[ z_k = C_k x_k + g(x_k, \zeta_k) + v_k, \]

where \( \tilde{A}_k = A_k + I. \)

Consider the above model with parameters:

\[
\tilde{A}_k = \begin{bmatrix}
\cos(w) & -\sin(w) \\
\sin(w) & \cos(w)
\end{bmatrix}, \quad I = \begin{bmatrix} 0 \\ 0 \end{bmatrix}, \quad w = 0.052\pi,
\]

\[ C_k = I_2, R_k = I_2, Q_k = \text{diag}\{0.0025, 0.0025\}, P_0 = I_2. \]
The oscillator example concerns about the estimation of ballistic roll rates in case of the noisy dynamic model and uncertain magnetometer data. The stochastic nonlinearities \( f(x_k, \eta_k) \) and \( g(x_k, \zeta_k) \) are the same as that adopted in the first numerical example. In order to further verify the effectiveness of our developed fractional-order filtering algorithm in case of \( q = 1 \), the performance comparison is made between the protocol-based Tobit Kalman filter (which is named as PBTKF and is capable of tackling censored observations under the RRP) and our developed protocol-based fractional TKF with stochastic nonlinearities (which is named as PBFTKF-SN and is capable of simultaneously addressing fractional dynamics, censoring and stochastic nonlinearities under the RRP).

Fig. 8 depicts the true state values and associated estimates generated by the PBTKF and our PBFTKF-SN, and Fig. 9 plots the comparison result in RMSE between the PBTKF and PBFTKF-SN after 1000 independent Monte Carlo trials. It is witnessed from Fig. 8 that, the PBFTKF-SN manages to track the true state values precisely, whilst the PBTKF appears to have considerable deviations from the true state values. Besides, it is sketched in Fig. 9 that, the RMSE curve of the PBFTKF-SN resides lower than that of the PBTKF, indicating that issues of sensor nonlinearities are suitably addressed in the PBFTKF-SN, whilst they are not settled in the PBTKF.

V. Conclusion

In this paper, we have dealt with the recursive filtering problem for fractional-order systems in the conjunction of censored observations and stochastic nonlinearities under the RRP. The fractional-order difference equation (stemming from the Grunwald-Letnikov definition of the fractional-order derivative) has been utilized to sketch the evolution of the involved dynamic model. The stochastic nonlinearities have been characterized in a generic form which includes some well-investigated nonlinearities as special cases, and the RRP is selected to schedule the data transmission in the network. These phenomena have been elaborately addressed via developing enhanced Tobit regression model, which generates a bank of new terms in the desired filter and adds extra computations in the algorithm implementation. Luckily, these computations are all recursive or off-line, and therefore the designed TKF is propitious for online scenarios. Finally, the feasibility of the proposed TKF has been verified by a numerical example.

Appendix

A. Proof of Lemma 1

Proof: It is not difficult to learn from the definition \( \nu_{m,k} \triangleq \sum_{l=0}^{p-1} \Gamma_{m,h_{k-l}} \nu_{m,k-l} \) that, \( \nu_{m,k} \) is a Gaussian distributed noise with mean zero and covariance \( R_{m,k} \triangleq \sum_{l=0}^{p-1} \Gamma_{m,h_{k-l}}^2 R_{m,k-l} \). Referring to \( \bar{y}_{m,k} \triangleq x_{m,k} + \nu_{m,k} \) and (7), the PDF of \( y_{m,k} \) is

\[
    f(y_{m,k}|x_k) = \frac{1}{\sqrt{R_{m,k}}} \phi \left( \frac{y_{m,k} - \xi_{m,k}}{\sqrt{R_{m,k}}} \right) u(y_{m,k} - \mathcal{I}_m) \\
    + \delta(\mathcal{I}_m - y_{m,k}) \Phi(\vartheta_{m,k}),
\]

(25)
where \( u(y_m,k - I_m) \) is the unit step function, and \( \phi \left( \frac{y_m,k - \xi_{m,k}}{\sqrt{R_{m,k}}} \right) \) and \( \Phi(\vartheta_{m,k}) \) are calculated via (14)–(15). In line with (25), the mean of \( y_{m,k} \) is derived as

\[
E\{y_{m,k}|x_k\} = \text{Prob}\{y_{m,k} > I_m|x_k\}E\{y_{m,k}|y_{m,k} > I_m, x_k\} + \text{Prob}\{y_{m,k} = I_m|x_k\}E\{y_{m,k}|y_{m,k} = I_m, x_k\}.
\]  

(26)

In order to compute \( E\{y_{m,k}|x_k\} \), the probabilities and means on the right-hand side of (26) should be provided first.

\[
\text{Prob}\{y_{m,k} > I_m|x_k\} = \text{Prob}\{\bar{y}_{m,k} > I_m|x_k\} = 1 - \Phi(\vartheta_{m,k}).
\]  

(27)

In the light of (25), we have

\[
E\{y_{m,k}|y_{m,k} > I_m, x_k\} = 1 - \Phi(\vartheta_{m,k}).
\]  

(28)

where \( \lambda(\vartheta_{m,k}) \) is given by (13). Parallel to (27)-(28), we arrive at

\[
\text{Prob}\{y_{m,k} = I_m|x_k\} = \Phi(\vartheta_{m,k}),
\]  

(29)

\[
E\{y_{m,k}|y_{m,k} = I_m, x_k\} = I_m.
\]  

(30)

Inserting (27)-(30) into (26) yields (10). Making reference to (25), (28) and (30), we have \( \text{var}\{y_{m,k}|y_{m,k} = I_m, x_k\} = 0 \) and

\[
\text{var}\{y_{m,k}|x_k\} = \text{var}\{y_{m,k}|y_{m,k} > I_m, x_k\}
\]  

\[
= E\{y_{m,k}^2|y_{m,k} > I_m, x_k\} - \left( E\{y_{m,k}|y_{m,k} > I_m, x_k\} \right)^2
\]  

\[
= R_{m,k}\left[ 1 - \varphi(\vartheta_{m,k}) \right],
\]

which is exactly the same as (11), where \( \varphi(\vartheta_{m,k}) \) is given by (12).

\[\blacksquare\]

**B. Proof of Theorem 1**

**Proof:** Denoting \( \Delta^q\hat{x}_{k}^- \triangleq E\{\Delta^q\hat{x}_{k}|y_{1:k-1}\} \) and noting Assumption 2, we attain \( \Delta^q\hat{x}_{k}^- = A_{k-1}\hat{x}_{k-1} \).

Adopting Assumption 1 in [36] that past estimates \( \hat{x}_{k-j} \) \( (j = 1, 2, \ldots, k) \) will not be updated with later arrived measurements \( y_{k-j+1} \), we have

\[
E\{x_{k-j}|y_{1:k}\} = E\{x_{k-j}|y_{1:k-j}\} = \hat{x}_{k-j|k-j}.
\]  

(31)
Substituting (1) into (2) produces
\[
x_k = (A_{k-1} + \Upsilon_1)x_{k-1} + f(x_{k-1}, \eta_{k-1}) + \omega_{k-1} - \sum_{j=2}^{k} (-1)^j \Upsilon_j x_{k-j},
\] (32)

Making use of Assumption 2 and (32) produces (24). Putting (2) and (31) into \( \hat{x}_k^- \triangleq \mathbb{E}\{x_k | y_{1:k-1}\} \) and noticing \( \Delta^q \hat{x}_k^- = A_{k-1} \hat{x}_{k-1}^- \), we obtain
\[
\hat{x}_k^- = \mathbb{E} \left\{ \Delta^q x_k - \sum_{j=1}^{k} (-1)^j \Upsilon_j x_{k-j} \bigg| y_{1:k-1} \right\}
\]
\[
= (A_{k-1} + \Upsilon_1) \hat{x}_{k-1}^- - \sum_{j=2}^{k} (-1)^j \Upsilon_j \hat{x}_{k-j},
\]
which is exactly the same as (16). (16) together with (32) generates
\[
\hat{x}_k^- = (A_{k-1} + \Upsilon_1) \hat{x}_{k-1}^- + f(x_{k-1}, \eta_{k-1}) + \omega_{k-1} - \sum_{j=2}^{k} (-1)^j \Upsilon_j \hat{x}_{k-j}.
\] (33)

Substituting (33) into \( P_{\hat{x}_k}^- \triangleq \mathbb{E}\{\hat{x}_k^- (\hat{x}_k^-)^T\} \) and utilizing Assumption 2 in [36] that \( \mathbb{E}\{\hat{x}_k^- (\hat{x}_t^-)^T\} = 0 \) for \( k \neq t \), we have (18).

A direct employment of the orthogonality projection principle to system (1)–(8) yields (17) where the filter gain is given by (21).

(32), along with (17), brings \( \bar{x}_k = \hat{x}_k^- - K_k \bar{y}_k^- \). As a result, \( P_{\bar{x}_k}^- \triangleq \mathbb{E}\{\bar{x}_k^- (\bar{x}_k^-)^T\} \) can be expressed as
\[
P_{\bar{x}_k}^- = \mathbb{E}\{ (\hat{x}_k^- - K_k \bar{y}_k^-) (\hat{x}_k^- - K_k \bar{y}_k^-)^T \}
\]
\[
= P_{\hat{x}_k}^- - P_{\hat{x}_k}^- K_k^T - K_k P_{\hat{x}_k}^- K_k + K_k P_{\bar{y}_k}^- K_k^T.
\] (34)

The combination of (21) and (34) leads to (19).

Making reference to Lemma 1, we obtain
\[
\hat{y}_k^- = [I - \Phi (\hat{\vartheta}_k)] \left[ \hat{\xi}_k^- + \sqrt{R_k} \lambda (\hat{\vartheta}_k) \right] + \Phi (\hat{\vartheta}_k) I
\]
\[
= \tilde{\gamma}_k \bigg[ \hat{\xi}_k^- + \sqrt{R_k} \lambda (\hat{\vartheta}_k) \bigg] + [I - \tilde{\gamma}_k] I,
\]
which is exactly the same as (20). Form (7) and the definition \( y_k \triangleq \text{vec}\{y_{m,k}\} \), we have \( y_k = \gamma_k \bar{y}_k^- + (I - \gamma_k) I \). Subtracting (20) from \( y_k \) induces
\[
\bar{y}_k^- = \gamma_k \bar{y}_k^- + (I - \gamma_k) I - \bar{y}_k^-
\] (35)

Denote \( \bar{y}_k \triangleq \xi_k + \nu_k, \nu_k \triangleq \sum_{l=0}^{p-1} \Gamma_{h_{k-l}} u_{k-l}, \xi_k \triangleq \sum_{l=0}^{p-1} \Gamma_{h_{k-l}} C_{k-l} x_{k-l} + g(x_{k-l}, \zeta_{k-l}) \) and \( \hat{\xi}_k^- \triangleq \xi_k - \hat{\xi}_k^- \). Keeping (33) and (35) in mind, we attain
\[
P_{\bar{x}_k}^- \bar{y}_k^- = \mathbb{E}\{ (x_k - \bar{x}_k^-) (\gamma_k \bar{y}_k^- + (I - \gamma_k) I - \bar{y}_k^-)^T \}
\]
\[
= \mathbb{E}\{ (x_k (\xi_k + \nu_k)^T \gamma_k) - \mathbb{E}\{ \hat{x}_k^- (\xi_k + \nu_k)^T \gamma_k \} \}
\]
\[
= \mathbb{E}\left\{ x_k \left( \sum_{l=0}^{p-1} \Gamma_{h_{k-l}} C_{k-l} x_{k-l} \right)^T \gamma_k^T \right\} - \mathbb{E}\left\{ \hat{x}_k^- \left( \sum_{l=0}^{p-1} \Gamma_{h_{k-l}} C_{k-l} x_{k-l} \right)^T \gamma_k^T \right\}.
\]
where the second equality holds from the independence of \( x_k \) and \( \gamma_k \) and the independence of \( x_k, g(x_{k-l}, \zeta_{k-l}) \) and \( \nu_k \), and the last equality holds from Assumption 2 in [36] that \( \mathbb{E}\{\bar{x}_k x_k^T\} = 0 \) for \( k \neq t \). The notice of

\[
\gamma_k \Gamma_h b_k C_k \gamma_k^T - \hat{x}_k^T (\gamma_k \Gamma_h b_k C_k)^T + \sum_{l=1}^{p-1} \mathbb{E}\{\hat{x}_k x_{k-l}^T\} (\gamma_k \Gamma_h b_{k-l} C_{k-l})^T
\]

(36)

where the second equality holds from the independence of \( x_k \) and \( \gamma_k \) and the independence of \( x_k, g(x_{k-l}, \zeta_{k-l}) \) and \( \nu_k \), and the last equality holds from Assumption 2 in [36] that \( \mathbb{E}\{\bar{x}_k x_k^T\} = 0 \) for \( k \neq t \). The notice of

\[
P_{x_k} = \mathbb{E}\{(\bar{x}_k^T + \hat{x}_k^T)(\bar{x}_k^T + \hat{x}_k^T)\}
\]

\[
P_{x_k} = \mathbb{E}\{\bar{x}_k^T + \hat{x}_k^T\},
\]

together with (36) yields (22).

Recalling the expressions of \( \xi_k \) and \( \hat{\xi}_k \) leads to

\[
\hat{\xi}_k = \sum_{l=0}^{p-1} \Gamma_{h_{k-l}} (C_{k-l} x_{k-l} + g(x_{k-l}, \zeta_{k-l})) - \Gamma_{h_k} C_k \hat{x}_k + \sum_{l=1}^{p-1} \Gamma_{h_{k-l}} C_{k-l} \bar{x}_{k-l}
\]

\[
- \Gamma_{h_k} C_k \bar{x}_k + \sum_{l=0}^{p-1} \sum_{l=1}^{p-1} \Gamma_{h_{k-l}} C_{k-l} \bar{x}_{k-l} + \sum_{l=0}^{p-1} g(x_{k-l}, \zeta_{k-l}),
\]

which gives rise to

\[
\mathbb{E}\{\hat{\xi}_k (\hat{\xi}_k)^T\} = \Gamma_{h_k} C_k P_{x_k} C_k^T \Gamma_{h_k} + \sum_{i=1}^{r} \Pi_2 \text{tr}(P_{x_k} \Gamma_i) + \sum_{l=1}^{p-1} \Gamma_{h_{k-l}} C_{k-l} P_{x_{k-l}} C_{k-l}^T \Gamma_{h_{k-l}},
\]

(37)

where the equality holds from the independence of \( \bar{x}_k \) and \( \bar{x}_{k-l} \). Parallel to the derivation of \( P_{\bar{x}_k \bar{y}_k} \), one has

\[
P_{\tilde{y}_k} = \mathbb{E}\{\tilde{y}_k (\tilde{y}_k)^T\}
\]

\[
= \gamma_k \mathbb{E}\{\hat{\xi}_k (\hat{\xi}_k)^T\} \quad \gamma_k + \mathbb{E}\{\gamma_k \tilde{v}_k \tilde{v}_k^T \gamma_k\},
\]

(38)

where \( \tilde{v}_k = \nu_k - \sqrt{R_k} \lambda (\tilde{v}_k) \). For the sake of derivation brevity, we suppose that \( \text{cov}\{y_{m,k}, y_{s,k}\} = 0 \) for \( m \neq s \) \((m, s = 1, 2, \ldots, p)\). The extension of such a result to the case where \( \text{cov}\{y_{m,k}, y_{s,k}\} \neq 0 \) for \( m \neq s \) is straightforward but rotationally cumbersome. Abiding by such a supposition, we have

\[
\mathbb{E}\{\gamma_k \tilde{v}_k \tilde{v}_k^T \gamma_k\} = \text{diag}\{\text{var}\{\gamma_{m,k} \tilde{v}_{m,k} \tilde{v}_{m,k}^T \gamma_{m,k}\}\}
\]

\[
= \text{diag}\{\text{var}\{y_{m,k} x_k\}\}
\]

\[
= \text{diag}\{\mathcal{R}_{m,k} [1 - \varphi (\tilde{v}_{m,k})]\}
\]

\[
= \mathcal{R}_k [I - \varphi (\tilde{v}_k)].
\]

(39)

Finally, putting (37) and (39) into (38) results in (23).
C. Proof of Theorem 2

Proof: Denote $\mathcal{K}_k \triangleq -\overline{\gamma}_k K_k$ and $u_k \triangleq -\sum_{j=1}^{k} (-1)^j \gamma_j \tilde{x}_{k-j} + f(x_k, \eta_k) + \mathcal{K}_k \Gamma_{h_k}^{-1} \tilde{y}_{k+1}$. Based on 1–(8), the filtering error dynamics is obtained as

\begin{equation}
\tilde{x}_{k+1} = x_{k+1} - \tilde{x}_{k+1} - K_{k+1} \tilde{y}_{k+1} = A_k x_k + f(x_k, \eta_k) + \omega_k - \sum_{j=1}^{k+1} (-1)^j \gamma_j \tilde{x}_{k+1-j} - (A_k + \gamma_1) \tilde{x}_k + \sum_{j=1}^{k} (-1)^j \gamma_j \tilde{x}_{k-j} - K_{k+1} \tilde{y}_{k+1}
\end{equation}

\begin{equation}
= (A_k + \gamma_1) \tilde{x}_k - \sum_{j=1}^{k} (-1)^j \gamma_j \tilde{x}_{k-j} + f(x_k, \eta_k) + \omega_k - K_{k+1} \tilde{y}_{k+1}
\end{equation}

\begin{equation}
= (A_k + \gamma_1) \tilde{x}_k + \omega_k + u_k.
\end{equation}

It follows from 22–(23) that

\begin{equation}
\left\| P_{\tilde{x}_k \tilde{y}_k} \right\| = \left\| P_{\tilde{x}_k} (\gamma_k \Gamma_{h_k} C_k) \right\| \leq \gamma \left\| P_{\tilde{x}_k} \right\| \left\| \Gamma_{h_k} \right\| \left\| C_k \right\| \leq \gamma \bar{c} \left\| P_{\tilde{x}_k} \right\|,
\end{equation}

and

\begin{equation}
\left\| P_{\tilde{x}_k}^{-1} \right\| \leq \left\| [\gamma_k \Gamma_{h_k} C_k P_{\tilde{x}_k}^T (\gamma_k \Gamma_{h_k} C_k)^T]^{-1} \right\| \leq \frac{\gamma^2 \bar{c}^2}{\gamma^2 \bar{c}^2} \left\| P_{\tilde{x}_k}^{-1} \right\|^{-1}.
\end{equation}

Taking the norm on both sides of $\mathcal{K}_k \triangleq -\overline{\gamma}_k K_k$ and keeping (22) and (23) in mind, we have

\begin{equation}
\| \mathcal{K}_k \| \leq \| \overline{\gamma}_k \| \left\| P_{\tilde{x}_k \tilde{y}_k} \right\| \left\| P_{\tilde{x}_k}^{-1} \right\| \leq \frac{\gamma \bar{c}}{\gamma \bar{c}} \triangleq \bar{\kappa}.
\end{equation}

Noting

\begin{equation}
u_k \triangleq -\sum_{j=1}^{k} (-1)^j \gamma_j \tilde{x}_{k-j} + f(x_k, \eta_k) + \gamma_1^{-1} \mathcal{K}_k \tilde{y}_{k+1}
\end{equation}

\begin{equation}= -\sum_{j=1}^{k} (-1)^j \gamma_j (x_{k-j} - \tilde{x}_{k-j}) + f(x_k, \eta_k) + \gamma_1^{-1} \mathcal{K}_k (y_k - \tilde{y}_k),
\end{equation}

and making use of the trace property, we have

\begin{equation}
\mathbb{E} \left\{ u_k^T u_k \right\} \leq \mathbb{E} \left\{ \sum_{j=1}^{k} x_{k-j}^T \gamma_j T_{j} x_{k-j} \right\} + \mathbb{E} \left\{ \sum_{j=1}^{k} \tilde{x}_{k-j}^T \gamma_j T_{j} x_{k-j} \right\} + \mathbb{E} \left\{ f^T (x_k, \eta_k) f (x_k, \eta_k) \right\} + \gamma^{-2} \mathbb{E} \left\{ \tilde{y}_{k+1}^T \gamma_{k+1} K_{k+1} \tilde{y}_{k+1} \right\}
\end{equation}

\begin{equation}= \text{tr} \left\{ \sum_{j=1}^{k} \gamma_j T_{j} P_{x_{k-j}} \right\} + \text{tr} \left\{ \sum_{j=1}^{k} \tilde{x}_{k-j}^T \gamma_j T_{j} \tilde{x}_{k-j} \right\} + \text{tr} \left\{ \sum_{i=1}^{r} \Pi_{i} \text{tr} (P_{x_{k-i}} \Gamma_{i}) \right\} + \text{tr} \left\{ \mathcal{K}_{k+1} P_{\tilde{y}_k} \gamma_{k+1} \right\}
\end{equation}

\begin{equation}\leq k \bar{r} \bar{r}^2 + k \bar{\kappa}^2 \bar{r}^2 + r \bar{r}^2 \bar{r}^2 + \bar{\kappa}^2 \bar{\eta} \triangleq \bar{u}.
\end{equation}
Subsequently, consider the following iterative matrix equation with respect to $\Psi_k$

$$\Psi_{k+1} \triangleq A_k \Psi_k A_k^T + Q_k + \kappa I,$$  \hfill (46)

with the initial value

$$\Psi_0 \triangleq Q_0 + \kappa I,$$

where $\kappa > 0$ is a scalar. Then, it is easy to know that

$$\|\Psi_{k+1}\| \leq \|A_k\|^2 \|\Psi_k\| + \|Q_k\| + \|\kappa I\| \leq \bar{a}^2 \|\Psi_k\| + \bar{q} + \kappa,$$

which, by iteration, leads to

$$\|\Psi_k\| \leq \bar{a}^{2k} \|\Psi_0\| + (\bar{q} + \kappa) \sum_{t=0}^{k-1} \bar{a}^{2t} \leq \bar{a}^{2k} \|\Psi_0\| + (\bar{q} + \kappa) \sum_{t=0}^{\infty} \bar{a}^{2t}. \hfill (47)$$

From Assumption 3, we know $0 < \bar{a}^2 < 1$ and therefore

$$\|\Psi_k\| \leq \|\Psi_0\| + \frac{(\bar{q} + \kappa)}{1 - \bar{a}^2},$$

$$\Psi_k \geq \kappa I.$$

In view of (46) and (47), there exists a positive scalar $\bar{\psi} \triangleq \|\Psi_0\| + \frac{(\bar{q} + \kappa)}{1 - \bar{a}^2}$ such that $\kappa I \leq \Psi_k \leq \bar{\psi} I$ holds for all $k \geq 0$.

Denote $V_k \triangleq x_k^T \Psi_k^{-1} x_k$ and remember $x_{k+1} = A_k x_k + \omega_k + u_k$. For any scalar $\rho > 0$, we have

$$\mathbb{E}\{V_{k+1}\} = (1 + \rho) V_k
\begin{aligned}
= & \mathbb{E}\{(A_k x_k + \omega_k + u_k)^T \Psi_k^{-1} (A_k x_k + \omega_k + u_k)\} - (1 + \rho) x_k^T \Psi_k^{-1} x_k \\
= & \mathbb{E}\{(A_k x_k)^T \Psi_k^{-1} A_k x_k\} + \mathbb{E}\{(A_k x_k)^T \Psi_k^{-1} u_k\} + \mathbb{E}\{(A_k x_k)^T \Psi_k^{-1} u_k\}^T \\
& + \mathbb{E}\{\omega_k^T \Psi_k^{-1} \omega_k\} + \mathbb{E}\{u_k^T \Psi_k^{-1} u_k\} - (1 + \rho) x_k^T \Psi_k^{-1} x_k \\
\leq & \mathbb{E}\{(A_k x_k)^T \Psi_k^{-1} A_k x_k\} + \rho \mathbb{E}\{(A_k x_k)^T \Psi_k^{-1} A_k x_k\} \\
& + \rho^{-1}\{u_k^T \Psi_k^{-1} u_k\} + \mathbb{E}\{\omega_k^T \Psi_k^{-1} \omega_k\} + \mathbb{E}\{u_k^T \Psi_k^{-1} u_k\} - (1 + \rho) x_k^T \Psi_k^{-1} x_k \\
= & (1 + \rho) \mathbb{E}\{x_k^T (A_k^T \Psi_k^{-1} A_k - \Psi_k^{-1}) x_k\} + (1 + \rho^{-1}) \mathbb{E}\{u_k^T \Psi_k^{-1} u_k\} + \mathbb{E}\{\omega_k^T \Psi_k^{-1} \omega_k\}. \hfill (48)
\end{aligned}
$$

By employing the matrix inversion lemma, it follows that

$$A_k^T \Psi_k^{-1} A_k - \Psi_k^{-1}
\begin{aligned}
= & A_k^T (A_k \Psi_k A_k^T + Q_k + \kappa I)^{-1} A_k - \Psi_k^{-1} \\
= & -[\Psi_k + \Psi_k A_k^T (Q_k + \kappa I)^{-1} A_k \Psi_k]^{-1} \\
= & -[I + A_k^T (Q_k + \kappa I)^{-1} A_k \Psi_k]^{-1} \Psi_k^{-1} \\
\leq & -\left(1 + \frac{\bar{a}^2 \bar{\psi}}{\bar{q}}\right)^{-1} \Psi_k^{-1}. \hfill (49)
\end{aligned}
$$

Combining (48) and (49), it is derived that

$$\mathbb{E}\{V_{k+1}\} - (1 + \rho) V_k$$
\[
\begin{align*}
&\leq - (1 + \rho) \left( 1 + \frac{\bar{a}^2 \psi}{q} \right) \mathbb{E} \left\{ \bar{x}_k^T \Psi_k^{-1} \bar{x}_k \right\} + (1 + \rho^{-1}) \mathbb{E} \left\{ u_k^T \Psi_k^{-1} u_k \right\} + \mathbb{E} \left\{ \omega_k^T \Psi_k^{-1} \omega_k \right\} \\
&\leq - (1 + \rho) \left( 1 + \frac{\bar{a}^2 \psi}{q} \right) \mathbb{E}\{ V_k \} + (1 + \rho^{-1}) \frac{\bar{a}}{\kappa} + \frac{\bar{q}}{\kappa},
\end{align*}
\]

Define \( \chi \triangleq (1 + \rho) \left( 1 - \left( 1 + \frac{\bar{a}^2 \psi}{q} \right)^{-1} \right) \) and \( \varsigma \triangleq (1 + \rho^{-1}) \frac{\bar{a}}{\kappa} + \frac{\bar{q}}{\kappa} \). It is no doubt that there always exist positive scalars \( \bar{a}, \bar{\psi}, q \) and \( \rho \) such that \( 0 < \chi < 1 \) and \( 0 < \frac{\bar{\psi} \varsigma}{1 - \chi} < 1 \). Thus, (50) is converted into

\[
\mathbb{E}\{ V_{k+1} \} \leq \chi \mathbb{E}\{ V_k \} + \varsigma \leq \chi^{k+1} \mathbb{E}\{ V_0 \} + \varsigma \sum_{i=0}^{k} \chi^i.
\]

Keeping \( V_k \triangleq \bar{x}_k^T \Psi_k^{-1} \bar{x}_k \) and \( \kappa I \leq \Psi_k \) in mind, we have

\[
\mathbb{E}\{ V_{k+1} \} = \mathbb{E}\{ \bar{x}_{k+1}^T \Psi_{k+1}^{-1} \bar{x}_{k+1} \} \leq \mathbb{E}\{ \bar{x}_{k+1}^T k^{-1} I \bar{x}_{k+1} \},
\]

and this leads to

\[
\mathbb{E}\{ V_0 \} \leq \mathbb{E}\{ \bar{x}_0^T k^{-1} I \bar{x}_0 \}.
\]

The combination of \( \bar{x}_{k+1} = A_k \bar{x}_k + \omega_k + u_k \), (51) and (52) yields

\[
\mathbb{E}\{ \| \bar{x}_{k+1} \|^2 \} = \mathbb{E}\{ \bar{x}_{k+1}^T \Psi_{k+1}^{-1} \bar{x}_{k+1} \}
\leq \bar{\psi} \mathbb{E}\{ \bar{x}_{k+1}^T \Psi_{k+1}^{-1} \bar{x}_{k+1} \}
= \bar{\psi} \mathbb{E}\{ V_{k+1} \}
\leq \bar{\psi} \chi^{k+1} \mathbb{E}\{ V_0 \} + \bar{\psi} \varsigma \sum_{i=0}^{k} \chi^i
= \bar{\psi} \chi^{k+1} \mathbb{E}\{ \bar{x}_0^T k^{-1} I \bar{x}_0 \} + \bar{\psi} \varsigma \sum_{i=0}^{k} \chi^i
= \frac{\bar{\psi}}{\kappa} \mathbb{E}\{ \| \bar{x}_0 \|^2 \} \chi^{k+1} + \bar{\psi} \varsigma \sum_{i=0}^{\infty} \chi^i
\leq \frac{\bar{\psi}}{\kappa} \mathbb{E}\{ \| \bar{x}_0 \|^2 \} \chi^{k+1} + \bar{\psi} \varsigma \frac{\chi}{1 - \chi}.
\]

Paying attention to \( 0 < \chi < 1 \) and \( 0 < \frac{\bar{\psi} \varsigma}{1 - \chi} < 1 \), (53) and Definition 1, we confirm that \( \bar{x}_k \) is exponentially bounded in mean square sense. This completes the proof.

References


[27] L. Ma, Z. Wang, Y. Liu and F. E. Alsaadi, Distributed filtering for nonlinear time-delay systems over sensor networks subject to...


