Neural-Network-Based State Estimation for Nonlinear Stochastic Systems Under Token Bucket Communication Protocol

Dong Wang, Zidong Wang, and Chuanbo Wen

***Abstract*—This paper is concerned with the recursive neural-network (NN)-based state estimation problem for a class of stochastic discrete time-varying systems subjected to both unknown nonlinear dynamics and the token bucket communication protocol. The token bucket protocol is utilized to determine whether the sensor signal is granted access to the network at each transmission instant, wherein the transmission may fail due to an insufficient number of tokens in the bucket. The objective of the addressed problem is to design a recursive NN-based state estimator such that, under the influence of the unknown nonlinear dynamics and the token bucket communication protocol, certain upper bounds of both the state estimation error covariance and the neural-network-weight (NNW) error covariance are guaranteed, while the explicit expressions of the NN-based estimator gain and the NN tuning parameters are derived. By employing two sets of matrix difference equations, two upper bounds for the state estimation error covariance and the NNW error covariance are established, and these upper bounds are subsequently minimized by parameterizing the NN-based estimator gain in terms of the solutions to the matrix difference equations. Finally, an illustrative example is provided to demonstrate the feasibility and effectiveness of the proposed estimation approach.**

***Index Terms*—Recursive state estimation, neural-network-based state estimation, unknown nonlinear dynamics, token bucket communication protocol, time-varying systems.**

I. INTRODUCTION

From a physical perspective, it has been widely recognized that almost all systems in actual industrial production contain nonlinear factors, such as clearances and dry friction in mechanical systems, elastic-plastic materials and large deformations of components in structural systems, saturation characteristics of components in systems, and variable structure control strategies [12], [14], [15], [29], [37], [45], [46]. The presence of nonlinearity increases the system complexity and introduces essential difficulties to both the analysis and synthesis processes. Consequently, the dynamic problems of nonlinear systems have been a subject of long-standing interest, and extensive research has been conducted on this topic. Traditional state estimation methods have generally been developed by constructing physical models of nonlinear systems based on empirical knowledge or mechanistic principles [25], [34], [39], [51]. Representative methods in state estimation include extended Kalman filtering (EKF), H_∞ filtering, unscented Kalman filtering (UKF), and particle filtering (PF), along with further extensive application-oriented research founded upon these technologies [1], [3], [6], [26], [31], [38]. For instance, EKF has been established as a classic approach for extending the traditional linear Kalman filtering method to nonlinear systems. Essentially, the system and measurement equations are linearized, thereby allowing the problem of nonlinear systems to be approximately addressed using the classical linear Kalman filtering framework. It should be noted that the traditional state estimation methods mentioned above are primarily predicated upon the assumption that the system nonlinearity is known. However, in engineering practice, it is often challenging to accurately characterize the nonlinear components. Therefore, it is necessary that the dynamic modeling methods of such systems be fundamentally improved to effectively describe both the identifiable and the difficult-to-identify nonlinear components. Hence, the design of a novel state estimator capable of meeting the specified performance indicators for an unknown nonlinear dynamic system remains a highly significant and challenging research topic from an engineering perspective.

Time-varying characteristics are ubiquitous in a vast number of practical engineering systems. The filtering problem for time-varying systems has continuously attracted extensive research attention, and numerous research results have been reported in the literature [2], [9], [18], [22], [50], [52], [54]. Two primary filtering problems for time-varying systems have been identified. The first is the H_∞ filtering problem over a finite horizon, which can be addressed by employing the recursive linear matrix inequality method or the backward recursive Riccati difference equation technique. The second is the Kalman filtering problem in the sense of minimum variance, where the objective is to design a filter that minimizes the filtering error covariance at each time step. Although the H_∞ filtering approach can effectively manage situations involving unknown yet bounded-in-energy disturbance noises, the Kalman filtering problem continues to be widely studied due to its acceptable computational complexity and superior filtering performance in practical applications.

It should be noted that several suboptimal approaches, such

This work was supported in part by the Engineering and Physical Sciences Research Council (EPSRC) of the UK, the Royal Society of the UK, and the Alexander von Humboldt Foundation of Germany. (*Corresponding author: Zidong Wang*).

Dong Wang and Chuanbo Wen are with the School of Electrical Engineering, Shanghai Dianji University, Shanghai 201306, China. (Emails: `dongwang@sdju.edu.cn; wencb@sdju.edu.cn`)

Zidong Wang is with the College of Electrical Engineering and Automation, Shandong University of Science and Technology, Qingdao 266590, China; and is also with the Department of Computer Science, Brunel University London, Uxbridge, Middlesex, UB8 3PH, United Kingdom. (Email: `Zidong.Wang@brunel.ac.uk`)

as the variance-constrained recursive state estimation, have been broadly adopted to extend the classical Kalman filter to time-varying nonlinear systems, and considerable efforts have been dedicated to this subject. For example, in [40], the recursive filtering problem was investigated for a class of discrete-time stochastic dynamical networks where data delivery from the sensors to the filter was conducted via a digital communication channel. The desired filter parameters were recursively computed by solving two coupled Riccati difference equations to optimize the obtained upper bound of the filtering error variance. Nevertheless, the issue of state estimation in the sense of minimized error covariance for time-varying systems with *unknown nonlinear dynamics* has, to the best of the authors' knowledge, not yet been adequately addressed in the existing literature, making the present work a first effective attempt to address such challenges caused by the unknown nonlinear dynamics.

With the rapid development of neural networks (NNs) and deep learning technologies, artificial intelligence technologies have been extensively studied and applied in various fields. Neural networks possess strong nonlinear fitting capabilities, enabling the automatic extraction of useful features from large volumes of data without manual feature extraction, thereby facilitating the processing of data collected by complex industrial control systems. Recently, neural networks have been applied to the research of state estimation in the control field, as reported in [4], [21], [23], [27], [28], [30], [49]. For instance, in [13], Deep Variational Bayes Filters (DVBFs) have been proposed, which primarily focus on the integration of the classical Kalman filtering method under the assumption that the transitions comply with the linear Gaussian state space condition. In [48], the neural-network-based secure state estimation problem has been addressed for a class of networked nonlinear systems under energy-constrained denial-of-service cyber-attacks combined with an encoding-decoding scheme. The estimator gains and the neural-network tuning parameters have been derived by solving certain matrix inequalities to ensure the ultimate boundedness of both the state estimation error and the neural-network-weight (NNW) estimation error.

Up to now, the existing methods based on artificial intelligence technologies have been unable to effectively address both the dynamic and time-varying characteristics simultaneously, particularly in complex industrial or safety-critical systems with high uncertainty and harsh environmental variations. Consequently, it is necessary that nonlinear systems be re-examined from the perspective of integrating modern control theory with artificial intelligence technologies, and that advanced mathematical tools such as system modeling, neural network techniques, and online adaptive learning methods be combined to establish a set of comprehensive dynamic theories and practical methods suitable for handling unknown nonlinear systems. In the currently available literature, few results have been reported concerning the research of state estimation problems of unknown nonlinear dynamic time-varying systems based on the online learning of neural networks. Therefore, this research area remains one of the key international frontier topics and critical challenges in the study of nonlinear dynamic systems.

On the other hand, measurement signals are inevitably affected by the limitations of finite transmission resources in real-world network control systems. Up to now, various communication protocols, including the Round-Robin (RR) protocol, the try-once-discard protocol, and the token bucket communication protocol, have been introduced in industrial applications due to the inherently limited bandwidth. In the industrial field, these protocols are typically provided in the form of textual descriptions or flowcharts. To enable theoretical analysis of these network protocols, it is necessary that precise mathematical models be developed to describe them, as noted in [5], [11], [16], [17], [32], [33], [36], [41]–[43], [47], [53], [55], [56]. It should be emphasized that the token bucket communication protocol (TBCP) has been widely adopted in various network control systems to avoid data congestion and reduce communication costs [10], [20], [24], [35].

In the context of TBCP, the token bucket determines whether information is transmitted in the network by evaluating whether the traffic rate exceeds a specified threshold. It features a container for storing tokens with a predefined capacity, into which tokens are placed at a preset rate. When the bucket becomes full, any additional tokens overflow and are discarded. If an information transmission request is to be processed, a token must first be obtained from the bucket. Once the bucket reaches capacity, newly added tokens are either discarded or rejected. Accordingly, based on the textual description of the token bucket, a measurement model framework can be established through mathematical formulation to accurately characterize the TBCP. Thus far, filtering and state estimation problems under the TBCP have begun to attract preliminary research interest. Nevertheless, in terms of mean square error, the recursive state estimation problem for a class of stochastic discrete time-varying systems subject to the TBCP constraints has not yet been adequately investigated, let alone the additional consideration of unknown nonlinear dynamics.

The addressed recursive NN-based state estimation problem for stochastic discrete time-varying systems under the unknown nonlinear dynamics and the TBCP involves several essential difficulties. Firstly, the design of a state estimator capable of meeting the corresponding performance indicators for an unknown nonlinear dynamic system remains an extremely significant and challenging research topic from an engineering perspective. Also, the challenge exists in employing precise mathematical models to describe the protocols, which are typically presented in the form of textual descriptions or flowcharts. Therefore, it is the primary aim of this paper to provide satisfactory solutions to the listed challenges. Accordingly, *the novelties of this work can be briefly summarized as follows: 1) the recursive NN-based state estimation problem is, for the first time, investigated in the presence of both the unknown nonlinear dynamics and the TBCP; 2) a novel measurement output model is proposed to characterize the TBCP; 3) the neural network is comprehensively utilized to model the unknown nonlinear dynamics; and 4) an effective recursive NN-based state estimation algorithm is developed, which can be iteratively computed online*.

The remainder of this paper is organized as follows. In Section II, the recursive NN-based state estimator design problem is formulated for stochastic discrete time-varying systems with the unknown nonlinear dynamics under the TBCP. The main results are presented in Section III, where the desired state estimator is provided in terms of two sets of matrix difference equations. A numerical example is given in Section IV to demonstrate the effectiveness of the designed state estimator. Finally, Section V concludes the paper.

Notation The notation used in this paper is fairly standard except where otherwise stated. $\mathbb{R}^n$ and $\mathbb{R}^{n\times m}$ are used to denote, respectively, the n-dimensional Euclidean space and the set of all $n \times m$ real matrices. The notation $X \geq Y$ (respectively, $X > Y$), where X and Y are real symmetric matrices, indicates that $X - Y$ is positive semi-definite (respectively, positive definite). M^T denotes the transpose of the matrix M. I_n is used to represent the identity matrix of $n \times n$ dimension. diag$\cdots$ stands for a block-diagonal matrix. $\mathbb{N}^+$ denotes the set of positive integers. $\mathbb{E}\{x\}$ is used to represent the expectation of the stochastic variable x. $|x|$ describes the Euclidean norm of a vector x. In symmetric block matrices, "$*$" is used to denote a term induced by symmetry. Unless otherwise specified, matrices are assumed to have compatible dimensions.

II. Problem Formulation

A. The Model

Consider the following stochastic discrete-time systems with unknown nonlinear dynamics:

$$\begin{cases} x_{k+1} = A_k x_k + f(x_k) + B_k w_k \\ y_k = C_k x_k + D_k w_k \end{cases} \tag{1}$$

where $x_k \in \mathbb{R}^{n_x}$ denotes the state vector, and $y_k \in \mathbb{R}^{n_y}$ represents the measurement output prior to being transmitted through the communication network. The term $w_k \in \mathbb{R}^{n_w}$ denotes the stochastic external disturbance signal, which satisfies $\mathbb{E}\{w_k^T w_k\} \leq \bar{w}^2$. A_k, B_k, C_k and D_k are known real time-varying matrices with appropriate dimensions. The unknown nonlinear function $f : \mathbb{R}^{n_x} \to \mathbb{R}^{n_x}$ is assumed to satisfy $\|f(\cdot)\| \leq \bar{f}$, where $\bar{f}$ is a known constant.

B. The Token Bucket Mechanism

In this paper, a novel token bucket mechanism is introduced with the objective of achieving an improved trade-off between communication cost and estimation performance, which is recognized as a prevalent network model within information theory. Specifically, new tokens are added to the bucket at a constant rate of $a \in \{s \in \mathbb{N}^+ \mid s \geq 1\}$. A transmission across the network incurs a specific cost $b \in \{s \in \mathbb{N}^+ \mid s \geq a\}$. The bucket is further assumed to possess a maximum capacity $S \in \{s \in \mathbb{N}^+ \mid s \geq b\}$, referred to as the bucket size, where any arriving tokens are discarded if the bucket is already full. The variables μ_k and z_k denote the signal transmission judgment variable and the current bucket level at time k, respectively, and are governed by

$$\begin{cases} \mu_k = \min\{z_k + a, S\} - b, \\ z_{k+1} = \min\{z_k + a, S\} - \mathbf{1}_{\{\mu_k \geq 0\}} b, \\ 0 \leq z_0 \leq S, \end{cases} \tag{2}$$

Here, $\mathbf{1}_{\{\mu_k \geq 0\}}$ is an indicator function satisfying $\mathbf{1}_{\{\mu_k \geq 0\}} = 1$ if $\mu_k \geq 0$ and 0 otherwise. The measurement received by the remote estimator can be expressed by

$$\bar{y}_k = \mathbf{1}_{\{\mu_k \geq 0\}} y_k. \tag{3}$$

Remark 1: It should be highlighted that the token bucket mechanism presented in this paper distinguishes itself from conventional implementations by its integration into the recursive state estimation framework for stochastic discrete time-varying systems with unknown nonlinear dynamics. Unlike traditional usage scenarios where the token bucket mechanism is primarily employed for traffic shaping and congestion control, the proposed approach leverages this protocol to explicitly balance communication cost and estimation performance within the context of state estimation under constrained network resources. Furthermore, the mathematical formulation provided in (2) and (3) enables the seamless incorporation of the token bucket constraints into the estimator design, which has not been adequately addressed in existing literature. This integration facilitates the development of a comprehensive theoretical framework capable of addressing both network-induced constraints and complex system dynamics simultaneously.

C. The Neural Network Approximation

In accordance with the universal approximation property, the unknown nonlinear function $f(\cdot)$ can be approximated by a NN with an approximation error that is sufficiently small. Specifically, $f(\cdot)$ can be represented as $H\varphi(x_k) + \varepsilon_k$, where $H \subseteq \mathbb{R}^{n_x \times n_x}$ designates the ideal NNW matrix, $\varphi(\cdot)$ indicates the activation function, and $\varepsilon_k \in \mathbb{R}^{n_x}$ represents the approximation error of NNs. We assume that the NNW matrix H, the NN activation function $\varphi(\cdot)$, and the NN approximate error ε_k satisfy $\|H\| \leq \bar{H}$, $\|\varphi(\cdot)\| \leq \bar{\varphi}$ and $\|\varepsilon_k\| \leq \bar{\varepsilon}$, respectively, where $\bar{H}$, $\bar{\varphi}$, and $\bar{\varepsilon}$ are known positive constants.

The system (1) can be rewritten into the following form:

$$\begin{cases} x_{k+1} = A_k x_k + H\varphi(x_k) + \varepsilon_k + B_k w_k \\ y_k = C_k x_k + D_k w_k \\ \bar{y}_k = \mathbf{1}_{\{\mu_k \geq 0\}} y_k \end{cases} \tag{4}$$

The NN-based estimator in conjunction with the TBCP is proposed as follows:

$$\hat{x}_{k+1} = A_k \hat{x}_k + \hat{H}_k \varphi(\hat{x}_k) + \mathbf{1}_{\{\mu_k \geq 0\}} L_k (y_k - C_k \hat{x}_k) \tag{5}$$

where $\hat{x}_k \in \mathbb{R}^{n_x}$ is the state estimate, and L_k is the estimator parameters to be designed.

The following adaptive tuning law $\hat{H}_k$ for the NNW is adopted

$$\hat{H}_{k+1} = \hat{H}_k + \theta_1 (\mathbf{1}_{\{\mu_{k+1} \geq 0\}} C_{k+1}^T \tilde{y}_{k+1} \tilde{\varphi}^T(\hat{x}_k) - \theta_2 \hat{H}_k) \tag{6}$$

where

$$\begin{aligned} &\tilde{y}_{k+1} \triangleq y_{k+1} - C_{k+1}\hat{x}_{k+1}, \\ &\tilde{\varphi}(\hat{x}_k) \triangleq \varphi(\hat{x}_k) / \left(\left\| 1 + \varphi^T(\hat{x}_k)\varphi(\hat{x}_k) \right\| \left\| C_{k+1}^T C_{k+1} \right\| \right), \end{aligned}$$

and θ_1 and θ_2 are two positive tuning parameters.

Now, we define the state estimation error as $e_k \triangleq x_k - \hat{x}_k$, the state estimation error covariance as $P_k \triangleq \mathbb{E}\{e_k e_k^T\}$, the NNW error as $\tilde{H}_k \triangleq H - \hat{H}_k$, and the NNW error covariance as $Q_k \triangleq \mathbb{E}\{\tilde{H}_k \tilde{H}_k^T\}$.

In this paper, the main objective is to determine the NN-based estimator gain L_k and the NN tuning parameters θ_1 and θ_2 in the form of (5) such that, under the measurements subjected to the TBCP, the following requirements are achieved.

1) Upper bounds for both the state estimation error covariance and the NNW error covariance are ensured, that is, there exist two sequences of positive-definite matrices, $\bar{P}_k$ and $\bar{Q}_k$, such that $P_k \le \bar{P}_k$ and $Q_k \le \bar{Q}_k$, respectively.
2) The designed NN-based estimator gain L_k minimizes the upper bound $\bar{R}_k$ where $\bar{R}_k \triangleq \bar{P}_k + \bar{Q}_k$ at each time instant k.

III. MAIN RESULTS

In this section, the NN-based estimator gains in the form presented in (5) will be devised for the system described in (1) and the transmitted measurements given in (3) when they are subjected to unknown nonlinear dynamics and the TBCP. Initially, upper bounds for the state estimation error covariance and the NNW error covariance will be determined, and subsequently, through parameterizing the NN-based estimator gain, these upper bounds will be minimized at each time point.

The following lemmas will be utilized to facilitate the presentation of the main results of this paper.

Lemma 1: For $0 \le k \le N$, suppose that $X_i = X_i^T \ge 0$, $Y_i = Y_i^T \ge 0$ and $g_{i,k}(\cdot,\cdot) : \mathbb{R}^{n\times n} \times \mathbb{R}^{n\times n} \to \mathbb{R}^{n\times n}$ $(i = 1, 2)$. If, for $\forall X_1 \le Y_1$ and $X_2 \le Y_2$, the following conditions hold:

$$g_{1,k}(X_1, X_2) \le g_{1,k}(Y_1, Y_2), \tag{7}$$

$$g_{2,k}(X_1, X_2) \le g_{2,k}(Y_1, Y_2), \tag{8}$$

then the solutions $X_{1,k+1}$, $X_{2,k+1}$, $Y_{1,k+1}$ and $Y_{2,k+1}$ to the following difference equations

$$\begin{aligned} Y_{1,k+1} &= g_{1,k}(Y_{1,k}, Y_{2,k}), \\ Y_{2,k+1} &= g_{2,k}(Y_{1,k}, Y_{2,k}), \\ X_{1,k+1} &\le g_{1,k}(X_{1,k}, X_{2,k}), \\ X_{2,k+1} &\le g_{2,k}(X_{1,k}, X_{2,k}), \\ X_{1,0} &= Y_{1,0}, \\ X_{2,0} &= Y_{2,0} \end{aligned} \tag{9}$$

satisfy $X_{1,k+1} \le Y_{1,k+1}$ and $X_{2,k+1} \le Y_{2,k+1}$.

Proof: It follows from (9) that $X_{1,0} \le Y_{1,0}$ and $X_{2,0} \le Y_{2,0}$. Assuming, by induction, that $X_{1,k} \le Y_{1,k}$ and $X_{2,k} \le Y_{2,k}$, it can be obtained that

$$\begin{aligned} X_{1,k+1} = g_{1,k}(X_{1,k}, X_{2,k}) \le g_{1,k}(Y_{1,k}, Y_{2,k}) = Y_{1,k+1}, \\ X_{2,k+1} = g_{2,k}(X_{1,k}, X_{2,k}) \le g_{2,k}(Y_{1,k}, Y_{2,k}) = Y_{2,k+1}. \end{aligned} \tag{10}$$

The inductive hypothesis, therefore, implies that $X_{1,k+1} \le Y_{1,k+1}$ and $X_{2,k+1} \le Y_{2,k+1}$ always hold, which completes the proof. ■

Lemma 2: For matrices M, N, X and P with appropriate dimensions, the following relationships hold:

$$\begin{aligned} &\frac{\partial \mathrm{tr}(MXN)}{\partial X} = M^T N^T, \quad \frac{\partial \mathrm{tr}(MX^T N)}{\partial X} = NM, \\ &\frac{\partial \mathrm{tr}[(MXN)P(MXN)^T]}{\partial X} = 2M^T MXNPN^T. \end{aligned} \tag{11}$$

Definition 1: Let $A \subseteq \mathbb{R}^{m\times n}$ be a matrix where $a_{i,j}$ denotes the element located at the i-th row and the j-th column of A. The notation vec(A) is used to represent the vectorization of A, which is defined as the $mn \times 1$ column vector obtained by stacking the columns of the matrix A on top of one another:

$$\mathrm{vec}(A) \triangleq \left[a_{1,1}, \ldots, a_{m,1}, a_{1,2}, \ldots, a_{m,2}, \ldots, a_{1,n}, \ldots, a_{m,n}\right]. \tag{12}$$

Lemma 3: For matrices $A \subseteq \mathbb{R}^{k\times l}$, $X \subseteq \mathbb{R}^{l\times m}$, $Y \subseteq \mathbb{R}^{l\times m}$ and $B \subseteq \mathbb{R}^{m\times n}$, the vectorization operation satisfies the following two properties:

$$\begin{aligned} \mathrm{vec}(X+Y) =& \mathrm{vec}(X) + \mathrm{vec}(Y), \\ \mathrm{vec}(AXB) =& (B^T \otimes A)\mathrm{vec}(X). \end{aligned} \tag{13}$$

The following theorem demonstrates an effective algorithm by which upper bounds for the state estimation error covariance and the NNW error covariance can be computed at each time step.

Theorem 1: Let the positive scalars $\alpha_{i,k}$ $(i = 1, 2, 3, 4)$ and $\beta_{i,k}$ $(i = 1, 2, \ldots, 12)$, θ_i $(i = 1, 2)$ and the NN-based estimator gain L_k be given. Assume that there exist two sets of matrix sequences $\bar{P}_k$ and $\bar{Q}_k$ $(0 \le k \le M)$ satisfying the following recursions (with initial conditions $\bar{P}_0 = P_0$ and $\bar{Q}_0 = Q_0$):

$$\begin{aligned} \bar{P}_{k+1} \triangleq& (1 + \alpha_{1,k}^{1} + \alpha_{2,k}^{1})\bar{A}_k \bar{P}_k \bar{A}_k^T \\ &+ (1 + \alpha_{1,k}^{-1} + \alpha_{3,k}^{1} + \alpha_{4,k}^{1})\bar{\varphi}^2 \bar{Q}_k \\ &+ (1 + \alpha_{3,k}^{-1})\bar{w}^2 \bar{B}_k \bar{B}_k^T \\ &+ (1 + \alpha_{2,k}^{-1} + \alpha_{4,k}^{-1})(2\bar{\varphi}\bar{H} + \bar{\varepsilon})^2 I, \end{aligned} \tag{14}$$

$$\begin{aligned} \bar{Q}_{k+1} \triangleq& \delta_{1,k}\bar{Q}_k + \bar{\delta}_{2,k} C_{k+1}^T C_{k+1} \bar{A}_k \bar{P}_k \bar{A}_k^T C_{k+1}^T C_{k+1} \\ &+ \bar{\delta}_{3,k} C_{k+1}^T C_{k+1} \bar{Q}_k C_{k+1}^T C_{k+1} \\ &+ \bar{\delta}_{4,k} C_{k+1}^T C_{k+1} \bar{B}_k \bar{B}_k^T C_{k+1}^T C_{k+1} \\ &+ \bar{\delta}_{5,k} I + \bar{\delta}_{6,k} C_{k+1}^T D_{k+1} D_{k+1}^T C_{k+1} + \bar{\delta}_{7,k} I \end{aligned} \tag{15}$$

where

$$\begin{aligned} \delta_{1,k} \triangleq& (1 + \beta_{1,k}^{1} + \beta_{2,k}^{1} + \beta_{3,k}^{1} + \beta_{4,k}^{1} + \beta_{5,k}^{1})(1 - \theta_1\theta_2)^2, \\ \delta_{2,k} \triangleq& (1 + \beta_{1,k}^{-1} + \beta_{6,k}^{1} + \beta_{7,k}^{1} + \beta_{8,k}^{1})\theta_1^2 \mathbf{1}_{\{\mu_{k+1} \ge 0\}}, \\ \delta_{3,k} \triangleq& (1 + \beta_{2,k}^{-1} + \beta_{6,k}^{-1} + \beta_{9,k}^{1} + \beta_{10,k}^{1} + \beta_{11,k}^{1})\theta_1^2 \mathbf{1}_{\{\mu_{k+1} \ge 0\}}, \\ \delta_{4,k} \triangleq& (1 + \beta_{3,k}^{-1} + \beta_{9,k}^{-1})\theta_1^2 \mathbf{1}_{\{\mu_{k+1} \ge 0\}}, \\ \delta_{5,k} \triangleq& (1 + \beta_{4,k}^{-1} + \beta_{7,k}^{-1} + \beta_{10,k}^{-1} + \beta_{12,k}^{1})\theta_1^2 \mathbf{1}_{\{\mu_{k+1} \ge 0\}}, \\ \delta_{6,k} \triangleq& \theta_1^2 \mathbf{1}_{\{\mu_{k+1} \ge 0\}}, \\ \delta_{7,k} \triangleq& (1 + \beta_{5,k}^{-1} + \beta_{8,k}^{-1} + \beta_{11,k}^{-1} + \beta_{12,k}^{-1})\theta_1^2\theta_2^2, \end{aligned} \tag{16}$$

and

$$\begin{aligned} \bar{\delta}_{2,k} \triangleq& \delta_{2,k} / \left\| C_{k+1}^T C_{k+1} \right\|^2, \\ \bar{\delta}_{3,k} \triangleq& \delta_{3,k}\bar{\varphi}^2 / \left\| C_{k+1}^T C_{k+1} \right\|^2, \end{aligned}$$

$$\begin{aligned}
\bar{\delta}_{4,k} \triangleq& \delta_{4,k}\bar{w}^2/\left\|C_{k+1}^TC_{k+1}\right\|^2,\\
\bar{\delta}_{5,k} \triangleq& \delta_{5,k}(2\bar{\varphi}\bar{H}+\bar{\varepsilon})^2,\\
\bar{\delta}_{6,k} \triangleq& \delta_{6,k}\bar{w}^2/\left\|C_{k+1}^TC_{k+1}\right\|^2,\\
\bar{\delta}_{7,k} \triangleq& \delta_{7,k}\bar{H}^2.
\end{aligned} \tag{17}$$

Then, the matrices $\bar{P}_k$ and $\bar{Q}_k$ are upper bounds of the state estimation error covariance P_k and the NNW error covariance Q_k, which satisfy $P_k \leq \bar{P}_k$ and $Q_k \leq \bar{Q}_k$, respectively.

Proof: For the state estimation error e_{k+1}, it is easily obtained that

$$\begin{aligned}
e_{k+1} \triangleq& x_{k+1} - \hat{x}_{k+1}\\
=& A_kx_k + H\varphi(x_k) + \varepsilon_k + B_kw_k\\
&- (A_k\hat{x}_k + \hat{H}_k\varphi(\hat{x}_k) + \mathbf{1}_{\{\mu_k\geq 0\}}L_k(y_k - C_k\hat{x}_k))\\
=& \bar{A}_ke_k + \tilde{H}_k\varphi(\hat{x}_k) + \bar{B}_kw_k + \bar{\varepsilon}_k
\end{aligned} \tag{18}$$

where

$$\begin{aligned}
\bar{A}_k &\triangleq A_k - \mathbf{1}_{\{\mu_k\geq 0\}}L_kC_k,\\
\bar{B}_k &\triangleq B_k - \mathbf{1}_{\{\mu_k\geq 0\}}L_kD_k,\\
\bar{\varepsilon}_k &\triangleq H(\varphi(x_k) - \varphi(\hat{x}_k)) + \varepsilon_k.
\end{aligned}$$

On the other hand, the NNW error $\tilde{H}_{k+1}$ can be expressed as follows:

$$\begin{aligned}
\tilde{H}_{k+1} \triangleq& H - \hat{H}_{k+1}\\
=& H - \hat{H}_k - \theta_1(\mathbf{1}_{\{\mu_{k+1}\geq 0\}}C_{k+1}^T(y_{k+1} - C_{k+1}\hat{x}_{k+1})\\
&\times\tilde{\varphi}^T(\hat{x}_k) - \theta_2\hat{H}_k)\\
=& H - \hat{H}_k - \theta_1(\mathbf{1}_{\{\mu_{k+1}\geq 0\}}C_{k+1}^T(C_{k+1}x_{k+1}\\
&+ D_{k+1}w_{k+1} - C_{k+1}\hat{x}_{k+1})\tilde{\varphi}^T(\hat{x}_k) - \theta_2\hat{H}_k)\\
=& H - \hat{H}_k - \theta_1(\mathbf{1}_{\{\mu_{k+1}\geq 0\}}C_{k+1}^T(C_{k+1}e_{k+1}\\
&\times + D_{k+1}w_{k+1})\tilde{\varphi}^T(\hat{x}_k) - \theta_2\hat{H}_k)\\
=& (1-\theta_1\theta_2)\tilde{H}_k - \theta_1\mathbf{1}_{\{\mu_{k+1}\geq 0\}}C_{k+1}^TC_{k+1}e_{k+1}\tilde{\varphi}^T(\hat{x}_k)\\
&- \theta_1\mathbf{1}_{\{\mu_{k+1}\geq 0\}}C_{k+1}^TD_{k+1}w_{k+1}\tilde{\varphi}^T(\hat{x}_k) + \theta_1\theta_2H\\
=& (1-\theta_1\theta_2)\tilde{H}_k - \theta_1\mathbf{1}_{\{\mu_{k+1}\geq 0\}}C_{k+1}^TC_{k+1}\\
&\times(\bar{A}_ke_k + \tilde{H}_k\varphi(\hat{x}_k) + \bar{B}_kw_k + \bar{\varepsilon}_k)\tilde{\varphi}^T(\hat{x}_k)\\
&- \theta_1\mathbf{1}_{\{\mu_{k+1}\geq 0\}}C_{k+1}^TD_{k+1}w_{k+1}\tilde{\varphi}^T(\hat{x}_k) + \theta_1\theta_2H\\
=& (1-\theta_1\theta_2)\tilde{H}_k - \theta_1\mathbf{1}_{\{\mu_{k+1}\geq 0\}}C_{k+1}^TC_{k+1}\bar{A}_ke_k\tilde{\varphi}^T(\hat{x}_k)\\
&- \theta_1\mathbf{1}_{\{\mu_{k+1}\geq 0\}}C_{k+1}^TC_{k+1}\tilde{H}_k\varphi(\hat{x}_k)\tilde{\varphi}^T(\hat{x}_k)\\
&- \theta_1\mathbf{1}_{\{\mu_{k+1}\geq 0\}}C_{k+1}^TC_{k+1}\bar{B}_kw_k\tilde{\varphi}^T(\hat{x}_k)\\
&- \theta_1\mathbf{1}_{\{\mu_{k+1}\geq 0\}}C_{k+1}^TC_{k+1}\bar{\varepsilon}_k\tilde{\varphi}^T(\hat{x}_k)\\
&- \theta_1\mathbf{1}_{\{\mu_{k+1}\geq 0\}}C_{k+1}^TD_{k+1}w_{k+1}\tilde{\varphi}^T(\hat{x}_k)\\
&+ \theta_1\theta_2H.
\end{aligned} \tag{19}$$

To continue, the following notations are introduced for simplicity:

$$\begin{aligned}
\Omega_{1,k} \triangleq& \bar{A}_k\mathbb{E}\{e_k\varphi(\hat{x}_k)^T\tilde{H}_k^T\},\\
\Omega_{2,k} \triangleq& \bar{A}_k\mathbb{E}\{e_kw_k^T\}\bar{B}_k^T,\\
\Omega_{3,k} \triangleq& \bar{A}_k\mathbb{E}\{e_k\bar{\varepsilon}_k^T\},\\
\Omega_{4,k} \triangleq& \mathbb{E}\{\tilde{H}_k\varphi(\hat{x}_k)w_k^T\}\bar{B}_k^T,\\
\Omega_{5,k} \triangleq& \mathbb{E}\{\tilde{H}_k\varphi(\hat{x}_k)\bar{\varepsilon}_k^T\},\\
\Omega_{6,k} \triangleq& \bar{B}_k\mathbb{E}\{w_k\bar{\varepsilon}_k^T\},
\end{aligned} \tag{20}$$

and

$$\begin{aligned}
\Pi_{1,k} \triangleq& -(1-\theta_1\theta_2)\theta_1\mathbf{1}_{\{\mu_{k+1}\geq 0\}}\\
&\times\mathbb{E}\{\tilde{H}_k\tilde{\varphi}(\hat{x}_k)e_k^T\}\bar{A}_k^TC_{k+1}^TC_{k+1},\\
\Pi_{2,k} \triangleq& -(1-\theta_1\theta_2)\theta_1\mathbf{1}_{\{\mu_{k+1}\geq 0\}}\\
&\times\mathbb{E}\{\tilde{H}_k\tilde{\varphi}(\hat{x}_k)\varphi^T(\hat{x}_k)\tilde{H}_k^T\}C_{k+1}^TC_{k+1},\\
\Pi_{3,k} \triangleq& -(1-\theta_1\theta_2)\theta_1\mathbf{1}_{\{\mu_{k+1}\geq 0\}}\\
&\times\mathbb{E}\{\tilde{H}_k\tilde{\varphi}(\hat{x}_k)w_k^T\}\bar{B}_k^TC_{k+1}^TC_{k+1},\\
\Pi_{4,k} \triangleq& -(1-\theta_1\theta_2)\theta_1\mathbf{1}_{\{\mu_{k+1}\geq 0\}}\\
&\times\mathbb{E}\{\tilde{H}_k\tilde{\varphi}(\hat{x}_k)\bar{\varepsilon}_k^T\}C_{k+1}^TC_{k+1},\\
\Pi_{5,k} \triangleq& -(1-\theta_1\theta_2)\theta_1\mathbf{1}_{\{\mu_{k+1}\geq 0\}}\\
&\times\mathbb{E}\{\tilde{H}_k\tilde{\varphi}(\hat{x}_k)w_{k+1}^T\}D_{k+1}^TC_{k+1},\\
\Pi_{6,k} \triangleq& (1-\theta_1\theta_2)\theta_1\theta_2\mathbb{E}\{\tilde{H}_kH^T\},\\
\Pi_{7,k} \triangleq& \theta_1^2\mathbf{1}_{\{\mu_{k+1}\geq 0\}}C_{k+1}^TC_{k+1}\bar{A}_k\\
&\times\mathbb{E}\{e_k\tilde{\varphi}^T(\hat{x}_k)\tilde{\varphi}(\hat{x}_k)\varphi^T(\hat{x}_k)\tilde{H}_k^T\}C_{k+1}^TC_{k+1},\\
\Pi_{8,k} \triangleq& \theta_1^2\mathbf{1}_{\{\mu_{k+1}\geq 0\}}C_{k+1}^TC_{k+1}\bar{A}_k\\
&\times\mathbb{E}\{e_k\tilde{\varphi}^T(\hat{x}_k)\tilde{\varphi}(\hat{x}_k)w_k^T\}\bar{B}_k^TC_{k+1}^TC_{k+1},\\
\Pi_{9,k} \triangleq& \theta_1^2\mathbf{1}_{\{\mu_{k+1}\geq 0\}}C_{k+1}^TC_{k+1}\bar{A}_k\\
&\times\mathbb{E}\{e_k\tilde{\varphi}^T(\hat{x}_k)\tilde{\varphi}(\hat{x}_k)\bar{\varepsilon}_k^T\}C_{k+1}^TC_{k+1},\\
\Pi_{10,k} \triangleq& \theta_1^2\mathbf{1}_{\{\mu_{k+1}\geq 0\}}C_{k+1}^TC_{k+1}\bar{A}_k\\
&\times\mathbb{E}\{e_k\tilde{\varphi}^T(\hat{x}_k)\tilde{\varphi}(\hat{x}_k)w_{k+1}^T\}D_{k+1}^TC_{k+1},\\
\Pi_{11,k} \triangleq& -\theta_1^2\theta_2\mathbf{1}_{\{\mu_{k+1}\geq 0\}}C_{k+1}^TC_{k+1}\bar{A}_k\mathbb{E}\{e_k\tilde{\varphi}^T(\hat{x}_k)H^T\},\\
\Pi_{12,k} \triangleq& \theta_1^2\mathbf{1}_{\{\mu_{k+1}\geq 0\}}C_{k+1}^TC_{k+1}\\
&\times\mathbb{E}\{\tilde{H}_k\varphi(\hat{x}_k)\tilde{\varphi}^T(\hat{x}_k)\tilde{\varphi}(\hat{x}_k)w_k^T\}\bar{B}_k^TC_{k+1}^TC_{k+1},\\
\Pi_{13,k} \triangleq& \theta_1^2\mathbf{1}_{\{\mu_{k+1}\geq 0\}}C_{k+1}^TC_{k+1}\\
&\times\mathbb{E}\{\tilde{H}_k\varphi(\hat{x}_k)\tilde{\varphi}^T(\hat{x}_k)\tilde{\varphi}(\hat{x}_k)\bar{\varepsilon}_k^T\}C_{k+1}^TC_{k+1},\\
\Pi_{14,k} \triangleq& \theta_1^2\mathbf{1}_{\{\mu_{k+1}\geq 0\}}C_{k+1}^TC_{k+1}\\
&\times\mathbb{E}\{\tilde{H}_k\varphi(\hat{x}_k)\tilde{\varphi}^T(\hat{x}_k)\tilde{\varphi}(\hat{x}_k)w_{k+1}^T\}D_{k+1}^TC_{k+1},\\
\Pi_{15,k} \triangleq& -\theta_1^2\theta_2\mathbf{1}_{\{\mu_{k+1}\geq 0\}}C_{k+1}^TC_{k+1}\\
&\times\mathbb{E}\{\tilde{H}_k\varphi(\hat{x}_k)\tilde{\varphi}^T(\hat{x}_k)H^T\},\\
\Pi_{16,k} \triangleq& \theta_1^2\mathbf{1}_{\{\mu_{k+1}\geq 0\}}C_{k+1}^TC_{k+1}\bar{B}_k\\
&\times\mathbb{E}\{w_k\tilde{\varphi}^T(\hat{x}_k)\tilde{\varphi}(\hat{x}_k)\bar{\varepsilon}_k^T\}C_{k+1}^TC_{k+1},\\
\Pi_{17,k} \triangleq& \theta_1^2\mathbf{1}_{\{\mu_{k+1}\geq 0\}}C_{k+1}^TC_{k+1}\bar{B}_k\\
&\times\mathbb{E}\{w_k\tilde{\varphi}^T(\hat{x}_k)\tilde{\varphi}(\hat{x}_k)w_{k+1}^T\}D_{k+1}^TC_{k+1},\\
\Pi_{18,k} \triangleq& -\theta_1^2\theta_2\mathbf{1}_{\{\mu_{k+1}\geq 0\}}C_{k+1}^TC_{k+1}\bar{B}_k\mathbb{E}\{w_k\tilde{\varphi}^T(\hat{x}_k)H^T\},\\
\Pi_{19,k} \triangleq& \theta_1^2\mathbf{1}_{\{\mu_{k+1}\geq 0\}}C_{k+1}^TC_{k+1}\\
&\times\mathbb{E}\{\bar{\varepsilon}_k\tilde{\varphi}^T(\hat{x}_k)\tilde{\varphi}(\hat{x}_k)w_{k+1}^T\}D_{k+1}^TC_{k+1},\\
\Pi_{20,k} \triangleq& -\theta_1^2\theta_2\mathbf{1}_{\{\mu_{k+1}\geq 0\}}C_{k+1}^TC_{k+1}\mathbb{E}\{\bar{\varepsilon}_k\tilde{\varphi}^T(\hat{x}_k)H^T\},\\
\Pi_{21,k} \triangleq& -\theta_1^2\theta_2\mathbf{1}_{\{\mu_{k+1}\geq 0\}}C_{k+1}^TD_{k+1}\\
&\times\mathbb{E}\{w_{k+1}\tilde{\varphi}^T(\hat{x}_k)H^T\}.
\end{aligned} \tag{21}$$

Now, the state estimation error covariance P_{k+1} is calculated in conjunction with (20) as follows:

$$\begin{aligned}
P_{k+1} \triangleq& \mathbb{E}\{e_{k+1}e_{k+1}^T\}\\
=& \mathbb{E}\{(\bar{A}_ke_k + \tilde{H}_k\varphi(\hat{x}_k) + \bar{B}_kw_k + \bar{\varepsilon}_k)
\end{aligned}$$

$$
\begin{aligned}
&\times (\bar{A}_k e_k + \tilde{H}_k \varphi(\hat{x}_k) + \bar{B}_k w_k + \bar{\varepsilon}_k)^T\} \\
=&\bar{A}_k \mathbb{E}\{e_k e_k^T\} \bar{A}_k^T + \mathbb{E}\{\tilde{H}_k \varphi(\hat{x}_k) \varphi(\hat{x}_k)^T \tilde{H}_k^T\} \\
&+ \bar{B}_k \mathbb{E}\{w_k w_k^T\} \bar{B}_k^T + \bar{\varepsilon}_k \bar{\varepsilon}_k^T \\
&+ \Omega_{1,k} + \Omega_{1,k}^T + \Omega_{2,k} + \Omega_{2,k}^T + \Omega_{3,k} + \Omega_{3,k}^T \\
&+ \Omega_{4,k} + \Omega_{4,k}^T + \Omega_{5,k} + \Omega_{5,k}^T + \Omega_{6,k} + \Omega_{6,k}^T. \qquad (22)
\end{aligned}
$$

By employing the elementary inequality $(\alpha^{1/2}M - \alpha^{-1/2}N)(\alpha^{1/2}M - \alpha^{-1/2}N)^T \geq 0$ where M and N are arbitrary matrices with compatible dimensions, it can be immediately obtained that

$$
\begin{aligned}
\Omega_{1,k} + \Omega_{1,k}^T \leq& \alpha_{1,k}^1 \bar{A}_k \mathbb{E}\{e_k e_k^T\} \bar{A}_k^T \\
&+ \alpha_{1,k}^{-1} \mathbb{E}\{\tilde{H}_k \varphi(\hat{x}_k) \varphi(\hat{x}_k)^T \tilde{H}_k^T\} \\
\Omega_{3,k} + \Omega_{3,k}^T \leq& \alpha_{2,k}^1 \bar{A}_k \mathbb{E}\{e_k e_k^T\} \bar{A}_k^T + \alpha_{2,k}^{-1} \bar{\varepsilon}_k \bar{\varepsilon}_k^T \\
\Omega_{4,k} + \Omega_{4,k}^T \leq& \alpha_{3,k}^1 \mathbb{E}\{\tilde{H}_k \varphi(\hat{x}_k) \varphi(\hat{x}_k)^T \tilde{H}_k^T\} \\
&+ \alpha_{3,k}^{-1} \bar{B}_k \mathbb{E}\{w_k w_k^T\} \bar{B}_k^T \\
\Omega_{5,k} + \Omega_{5,k}^T \leq& \alpha_{4,k}^1 \mathbb{E}\{\tilde{H}_k \varphi(\hat{x}_k) \varphi(\hat{x}_k)^T \tilde{H}_k^T\} \\
&+ \alpha_{4,k}^{-1} \bar{\varepsilon}_k \bar{\varepsilon}_k^T. \qquad (23)
\end{aligned}
$$

Noting that $\Omega_{2,k} = 0$ and $\Omega_{6,k} = 0$, and by referring to (23), the state estimation error covariance P_{k+1} in (22) can be further manipulated as follows:

$$
\begin{aligned}
P_{k+1} \leq& (1 + \alpha_{1,k}^1 + \alpha_{2,k}^1) \bar{A}_k \mathbb{E}\{e_k e_k^T\} \bar{A}_k^T \\
&+ (1 + \alpha_{1,k}^{-1} + \alpha_{3,k}^1 + \alpha_{4,k}^1) \mathbb{E}\{\tilde{H}_k \varphi(\hat{x}_k) \varphi(\hat{x}_k)^T \tilde{H}_k^T\} \\
&+ (1 + \alpha_{3,k}^{-1}) \bar{B}_k \mathbb{E}\{w_k w_k^T\} \bar{B}_k^T \\
&+ (1 + \alpha_{2,k}^{-1} + \alpha_{4,k}^{-1}) \bar{\varepsilon}_k \bar{\varepsilon}_k^T \\
\leq& (1 + \alpha_{1,k}^1 + \alpha_{2,k}^1) \bar{A}_k P_k \bar{A}_k^T \\
&+ (1 + \alpha_{1,k}^{-1} + \alpha_{3,k}^1 + \alpha_{4,k}^1) \bar{\varphi}^2 Q_k \\
&+ (1 + \alpha_{3,k}^{-1}) \bar{w}^2 \bar{B}_k \bar{B}_k^T \\
&+ (1 + \alpha_{2,k}^{-1} + \alpha_{4,k}^{-1}) (2\bar{\varphi}\bar{H} + \bar{\varepsilon})^2 I. \qquad (24)
\end{aligned}
$$

On the other hand, according to (21), the NNW error covariance Q_{k+1} can be readily concluded as follows:

$$
\begin{aligned}
&Q_{k+1} \\
\triangleq& \mathbb{E}\{\tilde{H}_{k+1} \tilde{H}_{k+1}^T\} \\
=& \mathbb{E}\{((1 - \theta_1\theta_2)\tilde{H}_k - \theta_1 \mathbf{1}_{\{\mu_{k+1} \geq 0\}} C_{k+1}^T C_{k+1} \bar{A}_k e_k \tilde{\varphi}^T(\hat{x}_k) \\
&- \theta_1 \mathbf{1}_{\{\mu_{k+1} \geq 0\}} C_{k+1}^T C_{k+1} \tilde{H}_k \varphi(\hat{x}_k) \tilde{\varphi}^T(\hat{x}_k) \\
&- \theta_1 \mathbf{1}_{\{\mu_{k+1} \geq 0\}} C_{k+1}^T C_{k+1} \bar{B}_k w_k \tilde{\varphi}^T(\hat{x}_k) \\
&- \theta_1 \mathbf{1}_{\{\mu_{k+1} \geq 0\}} C_{k+1}^T C_{k+1} \bar{\varepsilon}_k \tilde{\varphi}^T(\hat{x}_k) \\
&- \theta_1 \mathbf{1}_{\{\mu_{k+1} \geq 0\}} C_{k+1}^T D_{k+1} w_{k+1} \tilde{\varphi}^T(\hat{x}_k) + \theta_1\theta_2 H) \\
&\times ((1 - \theta_1\theta_2)\tilde{H}_k - \theta_1 \mathbf{1}_{\{\mu_{k+1} \geq 0\}} C_{k+1}^T C_{k+1} \bar{A}_k e_k \tilde{\varphi}^T(\hat{x}_k) \\
&- \theta_1 \mathbf{1}_{\{\mu_{k+1} \geq 0\}} C_{k+1}^T C_{k+1} \tilde{H}_k \varphi(\hat{x}_k) \tilde{\varphi}^T(\hat{x}_k) \\
&- \theta_1 \mathbf{1}_{\{\mu_{k+1} \geq 0\}} C_{k+1}^T C_{k+1} \bar{B}_k w_k \tilde{\varphi}^T(\hat{x}_k) \\
&- \theta_1 \mathbf{1}_{\{\mu_{k+1} \geq 0\}} C_{k+1}^T C_{k+1} \bar{\varepsilon}_k \tilde{\varphi}^T(\hat{x}_k) \\
&- \theta_1 \mathbf{1}_{\{\mu_{k+1} \geq 0\}} C_{k+1}^T D_{k+1} w_{k+1} \tilde{\varphi}^T(\hat{x}_k) + \theta_1\theta_2 H)^T\} \\
=& (1 - \theta_1\theta_2)^2 \mathbb{E}\{\tilde{H}_k \tilde{H}_k^T\} + \theta_1^2 \mathbf{1}_{\{\mu_{k+1} \geq 0\}} C_{k+1}^T C_{k+1} \\
&\times \bar{A}_k \mathbb{E}\{e_k \tilde{\varphi}^T(\hat{x}_k) \tilde{\varphi}(\hat{x}_k) e_k^T\} \bar{A}_k^T C_{k+1}^T C_{k+1} \\
&+ \theta_1^2 \mathbf{1}_{\{\mu_{k+1} \geq 0\}} C_{k+1}^T C_{k+1} \\
&\times \mathbb{E}\{\tilde{H}_k \varphi(\hat{x}_k) \tilde{\varphi}^T(\hat{x}_k) \tilde{\varphi}(\hat{x}_k) \varphi^T(\hat{x}_k) \tilde{H}_k^T\} C_{k+1}^T C_{k+1} \\
&+ \theta_1^2 \mathbf{1}_{\{\mu_{k+1} \geq 0\}} C_{k+1}^T C_{k+1} \bar{B}_k \mathbb{E}\{w_k \tilde{\varphi}^T(\hat{x}_k) \tilde{\varphi}(\hat{x}_k) w_k^T\} \\
&\times \bar{B}_k^T C_{k+1}^T C_{k+1} + \theta_1^2 \mathbf{1}_{\{\mu_{k+1} \geq 0\}} C_{k+1}^T C_{k+1} \\
&\times \mathbb{E}\{\bar{\varepsilon}_k \tilde{\varphi}^T(\hat{x}_k) \tilde{\varphi}(\hat{x}_k) \bar{\varepsilon}_k^T\} C_{k+1}^T C_{k+1} \\
&+ \theta_1^2 \mathbf{1}_{\{\mu_{k+1} \geq 0\}} C_{k+1}^T D_{k+1} \mathbb{E}\{w_{k+1} \tilde{\varphi}^T(\hat{x}_k) \tilde{\varphi}(\hat{x}_k) w_{k+1}^T\} \\
&\times D_{k+1}^T C_{k+1} + \theta_1^2 \theta_2^2 H H^T \\
&+ \Pi_{1,k} + \Pi_{1,k}^T + \Pi_{2,k} + \Pi_{2,k}^T + \Pi_{3,k} + \Pi_{3,k}^T + \Pi_{4,k} \\
&+ \Pi_{4,k}^T + \Pi_{5,k} + \Pi_{5,k}^T + \Pi_{6,k} + \Pi_{6,k}^T + \Pi_{7,k} + \Pi_{7,k}^T \\
&+ \Pi_{8,k} + \Pi_{8,k}^T + \Pi_{9,k} + \Pi_{9,k}^T + \Pi_{10,k} + \Pi_{10,k}^T \\
&+ \Pi_{11,k} + \Pi_{11,k}^T + \Pi_{12,k} + \Pi_{12,k}^T + \Pi_{13,k} + \Pi_{13,k}^T \\
&+ \Pi_{14,k} + \Pi_{14,k}^T + \Pi_{15,k} + \Pi_{15,k}^T + \Pi_{16,k} + \Pi_{16,k}^T \\
&+ \Pi_{17,k} + \Pi_{17,k}^T + \Pi_{18,k} + \Pi_{18,k}^T + \Pi_{19,k} + \Pi_{19,k}^T \\
&+ \Pi_{20,k} + \Pi_{20,k}^T + \Pi_{21,k} + \Pi_{21,k}^T. \qquad (25)
\end{aligned}
$$

By employing a similar approach as in (23), one has

$$
\begin{aligned}
&\Pi_{1,k} + \Pi_{1,k}^T \\
\leq& \beta_{1,k}^1 (1 - \theta_1\theta_2)^2 \mathbb{E}\{\tilde{H}_k \tilde{H}_k^T\} + \beta_{1,k}^{-1} \theta_1^2 \mathbf{1}_{\{\mu_{k+1} \geq 0\}} C_{k+1}^T C_{k+1} \\
&\times \bar{A}_k \mathbb{E}\{e_k \tilde{\varphi}^T(\hat{x}_k) \tilde{\varphi}(\hat{x}_k) e_k^T\} \bar{A}_k^T C_{k+1}^T C_{k+1}, \\
&\Pi_{2,k} + \Pi_{2,k}^T \\
\leq& \beta_{2,k}^1 (1 - \theta_1\theta_2)^2 \mathbb{E}\{\tilde{H}_k \tilde{H}_k^T\} + \beta_{2,k}^{-1} \theta_1^2 \mathbf{1}_{\{\mu_{k+1} \geq 0\}} C_{k+1}^T C_{k+1} \\
&\times \mathbb{E}\{\tilde{H}_k \varphi(\hat{x}_k) \tilde{\varphi}^T(\hat{x}_k) \tilde{\varphi}(\hat{x}_k) \varphi^T(\hat{x}_k) \tilde{H}_k^T\} C_{k+1}^T C_{k+1}, \\
&\Pi_{3,k} + \Pi_{3,k}^T \\
\leq& \beta_{3,k}^1 (1 - \theta_1\theta_2)^2 \mathbb{E}\{\tilde{H}_k \tilde{H}_k^T\} + \beta_{3,k}^{-1} \theta_1^2 \mathbf{1}_{\{\mu_{k+1} \geq 0\}} C_{k+1}^T C_{k+1} \\
&\times \bar{B}_k \mathbb{E}\{w_k \tilde{\varphi}^T(\hat{x}_k) \tilde{\varphi}(\hat{x}_k) w_k^T\} \bar{B}_k^T C_{k+1}^T C_{k+1}, \\
&\Pi_{4,k} + \Pi_{4,k}^T \\
\leq& \beta_{4,k}^1 (1 - \theta_1\theta_2)^2 \mathbb{E}\{\tilde{H}_k \tilde{H}_k^T\} + \beta_{4,k}^{-1} \theta_1^2 \mathbf{1}_{\{\mu_{k+1} \geq 0\}} C_{k+1}^T C_{k+1} \\
&\times \mathbb{E}\{\bar{\varepsilon}_k \tilde{\varphi}^T(\hat{x}_k) \tilde{\varphi}(\hat{x}_k) \bar{\varepsilon}_k^T\} C_{k+1}^T C_{k+1}, \\
&\Pi_{6,k} + \Pi_{6,k}^T \\
\leq& \beta_{5,k}^1 (1 - \theta_1\theta_2)^2 \mathbb{E}\{\tilde{H}_k \tilde{H}_k^T\} + \beta_{5,k}^{-1} \theta_1^2 \theta_2^2 H H^T, \\
&\Pi_{7,k} + \Pi_{7,k}^T \\
\leq& \beta_{6,k}^1 \theta_1^2 \mathbf{1}_{\{\mu_{k+1} \geq 0\}} C_{k+1}^T C_{k+1} \bar{A}_k \mathbb{E}\{e_k \tilde{\varphi}^T(\hat{x}_k) \tilde{\varphi}(\hat{x}_k) e_k^T\} \\
&\times \bar{A}_k^T C_{k+1}^T C_{k+1} + \beta_{6,k}^{-1} \theta_1^2 \mathbf{1}_{\{\mu_{k+1} \geq 0\}} C_{k+1}^T C_{k+1} \\
&\times \mathbb{E}\{\tilde{H}_k \varphi(\hat{x}_k) \tilde{\varphi}^T(\hat{x}_k) \tilde{\varphi}(\hat{x}_k) \varphi^T(\hat{x}_k) \tilde{H}_k^T\} C_{k+1}^T C_{k+1}, \\
&\Pi_{9,k} + \Pi_{9,k}^T \\
\leq& \beta_{7,k}^1 \theta_1^2 \mathbf{1}_{\{\mu_{k+1} \geq 0\}} C_{k+1}^T C_{k+1} \bar{A}_k \mathbb{E}\{e_k \tilde{\varphi}^T(\hat{x}_k) \tilde{\varphi}(\hat{x}_k) e_k^T\} \\
&\times \bar{A}_k^T C_{k+1}^T C_{k+1} + \beta_{7,k}^{-1} \theta_1^2 \mathbf{1}_{\{\mu_{k+1} \geq 0\}} C_{k+1}^T C_{k+1} \\
&\times \mathbb{E}\{\bar{\varepsilon}_k \tilde{\varphi}^T(\hat{x}_k) \tilde{\varphi}(\hat{x}_k) \bar{\varepsilon}_k^T\} C_{k+1}^T C_{k+1}, \\
&\Pi_{11,k} + \Pi_{11,k}^T \\
\leq& \beta_{8,k}^1 \theta_1^2 \mathbf{1}_{\{\mu_{k+1} \geq 0\}} C_{k+1}^T C_{k+1} \bar{A}_k \mathbb{E}\{e_k \tilde{\varphi}^T(\hat{x}_k) \tilde{\varphi}(\hat{x}_k) e_k^T\} \\
&\times \bar{A}_k^T C_{k+1}^T C_{k+1} + \beta_{8,k}^{-1} \theta_1^2 \theta_2^2 H H^T, \\
&\Pi_{12,k} + \Pi_{12,k}^T \\
\leq& \beta_{9,k}^1 \theta_1^2 \mathbf{1}_{\{\mu_{k+1} \geq 0\}} C_{k+1}^T C_{k+1} \\
&\times \mathbb{E}\{\tilde{H}_k \varphi(\hat{x}_k) \tilde{\varphi}^T(\hat{x}_k) \tilde{\varphi}(\hat{x}_k) \varphi^T(\hat{x}_k) \tilde{H}_k^T\} C_{k+1}^T C_{k+1} \\
&+ \beta_{9,k}^{-1} \theta_1^2 \mathbf{1}_{\{\mu_{k+1} \geq 0\}} C_{k+1}^T C_{k+1} \bar{B}_k \\
&\times \mathbb{E}\{w_k \tilde{\varphi}^T(\hat{x}_k) \tilde{\varphi}(\hat{x}_k) w_k^T\} \bar{B}_k^T C_{k+1}^T C_{k+1},
\end{aligned}
$$

$$
\begin{aligned}
&\Pi_{13,k}+\Pi_{13,k}^T\\
\leq&\beta_{10,k}^1\theta_1^2\mathbf{1}_{\{\mu_{k+1}\geq 0\}}C_{k+1}^TC_{k+1}\\
&\times\mathbb{E}\{\tilde{H}_k\varphi(\hat{x}_k)\tilde{\varphi}^T(\hat{x}_k)\tilde{\varphi}(\hat{x}_k)\varphi^T(\hat{x}_k)\tilde{H}_k^T\}C_{k+1}^TC_{k+1}\\
&+\beta_{10,k}^{-1}\theta_1^2\mathbf{1}_{\{\mu_{k+1}\geq 0\}}C_{k+1}^TC_{k+1}\\
&\times\mathbb{E}\{\bar{\varepsilon}_k\tilde{\varphi}^T(\hat{x}_k)\tilde{\varphi}(\hat{x}_k)\bar{\varepsilon}_k^T\}C_{k+1}^TC_{k+1},\\
&\Pi_{15,k}+\Pi_{15,k}^T\\
\leq&\beta_{11,k}^1\theta_1^2\mathbf{1}_{\{\mu_{k+1}\geq 0\}}C_{k+1}^TC_{k+1}\\
&\times\mathbb{E}\{\tilde{H}_k\varphi(\hat{x}_k)\tilde{\varphi}^T(\hat{x}_k)\tilde{\varphi}(\hat{x}_k)\varphi^T(\hat{x}_k)\tilde{H}_k^T\}C_{k+1}^TC_{k+1}\\
&+\beta_{11,k}^{-1}\theta_1^2\theta_2^2HH^T,\\
&\Pi_{20,k}+\Pi_{20,k}^T\\
\leq&\beta_{12,k}^1\theta_1^2\mathbf{1}_{\{\mu_{k+1}\geq 0\}}C_{k+1}^TC_{k+1}\mathbb{E}\{\bar{\varepsilon}_k\tilde{\varphi}^T(\hat{x}_k)\tilde{\varphi}(\hat{x}_k)\bar{\varepsilon}_k^T\}\\
&\times C_{k+1}^TC_{k+1}+\beta_{12,k}^{-1}\theta_1^2\theta_2^2HH^T. \qquad (26)
\end{aligned}
$$

By noting the following facts

$$
\begin{aligned}
&\mathbb{E}\left\{\tilde{\varphi}(\hat{x}_k)w_{k+1}^T\right\}=0,\quad \mathbb{E}\left\{w_{k+1}\tilde{\varphi}(\hat{x}_k)^T\right\}=0,\\
&\mathbb{E}\left\{\tilde{\varphi}(\hat{x}_k)w_k^T\right\}=0,\quad \mathbb{E}\left\{w_k\tilde{\varphi}(\hat{x}_k)^T\right\}=0, \qquad (27)
\end{aligned}
$$

it can be easily concluded that

$$
\begin{aligned}
&\Pi_{5,k}=0,\quad \Pi_{8,k}=0,\quad \Pi_{10,k}=0,\\
&\Pi_{14,k}=0,\quad \Pi_{16,k}=0,\quad \Pi_{17,k}=0,\\
&\Pi_{18,k}=0,\quad \Pi_{19,k}=0,\quad \Pi_{21,k}=0. \qquad (28)
\end{aligned}
$$

Next, by substituting (16), (26), and (28) into (25), it can be obtained that

$$
\begin{aligned}
&Q_{k+1}\\
\leq&\delta_{1,k}\mathbb{E}\{\tilde{H}_k\tilde{H}_k^T\}+\delta_{2,k}C_{k+1}^TC_{k+1}\bar{A}_k\mathbb{E}\{e_k\tilde{\varphi}^T(\hat{x}_k)\tilde{\varphi}(\hat{x}_k)e_k^T\}\\
&\times\bar{A}_k^TC_{k+1}^TC_{k+1}+\delta_{3,k}C_{k+1}^TC_{k+1}\\
&\times\mathbb{E}\{\tilde{H}_k\varphi(\hat{x}_k)\tilde{\varphi}^T(\hat{x}_k)\tilde{\varphi}(\hat{x}_k)\varphi^T(\hat{x}_k)\tilde{H}_k^T\}C_{k+1}^TC_{k+1}\\
&+\delta_{4,k}C_{k+1}^TC_{k+1}\bar{B}_k\mathbb{E}\{w_k\tilde{\varphi}^T(\hat{x}_k)\tilde{\varphi}(\hat{x}_k)w_k^T\}\bar{B}_k^TC_{k+1}^TC_{k+1}\\
&+\delta_{5,k}C_{k+1}^TC_{k+1}\mathbb{E}\{\bar{\varepsilon}_k\tilde{\varphi}^T(\hat{x}_k)\tilde{\varphi}(\hat{x}_k)\bar{\varepsilon}_k^T\}C_{k+1}^TC_{k+1}\\
&+\delta_{6,k}C_{k+1}^TD_{k+1}\mathbb{E}\{w_{k+1}\tilde{\varphi}^T(\hat{x}_k)\tilde{\varphi}(\hat{x}_k)w_{k+1}^T\}D_{k+1}^TC_{k+1}\\
&+\delta_{7,k}HH^T. \qquad (29)
\end{aligned}
$$

According to (17) and the following fact:

$$
\begin{aligned}
&\tilde{\varphi}^T(\hat{x}_k)\tilde{\varphi}(\hat{x}_k)\\
=&\left(\varphi(\hat{x}_k)/\left(\left\|1+\varphi^T(\hat{x}_k)\varphi(\hat{x}_k)\right\|\left\|C_{k+1}^TC_{k+1}\right\|\right)\right)^T\\
&\times\left(\varphi(\hat{x}_k)/\left(\left\|1+\varphi^T(\hat{x}_k)\varphi(\hat{x}_k)\right\|\left\|C_{k+1}^TC_{k+1}\right\|\right)\right)\\
\leq&1/\left\|C_{k+1}^TC_{k+1}\right\|^2, \qquad (30)
\end{aligned}
$$

the inequality (29) can be further derived as follows:

$$
\begin{aligned}
Q_{k+1}\leq&\delta_{1,k}Q_k+\bar{\delta}_{2,k}C_{k+1}^TC_{k+1}\bar{A}_kP_k\bar{A}_k^TC_{k+1}^TC_{k+1}\\
&+\bar{\delta}_{3,k}C_{k+1}^TC_{k+1}Q_kC_{k+1}^TC_{k+1}\\
&+\bar{\delta}_{4,k}C_{k+1}^TC_{k+1}\bar{B}_k\bar{B}_k^TC_{k+1}^TC_{k+1}\\
&+\bar{\delta}_{5,k}I+\bar{\delta}_{6,k}C_{k+1}^TD_{k+1}D_{k+1}^TC_{k+1}+\bar{\delta}_{7,k}I. \quad (31)
\end{aligned}
$$

Obviously, by noting the facts in (14), (24), and (31), it can be seen that the conditions in Lemma 1 are satisfied. Therefore, based on the above analysis, the following is obtained:

$$
P_k\leq\bar{P}_k,\quad Q_k\leq\bar{Q}_k \qquad (32)
$$

which completes the proof of Theorem 1. ■

We are now in a position to introduce the following theorem, which demonstrates how the NN-based estimator gain L_k can be parameterized by minimizing the upper bound $\bar{R}_{k+1}\triangleq\bar{P}_{k+1}+\bar{Q}_{k+1}$ as presented in Theorem 1.

Theorem 2: For the given positive scalars $\alpha_{i,k}$ $(i=1,2,3,4)$, $\beta_{i,k}$ $(i=1,2,\ldots,12)$, θ_i $(i=1,2)$, along with the definitions (16) and (17) in Theorem 1, the real-valued matrices $\bar{\Theta}_k$ and Φ_k are defined as follows:

$$
\begin{aligned}
\bar{\Theta}_k\triangleq&(\Theta_{2,k}^T\otimes\Theta_{1,k})+(\Theta_{4,k}^T\otimes\Theta_{3,k})\\
\Phi_k\triangleq&(1+\alpha_{1,k}^1+\alpha_{2,k}^1)A_k\bar{P}_kC_k^T+(1+\alpha_{3,k}^{-1})\bar{w}^2B_kD_k^T\\
&+\bar{\delta}_{2,k}C_{k+1}^TC_{k+1}C_{k+1}^TC_{k+1}A_k\bar{P}_kC_k^T\\
&+\bar{\delta}_{4,k}C_{k+1}^TC_{k+1}C_{k+1}^TC_{k+1}B_kD_k^T \qquad (33)
\end{aligned}
$$

where

$$
\begin{aligned}
\Theta_{1,k}\triangleq&(1+\alpha_{1,k}^1+\alpha_{2,k}^1)I+\bar{\delta}_{2,k}C_{k+1}^TC_{k+1}C_{k+1}^TC_{k+1}\\
\Theta_{2,k}\triangleq&C_k\bar{P}_kC_k^T\\
\Theta_{3,k}\triangleq&(1+\alpha_{3,k}^{-1})\bar{w}^2I+\bar{\delta}_{4,k}C_{k+1}^TC_{k+1}C_{k+1}^TC_{k+1}\\
\Theta_{4,k}\triangleq&D_kD_k^T. \qquad (34)
\end{aligned}
$$

Then, the upper bounds $\bar{P}_k$ and $\bar{Q}_k$ to (14) in Theorem 1 can be minimized iteratively with the NN-based estimator gain L_k, which satisfies the following form:

$$
\text{vec}(L_k)=\bar{\Theta}_k^{-1}\text{vec}(\Phi_k). \qquad (35)
$$

Proof: It is derived from Lemma 2 that

$$
\begin{aligned}
&\frac{\partial\text{tr}\{\bar{A}_k\bar{P}_k\bar{A}_k^T\}}{\partial L_k}=\frac{\partial\text{tr}\{(A_k-L_kC_k)\bar{P}_k(A_k-L_kC_k)^T\}}{\partial L_k}\\
&=-2A_k\bar{P}_kC_k^T+2L_kC_k\bar{P}_kC_k^T,\\
&\frac{\partial\text{tr}\{\bar{B}_k\bar{B}_k^T\}}{\partial L_k}=\frac{\partial\text{tr}\{(B_k-L_kD_k)(B_k-L_kD_k)^T\}}{\partial L_k}\\
&=-2B_kD_k^T+2L_kD_kD_k^T,\\
&\frac{\partial\text{tr}\{C_{k+1}^TC_{k+1}\bar{A}_k\bar{P}_k\bar{A}_k^TC_{k+1}^TC_{k+1}\}}{\partial L_k}\\
=&-2C_{k+1}^TC_{k+1}C_{k+1}^TC_{k+1}A_k\bar{P}_kC_k^T+2C_{k+1}^TC_{k+1}\\
&\times C_{k+1}^TC_{k+1}L_kC_k\bar{P}_kC_k^T,\\
&\frac{\partial\text{tr}\{C_{k+1}^TC_{k+1}\bar{B}_k\bar{B}_k^TC_{k+1}^TC_{k+1}\}}{\partial L_k}\\
=&-2C_{k+1}^TC_{k+1}C_{k+1}^TC_{k+1}B_kD_k^T+2C_{k+1}^TC_{k+1}\\
&\times C_{k+1}^TC_{k+1}L_kD_kD_k^T. \qquad (36)
\end{aligned}
$$

Then, we can easily have

$$
\begin{aligned}
\frac{\partial\text{tr}\{\bar{R}_{k+1}\}}{\partial L_k}=&-2(1+\alpha_{1,k}^1+\alpha_{2,k}^1)A_k\bar{P}_kC_k^T\\
&+2(1+\alpha_{1,k}^1+\alpha_{2,k}^1)L_kC_k\bar{P}_kC_k^T\\
&-2(1+\alpha_{3,k}^{-1})\bar{w}^2B_kD_k^T\\
&+2(1+\alpha_{3,k}^{-1})\bar{w}^2L_kD_kD_k^T
\end{aligned}
$$

$$
\begin{aligned}
&- 2\bar{\delta}_{2,k}C_{k+1}^TC_{k+1}C_{k+1}^TC_{k+1}A_k\bar{P}_kC_k^T\\
&+ 2\bar{\delta}_{2,k}C_{k+1}^TC_{k+1}C_{k+1}^TC_{k+1}L_kC_k\bar{P}_kC_k^T\\
&- 2\bar{\delta}_{4,k}C_{k+1}^TC_{k+1}C_{k+1}^TC_{k+1}B_kD_k^T\\
&+ 2\bar{\delta}_{4,k}C_{k+1}^TC_{k+1}C_{k+1}^TC_{k+1}L_kD_kD_k^T. \quad (37)
\end{aligned}
$$

Subsequently, letting

$$\frac{\partial \text{tr}\{\bar{R}_{k+1}\}}{\partial L_k} = 0,$$

it follows from (33) and (34) that

$$\Theta_{1,k}L_k\Theta_{2,k} + \Theta_{3,k}L_k\Theta_{4,k} = \Phi_k. \quad (38)$$

Next, by vectorizing both sides of equation (38) and applying Lemma 3, we obtain

$$(\Theta_{2,k}^T \otimes \Theta_{1,k})\text{vec}(L_k) + (\Theta_{4,k}^T \otimes \Theta_{3,k})\text{vec}(L_k) = \text{vec}(\Phi_k) \quad (39)$$

which, according to (33), can be rewritten as follows :

$$\bar{\Theta}_k\text{vec}(L_k) = \text{vec}(\Phi_k). \quad (40)$$

It is clear that the matrix $\bar{\Theta}_k$ is positive definite. Therefore, the NN-based estimator gain L_k can be computed as

$$\text{vec}(L_k) = \bar{\Theta}_k^{-1}\text{vec}(\Phi_k), \quad (41)$$

and the proof is now complete. ■

Remark 2: Theorems 1 and 2 have established the core results of this paper. Specifically, Theorem 1 has provided sufficient conditions under which the upper bounds of the state estimation error covariance and the NNW error covariance can be ensured for the addressed stochastic discrete time-varying systems with unknown nonlinear dynamics under the TBCP. Building upon these results, Theorem 2 has further presented an explicit parameterization of the NN-based estimator gain L_k by minimizing the combined upper bound $\bar{R}_{k+1}$. The proposed recursive NN-based state estimation framework holds significant practical value for networked control systems operating under limited communication resources and unknown nonlinear dynamics. By integrating the TBCP into the estimator design and providing an explicit gain parameterization strategy, the proposed approach offers a viable and computationally efficient solution for enhancing estimation accuracy while ensuring communication efficiency, thereby facilitating its application in industrial cyber-physical systems and resource-constrained environments.

Remark 3: A potential future application of the proposed NN-based state estimation algorithm is in large-scale industrial cyber-physical systems deployed over wireless sensor networks with limited bandwidth and constrained energy resources. In modern smart manufacturing, intelligent robotics, and remote industrial monitoring platforms, system dynamics are often nonlinear, time-varying, and partially unknown due to environmental changes, component aging, and operational uncertainties. Meanwhile, continuous data transmission is impractical because of communication congestion, shared network infrastructure, and power limitations of sensor nodes. By integrating neural-network-based online approximation with a token bucket communication mechanism, the proposed recursive estimator can simultaneously address unknown nonlinear dynamics and intermittent measurement availability. The theoretical guarantees on the upper bounds of both the state estimation error covariance and the neural-network-weight error covariance further enhance reliability, which is particularly important in safety-critical or high-precision applications. Therefore, the algorithm provides a promising solution for intelligent, resource-aware state estimation in next-generation Industry 4.0 systems and networked control environments.

Remark 4: The distinctive novelties of this paper, when compared to the existing literature, can be outlined as follows.

1) Firstly, the recursive NN-based state estimation problem has been addressed for the first time under the joint consideration of unknown nonlinear dynamics and the TBCP, which has rarely been investigated in prior studies.
2) Secondly, a novel measurement model characterizing the TBCP has been mathematically formulated and seamlessly integrated into the estimator design, enabling the joint handling of communication constraints and system uncertainties.
3) Thirdly, an explicit parameterization method for the NN-based estimator gain has been developed in Theorem 2, where the estimator gain is designed to minimize the combined upper bound of the state estimation error covariance and the NNW error covariance.

These contributions provide a comprehensive and systematic framework that bridges the gap between neural network-based estimation techniques and networked control systems with constrained communication, which, to the best of the authors' knowledge, has not been reported in the existing literature.

IV. AN ILLUSTRATIVE EXAMPLE

In this section, a numerical example is provided to demonstrate the effectiveness of the NN-based state estimation methods addressed in this paper. For comparison, the Kalman Filter (KF) and Extended Kalman Filter (EKF) are also applied in this simulation example.

Consider the system (1) with the following parameters:

$$
\begin{aligned}
A_k &= \begin{bmatrix} 1.5\cos(k) & 0.01 & 0.02 \\ 0.02 & -1.7\sin(k) & 0.3 \\ 0.01 & 0.03 & 1.9\cos(k) \end{bmatrix},\\
B_k &= \begin{bmatrix} 1 \\ \cos(k) \\ 1 \end{bmatrix}, \quad C_k = \begin{bmatrix} 0.9 & 0.25\cos(k) & 1 \\ 0.9 & 0.25 & 1 \end{bmatrix},\\
D_k &= \begin{bmatrix} 0.95 \\ 0.65\sin(k) \end{bmatrix}.
\end{aligned}
$$

The noise w_k is zero-mean stochastic noise satisfying $\bar{w}^2 = 0.1$. The initial states and covariances x_0, $\hat{x}_0$, P_0 and Q_0 are chosen as

$$
\begin{aligned}
x_0 &= \begin{bmatrix} x_0^{(1)} & x_0^{(2)} & x_0^{(3)} \end{bmatrix}^T = \begin{bmatrix} 3.1 & -2.65 & 4 \end{bmatrix}^T,\\
\hat{x}_0 &= \begin{bmatrix} \hat{x}_0^{(1)} & \hat{x}_0^{(2)} & \hat{x}_0^{(3)} \end{bmatrix}^T = \begin{bmatrix} 0 & 0 & 0 \end{bmatrix}^T,\\
P_0 &= 10I, \quad Q_0 = 10I.
\end{aligned}
$$

The unknown nonlinear function in the system (1) is taken as

$$f(x_k) = 8\begin{bmatrix} \sin(x_k^{(1)})\cos(x_k^{(3)}) \\ \sin(x_k^{(2)}) \\ \sin(x_k^{(2)})\cos(x_k^{(3)}) \end{bmatrix}.$$

The NN activation function and the initial NNW matrix are selected as

$$\phi(\hat{x}_k) = \begin{bmatrix} \tanh(\hat{x}_k^{(1)}) & \tanh(\hat{x}_k^{(2)}) & \tanh(\hat{x}_k^{(3)}) \end{bmatrix}^T,$$
$$\hat{H}_0 = \begin{bmatrix} 1.5 & 1.65 & 1.5 \\ 1.5 & 1.8 & 1.5 \\ 0.75 & 1.2 & 1.35 \end{bmatrix}.$$

For the KF, the linear component in the state equation adopts A_k, while the unknown nonlinear component is directly neglected. For the EKF, since the classic EKF cannot directly handle scenarios involving unknown nonlinearity, we use the form of $A_k + F_k$ (F_k is the Jacobian matrix of $f(x_k)$, i.e., $F_k = \partial f(x_k)/\partial x_k|_{\hat{x}_{k|k-1}}$) for the linear component in its state equation here to facilitate comparison. In other words, the nonlinear information is actually incorporated into the EKF.

For the TBCP, we set the token added rate $a = 2$, the transmission cost $b = 3$, the maximum capacity $S = 10$, and the initial bucket level $z_0 = 6$. Let the positive scalars be $\alpha_{i,k} = 0.1$ $(i = 1, 2, 3, 4)$ and $\beta_{i,k} = 0.9$ $(i = 1, 2, \ldots, 12)$, $\theta_i = 0.001$ $(i = 1, 2)$, $\bar{H} = 8$, $\bar{\varphi} = 1$, and $\bar{\varepsilon} = 0.1$.

With the above parameters, an upper bound for the state estimation error covariance can be calculated iteratively and the NN-based estimator gain and the NN tuning parameters can be determined recursively by using Theorems 1 and 2.

Simulation results are presented in Figs. 1-5. Specifically, Fig. 1 plots the state trajectories and their estimates for three state estimators. Fig. 2 shows the state estimation error covariances for three state estimators. Fig. 3 depicts the trace of the obtained upper bound and the log of the state estimation mean square error (MSE) of our proposed algorithm, which is defined by

$$\log(\mathrm{MSE}_{x,k}^{(j)}) \triangleq \frac{1}{T}\sum_{t=1}^{T}(x_k^{(j)} - \hat{x}_k^{(j)})^2$$

with $T = 300$, $j = 1, 2, 3$. Fig. 4 shows the unknown nonlinear dynamics trajectories and their approximations by the designed neural network. Fig. 5 describes the transmitted indicator for the TBCP at each instant k. The simulation results demonstrate that the filtering algorithm addressed in this paper is indeed effective.

Remark 5: This simulation compares the state estimation performance of three algorithms: Kalman Filter (KF), Extended Kalman Filter (EKF), and our estimator for a stochastic discrete time-varying system with unknown nonlinear dynamics and token bucket communication constraints. The KF relies on a linear system assumption and directly ignores the nonlinear term, leading to significant estimation bias when the system nonlinearity is strong. The EKF mitigates this limitation by linearizing the nonlinear term via first-order Taylor expansion and Jacobian matrix calculation, but its performance is highly dependent on the accuracy of linearization and may degrade if the nonlinearity is severe or the operating point changes rapidly. In contrast, the our proposed algorithm exhibits distinct advantages: 1) it adaptively learns the unknown nonlinear dynamics through online weight updates, avoiding reliance on manual linearization or prior knowledge of nonlinear structures; 2) it effectively handles communication constraints (token bucket protocol) by integrating transmission indicators into the estimation framework, maintaining robustness even when observations are intermittent. Overall, our estimator provides a more reliable and adaptive solution for state estimation of unknown nonlinear systems under communication constraints, overcoming the inherent limitations of linearization-based filters (KF/EKF).

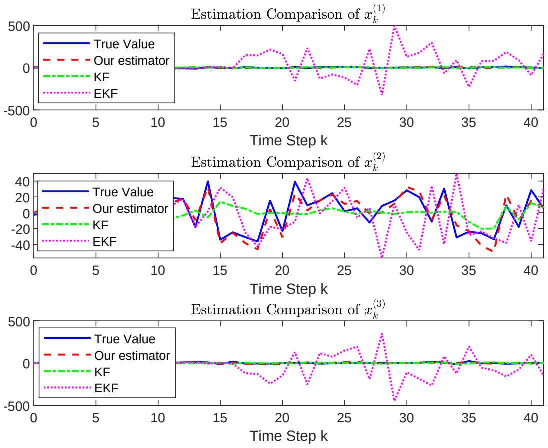


Fig. 1. State x_k and its estimations $\hat{x}_k$ for three state estimators.

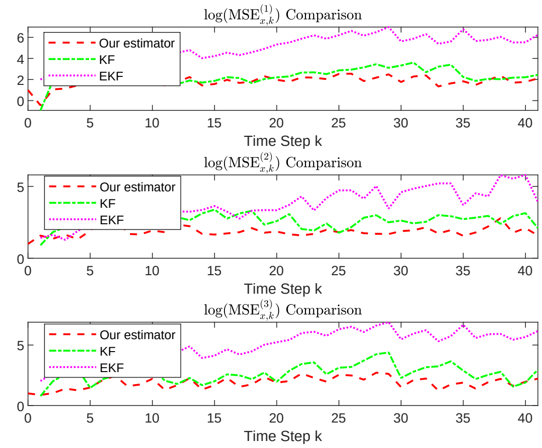


Fig. 2. State estimation error covariances for three state estimators.

V. CONCLUSIONS

In this paper, the recursive NN-based state estimation problem has been investigated under the joint consideration of unknown nonlinear dynamics and the TBCP, which has been incorporated to prevent data congestion and reduce the energy consumption associated with signal transmissions. The desired NN-based state estimation scheme has been developed, and the NN-based estimator gain parameters have been derived in terms of the solutions to two sets of matrix difference equations, aiming to minimize the upper bounds of both the state estimation error covariance and the NNW error covariance at each time instant. Finally, a numerical example has been provided to validate the effectiveness of the recursive NN-based state estimation algorithm proposed in this paper.

Future research will focus on extending the proposed methods to distributed systems, incorporating more complex network constraints, and enhancing robustness against model uncertainties [7], [8], [19], [44]. For example, the proposed framework can be extended to more general systems such as multi-agent systems and distributed sensor networks by incorporating distributed estimation and network interaction mechanisms. In multi-agent systems, each agent can implement a local NN-based estimator to approximate unknown nonlinear dynamics, while consensus or diffusion strategies can be introduced to coordinate state estimates across the network under communication constraints. For distributed sensor networks, local estimators can be deployed at individual sensor nodes, and distributed fusion techniques can be integrated to achieve global estimation objectives. By extending the matrix-recursion-based covariance analysis to block-structured or graph-dependent forms, and incorporating time-varying or switching communication topologies, the proposed communication-aware and learning-based estimation framework can provide a scalable solution for cooperative estimation in large-scale networked systems.

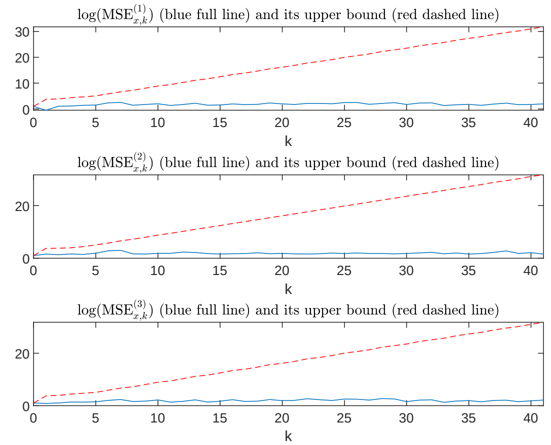


Fig. 3. State estimation error covariance of this paper and its upper bound.

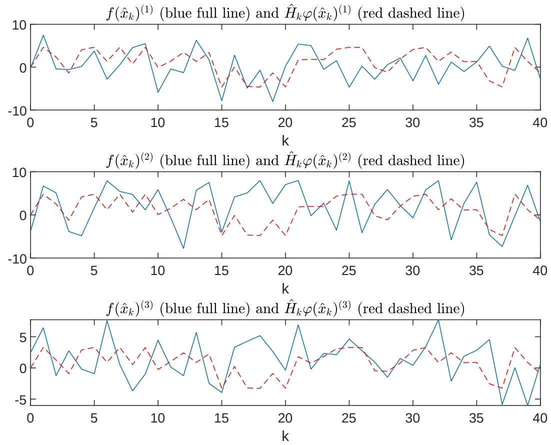


Fig. 4. Unknown nonlinear dynamics trajectories and their approximations.

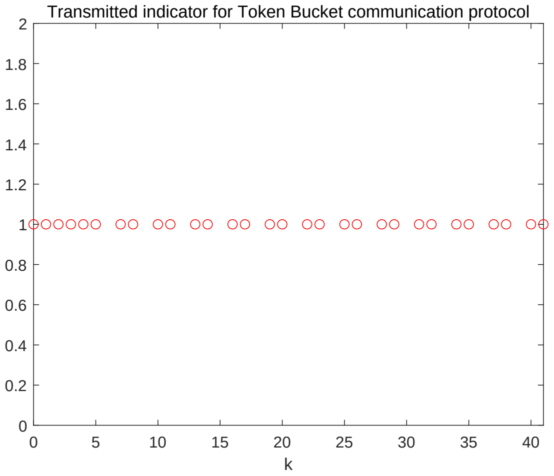


Fig. 5. Transmitted indicator for the TBCP at each instant k.

REFERENCES

[1] R. Caballero-Águila, M. P. Frías-Bustamante and A. Oya-Lechuga, Least-squares linear estimation for multirate uncertain systems subject to DoS attacks, *International Journal of Network Dynamics and Intelligence*, vol. 4, no. 2, art. no. 100014, Jun. 2025.

[2] S. Cai and J. Liang, Recursive filtering for nonlinear systems with relay communication, energy harvesting and correlated noises, *International Journal of Network Dynamics and Intelligence*, vol. 4, no. 3, art. no. 100021, Sept. 2025.

[3] D. Crisan and A. Doucet, A survey of convergence results on particle filtering methods for practitioners, *IEEE Transactions on Signal Processing*, Vol. 50, No. 3, pp. 736–746, 2002.

[4] M. Fraccaro, S. Sonderby, U. Paquet and O. Winther, Sequential neural models with stochastic layers, In *Advances in Neural Information Processing Systems*, 2016.

[5] Y. Gao, J. Hu, H. Yu, C. Jia and C. Chen, Variance-constrained H_∞ state estimation for time-varying delayed neural networks with random access protocol and sensor failures, *International Journal of Network Dynamics and Intelligence*, vol. 4, no. 3, art. no. 100019, Sept. 2025.

[6] H. Ge, L. Zou, H. Chen and J. Guo, Finite-time bounded state estimation for time-varying systems over FlexRay networks with probabilistic bit flips, *International Journal of Network Dynamics and Intelligence*, vol. 4, no. 4, art. no. 100023, Dec. 2025.

[7] Y. Guo, Z. Wang, J.-Y. Li and Y. Xu, An impulsive approach to state estimation for multirate singularly perturbed complex networks under bit rate constraints, *IEEE Transactions on Cybernetics*, vol. 55, no. 3, pp. 1197–1209, Mar. 2025.

[8] Y. Guo, Z. Wang, J.-Y. Li and Y. Xu, Pinning synchronization for stochastic complex networks with randomly occurring nonlinearities: tackling bit rate constraints and allocations, *IEEE Transactions on Cybernetics*, vol. 54, no. 12, pp. 7248–7259, Dec. 2024.

[9] A. Hasan, I. Kuncara, A. Widyotriatmo, O. Osen and R. T. Bye, Secure state estimation and control for autonomous ships under cyberattacks, *Systems Science & Control Engineering*, vol. 13, no. 1, art. no. 2518964, 2025.

[10] J. Kidambi, D. Ghosal, and B. Mukherjee, Dynamic token bucket: A fair bandwidth allocation algorithm for high-speed networks, *Journal of High Speed Networks*, Vol. 9, No. 2, pp. 67–87, 2000.

[11] J. Hu, S. Yang, R. Caballero-Águila, H. Dong and B. Wu, Mixed static-dynamic protocol-based tobit recursive filtering for stochastic nonlinear systems against random false data injection attacks, *IEEE Transactions on Signal and Information Processing over Networks*, Vol. 10, pp. 445–459, Apr. 2024.

[12] H. Kar and V. Singh, Elimination of overflow oscillations in fixed-point state-space digital filters with saturation arithmetic: An LMI Approach, *IEEE Transactions on Circuits and Systems-II*, Vol. 51, No. 1, pp. 40–42, Jan. 2004.

[13] M. Karl, M. Soelch, J. Bayer and P. Smagt, Deep variational bayes filters: Unsupervised learning of state space models from raw data, In *International Conference on Learning Representations*, 2017.

[14] Y. Lim and H. Ahn, Consensus under saturation constraints in interconnection states, *IEEE Transactions on Automatic Control*, Vol. 60, No. 11, pp. 3053–3058, Nov. 2015.

[15] D. Liu and A. Michel, Asymptotic stability of discrete-time systems with saturation nonlinearities with application to digital-filters, *IEEE Transactions on Circuits and Systems-I*, Vol. 39, No. 10, pp. 789–807, Oct. 1992.

[16] K. Liu, E. Fridman, K. Johansson and Y. Xia, Quantized control under Round-Robin communication protocol, *IEEE Transactions on Industrial Electronics*, Vol. 63, No. 7, pp. 4461–4471, Jul. 2016.

[17] S. Liu, L. Wang, Y. Zhang, Y.-A. Wang and H. Dong, Recursive filtering of networked systems with communication protocol scheduling: a survey, *International Journal of Systems Science*, vol. 56, no. 11, pp. 2499–2516, 2025.

[18] D. Lee, Y. Joo and M. Tak, Periodically time-varying H_∞ memory filter design for discrete-time LTI systems with polytopic uncertainty, *IEEE Transactions on Automatic Control*, Vol. 59, No. 5, pp. 1380–1385, May 2014.

[19] T. Liu, Z. Wang, Y. Liu and R. Wang, Unscented-Kalman-filter-based remote state estimation for complex networks with quantized measurements and amplify-and-forward relays, *IEEE Transactions on Cybernetics*, vol. 54, no. 11, pp. 6819–6831, Nov. 2024.

[20] K. Mathews, C. Kramer, and R. Gotzhein, Token bucket based traffic shaping and monitoring for WLAN-based control systems, *2017 IEEE 28th Annual International Symposium on Personal, Indoor, and Mobile Radio Communications*, pp. 1–7, 2017.

[21] W. Mao, M. Liu, M. Salzmann, and H. Li, Multi-level Motion Attention for Human Motion Prediction, *International Journal of Computer Vision*, Vol. 129, No. 9, pp. 2513–2535, Sep. 2021.

[22] X. Meng, H. Wang, Y. Li and Y. Shen, Unscented Kalman filtering for nonlinear systems with stochastic nonlinearities under FlexRay protocol, *International Journal of Network Dynamics and Intelligence*, vol. 4, no. 2, art. no. 100010, Jun. 2025.

[23] C. Naesseth, S. Linderman, R. Ranganath, and D. Blei, Variational sequential Monte Carlo, In: Proc. *International Conference on Artificial Intelligence and Statistics*, 2018.

[24] G. Procissi, A. Garg, M. Gerla, and M. Sanadidi, Token bucket characterization of long-range dependent traffic, *Computer Communications*, Vol. 25, No. 11, pp. 1009–1017, 2002.

[25] Z. Qiang and W. Tao, Enhancing visual SLAM localization accuracy through dynamic object detection and adaptive feature filtering, *International Journal of Network Dynamics and Intelligence*, vol. 4, no. 3, art. no. 100018, Sept. 2025.

[26] K. Reif, S. Gunther, E. Yaz and R. Unbehauen, Stochastic stability of the discrete-time extended Kalman filter, *IEEE Transactions on Automatic Control*, Vol. 44, No. 4, pp. 714–728, 1999.

[27] S. Rangapuram, M. Seeger, J. Gasthaus, L. Stella, Y. Wang and T. Januschowski, Deep state space models for time series forecasting, In *Advances in Neural Information Processing Systems*, 2018.

[28] X. Shao, J. Zhang, M. Lyu and Y. Lu, Event-based nonfragile state estimation for memristive neural networks with multiple time-delays and sensor saturations, *International Journal of Systems Science*, vol. 56, no. 3, pp. 618–637, 2025.

[29] C. C. Shen, C. Srisathapornphat and C. Jaikaeo, Sensor information networking architecture and applications, *IEEE Personal Communications*, Vol. 8, No. 4, pp. 52–59, Aug. 2001.

[30] X. Shu, L. Zhang, G. Qi, W. Liu and J. Tang, Spatiotemporal Co-Attention Recurrent Neural Networks for Human-Skeleton Motion Prediction, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 44, No. 6, pp. 3300–3315, 2022.

[31] D. Sivaraman, B. M. Pillai, C. Wiratkapun, J. Suthakorn and S. Ongwattanakul, Dynamic order selection analysis in adaptive polynomial Kalman filtering: implementation and integration of sensor data and hybrid image processing for bio-inspired needle systems, *Systems Science & Control Engineering*, vol. 13, no. 1, art. no. 2546839, 2025.

[32] J. Song, Z. Wang, Y. Niu, X. Yi and Q.-L. Han, Co-Design of dissipative deconvolution filter and Round-Robin protocol for networked 2-D digital systems: optimization and application, *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, Vol. 53, No. 10, pp. 6316–6328, Oct. 2023.

[33] M. Tabbara and D. Nesic, Input-output stability of networked control systems with stochastic protocols and channels, *IEEE Transactions on Automatic Control*, Vol. 53, No. 5, pp. 1160–1175, Jun. 2008.

[34] E. Tian, M. Fan, L. Ma and D. Yue, Stochastic important-data-based attack power allocation against remote state estimation in sensor networks, *IEEE Transactions on Automatic Control*, Vol. 70, No. 3, pp. 2012-2019, Mar. 2025.

[35] Y. Toroshanko, Y. Selepyna, N. Yakymchuk, and V. Cherevyk, Control of traffic streams with the multi-rate token bucket, *International Conference on Advanced Information and Communications Technologies*, pp. 352–355, 2019.

[36] V. Ugrinovskii and E. Fridman, A Round-Robin type protocol for distributed estimation with H_∞ consensus, *Systems and Control Letters*, Vol. 69, pp. 103–110, Jul. 2014.

[37] M. Van, R. Mandl, C. Stam, R. Kahn and H. Pol, Aberrant frontal and temporal complex network structure in schizophrenia: A graph theoretical analysis, *Journal of Neuroscience*, Vol. 30, No. 47, pp. 15915–15926, Nov. 2010.

[38] G. Valverde and V. Terzija, Unscented Kalman filter for power system dynamic state estimation, *IET Generation Transmission and Distribution*, Vol. 5, No. 1, pp. 29–37, 2011.

[39] X. Wan, C. Yang, C.-K. Zhang and M. Wu, Hybrid adjusting variables-dependent event-based finite-time state estimation for two-time-scale Markov jump complex networks, *IEEE Transactions on Neural Networks and Learning Systems*, Vol. 35, No. 2, pp. 1487–1500, Feb. 2024.

[40] L. Wang, Z. Wang, D. Zhao, Y. Liu and G. Wei, Recursive filtering for discrete-time stochastic complex networks under bit-rate constraints: a locally minimum variance approach, *IEEE Transactions on Automatic Control*, Vol. 69, No. 5, pp. 3441-3448, May 2024.

[41] R. Wang, T. Yu and S. He, Finite-time control for MJSs under protocol-based fading network and input saturation: a WOA-assisted method, *International Journal of Systems Science*, vol. 56, no. 16, pp. 3863–3877, 2025.

[42] Y. Wang, Z. Wang, L. Zou and H. Dong, Observer-based fuzzy PID tracking control under try-once-discard communication protocol: an affine fuzzy model approach, *IEEE Transactions on Fuzzy Systems*, Vol. 32, No. 4, pp. 2352–2365, Apr. 2024.

[43] Y. Wang, Z. Wang, L. Zou, L. Ma and H. Dong, Ultimately bounded PID control for T-S fuzzy systems under FlexRay communication protocol, *IEEE Transactions on Fuzzy Systems*, Vol. 31, No. 12, pp. 4308–4320, Dec. 2023.

[44] Y. Wang, Z. Wang, L. Zou, Y. Chen and D. Yue, Distributed proportional-integral fuzzy state estimation over sensor networks under energy-constrained denial-of-service attacks, *IEEE Transactions on Cybernetics*, vol. 54, no. 1, pp. 641–654, Jan. 2024.

[45] Z. Wang, M. Andrews, Z. Wu, L. Wang and C. Bauch, Coupled disease-behavior dynamics on complex networks: A review, *Physics of Life Reviews*, Vol. 15, pp. 1–29, Dec. 2015.

[46] L. Yu, J. Ding, H. Peng, Y. Liu and Y. Liu, Sampled-data based containment control for a class of nonlinear multiagent systems with dynamic leaders and control saturation, *International Journal of Network Dynamics and Intelligence*, vol. 4, no. 2, art. no. 100011, Jun. 2025.

[47] W. Zhang, Y. Zhao, H. Wang and Z. Yang, Robust model predictive control for polyhedral uncertain systems under relay and redundancy protocol, *International Journal of Systems Science*, vol. 56, no. 8, pp. 1617–1632, 2025.

[48] Y. Zhang, Z. Wang, L. Zou, H. Dong and X. Yi, Neural-network-based secure state estimation under energy-constrained denial-of-service attacks: an encoding-decoding scheme, *IEEE Transactions on Network Science and Engineering*, Vol. 10, No. 4, pp. 2002-2015, Jul. 2023.

[49] Y. Zhang, Z. Wang, L. Zou, W. Qian and S. Du, Neural-network-based recursive state estimation for nonlinear networked systems with

binary-encoding mechanisms, *IEEE Transactions on Neural Networks and Learning Systems*, Vol. 36, No. 6, pp. 10072-10083, Jun. 2025.

[50] D. Zhao, C. Gao, J. Li, H. Fu and D. Ding, PID control and PI state estimation for complex networked systems: a survey, *International Journal of Systems Science*, vol. 56, no. 11, pp. 2735–2750, 2025.

[51] P. Zhao, Y. Chen, D. Ding and H. Liu, Privacy-preserving distributed state estimation for microgrids based on encrypted measurements under bit-rate constraints, *International Journal of Systems Science*, vol. 56, no. 10, pp. 2481–2498, 2025.

[52] Y. Zhao, H. Gao and J. Lam, New results on H_∞ filtering for fuzzy systems with interval time-varying delays, *Information Sciences*, Vol. 181, No. 11, pp. 2356–2369, Jun. 2011.

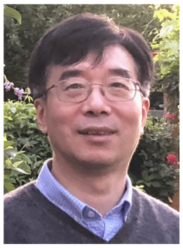
[53] Z. Zhao, Z. Wang, L. Zou, H. Liu and F. E. Alsaadi, Zonotopic multi-sensor fusion estimation with mixed delays under try-once-discard protocol: a set-membership framework, *Information Fusion*, Vol. 91, pp. 681–693, Mar. 2023.

[54] S. Zhou and J. Lam, H_∞ filtering for systems with delays and time-varying nonlinear parameters, *Circuits Systems and Signal Processing*, Vol. 29, No. 4, pp. 601–627, Aug. 2010.

[55] L. Zou, Z. Wang and H. Gao, Observer-based H_∞ control of networked systems with stochastic communication protocol: The finite-horizon case, *Automatica*, Vol. 63, pp. 366–373, Jan. 2016.

[56] L. Zou, Z. Wang, H. Gao and X. Liu, State estimation for discrete-time dynamical networks with time-varying delays and stochastic disturbances under the Round-Robin protocol, *IEEE Transactions on Neural Networks and Learning Systems*, Vol. 28, No. 5, pp. 1139–1151, May 2017.

Dong Wang was born in Changzhi, Shanxi, China, in May 1991. He received his B.S. degree in Mathematics and Applied Mathematics (Financial Engineering) and Ph.D. degree in Control Science and Engineering from Donghua University, Shanghai, China, in 2013 and 2018, respectively. From April 2016 to June 2016, he was a Visiting Scholar with the Department of Mechanical Engineering, The University of Hong Kong, Hong Kong, China. From January 2017 to January 2018, he was a Visiting Scholar with the Department of Computer Science, Brunel University London, U.K. He is currently a Lecturer with the School of Electrical Engineering, Shanghai Dianji University, Shanghai, China. His research interests include artificial intelligence, computer vision, and intelligent manufacturing.



Zidong Wang (SM'03-F'14) received the B.Sc. degree in mathematics in 1986 from Suzhou University, Suzhou, China, and the M.Sc. degree in applied mathematics in 1990 and the Ph.D. degree in electrical engineering in 1994, both from Nanjing University of Science and Technology, Nanjing, China.

He is currently a Professor of Dynamical Systems and Computing in the Department of Computer Science, Brunel University London, Uxbridge, U.K. From 1990 to 2002, he held teaching and research appointments in universities in China, Germany and the U.K. Prof. Wang's research interests include dynamical systems, signal processing, bioinformatics, control theory and applications. He has published a number of papers in international journals. He is a holder of the Alexander von Humboldt Research Fellowship of Germany, the JSPS Research Fellowship of Japan, William Mong Visiting Research Fellowship of Hong Kong.

Prof. Wang serves (or has served) as the Editor-in-Chief for *International Journal of Systems Science*, the Editor-in-Chief for *Neurocomputing*, the Editor-in-Chief for *Systems Science & Control Engineering*, and an Associate Editor for 12 international journals including IEEE Transactions on Automatic Control, IEEE Transactions on Control Systems Technology, IEEE Transactions on Neural Networks, IEEE Transactions on Signal Processing, and IEEE Transactions on Systems, Man, and Cybernetics-Part C. He is a Member of the Academia Europaea, a Member of the European Academy of Sciences and Arts, an Academician of the International Academy for Systems and Cybernetic Sciences, a Fellow of the IEEE, a Fellow of the Royal Statistical Society and a member of program committee for many international conferences.

Chuanbo Wen was born in Henan, China, in August 1981. He received the bachelor's degree in mathematics and the master's degree in applied mathematics from Henan University, Kaifeng, China, in 2004 and 2007, respectively, and the Ph.D. degree in control theory and control engineering from Shanghai Jiao Tong University, Shanghai, China, in 2015. From 2015 to 2016, he was a Senior Visiting Scholar with Brunel University London, London, U.K. He is currently a Professor with the School of Electrical Engineering, Shanghai Dianji University, Shanghai, China. His research interests include complex equipment fault diagnosis and life prediction, multisource information fusion, and artificial intelligence. Dr. Wen is currently the Vice President of the Shanghai Electric Machinery Association, Shanghai, China.