Delay-Distribution-Dependent H_∞ State Estimation for Delayed Neural Networks with (x, v)-Dependent Noises and Fading Channels

Li Sheng[a], Zidong Wang[b,c,*], Engang Tian[d] and Fuad E. Alsaadi[e]

Abstract

This paper deals with the H_∞ state estimation problem for a class of discrete-time neural networks with stochastic delays subject to state- and disturbance-dependent noises (also called (x, v)-dependent noises) and fading channels. The time-varying stochastic delay takes values on certain intervals with known probability distributions. The system measurement is transmitted through fading channels described by the Rice fading model. The aim of the addressed problem is to design a state estimator such that the estimation performance is guaranteed in the mean-square sense against admissible stochastic time-delays, stochastic noises as well as stochastic fading signals. By employing the stochastic analysis approach combined with the Kronecker product, several delay-distribution-dependent conditions are derived to ensure that the error dynamics of the neuron states is stochastically stable with prescribed H_∞ performance. Finally, a numerical example is provided to illustrate the effectiveness of the obtained results.

Index Terms

Delayed neural networks, H_∞ state estimation, delay-distribution-dependent condition, random delay, (x, v)-dependent noises, fading channels.

I. Introduction

The past few decades have witnessed the successful applications of recurrent neural networks (RNNs) in many areas including image processing [40], pattern recognition [3], combinatorial optimization [24], associative memories [32] and signal processing [35]. In general, these applications are heavily dependent on 1) the dynamic behaviors (e.g. stability and synchronization) of the RNNs; and 2) the true states of the neurons in a noisy environment. Therefore, the analysis issues of neural networks such as synchronization, stability and state estimation have attracted considerable attention, and a rich body of results have been reported in the literature [20], [39].

It is well known that the time-delay, which is inevitable during signal transmission between the neurons and in the implementation of neural networks, is one of the important sources which may cause instability and oscillation of the networks. As such, it is of great significance to investigate the mathematical properties of delayed neural networks, and much effort has been devoted to various types of time-delays (continuous, discrete, distributed or

This work was supported in part by the Royal Society of the UK, the Research Fund for the Taishan Scholar Project of Shandong Province of China, National Natural Science Foundation of China under Grants 61403420, 61573377 and 61329301, Fundamental Research Fund for the Central Universities of China under Grant 15CX08014A, and the Alexander von Humboldt Foundation of Germany.

[a] College of Information and Control Engineering, China University of Petroleum (East China), Qingdao 266580, China.

[b] College of Electrical Engineering and Automation, Shandong University of Science and Technology, Qingdao 266590, China.

[c] Department of Computer Science, Brunel University London, Uxbridge, Middlesex, UB8 3PH, United Kingdom.

[d] School of Electrical and Automation Engineering, Nanjing Normal University, Nanjing 210023, China.

[e] F. E. Alsaadi is with the Faculty of Engineering, King Abdulaziz University, Jeddah 21589, Saudi Arabia.

[*] Corresponding author. E-mail: Zidong.Wang@brunel.ac.uk

mixed), see [17], [21], [23], [30], [31], [34], [37], [39]. Nevertheless, an important class of time-delays, namely, probabilistic delays, have not gained sufficient attention in the context of dynamics analysis for RNNs. Probabilistic delays occur frequently in practice, for example, by using the statistical method, it has been found in [31] that a large delay occurs with a low probability in networked control systems. The time delay in neural networks may randomly appear as well due to synaptic voltage and temporal noise associated with transmitter release. In [39], the Bernoulli variable has been introduced to characterize the random delay and several less conservative stability conditions have been derived for delayed neural networks. By employing the information of both the probability distribution and the variation range of the time delay, the exponential H_∞ filtering problem has been addressed in [23] for switched neural networks with random delays.

Due to random fluctuations from the release of neurotransmitters as well as thermal noises in the electronic equipments, various stochastic perturbations are unavoidable with both biological and artificial neural networks. Up to now, most literature has focused on the stochastic neural networks with *state-dependent noises* only for the purpose of simplicity [1], [17]. As pointed out in [9], not only system states but also external disturbances may be corrupted by stochastic noises in the engineering practice. By means of Hamilton-Jacobi inequalities, the stochastic H_∞ control problem has been studied in [18] for nonlinear Markovian jump systems with state- and disturbance-dependent noises ((x, v)-dependent noises for short). For the stochastic H_2/H_∞ control problem, it has been revealed in [29] that there exist essential differences between the system with state-dependent noises and that with (x, v)-dependent noises. Note that (x, v)-dependent noises are typical phenomena for RNNs because 1) the neurotransmitter-induced noises are naturally neuron-state-dependent; 2) the thermal noises are usually external-disturbance-dependent; and 3) both kinds of noises tend to occur simultaneously in practice. Nevertheless, the dynamics analysis issue for neural networks with (x, v)-dependent noises has not been addressed and remains open.

In reality, the information of the neuron states of RNNs is crucial for some specific applications such as associative memories, optimization and state feedback control. Unfortunately, such information may not be fully accessible because of the complexity of neural networks and it is necessary to estimate the neuron states via available measurements. As such, the problem of state estimation for neural networks has stirred particular research interest and a wealth of literature has appeared [14], [16], [36], [38], [42]. By constructing a new Lyapunov-Krasovskii functional, the delay-distribution-dependent state estimator has been designed in [1] for discrete-time neural networks with time-varying delays. In [16], the Arcak-type state estimator, which is more general than the widely used Luenberger-type one, has been designed for the static neural networks with time delays. Recalling these existing methods, there has been a common assumption that communication channel is ideal such that the measurements of neural networks can be transmitted to the estimator in an *instantaneous way*. Such an assumption is, however, not always true when the RNNs and the estimator are connected via unreliable channels (i.e., wireless connection) in the case of hardware implementation. As such, it makes practical sense to study the state estimation problem for neural networks in a networked environment.

Recently, there have been some results on state estimation problems of neural networks against network-induced phenomena such as communication delays [17], missing measurements [20], quantization effects [42], and event-triggered strategy [34]. However, another network-induced phenomenon, i.e., fading channel, has gained relatively less attention in the context of state estimator design despite its practical significance in wireless communication networks. Generally, when signals are transmitted through wireless channels, they are often subject to several phenomena such as scattering and reflection due probably to shadowing effects from obstacles, the multipath propagation and the path loss. Therefore, the channel fading phenomenon is unavoidable in wireless networks and it could deteriorate the performance of networked systems if not handled properly [10], [26]. In order to reflect the changes of the transmitted signals in both the amplitude and the phase, fading can be modeled by a time-varying stochastic mathematical model such as Rice fading channel model [10] and Rayleigh fading channel model [28]. So far, some initial results have been reported for the problems of stabilization [10], H_∞ filtering [6], H_∞ control [4] and Kalman filtering [25] with fading channels. Nevertheless, to the best of the authors' knowledge, the state estimation problem for delayed neural networks with fading channels has not been adequately studied, not to mention the case when the (x, v)-dependent noises are also a concern.

Motivated by the above discussions, we aim to investigate the H_∞ state estimation problem for a class of delayed stochastic neural networks. *The main contributions of this paper can be summarized as follows. 1) The neural network addressed is comprehensive to cover random delays and (x, v)-dependent noises, which may reflect the reality more closely. 2) This paper represents the first of few attempts to study the problem of state estimation for neural networks with fading channels. 3) Based on the stochastic analysis approach and the Kronecker product, several delay-distribution-dependent conditions are derived under which the dynamics of the estimation error is stochastically stable with the prespecified H_∞ constraint.*

The rest of this paper is organized as follows. In Section II, the neural networks with random delays, (x, v)-dependent noises and fading channels are introduced and some preliminaries are briefly outlined. In Section III, the H_∞ state estimation problem is investigated by applying the stochastic analysis approach and the Kronecker product, and the estimator gains are obtained by solving a linear matrix inequality (LMI). A numerical example is provided to show the effectiveness of the main results in Section IV. Finally, conclusions are drawn in Section V.

Notations. Throughout this paper, $\mathbb{R}$ (respectively, $\mathbb{N}^+$) is the set of all real numbers (respectively, non-negative integers). $\mathbb{R}^n$ is the set of all real n-dimensional vectors and $\mathbb{R}^{m\times n}$ is the set of all $m \times n$ real matrices. $A > 0$ (respectively, $A \geq 0$) is a real symmetric positive definite (respectively, positive semi-definite) matrix. A^T denotes the transpose of a matrix A. $[a : b]$ is a set involving all integers between a and b. $\mathcal{CK}$ denotes the class of all continuous non-decreasing convex functions $\mu : \mathbb{R}^+ \to \mathbb{R}^+$ such that $\mu(0) = 0$ and $\mu(r) > 0$ for $r > 0$. $\mathcal{C}^m(\mathbb{R}^n)$ denotes the class of functions $V(x)$ that is m times continuously differentiable with respect to $x \in \mathbb{R}^n$. $\mathbb{E}\{x\}$ stands for the mathematical expectation of x. $\mathrm{diag}\{\cdots\}$ is a block-diagonal matrix. The symbol $\otimes$ denotes the Kronecker product. The asterisk $*$ in a matrix is used to denote the term that is induced by symmetry. Matrices, if they are not explicitly specified, are assumed to have compatible dimensions.

II. Problem Formulation and Preliminaries

In this section, we introduce some preliminaries related to H_∞ state estimation for neural networks and then give the problem formulation.

A. Neural Networks with (x, v)-dependent noises

Consider the following discrete-time neural network with time-varying delays and (x, v)-dependent noises:

$$\begin{cases} x(k+1) = Ax(k) + W_1\bar{f}(x(k)) + W_2\bar{g}(x(k-d(k))) + Cv(k) + [\bar{A}x(k) + \bar{B}x(k-d(k)) + \bar{C}v(k)]w(k), \\ \quad y(k) = Dx(k) + Ev(k), \\ \quad z(k) = Fx(k), \\ \quad x(j) = \phi(j), \quad -d_M \leq j \leq 0 \end{cases} \tag{1}$$

where $x(k) = [x_1(k), x_2(k), \ldots, x_n(k)]^T \in \mathbb{R}^n$ is the state vector associated with n neurons, $y(k) \in \mathbb{R}^{n_y}$ is the measurement output, $z(k) \in \mathbb{R}^{n_z}$ is the neural signal to be estimated and $w(k)$ is a one-dimensional zero-mean Gaussian white noise sequence on a probability space $(\Omega, \mathcal{F}, \mathrm{Prob})$ with $\mathbb{E}\{w^2(k)\} = 1$. Let $(\Omega, \mathcal{F}, \{\mathcal{F}_k\}_{k\in\mathbb{N}^+}, \mathrm{Prob})$ be a filtered probability space where $\{\mathcal{F}_k\}_{k\in\mathbb{N}^+}$ is the family of sub σ-algebras of $\mathcal{F}$ generated by $\{\omega(k)\}_{k\in\mathbb{N}^+}$.

$v(k) \in \mathbb{R}^{n_v}$ is the disturbance input which belongs to $l_2([0,\infty),\mathbb{R}^{n_v})$, where $l_2([0,\infty),\mathbb{R}^{n_v})$ is the space of nonanticipatory square-summable stochastic process $\{v(k) \in \mathbb{R}^{n_v}\}_{k\in\mathbb{N}^+}$ with respect to $\{\mathcal{F}_k\}_{k\in\mathbb{N}^+}$. $d(k) \in [d_m : d_M]$ ($d_M \geq d_m \geq 0$) is the time-varying *random* delay. $\phi(j)$, $-d_M < j \leq 0$ is the initial condition which is assumed to be independent of the process $\{w(\cdot)\}$. $A = \mathrm{diag}\{a_1, a_2, \ldots, a_n\}$ with $|a_i| < 1$ describes the rate with which the i-th neuron will reset its potential to the resting state in isolation when disconnected from the networks and external inputs. W_1 and W_2 are the connection weight matrix and the delayed connection weight matrix, respectively. C, $\bar{A}$, $\bar{B}$, $\bar{C}$, D, E and F are known real constant matrices with appropriate dimensions.

The neuron activation functions $\bar{f}(x(k)) = [\bar{f}_1(x_1(k)), \bar{f}_2(x_2(k)), \ldots, \bar{f}_n(x_n(k))]^T$, $\bar{g}(x(k)) = [\bar{g}_1(x_1(k)), \bar{g}_2(x_2(k)), \ldots, \bar{g}_n(x_n(k))]^T$ are continuous, and satisfy $\bar{f}(0) = 0$, $\bar{g}(0) = 0$ and the following sector-bounded condition:

$$\begin{aligned}&[\bar{f}(x) - \bar{f}(y) - \Phi_f(x-y)]^T[\bar{f}(x) - \bar{f}(y) - \Psi_f(x-y)] \leq 0,\\ &[\bar{g}(x) - \bar{g}(y) - \Phi_g(x-y)]^T[\bar{g}(x) - \bar{g}(y) - \Psi_g(x-y)] \leq 0\end{aligned} \tag{2}$$

for all $x, y \in \mathbb{R}^n$, where Φ_f, Ψ_f, Φ_g and Ψ_g are real matrices with appropriate dimensions.

The system (1) is described by a discrete-time stochastic difference equation (SDE). Since the difference equation is a recursive relation, the solution to the SDE is obtained iteratively by beginning with any initial condition. According to Theorem 2.2 of [41], a solution of such kind of stochastic difference equation exists if the diffusion and drift terms are measurable. Since the neuron activation functions in (1) are assumed to be continuous, both the diffusion and drift terms are therefore continuous. Based on Ex. 11.14 of [27], a continuous function is measurable, and we can conclude that the solution of (1) exists with any initial condition.

B. Random Delay

In the system (1), it is assumed that the random delay $d(k)$ is bounded and its probability distribution can be observed. For a given number $N \leq \lfloor \frac{d_M - d_m}{2} \rfloor$ where $\lfloor \cdot \rfloor$ means the rounding down function, suppose that $d(k)$ takes values in $[d_1^m : d_1^M]$, or $[d_2^m : d_2^M]$, or ..., or $[d_N^m : d_N^M]$ with $d_m \leq d_1^m < d_1^M < d_2^m < d_2^M < \cdots < d_N^m < d_N^M \leq d_M$, and

$$\mathrm{Prob}\{d(k) \in [d_1^m : d_1^M]\} = \bar{\alpha}_1, \;\; \mathrm{Prob}\{d(k) \in [d_2^m : d_2^M]\} = \bar{\alpha}_2, \;\; \ldots, \;\; \mathrm{Prob}\{d(k) \in [d_N^m : d_N^M]\} = \bar{\alpha}_N, \tag{3}$$

where $0 \leq \bar{\alpha}_i \leq 1$, $i = 1, 2, \ldots, N$ and $\sum_{i=1}^N \bar{\alpha}_i = 1$. In order to describe the probability distribution of the time delay, we define the following sets

$$\mathbb{D}_1 = \{k|d(k) \in [d_1^m : d_1^M]\}, \;\; \mathbb{D}_2 = \{k|d(k) \in [d_2^m : d_2^M]\}, \;\; \ldots, \;\; \mathbb{D}_N = \{k|d(k) \in [d_N^m : d_N^M]\}, \tag{4}$$

which imply that $\mathbb{D}_1 \cup \mathbb{D}_2 \cup \ldots \cup \mathbb{D}_N = \mathbb{N}^+$ and $\mathbb{D}_i \cap \mathbb{D}_j = \varnothing$, $\forall i \neq j$, $i, j = 1, 2, \ldots, N$.

Define N mapping functions

$$d_1(k) = \begin{cases} d(k), & k \in \mathbb{D}_1 \\ d_2^m, & \text{else} \end{cases} \quad d_2(k) = \begin{cases} d(k), & k \in \mathbb{D}_2 \\ d_3^m, & \text{else} \end{cases} \quad \ldots \quad d_N(k) = \begin{cases} d(k), & k \in \mathbb{D}_N \\ d_{N+1}^m, & \text{else.} \end{cases} \tag{5}$$

From (5), it can be found that $k \in \mathbb{D}_i$ implies the event $d(k) \in [d_i^m : d_i^M]$, $i = 1, 2, \ldots, N$ occurs. Define the following stochastic variables

$$\alpha_1(k) = \begin{cases} 1, & k \in \mathbb{D}_1 \\ 0, & \text{else} \end{cases} \quad \alpha_2(k) = \begin{cases} 1, & k \in \mathbb{D}_2 \\ 0, & \text{else} \end{cases} \quad \ldots \quad \alpha_N(k) = \begin{cases} 1, & k \in \mathbb{D}_N \\ 0, & \text{else.} \end{cases} \tag{6}$$

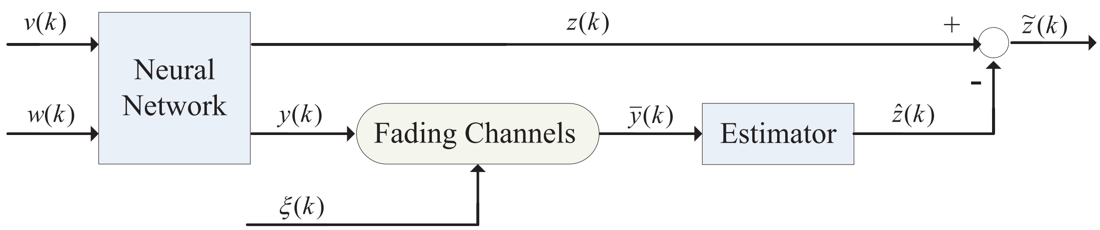

Fig. 1. Framework of the H_∞ state estimator design for neural networks over network environments.

According to (3), we have $\mathrm{Prob}\{\alpha_i(k)=1\}=\mathbb{E}\{\alpha_i(k)\}=\bar{\alpha}_i$, $i=1,2,\ldots,N$. Then, the original system (1) can be rewritten as

$$\begin{cases} x(k+1) = Ax(k) + W_1\bar{f}(x(k)) + \sum_{i=1}^{N}\alpha_i(k)W_2\bar{g}\left(x\left(k-d_i(k)\right)\right) + Cv(k) \\ \qquad + \left[\bar{A}x(k) + \sum_{i=1}^{N}\alpha_i(k)\bar{B}x(k-d_i(k)) + \bar{C}v(k)\right]w(k), \\ y(k) = Dx(k) + Ev(k), \\ z(k) = Fx(k), \\ x(j) = \phi(j), \quad -d_M \le j \le 0. \end{cases} \tag{7}$$

C. Fading Channels

In this paper, we consider the phenomenon of fading channels in the signal transmission which could be caused by the unreliable wireless network medium. The measurement of the neural network is no longer equivalent to the input of the estimator when there exist fading channels between the neural network and the estimator. Considering the L-th order Rice fading model in [10], the measurement signal received by the estimator is described by

$$\bar{y}(k) = \sum_{j=0}^{l_k}\beta_j(k)y(k-j) + G\xi(k) \tag{8}$$

where $l_k=\min\{L,k\}$, L is the given number of paths, $\beta_j(k)$ $(j=0,1,\ldots,l_k)$ are the channel coefficients that are random variables taking values on the interval $[0,1]$ with mathematical expectations $\bar{\beta}_j$ and variances $\tilde{\beta}_j$. $\xi(k)\in l_2([0,\infty),\mathbb{R})$ is an external disturbance and G is a constant matrix. For simplicity, we set $\{y(k)\}_{k\in[-L,-1]}=0$ and $\{v^T(k)\ \xi^T(k)\}_{k\in[-L,-1]}=0$. It is assumed that the random variables $w(k)$, $\alpha_i(k)$, $i=1,2,\ldots,N$ and $\beta_j(k)$, $j=0,1,\ldots,l_k$ are mutually independent in this paper.

D. H_∞ State Estimator

We will investigate the problem of H_∞ state estimation for a class of neural networks with (x,v)-dependent noises and fading channels, where the framework is shown in Fig. 1. For the system (7), we are interested in constructing a full-order estimator of the form:

$$\begin{cases} \hat{x}(k+1) = A_f\hat{x}(k) + B_f\bar{y}(k), \\ \hat{z}(k) = F\hat{x}(k) \end{cases} \tag{9}$$

where $\hat{x}(k)\in\mathbb{R}^n$ is the estimated state, $\hat{z}(k)\in\mathbb{R}^{n_z}$ is the estimated output, A_f and B_f are the estimator gain matrices to be designed.

Setting $\eta(k)=[x^T(k)\ \ \hat{x}^T(k)]^T$, $\zeta(k)=[v^T(k)\ \ \xi^T(k)]^T$ and $\tilde{z}(k)=z(k)-\hat{z}(k)$, the estimation error system connecting the neural network (7) with the estimator (9) is obtained as follows:

$$\begin{cases}\eta(k+1)=&\mathcal{A}\eta(k)+(\beta_0(k)-\bar{\beta}_0)\mathcal{D}\eta(k)+\sum_{j=1}^{l_k}\bar{\beta}_j\mathcal{D}\eta(k-j)+\sum_{j=1}^{l_k}(\beta_j(k)-\bar{\beta}_j)\mathcal{D}\eta(k-j)+\mathcal{W}_1f(\eta(k))\\ &+\sum_{i=1}^{N}\bar{\alpha}_i\mathcal{W}_2g(\eta(k-d_i(k)))+\sum_{i=1}^{N}(\alpha_i(k)-\bar{\alpha}_i)\mathcal{W}_2g(\eta(k-d_i(k)))+\mathcal{C}\zeta(k)+(\beta_0(k)-\bar{\beta}_0)\mathcal{E}\zeta(k)\\ &+\sum_{j=1}^{l_k}\bar{\beta}_j\mathcal{E}\zeta(k-j)+\sum_{j=1}^{l_k}(\beta_j(k)-\bar{\beta}_j)\mathcal{E}\zeta(k-j)+\Big[\bar{\mathcal{A}}\eta(k)+\sum_{i=1}^{N}\bar{\alpha}_i\bar{\mathcal{B}}\eta(k-d_i(k))\\ &+\sum_{i=1}^{N}(\alpha_i(k)-\bar{\alpha}_i)\bar{\mathcal{B}}\eta(k-d_i(k))+\bar{\mathcal{C}}\zeta(k)\Big]w(k)\\ \tilde{z}(k)=&\mathcal{F}\eta(k)\end{cases}\tag{10}$$

where

$$\mathcal{A}=\begin{bmatrix}A&0\\ \bar{\beta}_0B_fD&A_f\end{bmatrix},\quad \mathcal{D}=\begin{bmatrix}0&0\\ B_fD&0\end{bmatrix},\mathcal{W}_1=\begin{bmatrix}W_1&0\\0&0\end{bmatrix},\quad \mathcal{W}_2=\begin{bmatrix}W_2&0\\0&0\end{bmatrix},$$
$$\mathcal{C}=\begin{bmatrix}C&0\\ \bar{\beta}_0B_fE&B_fG\end{bmatrix},\quad \mathcal{E}=\begin{bmatrix}0&0\\ B_fE&0\end{bmatrix},\quad \bar{\mathcal{A}}=\begin{bmatrix}\bar{A}&0\\0&0\end{bmatrix},\quad \bar{\mathcal{B}}=\begin{bmatrix}\bar{B}&0\\0&0\end{bmatrix},\quad \bar{\mathcal{C}}=\begin{bmatrix}\bar{C}&0\\0&0\end{bmatrix},$$
$$\mathcal{F}=[F\ \ -F],\quad f(\eta(k))=\begin{bmatrix}\bar{f}(x(k))\\ \bar{f}(\hat{x}(k))\end{bmatrix},\quad g(\eta(k-d_i(k)))=\begin{bmatrix}\bar{g}(x(k-d_i(k)))\\ \bar{g}(\hat{x}(k-d_i(k)))\end{bmatrix},\ i=1,2,\dots,N.$$

Denote

$$\Gamma=[\bar{\alpha}_1I\ \ \bar{\alpha}_2I\ \ \cdots\ \ \bar{\alpha}_NI],\quad \tilde{\Gamma}=\mathrm{diag}\{\tilde{\alpha}_1,\ \tilde{\alpha}_2,\ \cdots,\ \tilde{\alpha}_N\},\quad \tilde{\alpha}_i=\bar{\alpha}_i(1-\bar{\alpha}_i),\ i=1,2,\dots,N,$$
$$\hat{\Gamma}(k)=[\hat{\alpha}_1(k)I\ \ \hat{\alpha}_2(k)I\ \ \cdots\ \ \hat{\alpha}_N(k)I],\quad \hat{\alpha}_i(k)=\alpha_i(k)-\bar{\alpha}_i,\ i=1,2,\dots,N,$$
$$\Theta=[\underbrace{\bar{\beta}_1I\ \ \bar{\beta}_2I\ \ \cdots\ \ \bar{\beta}_{l_k}I\ \ 0\ \ \cdots\ \ 0}_{L}],\quad \tilde{\Theta}=\mathrm{diag}\{\underbrace{\tilde{\beta}_1,\ \tilde{\beta}_2,\ \cdots,\ \tilde{\beta}_{l_k},\ 0,\ \cdots,\ 0}_{L}\},$$
$$\hat{\Theta}(k)=[\underbrace{\hat{\beta}_1(k)I\ \ \hat{\beta}_2(k)I\ \ \cdots\ \ \hat{\beta}_{l_k}(k)I\ \ 0\ \ \cdots\ \ 0}_{L}],\quad \hat{\beta}_j(k)=\beta_j(k)-\bar{\beta}_j,\ j=0,1,2,\dots,l_k,$$
$$\eta_L(k)=[\eta^T(k-1)\ \ \eta^T(k-2)\ \ \cdots\ \ \eta^T(k-L)]^T,$$
$$\zeta_L(k)=[\zeta^T(k-1)\ \ \zeta^T(k-1)\ \ \cdots\ \ \zeta^T(k-L)]^T,$$
$$\eta_N(k)=\left[\eta^T(k-d_1(k))\ \ \eta^T(k-d_2(k))\ \ \cdots\ \ \eta^T(k-d_N(k))\right]^T,$$
$$g_N(\eta(k))=\left[g^T(\eta(k-d_1(k)))\ \ g^T(\eta(k-d_2(k)))\ \ \cdots\ \ g^T(\eta(k-d_N(k)))\right]^T.$$

Then, the system (10) can be written as

$$\begin{cases}\eta(k+1)=&\mathcal{A}\eta(k)+\hat{\beta}_0(k)\mathcal{D}\eta(k)+\mathcal{D}\Theta\eta_L(k)+\mathcal{D}\hat{\Theta}(k)\eta_L(k)+\mathcal{W}_1f(\eta(k))+\mathcal{W}_2\Gamma g_N(\eta(k))\\ &+\mathcal{W}_2\hat{\Gamma}(k)g_N(\eta(k))+\mathcal{C}\zeta(k)+\hat{\beta}_0(k)\mathcal{E}\zeta(k)+\mathcal{E}\Theta\zeta_L(k)+\mathcal{E}\hat{\Theta}(k)\zeta_L(k)\\ &+\Big[\bar{\mathcal{A}}\eta(k)+\bar{\mathcal{B}}\Gamma\eta_N(k)+\bar{\mathcal{B}}\hat{\Gamma}(k)\eta_N(k)+\bar{\mathcal{C}}\zeta(k)\Big]w(k)\\ \tilde{z}(k)=&\mathcal{F}\eta(k).\end{cases}\tag{11}$$

Definition 1: [33] The zero solution of the estimation error system (11) with $\zeta(k)=0$ is said to be stochastically stable if, for any $\varepsilon>0$, there exists a $\delta>0$ such that

$$\mathbb{E}\{\|\eta(k)\|\}<\varepsilon\tag{12}$$

whenever $k \in \mathbb{N}^+$ and $\max_{j \in \{-d_M^L, -d_M^L+1, \ldots, 0\}} \{\|\bar{\phi}(j)\|\} < \delta$ with $d_M^L = \max\{d_M, L\}$, where $\bar{\phi}(j) = [\phi^T(j) \ 0]^T$ for $j = -d_M^L, -d_M^L + 1, \ldots, 0$.

Lemma 1: [33] If there exists a Lyapunov function $V(\rho(k)) \in \mathcal{C}^1\left(\mathbb{R}^{2(d_M^L+1)n}\right)$ and a function $a(r) \in \mathcal{CK}$ satisfying the following conditions:

$$V(0) = 0, \tag{13}$$

$$a(\|\rho(k)\|) \leq V(\rho(k)), \tag{14}$$

$$\mathbb{E}\{V(\rho(k+1))\} \leq \mathbb{E}\{V(\rho(k))\}, \quad k \in \mathbb{N}^+ \tag{15}$$

where $\rho(k) = \left[\eta^T(k) \ \eta^T(k-1) \ \cdots \ \eta^T(k - d_M^L)\right]^T$, then the zero solution of the system (11) with $\zeta(k) = 0$ is stochastically stable.

Lemma 2: [13] For a scalar α and arbitrary matrices A, B, C, D with appropriate dimensions, the Kronecker product $\otimes$ satisfies

$$\begin{aligned}
&\text{(i)} && \alpha(A \otimes B) = (\alpha A) \otimes B = A \otimes (\alpha B),\\
&\text{(ii)} && (A + B) \otimes C = A \otimes C + B \otimes C,\\
&\text{(iii)} && (A \otimes B)(C \otimes D) = (AC) \otimes (BD),\\
&\text{(iv)} && (A \otimes B)^T = A^T \otimes B^T,\\
&\text{(v)} && (A \otimes B)^{-1} = A^{-1} \otimes B^{-1}.
\end{aligned}$$

Lemma 3: Considering $\hat{\Gamma}(k)$, $\hat{\Theta}(k)$, $\tilde{\Gamma}$ and $\tilde{\Theta}$ in (11), the following equalities hold for any matrix X with appropriate dimension

$$\mathbb{E}\{\hat{\Gamma}^T(k) X \hat{\Gamma}(k)\} = \tilde{\Gamma} \otimes X, \tag{16}$$

$$\mathbb{E}\{\hat{\Theta}^T(k) X \hat{\Theta}(k)\} = \tilde{\Theta} \otimes X. \tag{17}$$

Proof: According to definitions of the matrix $\hat{\Gamma}(k)$ and stochastic variables $\alpha_i(k)$, $i = 1, 2, \ldots, N$, we have

$$\begin{aligned}
\mathbb{E}\{\hat{\Gamma}^T(k) X \hat{\Gamma}(k)\} =& \mathbb{E}\left\{[\hat{\alpha}_1(k)I \ \ \hat{\alpha}_2(k)I \ \ \cdots \ \ \hat{\alpha}_N(k)I]^T X [\hat{\alpha}_1(k)I \ \ \hat{\alpha}_2(k)I \ \ \cdots \ \ \hat{\alpha}_N(k)I]\right\}\\
=& \operatorname{diag}\{\tilde{\alpha}_1 X, \ \ \tilde{\alpha}_2 X, \ \ \cdots, \ \ \tilde{\alpha}_N X\} = \tilde{\Gamma} \otimes X
\end{aligned}$$

which is equivalent to (16). The equality (17) can be proven in a similar way and the details are omitted here. ■

Our aim in this paper is to design an H_∞ estimator of the form (9) such that the following requirements are satisfied simultaneously:

(i) The zero-solution of the estimation error system (11) with $\zeta(k) = 0$ is stochastically stable.

(ii) Under the zero-initial condition, the estimator error $\tilde{z}(k)$ satisfies

$$\sum_{k=0}^{\infty} \mathbb{E}\{\|\tilde{z}(k)\|^2\} < \gamma^2 \sum_{k=0}^{\infty} \mathbb{E}\{\|\tilde{\zeta}(k)\|^2\} \tag{18}$$

for any nonzero $\tilde{\zeta}(k) = \left[\zeta^T(k) \ \zeta_L^T(k)\right]^T \in l_2[0, \infty)$, where $\gamma > 0$ is a given disturbance attenuation level.

Remark 1: In this paper, we consider a general class of delayed stochastic neural networks (1). The time-varying delay is characterized by introducing a sequence of Bernoulli stochastic variable, and the system state and disturbance input are both subject to noises. The neural network model is comprehensive to describe the practical phenomena more precisely. The nonlinear description in (2) is quite general that includes the usual Lipschitz condition as a special case, and it provides a vector-based sector-bounded condition that would facilitate the mathematical analysis on the dynamic behaviors of neural networks.

Remark 2: Due to the shadowing problem and multipath transmission, the L-th order Rice fading model (8) has been widely used in areas of remote control and signal processing. Such a model is capable of accounting for packet dropouts, channel fading and time-delays simultaneously, and may reflect the reality of measurement transmission especially through wireless networks. Moreover, such a signal propagation process will lead to substantial difficulties in subsequent analysis and design.

III. Main Results

In this section, we deal with the H_∞ state estimation problem for the neural network (11) by using the stochastic analysis approach and the Kronecker product.

Theorem 1: Consider the delayed neural network (1) and assume that estimator gains A_f and B_f in (9) are given. The estimation error system (11) is stochastically stable with a prescribed H_∞ performance $\gamma > 0$ if there exist positive definite matrices $\mathcal{P} > 0$, $\mathcal{Q}_j > 0$, $\mathcal{R}_i > 0$ and positive constant scalars $\delta > 0$, $\lambda_i > 0$ ($i = 1, 2, \ldots, N$, $j = 1, 2, \ldots, L$) such that the following LMI holds:

$$\Pi = \begin{bmatrix} \Pi_{11} & \Pi_{12} \\ * & \Pi_{22} \end{bmatrix} < 0 \tag{19}$$

where

$$\Pi_{11} = \begin{bmatrix} \Omega_{11} + \mathcal{F}^T\mathcal{F} & \Omega_{12} & \Omega_{13} & \Omega_{14} & \Omega_{15} \\ * & \Omega_{22} & 0 & \Omega_{24} & \Omega_{25} \\ * & * & \Omega_{33} & 0 & \Omega_{35} \\ * & * & * & \Omega_{44} & \Omega_{45} \\ * & * & * & * & \Omega_{55} \end{bmatrix},$$

$$\Pi_{12} = \begin{bmatrix} \Omega_{16}^T & \Omega_{26}^T & \Omega_{36}^T & \Omega_{46}^T & \Omega_{56}^T \\ \Omega_{17}^T & \Omega_{27}^T & 0 & \Omega_{47}^T & \Omega_{57}^T \end{bmatrix}^T, \quad \Pi_{22} = \begin{bmatrix} \Omega_{66} - \gamma^2 I & \Omega_{67} \\ * & \Omega_{77} - \gamma^2 I \end{bmatrix},$$

with

$$\Omega_{11} = \mathcal{A}^T\mathcal{P}\mathcal{A} - \mathcal{P} + \tilde{\beta}_0\mathcal{D}^T\mathcal{P}\mathcal{D} + \bar{\mathcal{A}}^T\mathcal{P}\bar{\mathcal{A}} + \sum_{j=1}^{L}\mathcal{Q}_j + \sum_{i=1}^{N}\left(d_i^M - d_i^m + 1\right)\mathcal{R}_i - \delta\mathscr{F}_1,$$

$$\Omega_{12} = \mathcal{A}^T\mathcal{P}\mathcal{D}\Theta, \quad \Omega_{13} = \bar{\mathcal{A}}^T\mathcal{P}\bar{\mathcal{B}}\Gamma, \quad \Omega_{14} = \mathcal{A}^T\mathcal{P}\mathcal{W}_1 + \delta\mathscr{F}_2, \quad \Omega_{15} = \mathcal{A}^T\mathcal{P}\mathcal{W}_2\Gamma,$$

$$\Omega_{16} = \mathcal{A}^T\mathcal{P}\mathcal{C} + \tilde{\beta}_0\mathcal{D}^T\mathcal{P}\mathcal{E} + \bar{\mathcal{A}}^T\mathcal{P}\bar{\mathcal{C}}, \quad \Omega_{17} = \mathcal{A}^T\mathcal{P}\mathcal{E}\Theta,$$

$$\Omega_{22} = \Theta^T\mathcal{D}^T\mathcal{P}\mathcal{D}\Theta + \tilde{\Theta} \otimes (\mathcal{D}^T\mathcal{P}\mathcal{D}) - \mathrm{diag}\{\mathcal{Q}_1, \mathcal{Q}_2, \cdots, \mathcal{Q}_L\}, \quad \Omega_{24} = \Theta^T\mathcal{D}^T\mathcal{P}\mathcal{W}_1,$$

$$\Omega_{25} = \Theta^T\mathcal{D}^T\mathcal{P}\mathcal{W}_2\Gamma, \quad \Omega_{26} = \Theta^T\mathcal{D}^T\mathcal{P}\mathcal{C}, \quad \Omega_{27} = \Theta^T\mathcal{D}^T\mathcal{P}\mathcal{E}\Theta + \tilde{\Theta} \otimes (\mathcal{D}^T\mathcal{P}\mathcal{E}),$$

$$\Omega_{33} = \Gamma^T\bar{\mathcal{B}}^T\mathcal{P}\bar{\mathcal{B}}\Gamma + \tilde{\Gamma} \otimes (\bar{\mathcal{B}}^T\mathcal{P}\bar{\mathcal{B}}) - \mathrm{diag}\{\mathcal{R}_1, \mathcal{R}_2, \cdots, \mathcal{R}_N\} - \Lambda \otimes \mathscr{G}_1, \ \Omega_{35} = \Lambda \otimes \mathscr{G}_2, \ \Omega_{36} = \Gamma^T\bar{\mathcal{B}}^T\mathcal{P}\bar{\mathcal{C}},$$

$$\Omega_{44} = \mathcal{W}_1^T\mathcal{P}\mathcal{W}_1 - \delta I, \quad \Omega_{45} = \mathcal{W}_1^T\mathcal{P}\mathcal{W}_2\Gamma, \quad \Omega_{46} = \mathcal{W}_1^T\mathcal{P}\mathcal{C}, \quad \Omega_{47} = \mathcal{W}_1^T\mathcal{P}\mathcal{E}\Theta,$$

$$\Omega_{55} = \Gamma^T\mathcal{W}_2^T\mathcal{P}\mathcal{W}_2\Gamma + \tilde{\Gamma} \otimes (\mathcal{W}_2^T\mathcal{P}\mathcal{W}_2) - \Lambda \otimes I, \quad \Omega_{56} = \Gamma^T\mathcal{W}_2^T\mathcal{P}\mathcal{C}, \quad \Omega_{57} = \Gamma^T\mathcal{W}_2^T\mathcal{P}\mathcal{E}\Theta,$$

$$\Omega_{66} = \mathcal{C}^T\mathcal{P}\mathcal{C} + \tilde{\beta}_0\mathcal{E}^T\mathcal{P}\mathcal{E} + \bar{\mathcal{C}}^T\mathcal{P}\bar{\mathcal{C}}, \quad \Omega_{67} = \mathcal{C}^T\mathcal{P}\mathcal{E}\Theta, \quad \Omega_{77} = \Theta^T\mathcal{E}^T\mathcal{P}\mathcal{E}\Theta + \tilde{\Theta} \otimes (\mathcal{E}^T\mathcal{P}\mathcal{E}),$$

$$\mathscr{F}_1 = I \otimes \frac{\Phi_f^T\Psi_f + \Psi_f^T\Phi_f}{2}, \quad \mathscr{F}_2 = \frac{I \otimes (\Phi_f + \Psi_f)^T}{2}, \quad \Lambda = \mathrm{diag}\{\lambda_1, \lambda_2, \cdots, \lambda_N\},$$

$$\mathscr{G}_1 = I \otimes \frac{\Phi_g^T\Psi_g + \Psi_g^T\Phi_g}{2}, \quad \mathscr{G}_2 = \frac{I \otimes (\Phi_g + \Psi_g)^T}{2}.$$

Proof: Choose the Lyapunov functional $V(\rho(k))$ as

$$V(\rho(k)) = V_1(k) + V_2(k) + \sum_{i=1}^{N}\tilde{V}_i(k), \tag{20}$$

where $\rho(k)$ is defined in Lemma 1 and

$$V_1(k) = \eta^T(k)\mathcal{P}\eta(k), \tag{21}$$

$$V_2(k) = \sum_{j=1}^{L}\sum_{s=k-j}^{k-1} \eta^T(s)\mathcal{Q}_j\eta(s), \tag{22}$$

$$\tilde{V}_i(k) = \sum_{j=k-d_i(k)}^{k-1} \eta^T(j)\mathcal{R}_i\eta(j) + \sum_{j=d_i^m}^{d_i^M-1}\sum_{s=k-j}^{k-1} \eta^T(s)\mathcal{R}_i\eta(s), \quad i = 1,2,\dots,N. \tag{23}$$

Obviously, the conditions (13) and (14) in Lemma 1 are satisfied when $V(\rho(k))$ is chosen as (20). Along the trajectory of the system (11), we calculate the expectation of the difference of $V(\rho(k))$ and have

$$\begin{aligned}
&\mathbb{E}\{V(\rho(k+1)) - V(\rho(k))\} \\
=&\mathbb{E}\left\{V_1(k+1) + V_2(k+1) + \sum_{i=1}^{N}\tilde{V}_i(k+1) - V_1(k) - V_2(k) - \sum_{i=1}^{N}\tilde{V}_i(k)\right\} \\
=&\mathbb{E}\left\{\Delta V_1(k) + \Delta V_2(k) + \sum_{i=1}^{N}\Delta\tilde{V}_i(k)\right\}.
\end{aligned} \tag{24}$$

Considering Lemma 3, one has

$$\begin{aligned}
&\mathbb{E}\{\Delta V_1(k)\} = \mathbb{E}\{V_1(k+1) - V_1(k)\} = \mathbb{E}\{\eta^T(k+1)\mathcal{P}\eta(k+1) - \eta^T(k)\mathcal{P}\eta(k)\} \\
=&\mathbb{E}\Big\{\Big[\mathcal{A}\eta(k) + \hat{\beta}_0(k)\mathcal{D}\eta(k) + \mathcal{D}\Theta\eta_L(k) + \mathcal{D}\hat{\Theta}(k)\eta_L(k) + \mathcal{W}_1 f(\eta(k)) + \mathcal{W}_2\Gamma g_N(\eta(k)) \\
&+ \mathcal{W}_2\hat{\Gamma}(k)g_N(\eta(k)) + \mathcal{C}\zeta(k) + \hat{\beta}_0(k)\mathcal{E}\zeta(k) + \mathcal{E}\Theta\zeta_L(k) + \mathcal{E}\hat{\Theta}(k)\zeta_L(k)\Big]^T\mathcal{P}\Big[\mathcal{A}\eta(k) + \hat{\beta}_0(k)\mathcal{D}\eta(k) \\
&+ \mathcal{D}\Theta\eta_L(k) + \mathcal{D}\hat{\Theta}(k)\eta_L(k) + \mathcal{W}_1 f(\eta(k)) + \mathcal{W}_2\Gamma g_N(\eta(k)) + \mathcal{W}_2\hat{\Gamma}(k)g_N(\eta(k)) + \mathcal{C}\zeta(k) \\
&+ \hat{\beta}_0(k)\mathcal{E}\zeta(k) + \mathcal{E}\Theta\zeta_L(k) + \mathcal{E}\hat{\Theta}(k)\zeta_L(k)\Big] + \Big[\bar{\mathcal{A}}\eta(k) + \bar{\mathcal{B}}\Gamma\eta_N(k) + \bar{\mathcal{B}}\hat{\Gamma}(k)\eta_N(k) + \bar{\mathcal{C}}\zeta(k)\Big]^T\mathcal{P} \\
&\times\Big[\bar{\mathcal{A}}\eta(k) + \bar{\mathcal{B}}\Gamma\eta_N(k) + \bar{\mathcal{B}}\hat{\Gamma}(k)\eta_N(k) + \bar{\mathcal{C}}\zeta(k)\Big] - \eta^T(k)\mathcal{P}\eta(k)\Big\} \\
=&\mathbb{E}\Big\{\eta^T(k)\mathcal{A}^T\mathcal{P}\mathcal{A}\eta(k) - \eta^T(k)\mathcal{P}\eta(k) + 2\eta^T(k)\mathcal{A}^T\mathcal{P}\mathcal{D}\Theta\eta_L(k) + 2\eta^T(k)\mathcal{A}^T\mathcal{P}\mathcal{W}_1 f(\eta(k)) \\
&+ 2\eta^T(k)\mathcal{A}^T\mathcal{P}\mathcal{W}_2\Gamma g_N(\eta(k)) + 2\eta^T(k)\mathcal{A}^T\mathcal{P}\mathcal{C}\zeta(k) + 2\eta^T(k)\mathcal{A}^T\mathcal{P}\mathcal{E}\Theta\zeta_L(k) \\
&+ \tilde{\beta}_0\eta^T(k)\mathcal{D}^T\mathcal{P}\mathcal{D}\eta(k) + 2\tilde{\beta}_0\eta^T(k)\mathcal{D}^T\mathcal{P}\mathcal{E}\zeta(k) + \eta_L^T(k)\Theta^T\mathcal{D}^T\mathcal{P}\mathcal{D}\Theta\eta_L(k) \\
&+ 2\eta_L^T(k)\Theta^T\mathcal{D}^T\mathcal{P}\mathcal{W}_1 f(\eta(k)) + 2\eta_L^T(k)\Theta^T\mathcal{D}^T\mathcal{P}\mathcal{W}_2\Gamma g_N(\eta(k)) + 2\eta_L^T(k)\Theta^T\mathcal{D}^T\mathcal{P}\mathcal{C}\zeta(k) \\
&+ 2\eta_L^T(k)\Theta^T\mathcal{D}^T\mathcal{P}\mathcal{E}\Theta\zeta_L(k) + \eta_L^T(k)\left[\tilde{\Theta}\otimes(\mathcal{D}^T\mathcal{P}\mathcal{D})\right]\eta_L(k) + 2\eta_L^T(k)\left[\tilde{\Theta}\otimes(\mathcal{D}^T\mathcal{P}\mathcal{E})\right]\zeta_L(k) \\
&+ f^T(\eta(k))\mathcal{W}_1^T\mathcal{P}\mathcal{W}_1 f(\eta(k)) + 2f^T(\eta(k))\mathcal{W}_1^T\mathcal{P}\mathcal{W}_2\Gamma g_N(\eta(k)) + 2f^T(\eta(k))\mathcal{W}_1^T\mathcal{P}\mathcal{C}\zeta(k) \\
&+ 2f^T(\eta(k))\mathcal{W}_1^T\mathcal{P}\mathcal{E}\Theta\zeta_L(k) + g_N^T(\eta(k))\Gamma^T\mathcal{W}_2^T\mathcal{P}\mathcal{W}_2\Gamma g_N(\eta(k)) + 2g_N^T(\eta(k))\Gamma^T\mathcal{W}_2^T\mathcal{P}\mathcal{C}\zeta(k) \\
&+ 2g_N^T(\eta(k))\Gamma^T\mathcal{W}_2^T\mathcal{P}\mathcal{E}\Theta\zeta_L(k) + g_N^T(\eta(k))\left[\tilde{\Gamma}\otimes(\mathcal{W}_2^T\mathcal{P}\mathcal{W}_2)\right]g_N(\eta(k)) \\
&+ \zeta^T(k)\mathcal{C}^T\mathcal{P}\mathcal{C}\zeta(k) + 2\zeta^T(k)\mathcal{C}^T\mathcal{P}\mathcal{E}\Theta\zeta_L(k) + \tilde{\beta}_0\zeta^T(k)\mathcal{E}^T\mathcal{P}\mathcal{E}\zeta(k) + \zeta_L^T(k)\Theta^T\mathcal{E}^T\mathcal{P}\mathcal{E}\Theta\zeta_L(k) \\
&+ \zeta_L^T(k)\left[\tilde{\Theta}\otimes(\mathcal{E}^T\mathcal{P}\mathcal{E})\right]\zeta_L(k) + \eta^T(k)\bar{\mathcal{A}}^T\mathcal{P}\bar{\mathcal{A}}\eta(k) + 2\eta^T(k)\bar{\mathcal{A}}^T\mathcal{P}\bar{\mathcal{B}}\Gamma\eta_N(k) + 2\eta^T(k)\bar{\mathcal{A}}^T\mathcal{P}\bar{\mathcal{C}}\zeta(k) \\
&+ \eta_N^T(k)\Gamma^T\bar{\mathcal{B}}^T\mathcal{P}\bar{\mathcal{B}}\Gamma\eta_N(k) + 2\eta_N^T(k)\Gamma^T\bar{\mathcal{B}}^T\mathcal{P}\bar{\mathcal{C}}\zeta(k) + \eta_N^T(k)\left[\tilde{\Gamma}\otimes(\bar{\mathcal{B}}^T\mathcal{P}\bar{\mathcal{B}})\right]\eta_N(k) + \zeta^T(k)\bar{\mathcal{C}}^T\mathcal{P}\bar{\mathcal{C}}\zeta(k)\Big\}.
\end{aligned} \tag{25}$$

On the other hand, it is not difficult to see that

$$\mathbb{E}\{\Delta V_2(k)\} =\mathbb{E}\{V_2(k+1) - V_2(k)\}$$

$$
\begin{aligned}
&=\mathbb{E}\left\{\sum_{j=1}^{L}\left(\sum_{s=k-j+1}^{k}\eta^T(s)\mathcal{Q}_j\eta(s)-\sum_{s=k-j}^{k-1}\eta^T(s)\mathcal{Q}_j\eta(s)\right)\right\}\\
&=\sum_{j=1}^{L}\mathbb{E}\left\{\eta^T(k)\mathcal{Q}_j\eta(k)-\eta^T(k-j)\mathcal{Q}_j\eta(k-j)\right\}\\
&=\sum_{j=1}^{L}\mathbb{E}\left\{\eta^T(k)\mathcal{Q}_j\eta(k)\right\}-\mathbb{E}\Big\{\eta_L^T(k)\mathrm{diag}\{\mathcal{Q}_1,\mathcal{Q}_2,\cdots,\mathcal{Q}_L\}\eta_L(k)\Big\}. \qquad (26)
\end{aligned}
$$

Moreover, we can show that

$$
\begin{aligned}
\mathbb{E}\{\Delta\tilde{V}_i(k)\} =&\mathbb{E}\{\tilde{V}_i(k+1)-\tilde{V}_i(k)\}\\
=&\mathbb{E}\Bigg\{\sum_{j=k-d_i(k+1)+1}^{k}\eta^T(j)\mathcal{R}_i\eta(j)-\sum_{j=k-d_i(k)}^{k-1}\eta^T(j)\mathcal{R}_i\eta(j)\\
&+\sum_{j=d_i^m}^{d_i^M-1}\left(\sum_{s=k-j+1}^{k}\eta^T(s)\mathcal{R}_i\eta(s)-\sum_{s=k-j}^{k-1}\eta^T(s)\mathcal{R}_i\eta(s)\right)\Bigg\}\\
=&\mathbb{E}\Bigg\{\left(d_i^M-d_i^m+1\right)\eta^T(k)\mathcal{R}_i\eta(k)-\eta^T(k-d_i(k))\mathcal{R}_i\eta(k-d_i(k))\\
&+\left(\sum_{j=k-d_i(k+1)+1}^{k-d_i^m}+\sum_{j=k-d_i^m+1}^{k-1}-\sum_{j=k-d_i(k)+1}^{k-1}-\sum_{j=k-d_i^M+1}^{k-d_i^m}\right)\eta^T(j)\mathcal{R}_i\eta(j)\Bigg\}\\
\le&\mathbb{E}\Big\{\left(d_i^M-d_i^m+1\right)\eta^T(k)\mathcal{R}_i\eta(k)-\eta^T(k-d_i(k))\mathcal{R}_i\eta(k-d_i(k))\Big\},\quad i=1,2,\ldots,N. \qquad (27)
\end{aligned}
$$

Therefore, we have

$$
\begin{aligned}
\mathbb{E}\left\{\sum_{i=1}^{N}\Delta\tilde{V}_i(k)\right\}\le&\sum_{i=1}^{N}\mathbb{E}\Big\{\left[\left(d_i^M-d_i^m+1\right)\eta^T(k)\mathcal{R}_i\eta(k)-\eta^T(k-d_i(k))\mathcal{R}_i\eta(k-d_i(k))\right]\Big\}\\
=&\sum_{i=1}^{N}\mathbb{E}\left\{\left(d_i^M-d_i^m+1\right)\eta^T(k)\mathcal{R}_i\eta(k)\right\}-\mathbb{E}\Big\{\eta_N^T(k)\mathrm{diag}\{\mathcal{R}_1,\mathcal{R}_2,\cdots,\mathcal{R}_N\}\eta_N(k)\Big\}. \qquad (28)
\end{aligned}
$$

Notice that (2) implies

$$
\begin{aligned}
&[f(\eta(k))-(I\otimes\Phi_f)\eta(k)]^T[f(\eta(k))-(I\otimes\Psi_f)\eta(k)]\le 0,\\
&[g(\eta(k-d_i(k)))-(I\otimes\Phi_g)\eta(k-d_i(k))]^T[g(\eta(k-d_i(k)))-(I\otimes\Psi_g)\eta(k-d_i(k))]\le 0,\quad i=1,2,\ldots,N.
\end{aligned}
$$

There exist scalars $\delta>0$ and $\lambda_i>0$ ($i=1,2,\ldots,N$) such that

$$
\delta[f(\eta(k))-(I\otimes\Phi_f)\eta(k)]^T[f(\eta(k))-(I\otimes\Psi_f)\eta(k)]\le 0, \qquad (29)
$$

$$
\sum_{i=1}^{N}\lambda_i[g(\eta(k-d_i(k)))-(I\otimes\Phi_g)\eta(k-d_i(k))]^T[g(\eta(k-d_i(k)))-(I\otimes\Psi_g)\eta(k-d_i(k))]\le 0. \qquad (30)
$$

By Lemma 2, the inequality (30) can be written as

$$
\begin{aligned}
&g_N^T(\eta(k))(\Lambda\otimes I)g_N(\eta(k))+\eta_N^T(k)\left(\Lambda\otimes\left(I\otimes\frac{\Phi_g^T\Psi_g+\Psi_g^T\Phi_g}{2}\right)\right)\eta_N(k)\\
&-\eta_N^T(k)\left(\Lambda\otimes\left(I\otimes(\Phi_g+\Psi_g)^T\right)\right)g_N(\eta(k))\le 0 \qquad (31)
\end{aligned}
$$

where $\Lambda=\mathrm{diag}\{\lambda_1,\ \lambda_2,\ \cdots\ \lambda_N\}$.

Denote $\tilde{\eta}(k) = \begin{bmatrix}\eta^T(k) & \eta_L^T(k) & \eta_N^T(k) & f^T(\eta(k)) & g_N^T(\eta(k)) & \zeta^T(k) & \zeta_L^T(k)\end{bmatrix}^T$. According to (25), (26), (28) and considering (29), (31), we have

$$\begin{aligned}
&\mathbb{E}\{V(\rho(k+1)) - V(\rho(k))\} \\
\leq &\mathbb{E}\left\{\Delta V_1(k) + \Delta V_2(k) + \sum_{i=1}^{N} \Delta \tilde{V}_i(k)\right\} - \delta[f(\eta(k)) - (I \otimes \Phi_f)\eta(k)]^T[f(\eta(k)) - (I \otimes \Psi_f)\eta(k)] \\
&- \sum_{i=1}^{N} \lambda_i[g(\eta(k - d_i(k))) - (I \otimes \Phi_g)\eta(k - d_i(k))]^T[g(\eta(k - d_i(k))) - (I \otimes \Psi_g)\eta(k - d_i(k))] \\
= &\mathbb{E}\{\tilde{\eta}^T(k)\Omega_1\tilde{\eta}(k)\} \qquad (32)
\end{aligned}$$

where

$$\Omega_1 = \begin{bmatrix}
\Omega_{11} & \Omega_{12} & \Omega_{13} & \Omega_{14} & \Omega_{15} & \Omega_{16} & \Omega_{17} \\
* & \Omega_{22} & 0 & \Omega_{24} & \Omega_{25} & \Omega_{26} & \Omega_{27} \\
* & * & \Omega_{33} & 0 & \Omega_{35} & \Omega_{36} & 0 \\
* & * & * & \Omega_{44} & \Omega_{45} & \Omega_{46} & \Omega_{47} \\
* & * & * & * & \Omega_{55} & \Omega_{56} & \Omega_{57} \\
* & * & * & * & * & \Omega_{66} & \Omega_{67} \\
* & * & * & * & * & * & \Omega_{77}
\end{bmatrix}.$$

Now, we first prove the stochastic stability of the estimation error system (11) with $\zeta(k) = 0$. From (32), one can easily obtain that

$$\mathbb{E}\{V(\rho(k+1)) - V(\rho(k))\}|_{\zeta(k)=0} \leq \mathbb{E}\{\bar{\eta}^T(k)\bar{\Omega}_1\bar{\eta}(k)\}$$

where

$$\bar{\eta}(k) = \begin{bmatrix}\eta^T(k) & \eta_L^T(k) & \eta_N^T(k) & f^T(\eta(k)) & g_N^T(\eta(k))\end{bmatrix}^T,$$

$$\bar{\Omega}_1 = \begin{bmatrix}
\Omega_{11} & \Omega_{12} & \Omega_{13} & \Omega_{14} & \Omega_{15} \\
* & \Omega_{22} & 0 & \Omega_{24} & \Omega_{25} \\
* & * & \Omega_{33} & 0 & \Omega_{35} \\
* & * & * & \Omega_{44} & \Omega_{45} \\
* & * & * & * & \Omega_{55}
\end{bmatrix}.$$

According to (19), it is easy to see that $\Pi_{11} < 0$ which implies $\bar{\Omega}_1 < 0$. By Lemma 1, the system (11) with $\zeta(k) = 0$ is stochastically stable.

Next, let us show that the estimation error system (11) satisfies the H_∞ performance for all nonzero exogenous disturbances under the zero-initial condition. Adding the zero term

$$\tilde{z}^T(k)\tilde{z}(k) - \gamma^2\tilde{\zeta}^T(k)\tilde{\zeta}(k) - [\tilde{z}^T(k)\tilde{z}(k) - \gamma^2\tilde{\zeta}^T(k)\tilde{\zeta}(k)]$$

to (32) results in

$$\begin{aligned}
&\mathbb{E}\{V(\rho(k+1)) - V(\rho(k))\} \\
\leq &\mathbb{E}\left\{\tilde{\eta}^T(k)\Omega_1\tilde{\eta}(k) + \eta^T(k)\mathcal{F}^T\mathcal{F}\eta(k) - \gamma^2\zeta^T(k)\zeta(k) - \gamma^2\zeta_L^T(k)\zeta_L(k) - [\tilde{z}^T(k)\tilde{z}(k) - \gamma^2\tilde{\zeta}^T(k)\tilde{\zeta}(k)]\right\}. \qquad (33)
\end{aligned}$$

Summing up (33) on both sides from 0 to n with respect to k, one gets

$$V(\rho(n+1)) - V(\rho(0))$$

$$\leq \mathbb{E}\sum_{k=0}^{n}\left\{\tilde{\eta}^T(k)\Pi\tilde{\eta}(k)\right\}-\mathbb{E}\sum_{k=0}^{n}\left\{\tilde{z}^T(k)\tilde{z}(k)-\gamma^2\tilde{\zeta}^T(k)\tilde{\zeta}(k)\right\}, \tag{34}$$

where Π is defined in (19). Letting $n\to\infty$ and considering the zero-initial condition, it can be obtained from (19) and (34) that

$$\mathbb{E}\sum_{k=0}^{n}\left\{\tilde{z}^T(k)\tilde{z}(k)-\gamma^2\tilde{\zeta}^T(k)\tilde{\zeta}(k)\right\}\leq \mathbb{E}\sum_{k=0}^{n}\left\{\tilde{\eta}^T(k)\Pi\tilde{\eta}(k)\right\}<0,$$

which is equivalent to (18), and the proof is now complete. ■

Having derived the analysis results, we are now ready to solve the state estimator design problem for the neural network (11) in the following theorem.

Theorem 2: Consider the delayed neural network (1) and the disturbance attenuation level $\gamma>0$. The addressed H_∞ state estimator design problem is solvable if there exist positive definite matrices $\mathcal{P}>0$, $\mathcal{Q}_j>0$, $\mathcal{R}_i>0$, positive constant scalars $\delta>0$, $\lambda_i>0$ ($i=1,2,\ldots,N$, $j=1,2,\ldots,L$) and a matrix $\mathcal{X}$ such that the following LMI holds:

$$\Xi=\begin{bmatrix} -\mathcal{P} & 0 & 0 & 0 & 0 & 0 & \Upsilon_1 \\ * & -\mathcal{P} & 0 & 0 & 0 & 0 & \Upsilon_2 \\ * & * & -I\otimes\mathcal{P} & 0 & 0 & 0 & \Upsilon_3 \\ * & * & * & -\mathcal{P} & 0 & 0 & \Upsilon_4 \\ * & * & * & * & -I\otimes\mathcal{P} & 0 & \Upsilon_5 \\ * & * & * & * & * & -I\otimes\mathcal{P} & \Upsilon_6 \\ * & * & * & * & * & * & \Sigma \end{bmatrix}<0 \tag{35}$$

where

$$\begin{aligned}
\Upsilon_1&=\begin{bmatrix} \mathcal{P}\hat{\mathcal{A}}+\mathcal{X}\mathcal{A}_1 & \mathcal{X}\mathcal{D}_1\Theta & 0 & \mathcal{P}\mathcal{W}_1 & \mathcal{P}\mathcal{W}_2\Gamma & \mathcal{P}\hat{\mathcal{C}}+\mathcal{X}\mathcal{C}_1 & \mathcal{X}\mathcal{E}_1\Theta \end{bmatrix},\\
\Upsilon_2&=\begin{bmatrix} \sqrt{\tilde{\beta}_0}\mathcal{X}\mathcal{D}_1 & 0 & 0 & 0 & 0 & \sqrt{\tilde{\beta}_0}\mathcal{X}\mathcal{E}_1 & 0 \end{bmatrix},\\
\Upsilon_3&=\begin{bmatrix} 0 & \breve{\Theta}\otimes\mathcal{X}\mathcal{D}_1 & 0 & 0 & 0 & 0 & \breve{\Theta}\otimes\mathcal{X}\mathcal{E}_1 \end{bmatrix},\quad \Upsilon_4=\begin{bmatrix} \mathcal{P}\bar{\mathcal{A}} & 0 & \mathcal{P}\bar{\mathcal{B}}\Gamma & 0 & 0 & \mathcal{P}\bar{\mathcal{C}} & 0 \end{bmatrix},\\
\Upsilon_5&=\begin{bmatrix} 0 & 0 & \breve{\Gamma}\otimes(\mathcal{P}\bar{\mathcal{B}}) & 0 & 0 & 0 & 0 \end{bmatrix},\quad \Upsilon_6=\begin{bmatrix} 0 & 0 & 0 & 0 & \breve{\Gamma}\otimes(\mathcal{P}\mathcal{W}_2) & 0 & 0 \end{bmatrix},\\
\Sigma&=\begin{bmatrix} \Sigma_{11} & 0 & 0 & \delta\mathscr{F}_2 & 0 & 0 & 0 \\ * & \Sigma_{22} & 0 & 0 & 0 & 0 & 0 \\ * & * & \Sigma_{33} & 0 & \Lambda\otimes\mathscr{G}_2 & 0 & 0 \\ * & * & * & \Sigma_{44} & 0 & 0 & 0 \\ * & * & * & * & \Sigma_{55} & 0 & 0 \\ * & * & * & * & * & \Sigma_{66} & 0 \\ * & * & * & * & * & * & \Sigma_{77} \end{bmatrix},
\end{aligned}$$

with

$$\hat{\mathcal{A}}=\begin{bmatrix} A & 0 \\ 0 & 0 \end{bmatrix},\quad \mathcal{A}_1=\begin{bmatrix} 0 & I \\ \bar{\beta}_0 D & 0 \end{bmatrix},\quad \mathcal{D}_1=\begin{bmatrix} 0 & 0 \\ D & 0 \end{bmatrix},\quad \hat{\mathcal{C}}=\begin{bmatrix} C & 0 \\ 0 & 0 \end{bmatrix},\quad \mathcal{C}_1=\begin{bmatrix} 0 & 0 \\ \bar{\beta}_0 E & G \end{bmatrix},$$

$$\mathcal{E}_1=\begin{bmatrix} 0 & 0 \\ E & 0 \end{bmatrix},\quad \breve{\Gamma}=\mathrm{diag}\left\{\sqrt{\tilde{\alpha}_1},\ \sqrt{\tilde{\alpha}_2},\ \cdots,\ \sqrt{\tilde{\alpha}_N}\right\},\quad \breve{\Theta}=\mathrm{diag}\Big\{\underbrace{\sqrt{\tilde{\beta}_1},\ \sqrt{\tilde{\beta}_2},\ \cdots,\ \sqrt{\tilde{\beta}_{l_k}},\ 0,\ \cdots,\ 0}_{L}\Big\}$$

$$\Sigma_{11}=-\mathcal{P}+\sum_{j=1}^{L}\mathcal{Q}_j+\sum_{i=1}^{N}\left(d_i^M-d_i^m+1\right)\mathcal{R}_i-\delta\mathscr{F}_1+\mathcal{F}^T\mathcal{F},\quad \Sigma_{22}=-\mathrm{diag}\{\mathcal{Q}_1,\mathcal{Q}_2,\cdots,\mathcal{Q}_L\},$$

$$\Sigma_{33}=-\mathrm{diag}\{\mathcal{R}_1,\mathcal{R}_2,\cdots,\mathcal{R}_N\}-\Lambda\otimes\mathscr{G}_1,\quad \Sigma_{44}=-\delta I,\quad \Sigma_{55}=-\Lambda\otimes I,\quad \Sigma_{66}=-\gamma^2 I,\quad \Sigma_{77}=-\gamma^2 I.$$

Moreover, if the aforementioned inequality is feasible, the desired estimator gains in (9) can be determined by

$$\begin{bmatrix} A_f & B_f \end{bmatrix} = \left(\mathcal{I}^T \mathcal{P} \mathcal{I}\right)^{-1} \mathcal{I}^T \mathcal{X} \tag{36}$$

where $\mathcal{I} = \begin{bmatrix} 0 & I \end{bmatrix}^T$.

Proof: Denoting

$$\begin{aligned}
\bar{\Upsilon}_1 &= \left[\begin{array}{ccccccc} \mathcal{A} & \mathcal{D}\Theta & 0 & \mathcal{W}_1 & \mathcal{W}_2\Gamma & \mathcal{C} & \mathcal{E}\Theta \end{array}\right], \quad \bar{\Upsilon}_2 = \left[\begin{array}{ccccccc} \sqrt{\tilde{\beta}_0}\mathcal{D} & 0 & 0 & 0 & 0 & \sqrt{\tilde{\beta}_0}\mathcal{E} & 0 \end{array}\right], \\
\bar{\Upsilon}_3 &= \left[\begin{array}{ccccccc} 0 & \breve{\Theta}\otimes\mathcal{D} & 0 & 0 & 0 & 0 & \breve{\Theta}\otimes\mathcal{E} \end{array}\right], \quad \bar{\Upsilon}_4 = \left[\begin{array}{ccccccc} \bar{\mathcal{A}} & 0 & \bar{\mathcal{B}}\Gamma & 0 & 0 & \bar{\mathcal{C}} & 0 \end{array}\right], \\
\bar{\Upsilon}_5 &= \left[\begin{array}{ccccccc} 0 & 0 & \breve{\Gamma}\otimes\bar{\mathcal{B}} & 0 & 0 & 0 & 0 \end{array}\right], \quad \bar{\Upsilon}_6 = \left[\begin{array}{ccccccc} 0 & 0 & 0 & 0 & \breve{\Gamma}\otimes\mathcal{W}_2 & 0 & 0 \end{array}\right],
\end{aligned}$$

the LMI (19) can be rewritten as follows

$$\Pi = \Sigma + \bar{\Upsilon}_1^T \mathcal{P} \bar{\Upsilon}_1 + \bar{\Upsilon}_2^T \mathcal{P} \bar{\Upsilon}_2 + \bar{\Upsilon}_3^T (I\otimes\mathcal{P}) \bar{\Upsilon}_3 + \bar{\Upsilon}_4^T \mathcal{P} \bar{\Upsilon}_4 + \bar{\Upsilon}_5^T (I\otimes\mathcal{P}) \bar{\Upsilon}_5 + \bar{\Upsilon}_6^T (I\otimes\mathcal{P}) \bar{\Upsilon}_6 < 0. \tag{37}$$

By applying the Schur Complement Lemma and Lemma 2, we know that the inequality (37) is equivalent to

$$\begin{bmatrix}
-\mathcal{P}^{-1} & 0 & 0 & 0 & 0 & 0 & \bar{\Upsilon}_1 \\
* & -\mathcal{P}^{-1} & 0 & 0 & 0 & 0 & \bar{\Upsilon}_2 \\
* & * & -I\otimes\mathcal{P}^{-1} & 0 & 0 & 0 & \bar{\Upsilon}_3 \\
* & * & * & -\mathcal{P}^{-1} & 0 & 0 & \bar{\Upsilon}_4 \\
* & * & * & * & -I\otimes\mathcal{P}^{-1} & 0 & \bar{\Upsilon}_5 \\
* & * & * & * & * & -I\otimes\mathcal{P}^{-1} & \bar{\Upsilon}_6 \\
* & * & * & * & * & * & \Sigma
\end{bmatrix} < 0. \tag{38}$$

In order to avoid partitioning the positive definite matrix $\mathcal{P}$, the parameters in (38) are rewritten in the following form:

$$\mathcal{A} = \hat{\mathcal{A}} + \mathcal{I}\mathcal{K}\mathcal{A}_1, \quad \mathcal{D} = \mathcal{I}\mathcal{K}\mathcal{D}_1, \quad \mathcal{C} = \hat{\mathcal{C}} + \mathcal{I}\mathcal{K}\mathcal{C}_1, \quad \mathcal{E} = \mathcal{I}\mathcal{K}\mathcal{E}_1,$$

where $\mathcal{K} = [A_f \ \ B_f]$. Pre- and post-multiplying the inequality (38) by

$$\mathrm{diag}\{\mathcal{P},\ \ \mathcal{P},\ \ I\otimes\mathcal{P},\ \ \mathcal{P},\ \ I\otimes\mathcal{P},\ \ I\otimes\mathcal{P},\ \ I\}$$

and setting $\mathcal{X} = \mathcal{P}\mathcal{I}\mathcal{K}$, it is easily known that the inequality (38) is equivalent to the inequality (35). Furthermore, the estimator gains can be derived by (36), which completes the proof. ■

Remark 3: Different from the existing results for state estimation of neural networks [16], [42], several delay-distribution-dependent conditions are derived in this paper, i.e., the conditions are dependent on both the probability distribution and the variation range of the time delay. It is well known that the more information of the time delay is employed, the less conservative results may be derived. On the other hand, as can be seen from our main results, the larger upper bounds d_i^M and the larger variation d_i^M-d_i^m would have more side effects to the feasibility of the obtained LMI conditions.

Remark 4: Assume that $\mathscr{M}$ denotes the row size of the LMI, $\mathscr{N}$ represents the number of decision variables and $\mathscr{V}$ stands for a scaling factor, then the number of flops needed to calculate an ε-accurate solution to the LMI is bounded by $\mathcal{O}(\mathscr{M}\mathscr{N}^3\log(\mathscr{V}/\varepsilon))$. For the neural network (1) with the measurement (8), the variable dimensions are as follows: $x(k)\in\mathbb{R}^n$, $y(k)\in\mathbb{R}^{n_y}$, $v(k)\in\mathbb{R}^{n_v}$ and $\xi(k)\in\mathbb{R}$. For the LMI condition in Theorem 2, we have $\mathscr{M} = (4L+8N+10)n + (L+1)(n_v+1)$ and $\mathscr{N} = (L+N+1)n(2n+1) + 2n(n+n_y) + N + 1$. Hence, the computational complexity of the derived LMI condition can be represented as $\mathcal{O}((L+N)^4 n^7)$. It is obvious that the computational complexity depends polynomially on parameters L, N and the variable dimension n.

Remark 5: In Theorem 2, the problem of H_∞ state estimation is solved for neural networks with random delays, (x, v)-dependent noises and fading channels. The optimal solution to derive the minimum disturbance attenuation level, i.e., $\gamma_{\min}$, can be found by solving the convex optimization problem as follows:

$$\min_{\mathcal{P}>0,\ \mathcal{Q}_j>0,\ \mathcal{R}_i>0,\ \delta>0,\ \lambda_i>0,\ \mathcal{X},\ i=1,2,\ldots,N,\ j=1,2,\ldots,L} \chi \quad \text{subject to the LMI (35) with } \chi=\gamma^2.$$

Remark 6: In the neural network (1), although $w(k)$ is supposed to be a one-dimensional white noise for simplicity, all our results can be extended to the following neural network with multiple noises without any difficulty:

$$\begin{cases} x(k+1) = Ax(k) + W_1\bar{f}(x(k)) + W_2\bar{g}(x(k-d(k))) + Cv(k) \\ \qquad\qquad + \sum_{s=1}^{r}[\bar{A}_s x(k) + \bar{B}_s x(k-d(k)) + \bar{C}_s v(k)]w_s(k), \\ y(k) = Dx(k) + Ev(k), \\ z(k) = Fx(k), \\ x(j) = \phi(j), \quad -d_M \le j \le 0 \end{cases}$$

where $w_s(k)$, $s=1,\ldots,r$ are independent, standard one-dimensional white noises on a probability space $(\Omega, \mathcal{F}, \mathrm{Prob})$, and $\bar{A}_s$, $\bar{B}_s$, $\bar{C}_s$, $s=1,\ldots,r$ are constant matrices with appropriate dimensions.

IV. A Numerical Example

In this section, a numerical example is presented so as to demonstrate the effectiveness of our main results. Consider the third-order delayed neural network (1) with the following parameters:

$$A = \begin{bmatrix} 0.2 & 0 & 0 \\ 0 & 0.3 & 0 \\ 0 & 0 & 0.1 \end{bmatrix}, \quad W_1 = \begin{bmatrix} 0.2 & -0.2 & 0 \\ 0 & -0.3 & 0 \\ 0 & 0 & 0.2 \end{bmatrix}, \quad W_2 = \begin{bmatrix} -0.2 & 0.1 & 0 \\ -0.2 & 0.3 & 0 \\ 0 & 0 & 0.1 \end{bmatrix}, \quad C = \begin{bmatrix} 0.2 \\ -0.3 \\ 0.1 \end{bmatrix},$$

$$\bar{A} = \begin{bmatrix} 0.2 & -0.05 & 0 \\ 0.1 & 0.1 & 0 \\ 0 & 0 & 0.05 \end{bmatrix}, \quad \bar{B} = \begin{bmatrix} 0.05 & -0.02 & 0 \\ 0.05 & 0.1 & 0 \\ 0 & 0 & -0.01 \end{bmatrix}, \quad \bar{C} = \begin{bmatrix} 0.05 \\ 0.1 \\ 0.05 \end{bmatrix},$$

$$D = \begin{bmatrix} 0.2 & 0.4 & 0.1 \end{bmatrix}, \quad E = 0.2, \quad F = \begin{bmatrix} 0.3 & 0.1 & 0.1 \end{bmatrix}.$$

The activation functions are taken as

$$\bar{f}(x(k)) = \begin{bmatrix} 0.4x_1(k) - \tanh(0.2x_1(k)) \\ 0.3x_2(k) - \tanh(0.2x_2(k)) \\ 0.4x_3(k) - \tanh(0.3x_3(k)) \end{bmatrix}, \quad \bar{g}(x(k)) = \begin{bmatrix} 0.4x_1(k) - \tanh(0.2x_1(k) + 0.08x_2(k)) \\ 0.2x_2(k) - \tanh(0.1x_2(k)) \\ 0.3x_3(k) - \tanh(0.2x_3(k)) \end{bmatrix}$$

where $x_s(k)$ $(s = 1, 2, 3)$ represents the s-th element of the system state $x(k)$. It is easy to see that the sector-bounded condition (2) can be met with

$$\Phi_f = \begin{bmatrix} 0.2 & 0 & 0 \\ 0 & 0.1 & 0 \\ 0 & 0 & 0.1 \end{bmatrix}, \quad \Psi_f = \begin{bmatrix} 0.4 & 0 & 0 \\ 0 & 0.3 & 0 \\ 0 & 0 & 0.3 \end{bmatrix}, \quad \Phi_g = \begin{bmatrix} 0.2 & 0.08 & 0 \\ 0 & 0.1 & 0 \\ 0 & 0 & 0.1 \end{bmatrix}, \quad \Psi_g = \begin{bmatrix} 0.4 & 0.08 & 0 \\ 0 & 0.2 & 0 \\ 0 & 0 & 0.3 \end{bmatrix}.$$

In this example, the variation of the time-varying delay $d(k)$ is shown in Fig. 2 from which it can be calculated that for $N = 3$, the probability distribution of the delay is

$$\mathrm{Prob}\{d(k) \in [1:2]\} = 0.7, \quad \mathrm{Prob}\{d(k) \in [3:5]\} = 0.2, \quad \mathrm{Prob}\{d(k) \in [6:8]\} = 0.1.$$

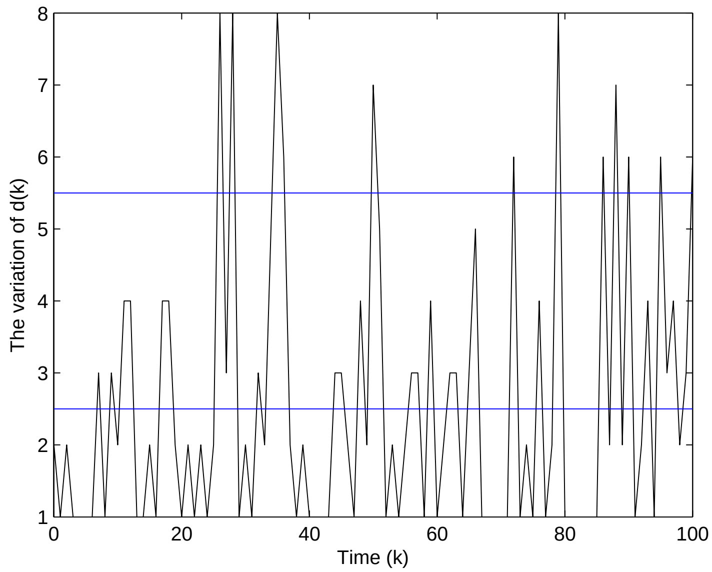

Fig. 2. The variation of the time-varying delay $d(k)$.

Finally, assume that the channel number of the model (8) is $L = 2$, the mathematical expectations of the channel coefficients are $\bar{\beta}_0 = 0.8$, $\bar{\beta}_1 = 0.3$, $\bar{\beta}_2 = 0.2$, the variances of the channel coefficients are $\tilde{\beta}_0 = 0.01$, $\tilde{\beta}_1 = 0.09$, $\tilde{\beta}_2 = 0.25$, and the constant matrix $G = 0.1$.

The H_∞ performance level γ is taken as 0.9. By using MATLAB software with YALMIP 3.0, we can obtain the desired estimator gains as follows:

$$A_f = \begin{bmatrix} 0.2519 & -0.0100 & -0.0096 \\ -0.0093 & 0.2570 & -0.0089 \\ -0.0090 & -0.0090 & 0.2549 \end{bmatrix}, \quad B_f = \begin{bmatrix} 0.3118 \\ 0.3154 \\ 0.3135 \end{bmatrix}.$$

In this simulation, the initial values are assumed to be $\{x(k)\}_{k\in[-8,-1]} = [0\ 0\ 0]^T$ and $x(0) = [0.5\ \ -0.5\ \ 0.3]^T$. The exogenous disturbance inputs are selected as

$$v(k) = e^{-\frac{k}{25}} \sin(k), \qquad \xi(k) = e^{-\frac{k}{25}} \cos(k).$$

The simulation results are shown in Figs. 3–5. Fig. 3 depicts the measurement output and the received signal by the estimator, respectively. Fig. 4 plots the plant and estimator outputs while Fig. 5 shows the estimation error. The simulation results have confirmed that the designed H_∞ estimator performs very well.

V. Conclusions

The problem of H_∞ state estimation for delayed neural networks with (x, v)-dependent noises and fading channels has been investigated in this paper. A sequence of random variables obeying the Bernoulli distribution has been employed to characterized the time-varying delay, and the Rice fading model has been utilized to describe the phenomenon of fading channels. Several delay-distribution-dependent conditions have been derived in terms of LMIs, which guarantee that the estimation error system is stochastically stable with the given H_∞ constraint. Finally, a numerical example has been presented to show the effectiveness of the results derived. It would be interesting to study the following future research topics: 1) development of less conservative conditions for the problem of H_∞ state estimation for delayed neural networks; 2) extension of the results obtained in this paper to neural networks with other network-induced phenomena [2], [5], [7], [8], [11], [12], [15], [19], [22], [43].

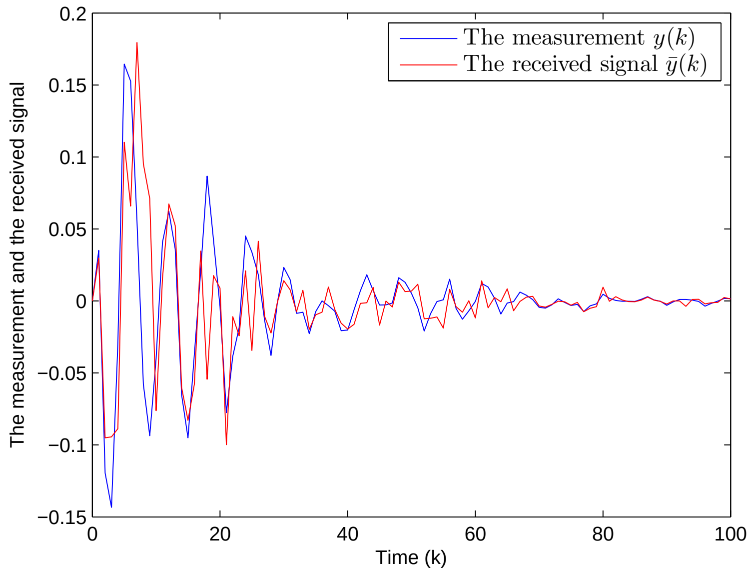

Fig. 3. The measurement $y(k)$ and the received signal $\bar{y}(k)$.

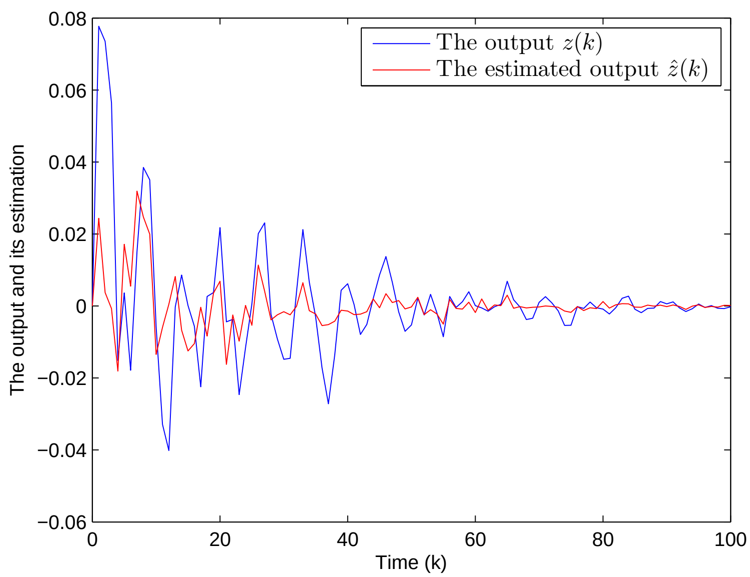

Fig. 4. The plant output $z(k)$ and the estimator output $\hat{z}(k)$.

References

[1] H. Bao, and J. Cao, "Delay-distribution-dependent state estimation for discrete-time stochastic neural networks with random delay," *Neural Networks*, vol. 24, no. 1, pp. 19-28, 2011.

[2] H. Chen, J. Liang and Z. Wang, Pinning controllability of autonomous Boolean control networks, *Science China Information Sciences*, Vol. 59, No. 7, Jul. 2016, Art. No. 070107, DOI: 10.1007/s11432-016-5579-8.

[3] Y. Chherawala, P. Pratim Roy, and M. Cheriet, "Feature set evaluation for offline handwriting recognition systems: application to the recurrent neural network model," *IEEE Transactions on Cybernetics*, doi: 10.1109/TCYB.2015.2490165.

[4] D. Ding, Z. Wang, J. Lam, and B. Shen, "Finite-horizon H_∞ control for discrete time-varying systems with randomly occurring nonlinearities and fading measurements," *IEEE Transactions on Automatic Control*, vol. 60, no. 9, pp. 2488-2493, 2015.

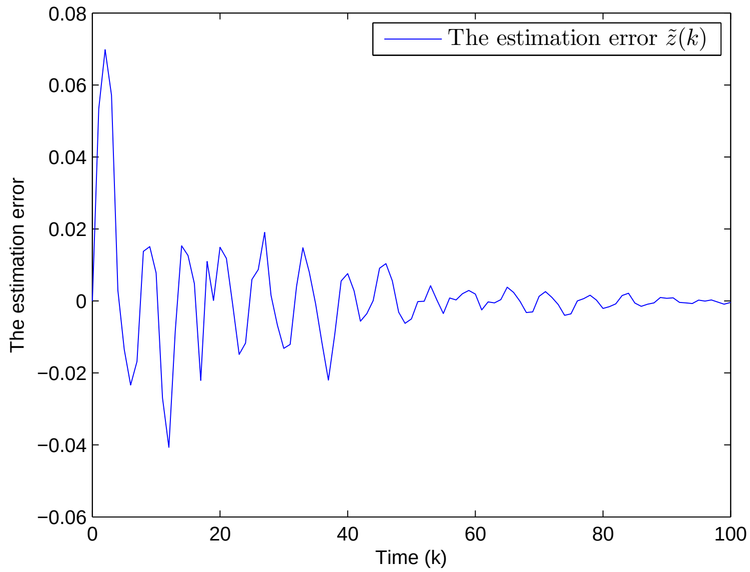

Fig. 5. The estimation error $\tilde{z}(k)$.

[5] D. Ding, Z. Wang, B. Shen and G. Wei, Event-triggered consensus control for discrete-time stochastic multi-agent systems: the input-to-state stability in probability, *Automatica*, Vol. 62, Dec. 2015, pp. 284-291.

[6] H. Dong, Z. Wang, S. X. Ding, and H. Gao, "Event-based H_∞ filter design for a class of nonlinear time-varying systems with fading channels and multiplicative noises," *IEEE Transactions on Signal Processing*, vol. 63, no. 13, pp. 3387-3395, 2015.

[7] H. Dong, Z. Wang, B. Shen and D. Ding, Variance-constrained H_∞ control for a class of nonlinear stochastic discrete time-varying systems: the event-triggered design, *Automatica*, Vol. 72, Oct. 2016, pp. 28-36.

[8] H. Dong, Z. Wang, X. Bu and F. E. Alsaadi, Distributed fault estimation with randomly occurring uncertainties over sensor networks, *International Journal of General Systems*, Vol. 45, No. 5, Jul. 2016, pp. 662-674.

[9] A. El Bouhtouri, D. Hinrichsen, and A. J. Pritchard, "H_∞-type control for discrete-time stochastic systems," *International Journal of Robust and Nonlinear Control*, vol. 9, no. 13, pp. 923-948, 1999.

[10] N. Elia, "Remote stabilization over fading channels," *Systems & Control Letters*, vol. 54, no. 3, pp. 237-249, 2005.

[11] M. Gao, L. Sheng, Y. Liu, and Z. Zhu, "Observer-based H_∞ fuzzy control for nonlinear stochastic systems with multiplicative noise and successive packet dropouts," *Neurocomputing*, vol. 173, pp. 2001-2008, 2016.

[12] M. Gao, L. Sheng, and Y. Liu, "Robust H_∞ control for T-S fuzzy systems subject to missing measurements with uncertain missing probabilities," *Neurocomputing*, vol. 193, pp. 235-241, 2016.

[13] R. A. Horn, and C. R. Johnson, *Topics in Matrix Analysis*, Cambridge, U.K.: Cambridge Univ. Press, 1994.

[14] N. Hou, H. Dong, Z. Wang, W. Ren and F. E. Alsaadi, Non-fragile state estimation for discrete Markovian jumping neural networks, *Neurocomputing*, Vol. 179, Feb. 2016, pp. 238-245.

[15] J. Hu, Z. Wang, S. Liu and H. Gao, "A variance-constrained approach to recursive state estimation for time-varying complex networks with missing measurements", *Automatica*, vol. 64, pp. 155-162, 2016.

[16] H. Huang, T. Huang, and X. Chen, "Further result on guaranteed H_∞ performance state estimation of delayed static neural networks," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 26, no. 6, pp. 1335-1341, 2015.

[17] J. Liang, Z. Wang, and X. Liu, "State estimation for coupled uncertain stochastic networks with missing measurements and time-varying delays: the discrete-time case," *IEEE Transactions on Neural Networks*, vol. 20, no. 5, pp. 781-793, 2009.

[18] Z. Lin, Y. Lin, and W. Zhang, "A unified design for state and output feedback H_∞ control of nonlinear stochastic Markovian jump systems with state and disturbance-dependent noise," *Automatica*, vol. 45, no. 12, pp. 2955-2962, 2009.

[19] H. Liu, Z. Wang, B. Shen and F. E. Alsaadi, State estimation for discrete-time memristive recurrent neural networks with stochastic time-delays, *International Journal of General Systems*, Vol. 45, No. 5, Jul. 2016, pp. 633-647.

[20] M. Liu, and H. Chen, "H_∞ state estimation for discrete-time delayed systems of the neural network type with multiple missing measurements," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 26, no. 12, pp. 2987-2998, 2015.

[21] Y. Liu, Z. Wang, J. Liang, and X. Liu, "Synchronization of coupled neutral-type neural networks with jumping-mode-dependent discrete and unbounded distributed delays," *IEEE Transactions on Cybernetics*, vol. 43, no. 1, pp. 102-114, 2013.

[22] Q. Liu, Z. Wang, X. He, and D. H. Zhou, “Event-based distributed filtering with stochastic measurement fading”, *IEEE Transactions on Industrial Informatics*, vol. 11, no. 6, pp. 1643-1652, 2015.

[23] K. Mathiyalagan, H. Su, P. Shi, and R. Sakthivel, “Exponential H_∞ filtering for discrete-time switched neural networks with random delays,” *IEEE Transactions on Cybernetics*, vol. 45, no. 4, pp. 676-687, 2015.

[24] M. Mestari, M. Benzirar, N. Saber, and M. Khouil, “Solving nonlinear equality constrained multiobjective optimization problems using neural networks,” *IEEE Transactions on Neural Networks and Learning Systems*, vol. 26, no. 10, pp. 2500-2520, 2015.

[25] D. E. Quevedo, A. Ahlén, A. S. Leong, and S. Dey, “On Kalman filtering over fading wireless channels with controlled transmission powers,” *Automatica*, vol. 48, no. 7, pp. 1306-1316, 2012.

[26] T. S. Rappaport. *Wireless Communications: Principles and Practice*, 2nd ed. New Jersey: Prentice Hall, 2002.

[27] W. Rudin, *Principles of Mathematical Analysis*, 3rd ed. New York: McGraw-Hill, 1976.

[28] B. Sklar, “Rayleigh fading channels in mobile digital communication systems part I: Characterization” *IEEE Communications Magazine*, vol. 35, no. 7, pp. 90-100, 1997.

[29] L. Sheng, W. Zhang, and M. Gao, “Relationship between Nash equilibrium strategies and H_2/H_∞ control of stochastic Markov jump systems with multiplicative noise,” *IEEE Transactions on Automatic Control*, vol. 59, no. 9, pp. 2592-2597, 2014.

[30] K. Shi, H. Zhu, S. Zhong, Y. Zeng, Y. Zhang, and W. Wang, “Stability analysis of neutral type neural networks with mixed time-varying delays using triple-integral and delay-partitioning methods,” *ISA Transactions*, vol. 58, pp. 85-95, 2015.

[31] Y. C. Tian, D. Levy, M. O. Tade, T. L. Gu, and C. Fidge, “Queuing packet in communication networks networked control systems,” In *Proceedings of the 6th World Congress on Intelligent Control and Automation*, Dalian, China, 210-214, 2006.

[32] A. R. Trivedi, S. Datta, and S. Mukhopadhyay, “Application of silicon-germanium source tunnel-FET to enable ultralow power cellular neural network-based associative memory,” *IEEE Transactions on Electron Devices*, vol. 61, no. 11, pp. 3707-3715, 2014.

[33] Z. Wang, B. Shen, H. Shu, and G. Wei, “Quantized H_∞ control for nonlinear stochastic time-delay systems with missing measurements,” *IEEE Transactions on Automatic Control*, vol. 57, no. 6, pp. 1431-1444, 2012.

[34] J. Wang, X. M. Zhang, and Q. L. Han, “Event-triggered generalized dissipativity filtering for neural networks with time-varying delays,” *IEEE Transactions on Neural Networks and Learning Systems*, vol. 27, no. 1, pp. 77-88, 2016.

[35] S. Wen, Z. Zeng, T. Huang, and Y. Zhang “Exponential adaptive lag synchronization of memristive neural networks via fuzzy method and applications in pseudorandom number generators,” *IEEE Transactions on Fuzzy Systems*, vol. 22, no. 6, pp. 1704-1713, 2014.

[36] F. Yang, H. Dong, Z. Wang, W. Ren and F. E. Alsaadi, A new approach to non-fragile state estimation for continuous neural networks with time-delays, *Neurocomputing*, Vol. 197, Jul. 2016, pp. 205-211.

[37] H. Yang, Z. Wang, H. Shu, F. E. Alsaadi and T. Hayat, Almost sure H_∞ sliding mode control for nonlinear stochastic systems with Markovian switching and time-delays, *Neurocomputing*, Vol. 175, Part A, Jan. 2016, pp. 392-400.

[38] Y. Yu, H. Dong, Z. Wang, W. Ren and F. E. Alsaadi, Design of non-fragile state estimators for discrete time-delayed neural networks with parameter uncertainties, *Neurocomputing*, Vol. 182, Mar. 2016, pp. 18-24.

[39] D. Yue, Y. Zhang, E. Tian, and C. Peng, “Delay-distribution-dependent exponential stability criteria for discrete-time recurrent neural networks with stochastic delay,” *IEEE Transactions on Neural Networks*, vol. 19, no. 7, pp. 1299-1306, 2008.

[40] N. Zeng, Z. Wang, B. Zineddin, Y. Li, M. Du, L. Xiao, X. Liu, and T. Young, “Image-based quantitative analysis of gold immunochromatographic strip via cellular neural network approach”, *IEEE Transactions on Medical Imaging*, vol. 33, no. 5, pp. 1129-1136, 2014.

[41] B. Zhang, J. Zhang, and D. Kannan, “Nonlinear stochastic difference equations driven by martingales,” *Stochastic Analysis and Applications*, vol. 23, pp. 1277-1304, 2005.

[42] J. Zhang, Z. Wang, D. Ding, and X. Liu, “H_∞ state estimation for discrete-time delayed neural networks with randomly occurring quantizations and missing measurements,” *Neurocomputing*, vol. 148, pp. 388-396, 2015.

[43] 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, 366-373, 2016.