This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication.
Citation information: DOI 10.1109/TNNLS.2024.3411484, IEEE Transactions on Neural Networks and Learning Systems

State Estimation for Markovian Jump Neural Networks Under Probabilistic Bit Flips: Allocating Constrained Bit Rates

Yuru Guo, Zidong Wang, Jun-Yi Li and Yong Xu

***Abstract*—In this paper, the state estimation problem is studied for Markovian jump neural networks within a digital network framework. The wireless communication channel with limited bandwidth is characterized by a constrained bit rate, and the occurrence of bit flips during wireless transmission is mathematically modeled. A transmission mechanism, which includes coding-decoding under bit-rate constraints and considers probabilistic bit flips, is introduced, providing a thorough characterization of the digital transmission process. A mode-dependent remote estimator is designed, which is capable of effectively capturing the internal state of the neural network. Furthermore, a sufficient condition is proposed to ensure the estimation error to remain bounded under challenging network conditions. Within this theoretical framework, the relationship between the neural network's estimation performance and the bit rate is explored. Finally, a simulation example is provided to validate the theoretical findings.**

***Index Terms*—Markovian jump neural networks, bit-rate constraint, probabilistic bit flips, state estimation.**

I. Introduction

Artificial neural networks (ANNs), which are inspired by the intricate and efficient structure of the human brain, serve as a cornerstone in the field of artificial intelligence. Consisting of interconnected neurons organized in layers, ANNs exhibit a remarkable ability to process information, learn from data, and make complex decisions [17], [21]–[23], [25], [31], [40]. The architecture of ANNs allows them to capture and represent intricate patterns and relationships within datasets, making them versatile and potent tools in a variety of applications. Specifically, ANNs are crucial in areas such as pattern recognition, model prediction, and disease diagnosis [1]–[3], [5], [13], [24], [34]. Furthermore, the use of ANNs extends to domains like robotics, control systems, and optimization problems, where their adaptive capabilities are effectively utilized.

In recent years, a significant focus has been placed on switching neural networks, among which Markovian jump neural networks (MJNNs) have emerged as a noteworthy subset. Characterized by their incorporation of Markovian jump parameters, MJNNs exhibit random transitions between various network states [8], [11], [28], [30], [35]. This attribute reflects the dynamic and probabilistic nature of these systems, as explored in various research works [6], [32], [36]–[38]. The complex nature of MJNNs makes them particularly suitable for simulating real-world scenarios where abrupt changes or transitions occur in the operational dynamics of ANNs. Current research on MJNNs primarily focuses on areas such as control theory and state estimation, and some key topics include exponential synchronization [47], finite-time H_∞ state estimation [16], and resilient asynchronous state estimation [43]. Despite these advances, it is important to recognize the existence of significant research gaps, especially regarding challenges associated with the digital network transmission process in MJNNs, which highlights the necessity for continued exploration and development in this specialized area of MJNNs research.

The historical emphasis in state estimation has predominantly been on networked systems within analog communication frameworks. Note that the rapid advancement in digital network technology has prompted a significant shift in communication methods used in control systems [4], [44]. Traditional analog communication techniques, which were once widespread, are increasingly seen as insufficient for the needs of contemporary control systems. Consequently, digital communication methods have gained prominence by offering distinct advantages over analog systems in terms of robustness, reliability, and energy efficiency [9], [12], [18], [26], [33]. Despite these developments, research specifically targeting the challenges associated with control and estimation in ANNs using digital communication networks is still relatively underexplored. As technological advancements continue, there is a growing opportunity to explore the complexities of digital communication within the context of ANNs.

In digital networks, the bit rate is a critical factor that defines the amount of data transmitted within a given time frame [10], [20]. A higher bit rate enables the transmission of more data per second, thereby enhancing the efficiency of communication in digital networks. However, practical scenarios often face limitations on bit rate due to constraints

This work was supported in part by the Natural Science Foundation of Guangdong Province of China under Grants 2021A1515011634 and 2021B1515420008, the National Natural Science Foundation of China under Grants 62121004, U22A2044, and 62006043, the Key Area Research and Development Program of Guangdong Province of China under Grant 2021B0101410005, the Local Innovative and Research Teams Project of Guangdong Special Support Program of China under Grant 2019BT02X353, and the China Scholarship Council under Grant 202208440312. *(Corresponding author: Jun-Yi Li)*

Yuru Guo, Jun-Yi Li and Yong Xu are with the Guangdong-Hong Kong Joint Laboratory for Intelligent Decision and Cooperative Control, Guangdong Provincial Key Laboratory of Intelligent Decision and Cooperative Control, School of Automation, Guangdong University of Technology, Guangzhou 510006, China. (E-mails: guo_yuru0626@163.com; jun-yi-li@foxmail.com; xuyong809@163.com)

Zidong Wang is with the Department of Computer Science, Brunel University London, Uxbridge UB8 3PH, United Kingdom. (E-mail: Zidong.Wang@brunel.ac.uk)

Citation information: DOI 10.1109/TNNLS.2024.3411484, IEEE Transactions on Neural Networks and Learning Systems

in wireless channel capacity and bandwidth availability. These limitations present significant challenges in achieving rapid and reliable data transmission in wireless digital networks [39], [46]. Addressing these challenges necessitates strategic allocation of bit rates, which is crucial for optimizing resource utilization across network nodes [7]. Therefore, investigating bit rate constraints and their allocation strategies is essential for a deeper understanding of the dynamics in wireless digital networks.

In wireless digital networks, the coding-decoding process is essential for data exchange [45], [48]. This process involves converting an analog signal into a digital format through sampling, quantization, and coding. The analog signal is first made discrete and then encoded into binary values 0 and 1 for digital transmission [41]. A significant challenge in this process arises from the phenomenon of bit flips during transmission, wherein a bit within the original binary data stream may change from 1 to 0 or from 0 to 1 with a certain probability. These phenomena can result in errors attributed to factors such as channel noise, signal attenuation, and equipment malfunctions [14], [19], [27], [29]. The occurrence of bit flips can result in decoding errors, particularly in scenarios with constrained bit rates, thereby impacting system performance. Therefore, understanding and mitigating the effects of probabilistic bit flipping becomes paramount to ensure reliable data transmission.

Motivated by the above discussions, in this paper, we tackle the state estimation problem for MJNNs with bit flips under conditions of constrained bit rates. This problem encompasses three primary challenges: 1) the development of a comprehensive mathematical model that accurately represents both the dynamics of MJNNs and the bit rate limitations characteristic of the communication network; 2) the formulation of a method to model and manage stochastic bit flips that occur during wireless transmission; and 3) the development of strategies for devising estimator gains that ensure error boundedness, even in the face of bit rate constraints. These challenges form the core of our research, aiming to enhance the robustness and efficiency of state estimation in MJNNs within the digital communication landscape.

In response to the identified challenges, our study makes several key contributions in addressing the state estimation problem for discrete-time ANNs with Markovian jump parameters in digital communication networks, particularly under constrained bit rate conditions.

1) We tackle the state estimation problem for discrete-time ANNs featuring Markovian jump parameters by taking into account the bit rate constraints that reflect the inherent bandwidth limitations in such networks.
2) We systematically address the occurrence of probabilistic bit flipping, caused by interference noise, within the context of constrained wireless digital transmission networks. This issue is comprehensively mathematically modeled, providing a robust framework for understanding and mitigating its impact.
3) We establish a significant connection between estimation performance and bit rate allocation. To this end, we introduce a collaborative optimization algorithm that effectively allocates bit rates while fine-tuning the estimator gains. This approach aims to optimize performance by balancing the demands of data transmission and processing efficiency in the network.

The structure of this paper is organized as follows. Section II offers a comprehensive description of the models used in our study, which includes detailed information on discrete-time MJNNs, the communication network with bit-rate constraints, the coding-decoding process under the influence of probabilistic bit flips, and an overview of the estimation error dynamics (EED). In Section III, the focus is on the analytical aspects of the study, which encompasses the analysis of the boundedness of estimation error, the design of estimator gains at an optimal delay rate, and a collaborative method for the co-design of bit rate allocation and estimator gains. In Section IV, a numerical example is presented to demonstrate the practical application of the theoretical findings, which serves to validate the correctness of the results derived in the study and includes explanatory notes for better comprehension. The paper concludes in Section V by summarizing the main findings and contributions of the research.

Notations: In this paper, we employ specific symbols to represent various mathematical concepts. $\mathbb{R}^m$, $\mathbb{R}^{m\times n}$, and $\mathbb{N}$ represent the m-dimensional Euclidean space, the $m\times n$ real matrices, and the non-negative integers, respectively. The symbol $\|\cdot\|$ refers to the Euclidean norm, and $|\cdot|$ stands for the absolute value. $\mathbb{E}\cdot$, $\mathbb{P}\cdot$ and $\mathrm{Var}\cdot$ depict the expectation, probability and variance for a stochastic variable. For a matrix X, its transpose is denoted by X^T, and $\underline{\lambda}(X)$ signifies its minimum eigenvalue. A column vector is expressed by $\mathrm{col}\cdot$. The diagonal matrix is articulated as $\mathrm{diag}\cdots$. The Kronecker product is represented by the symbol $\otimes$, and I denotes the identity matrix with proper dimensions.

II. Problem Formulation and Preliminaries

A. Markovian Jump Neural Networks

For a discrete-time homogeneous Markov chain $\tau(k)$ taking values in a finite set $\Pi \triangleq \{1,2,\dots,s\}$, let the transition probability matrix $\aleph \triangleq (\sigma_{\imath\jmath}) \in \mathbb{R}^{n\times n}$ be given by

$$\mathbb{P}\{\tau(k+1)=\jmath|\tau(k)=\imath\}=\sigma_{\imath\jmath},\ \forall \imath,\jmath\in\Pi \tag{1}$$

where $\sigma_{\imath\jmath}\geq 0$ and $\sum_{\jmath=1}^{s}\sigma_{\imath\jmath}=1$.

Consider a class of discrete-time MJNNs with noise disturbance represented by

$$\begin{cases} x_i(k+1)=h_{i,\tau(k)}x_i(k)+\sum\limits_{j=1}^{n}\varpi_{ij,\tau(k)}g_j(x_j(k)) \\ \qquad\qquad +a_{i,\tau(k)}v_i(k) \\ y_i(k)=c_{i,\tau(k)}x_i(k)+b_{i,\tau(k)}v_i(k) \end{cases} \tag{2}$$

for $i\in\mathcal{I}\triangleq\{1,2,\dots,n\}$, where $x_i(k)\in\mathbb{R}$ and $y_i(k)\in\mathbb{R}$ denote the state and measurement output of neuron i, respectively; $h_{i,\tau(k)}\in\mathbb{R}$ refers to the state feedback coefficient; scalars $a_{i,\tau(k)}$ and $b_{i,\tau(k)}$ are weight coefficients of the noise; $\varpi_{ij,\tau(k)}\in\mathbb{R}$ represents the interconnection strength between neurons i and j; the disturbance input $v_i(k)\in\mathbb{R}$ satisfies $|v_i(k)|\leq\bar{v}$ with a given scalar $\bar{v}$. In addition, $g_j(\cdot):\mathbb{R}\to\mathbb{R}$

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication.
Citation information: DOI 10.1109/TNNLS.2024.3411484, IEEE Transactions on Neural Networks and Learning Systems

is an activation function of the j-th neuron subjecting to the following assumption.

Assumption 1: The nonlinear activation function satisfies the following condition:

$$\big(g_j(d_1) - g_j(d_2) - \acute{m}_j(d_1 - d_2)\big)^T \times \big(g_j(d_1) - g_j(d_2) - \grave{m}_j(d_1 - d_2)\big) \le 0$$

where $d_1, d_2 \in \mathbb{R}$ are some scalars, and $\acute{m}_j$ and $\grave{m}_j$ are known constants.

B. Communication Network With Bit-Rate Constraints

In practical digital communication networks, especially those that are wireless, the original signals from sensors must be converted into binary characters by a coder for transmission. Due to the often limited bandwidth of these networks, there is a restriction on the number of bits that can be transmitted at any given time. Efficiently allocating the bit rate for each neuron is crucial to avoid data collisions during wireless transmission.

In our study, we consider a scenario where the total available bit rates for the entire network are represented by Λ ($\Lambda \in \mathbb{N}$). The measurement $y_i(k)$ of each neuron in the MJNNs is transmitted over this bit-rate constrained wireless network. The bit rates allocated to each coder in the network must adhere to the following specific condition for ensuring optimal use of the available bandwidth and efficient transmission of data [15]:

$$\Lambda \ge \sum_{i=1}^{n} \Lambda_i, \;\; \Lambda_i \in \mathbb{N} \tag{3}$$

where Λ_i ($i \in \mathcal{I}$) denotes the bit rates allocated to the neuron i, that is, each coder has limited bit rates to encode the data packet. As a result, the data compression is required, which can be realized by a uniform quantizer. To be specific, this quantizer segments the quantization region into a set number of uniformly spaced intervals. The quantization process involves mapping each input data to its corresponding interval.

Given a scalar $\delta_i > 0$ decided by the range of the sensors, the quantization region of the i-th sensor measurement is represented by

$$|y_i(k)| \le \delta_i. \tag{4}$$

Choosing a quantization level Δ_i (denoting the number of intervals) for the sensor i, the quantization region can be uniformly segmented into some subintervals, which are denoted by

$$\begin{cases} \mathcal{Q}_{1,i} : -\delta_i \le y_i(k) < -\delta_i + \dfrac{2\delta_i}{\Delta_i} \\ \mathcal{Q}_{2,i} : -\delta_i + \dfrac{2\delta_i}{\Delta_i} \le y_i(k) < -\delta_i + \dfrac{4\delta_i}{\Delta_i} \\ \quad\vdots \\ \mathcal{Q}_{\Delta_i,i} : \delta_i - \dfrac{2\delta_i}{\Delta_i} \le y_i(k) \le \delta_i. \end{cases} \tag{5}$$

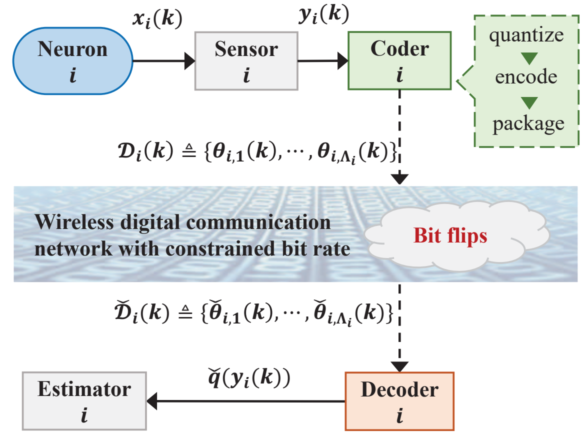


Fig. 1. Research framework for ANNs under bit flips.

The maximum quantization level $\widehat{\Delta}_i$ of sensor i is limited by allocated bit rates, which can be deduced by

$$\widehat{\Delta}_i \triangleq 2^{\Lambda_i}. \tag{6}$$

Let $\{\ell_1, \ell_2, \ldots, \ell_n\}$ be the corresponding quantization region for n sensors. We take the central value of the subinterval to approximate the original data, which is computed by

$$q(y_i(k)) = -\delta_i + \frac{(2\ell_i - 1)\delta_i}{\widehat{\Delta}_i}. \tag{7}$$

According to the above description, the quantization error of the measurement $y_i(k)$ is defined as:

$$e_i(k) \triangleq y_i(k) - q(y_i(k)).$$

The upper bound on the absolute value of the quantization error $|e_i(k)|$ is the distance from the center point to the end point of the subinterval. Then, we further obtain

$$|e_i(k)| \le \frac{\delta_i}{\widehat{\Delta}_i}. \tag{8}$$

Remark 1: In wireless communication networks, bit rate allocation can follow either dynamic or static protocols. Dynamic protocols adjust bit rates dynamically based on the fluctuating needs of user devices, aiming to optimize the efficiency of data transmission for each device. On the other hand, static protocols allocate bit rates based on predetermined criteria, regardless of individual device requirements. More specifically, static protocols are designed to ensure fair data transmission, particularly beneficial in scenarios where multiple users share limited bandwidth. In this paper, we have chosen to employ a static bit rate allocation protocol within the context of MJNNs so as to ensure a fair and equitable distribution of bandwidth among all neurons in the network. By doing so, we aim to promote consistent and unbiased data transmission across the network, which an important consideration in environments where bandwidth is a limited resource.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication.
Citation information: DOI 10.1109/TNNLS.2024.3411484, IEEE Transactions on Neural Networks and Learning Systems

C. Coding-Decoding Process Under Probabilistic Bit Flips

As shown in Fig. 1, the coder i plays a critical role by converting quantized data into a binary data stream $\mathcal{D}_i(k)$, which is subsequently transmitted over a wireless digital communication network. While numerous studies have extensively explored encoding mechanism, many tend to be overly idealized when it comes to the wireless transmission of signals. In practice, many factors such as channel noise, multi-user interference, and equipment failures introduce a non-negligible probability of bit flips (i.e., bit error) within the binary data stream during transmission. This, undoubtedly, exerts a certain influence on the decoding error.

In order to specifically analyze the impact of bit flips on the estimation performance, let us first introduce the mathematical model [19]. Denote the binary bit stream generated by the i-th coder as

$$\mathcal{D}_i(k) \triangleq \{\theta_{i,1}(k), \theta_{i,2}(k), \ldots, \theta_{i,\Lambda_i}(k)\} \tag{9}$$

where $\theta_{i,r}(k) \in \{0,1\}$ is the codeword and $r \in \{1,2,\ldots,\Lambda_i\}$ stands for the number of bits.

After the bit stream passes through the bandwidth-constrained communication network, the data received by the decoder i becomes

$$\check{\mathcal{D}}_i(k) \triangleq \{\check{\theta}_{i,1}(k), \check{\theta}_{i,2}(k), \ldots, \check{\theta}_{i,\Lambda_i}(k)\} \tag{10}$$

with the codeword $\check{\theta}_{i,r}(k) \in \{0,1\}$. In this case, each probabilistically flipped bit satisfies the following condition:

$$\check{\theta}_{i,r}(k) = p_{i,r}(k)\big(1-\theta_{i,r}(k)\big) + \big(1-p_{i,r}(k)\big)\theta_{i,r}(k) \tag{11}$$

where $p_{i,r}(k)$ represents the flip probability of the r-th bit for the i-th coder data, obeying the Bernoulli distribution. It is imperative to note that

$$p_{i,r}(k) = \begin{cases} 1, & \text{the } r\text{-th bit is flipped,} \\ 0, & \text{the } r\text{-th bit remains unchanged.} \end{cases}$$

For the sake of analysis, we assume that occurrence of the flipping of each bit is mutually independent of each other and satisfies

$$\mathbb{P}\{p_{i,r}(k) = 1\} = \bar{p}_i, \tag{12}$$

where $\bar{p}_i \in (0,1)$ $(i \in \mathcal{I})$ is a known constant.

The quantization output (7) is further expressed in concrete bits in the following form:

$$q(y_i(k)) = -\delta_i + \frac{\left(2\sum_{r=1}^{\Lambda_i}\theta_{i,r}(k)2^{r-1}+1\right)\delta_i}{\widehat{\Delta}_i}. \tag{13}$$

Similarly, after passing through the wireless digital network and experiencing probabilistic bit flipping, the decoding output is denoted as:

$$\check{q}(y_i(k)) = -\delta_i + \frac{\left(2\sum_{r=1}^{\Lambda_i}\check{\theta}_{i,r}(k)2^{r-1}+1\right)\delta_i}{\widehat{\Delta}_i}. \tag{14}$$

Example 1: Given $\Lambda_i = 3$ bits, $\delta_i = 2$, the maximum quantization level is deduced by $\widehat{\Delta}_i = 8$. We take the central value of subinterval as the quantization output, subsequently, the calculation unit is based on half the length $\frac{\delta_i}{\widehat{\Delta}_i}$ of each subinterval. According to the formula (13), the correspondence rule between the number of interval and codeword can be clearly obtained as shown in Table I. If the bit stream is flipped from '000' to '100', the corresponding data is changed from $q(y_i(k)) = -1.75$ to $\check{q}(y_i(k)) = -1.25$.

TABLE I
QUANTITATIVE INTERVALS AND CORRESPONDING CODEWORDS

Codeword	000	100	010	110	001	101	011	111
Number	1	2	3	4	5	6	7	8

In order to deal with the decoding output $\check{q}(y_i(k))$ containing uncertainty, a lemma is given to describe its statistical properties.

Lemma 1: Let the signal $q(y_i(k))$ be transmitted via a memoryless binary symmetric channel with bit flipping probability $\bar{p}_i$. Then, the decoding signal $\check{q}(y_i(k))$ has the mean and variance given by

$$\mathbb{E}\{\check{q}(y_i(k))\} = (1-2\bar{p}_i)q(y_i(k)) \tag{15}$$

and

$$\mathrm{Var}\{\check{q}(y_i(k))\} = \frac{4}{3}\bar{p}_i(1-\bar{p}_i)\frac{\delta_i^2(2^{2\Lambda_i}-1)}{2^{2\Lambda_i}} \triangleq \Psi_i \tag{16}$$

where the expectation is taken with respect to the stochastic variables $p_{i,r}(k)$.

Proof: Taking the expectation of (14), we have from (11) and (12) that

$$\begin{aligned}
&\mathbb{E}\{\check{q}(y_i(k))\} \\
&= -\delta_i + \frac{\left(2\sum_{r=1}^{\Lambda_i}\mathbb{E}\left\{\check{\theta}_{i,r}(k)\right\}2^{r-1}+1\right)\delta_i}{\widehat{\Delta}_i} \\
&= -\delta_i + \Bigg(2\sum_{r=1}^{\Lambda_i}\Big(\bar{p}_i\big(1-\theta_{i,r}(k)\big) \\
&\quad + (1-\bar{p}_i)\theta_{i,r}(k)\Big)2^{r-1}+1\Bigg)\delta_i/\widehat{\Delta}_i \\
&= -\delta_i + \left(2\sum_{r=1}^{\Lambda_i}\theta_{i,r}(k)2^{r-1}+1\right)\delta_i/\widehat{\Delta}_i \\
&\quad + 2\bar{p}_i\sum_{r=1}^{\Lambda_i}\big(1-2\theta_{i,r}(k)\big)2^{r-1}\delta_i/\widehat{\Delta}_i.
\end{aligned} \tag{17}$$

According to (14), the following relation is further obtained:

$$\begin{aligned}
&\mathbb{E}\{\check{q}(y_i(k))\} \\
&= q(y_i(k)) + 2\bar{p}_i\Bigg(\sum_{r=1}^{\Lambda_i}2^{r-1}\delta_i/\widehat{\Delta}_i + \delta_i/\widehat{\Delta}_i \\
&\quad - 2\sum_{r=1}^{\Lambda_i}\theta_{i,r}(k)2^{r-1}\delta_i/\widehat{\Delta}_i - \delta_i/\widehat{\Delta}_i\Bigg) \\
&= q(y_i(k)) + 2\bar{p}_i\left(\delta_i - \left(2\sum_{r=1}^{\Lambda_i}\theta_{i,r}(k)2^{r-1}+1\right)\frac{\delta_i}{\widehat{\Delta}_i}\right) \\
&= (1-2\bar{p}_i)q(y_i(k)).
\end{aligned} \tag{18}$$

Citation information: DOI 10.1109/TNNLS.2024.3411484, IEEE Transactions on Neural Networks and Learning Systems

Subsequently, we delve into the analysis of the variance of $\check{q}(y_i(k))$. To streamline the calculations, we obtain

$$\begin{aligned}
&-\delta_i+\frac{\left(2\sum_{r=1}^{\Lambda_i}\check{\theta}_{i,r}(k)2^{r-1}+1\right)\delta_i}{\widehat{\Delta}_i}\\
=&2\sum_{r=1}^{\Lambda_i}\check{\theta}_{i,r}(k)2^{r-1}\frac{\delta_i}{2^{\Lambda_i}}-\delta_i+\frac{\delta_i}{2^{\Lambda_i}}\\
=&2\sum_{r=1}^{\Lambda_i}\left(\check{\theta}_{i,r}(k)-\mathbb{E}\left\{\check{\theta}_{i,r}(k)\right\}\right)2^{r-1}\frac{\delta_i}{2^{\Lambda_i}}-\delta_i\\
&+\left(2\sum_{r=1}^{\Lambda_i}\mathbb{E}\left\{\check{\theta}_{i,r}(k)\right\}2^{r-1}+1\right)\frac{\delta_i}{2^{\Lambda_i}}\\
=&2\sum_{r=1}^{\Lambda_i}\left(\check{\theta}_{i,r}(k)-\mathbb{E}\left\{\check{\theta}_{i,r}(k)\right\}\right)2^{r-1}\frac{\delta_i}{2^{\Lambda_i}}-\mathbb{E}\{\check{q}(y_i(k))\}.
\end{aligned}\tag{19}$$

Based on (19), the variance of $\check{q}(y_i(k))$ is expressed as

$$\begin{aligned}
&\mathrm{Var}\{\check{q}(y_i(k))\}\\
=&\mathbb{E}\left\{\left(-\delta_i+\left(2\sum_{r=1}^{\Lambda_i}\check{\theta}_{i,r}(k)2^{r-1}+1\right)\delta_i/\widehat{\Delta}_i\right)^2\right\}\\
&-\left(\mathbb{E}\left\{\check{q}(y_i(k))\right\}\right)^2\\
=&\mathbb{E}\left\{\left(2\sum_{r=1}^{\Lambda_i}\left(\check{\theta}_{i,r}(k)-\mathbb{E}\left\{\check{\theta}_{i,r}(k)\right\}\right)2^{r-1}\frac{\delta_i}{2^{\Lambda_i}}\right)^2\right\}\\
=&4\sum_{r=1}^{\Lambda_i}\left(\mathbb{E}\left\{\check{\theta}^2_{i,r}(k)\right\}-\mathbb{E}\left\{\check{\theta}_{i,r}(k)\right\}^2\right)2^{2r-2}\frac{\delta_i^2}{2^{2\Lambda_i}}\\
=&4\sum_{r=1}^{\Lambda_i}\left((1-\theta_{i,r}(k))\bar{P}_i(1-\bar{P}_i)+\theta_{i,r}(k)\bar{P}_i(1-\bar{P}_i)\right)\\
&\times 2^{2r-2}\frac{\delta_i^2}{2^{2\Lambda_i}}\\
=&\frac{4}{3}\bar{p}_i(1-\bar{p}_i)\frac{\delta_i^2(2^{2\Lambda_i}-1)}{2^{2\Lambda_i}}=\Psi_i,
\end{aligned}\tag{20}$$

which ends the proof. ■

From Lemma 1, the decoding output is given by

$$\check{q}(y_i(k))=(1-2\bar{p}_i)q(y_i(k))+\phi_i(k)\tag{21}$$

where $\phi_i(k)$ is a stochastic variable with zero mean and variance Ψ_i.

D. Estimation Error Dynamics

To estimate the internal state of MJNNs efficiently by utilizing the decoded measurement output, we devise the remote estimator as follows:

$$\begin{aligned}
\hat{x}_i(k+1)=&h_{i,\imath}\hat{x}_i(k)+\sum_{j=1}^{n}\varpi_{ij,\imath}g_j(\hat{x}_j(k))\\
&+l_{i,\imath}\left(\check{q}(y_i(k))-(1-2\bar{p}_i)c_{i,\imath}\hat{x}_i(k)\right)
\end{aligned}\tag{22}$$

where $\hat{x}_i(k)$ denotes the state estimate for the neuron i, and $l_{i,\imath}\in\mathbb{R}$ represents the mode-dependent gain to be designed.

Define the estimation error as $\tilde{x}_i(k)\triangleq x_i(k)-\hat{x}_i(k)$ and the nonlinear function as $\tilde{g}_j(k)\triangleq g_j(x_j(k))-g_j(\hat{x}_j(k))$. We have the following EED:

$$\begin{aligned}
\tilde{x}_i(k+1)=&h_{i,\imath}\tilde{x}_i(k)+\sum_{j=1}^{n}\varpi_{ij,\imath}\tilde{g}_j(k)+a_{i,\imath}v_i(k)\\
&-l_{i,\imath}\Big((1-2\bar{p}_i)\big(y_i(k)-e_i(k)\big)+\phi_i(k)\Big)\\
&+l_{i,\imath}(1-2\bar{p}_i)c_{i,\imath}\hat{x}_i(k)\\
=&\big(h_{i,\imath}-l_{i,\imath}(1-2\bar{p}_i)c_{i,\imath}\big)\tilde{x}_i(k)+\sum_{j=1}^{n}\varpi_{ij,\imath}\tilde{g}_j(k)\\
&+\big(a_{i,\imath}-l_{i,\imath}(1-2\bar{p}_i)b_{i,\imath}\big)v_i(k)-l_{i,\imath}\phi_i(k)\\
&+(1-2\bar{p}_i)e_i(k).
\end{aligned}\tag{23}$$

Define the augmented error vector as

$$\tilde{x}(k)\triangleq\begin{bmatrix}\tilde{x}_1^T(k) & \tilde{x}_2^T(k) & \dots & \tilde{x}_n^T(k)\end{bmatrix}^T.$$

The augmented EED is presented by

$$\begin{aligned}
\tilde{x}(k+1)=&\tilde{\mathcal{H}}_\imath\tilde{x}(k)+W_\imath\tilde{g}(k)+\tilde{\mathcal{A}}_\imath v(k)\\
&-L_\imath\phi(k)+(I-2\bar{P})e(k)
\end{aligned}\tag{24}$$

where

$$\begin{aligned}
W_\imath&\triangleq[\varpi_{ij,\imath}],(i,j\in\mathcal{I},\ \imath\in\Pi),\\
\tilde{\mathcal{H}}_\imath&\triangleq H_\imath-L_\imath(I-2\bar{P})C_\imath,\\
\tilde{\mathcal{A}}_\imath&\triangleq A_\imath-L_\imath(I-2\bar{P})B_\imath,\\
H_\imath&\triangleq\mathrm{diag}\{h_{1,\imath},h_{2,\imath},\dots,h_{n,\imath}\},\\
L_\imath&\triangleq\mathrm{diag}\{l_{1,\imath},l_{2,\imath},\dots,l_{n,\imath}\},\\
A_\imath&\triangleq\mathrm{diag}\{a_{1,\imath},a_{2,\imath},\dots,a_{n,\imath}\},\\
B_\imath&\triangleq\mathrm{diag}\{b_{1,\imath},b_{2,\imath},\dots,b_{n,\imath}\},\\
C_\imath&\triangleq\mathrm{diag}\{c_{1,\imath},c_{2,\imath},\dots,c_{n,\imath}\},\\
\bar{P}&\triangleq\mathrm{diag}\{\bar{p}_1,\bar{p}_2,\dots,\bar{p}_n\},\\
v(k)&\triangleq\mathrm{col}\{v_1(k),v_2(k),\dots,v_n(k)\},\\
e(k)&\triangleq\mathrm{col}\{e_1(k),e_2(k),\dots,e_n(k)\},\\
\tilde{g}(k)&\triangleq\mathrm{col}\{\tilde{g}_1(k),\tilde{g}_2(k),\dots,\tilde{g}_n(k)\},\\
\phi(k)&\triangleq\mathrm{col}\{\phi_1(k),\phi_2(k),\dots,\phi_n(k)\}.
\end{aligned}$$

Definition 1: [42] The EED is said to be exponentially mean-square bounded if there exist constants $|\rho_1|<1$, $\rho_2>0$ and $\rho_3>0$ such that the following inequality holds:

$$\mathbb{E}\{\|\tilde{x}(k)\|^2\}\leq\rho_1^k\rho_2+\rho_3\tag{25}$$

where ρ_3 is an asymptotic upper bound of the error $\mathbb{E}\{\|\tilde{x}(k)\|^2\}$.

III. Main Result

In this section, the analysis of the mean-square boundedness for the EED is presented. Subsequently, we formulate designs for estimator gains by incorporating the optimized decay rate and the bit-rate allocation strategy.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication.
Citation information: DOI 10.1109/TNNLS.2024.3411484, IEEE Transactions on Neural Networks and Learning Systems

A. Boundedness Analysis

In the following theorem, a sufficient condition is given to ensure the exponential boundedness in the mean-square sense for the EED.

Theorem 1: Consider the MJNNs (2), the known positive integers Λ_i, and the remote estimator (22) with given estimator gain $l_{i,\imath}$ ($i \in \mathcal{I}$). Then, the augmented EED (24) is exponentially mean-square bounded if there exist positive scalars $\alpha_1,\ \gamma_1,\ \gamma_2,\ 0<\alpha_2<1$, and positive definite and symmetric matrices $\mathcal{P}_\imath$ such that the following inequality holds for all $\imath \in \Pi$:

$$\begin{bmatrix} \Omega_{11} & 0 & \Omega_{13} \\ * & \Omega_{22} & \Omega_{23} \\ * & * & -\hat{\mathcal{P}}_\jmath^{-1} \end{bmatrix} < 0 \tag{26}$$

where

$$\begin{aligned}
&\hat{\mathcal{P}}_\jmath \triangleq \sum_{\jmath=1}^{s} \sigma_{\imath\jmath}\mathcal{P}_\jmath,\ \jmath \in \Pi,\\
&\Omega_{11} \triangleq \begin{bmatrix} -\alpha_1 \acute{M}^T \grave{M} - (1-\alpha_2)\mathcal{P}_\imath & \alpha_1 \frac{\acute{M}^T + \grave{M}}{2} \\ * & -\alpha_1 I \end{bmatrix},\\
&\Omega_{13} \triangleq \begin{bmatrix} \tilde{\mathcal{H}}_\imath & W_\imath \end{bmatrix}^T,\ \Omega_{23} \triangleq \begin{bmatrix} \tilde{\mathcal{A}}_\imath & I - 2\bar{P} \end{bmatrix}^T,\\
&\Omega_{22} \triangleq \mathrm{diag}\{-\gamma_1, -\gamma_2\},\ \mathcal{P}_\imath \triangleq \mathrm{diag}\{\mathcal{P}_{1,\imath}, \mathcal{P}_{2,\imath}, \ldots, \mathcal{P}_{n,\imath}\},\\
&\acute{M} \triangleq \mathrm{diag}\{\acute{m}_1, \ldots, \acute{m}_n\},\ \grave{M} \triangleq \mathrm{diag}\{\grave{m}_1, \ldots, \grave{m}_n\}.
\end{aligned}$$

Proof: Choose the following Lyapunov-like function for boundedness analysis:

$$\mathcal{V}(k) = \tilde{x}^T(k)\mathcal{P}_{\sigma(k)}\tilde{x}(k). \tag{27}$$

Denote $\sigma(k) \triangleq \imath$, $\sigma(k+1) \triangleq \jmath$, and define an augmented vector as

$$\chi(k) \triangleq \begin{bmatrix} \tilde{x}^T(k) & \tilde{g}^T(k) & v^T(k) & e^T(k) \end{bmatrix}^T.$$

Calculating the mathematical expectation of the difference of $\mathcal{V}(k)$, we obtain

$$\begin{aligned}
\Delta\mathcal{V}(k) \triangleq& \mathbb{E}\{\mathcal{V}(k+1)|\mathcal{V}(k)\} - \mathcal{V}(k)\\
=& \chi(k)\Pi_1\chi^T(k) + \gamma_1 v^T(k)v(k) - \alpha_2\mathcal{V}(k)\\
&+ \gamma_2 e^T(k)e(k) + \mathbb{E}\left\{\phi^T(k)L_\imath^T\mathcal{P}_\jmath L_\imath\phi(k)\right\}
\end{aligned} \tag{28}$$

where

$$\Pi_1 \triangleq \begin{bmatrix} \Pi_{11} & \tilde{\mathcal{H}}_\imath^T\hat{\mathcal{P}}_\jmath W_\imath & \tilde{\mathcal{H}}_\imath^T\hat{\mathcal{P}}_\jmath\tilde{\mathcal{A}}_\imath & \Pi_{14} \\ * & W_\imath^T\hat{\mathcal{P}}_\jmath W_\imath & W_\imath^T\hat{\mathcal{P}}_\jmath\tilde{\mathcal{A}}_\imath & \Pi_{24} \\ * & * & \Pi_{33} & \Pi_{34} \\ * & * & * & \Pi_{44} \end{bmatrix},$$

$$\begin{aligned}
\Pi_{11} \triangleq& \tilde{\mathcal{H}}_\imath^T\hat{\mathcal{P}}_\jmath\tilde{\mathcal{H}}_\imath - (1-\alpha_2)\mathcal{P}_\imath,\\
\Pi_{33} \triangleq& \tilde{\mathcal{A}}_\imath^T\hat{\mathcal{P}}_\jmath\tilde{\mathcal{A}}_\imath - \gamma_1 I,\ \Pi_{14} \triangleq \tilde{\mathcal{H}}_\imath^T\hat{\mathcal{P}}_\jmath(I-2\bar{P}),\\
\Pi_{24} \triangleq& W_\imath^T\hat{\mathcal{P}}_\jmath(I-2\bar{P}),\ \Pi_{34} \triangleq \tilde{\mathcal{A}}_\imath^T\hat{\mathcal{P}}_\jmath(I-2\bar{P}),\\
\Pi_{44} \triangleq& (I-2\bar{P})^T\hat{\mathcal{P}}_\jmath(I-2\bar{P}) - \gamma_2 I.
\end{aligned}$$

From Assumption 1, we know that the nonlinearity $\tilde{g}(\cdot)$ fulfills

$$\left(\tilde{g}(k) - \acute{M}\tilde{x}(k)\right)^T\left(\tilde{g}(k) - \grave{M}\tilde{x}(k)\right) \le 0, \tag{29}$$

which can be further expressed as

$$\begin{bmatrix} \tilde{x}(k) \\ \tilde{g}(k) \end{bmatrix}^T \begin{bmatrix} \alpha_1\acute{M}^T\grave{M} & -\alpha_1\frac{\acute{M}^T + \grave{M}^T}{2} \\ * & \alpha_1 I \end{bmatrix} \begin{bmatrix} \tilde{x}(k) \\ \tilde{g}(k) \end{bmatrix} \le 0 \tag{30}$$

for any scalar $\alpha_1 > 0$.

Combining (28) with (30), one has

$$\begin{aligned}
\Delta\mathcal{V}(k) \le& \chi(k)\Pi_2\chi^T(k) + \gamma_1 v^T(k)v(k) - \alpha_2\mathcal{V}(k)\\
&+ \gamma_2 e^T(k)e(k) + \mathrm{tr}\left(\bar{\Psi}L_\imath^T\hat{\mathcal{P}}_\jmath L_\imath\bar{\Psi}\right)
\end{aligned} \tag{31}$$

where

$$\Pi_2 \triangleq \begin{bmatrix} \bar{\Pi}_{11} & \bar{\Pi}_{12} & \tilde{\mathcal{H}}_\imath^T\hat{\mathcal{P}}_\jmath\tilde{\mathcal{A}}_\imath & \Pi_{14} \\ * & \bar{\Pi}_{22} & W_\imath^T\hat{\mathcal{P}}_\jmath\tilde{\mathcal{A}}_\imath & \Pi_{24} \\ * & * & \Pi_{33} & \Pi_{34} \\ * & * & * & \Pi_{44} \end{bmatrix},$$

$$\begin{aligned}
\bar{\Pi}_{11} \triangleq& \tilde{\mathcal{H}}_\imath^T\hat{\mathcal{P}}_\jmath\tilde{\mathcal{H}}_\imath - (1-\alpha_2)\mathcal{P}_\imath - \alpha_1\acute{M}^T\grave{M},\\
\bar{\Pi}_{22} \triangleq& W_\imath^T\hat{\mathcal{P}}_\jmath W_\imath - \alpha_1 I,\\
\bar{\Pi}_{12} \triangleq& \tilde{\mathcal{H}}_\imath^T\hat{\mathcal{P}}_\jmath W_\imath + \alpha_1\frac{\acute{M}^T + \grave{M}}{2},\\
\Psi \triangleq& \mathrm{diag}\{\Psi_1, \Psi_2, \ldots, \Psi_n\} \triangleq \bar{\Psi}\bar{\Psi}^T.
\end{aligned}$$

Applying the Schur Complement Lemma to the above formula, one has $\Pi_2 < 0$ based on the condition (26) in Theorem 1. Then, we derive that

$$\begin{aligned}
\Delta\mathcal{V}(k) <& -\alpha_2\mathcal{V}(k) + \gamma_1 v^T(k)v(k)\\
&+ \gamma_2 e^T(k)e(k) + \max_{\imath\in\Pi}\left\{\mathrm{tr}\left(\bar{\Psi}L_\imath^T\hat{\mathcal{P}}_\jmath L_\imath\bar{\Psi}\right)\right\}.
\end{aligned} \tag{32}$$

In terms of the definitions of noise $v(k)$ and quantization error $e(k)$, we obtain

$$v^T(k)v(k) \le n\bar{v}^2, \tag{33}$$

$$e^T(k)e(k) \le \sum_{i=1}^{n}\left(\frac{\delta_i}{2^{\Lambda_i}}\right)^2. \tag{34}$$

Substituting (33) and (34) into inequality (32), and taking the mathematical expectation of it, we deduce the following inequality:

$$\mathbb{E}\{\mathcal{V}(k+1)\} < (1-\alpha_2)\mathbb{E}\{\mathcal{V}(k)\} + \Theta \tag{35}$$

where

$$\Theta \triangleq n\bar{v}^2 + \sum_{i=1}^{n}\left(\frac{\delta_i}{2^{\Lambda_i}}\right)^2 + \max_{\imath\in\Pi}\left\{\mathrm{tr}\left(\bar{\Psi}L_\imath^T\hat{\mathcal{P}}_\jmath L_\imath\bar{\Psi}\right)\right\}.$$

Furthermore, we obtain from (35) that

$$\begin{aligned}
\mathbb{E}\{\mathcal{V}(k)\} &< (1-\alpha_2)\mathbb{E}\{\mathcal{V}(k-1)\} + \Theta\\
&< (1-\alpha_2)^2\mathbb{E}\{\mathcal{V}(k-2)\} + \big(1 + (1-\alpha_2)\big)\Theta\\
&< \cdots\\
&< (1-\alpha_2)^k\mathbb{E}\{\mathcal{V}(0)\} + \Theta\sum_{\varsigma=0}^{k-1}(1-\alpha_2)^\varsigma,
\end{aligned} \tag{36}$$

which implies

$$\mathbb{E}\{\|\tilde{x}(k)\|^2\} < \frac{(1-\alpha_2)^k\mathbb{E}\{\mathcal{V}(0)\}}{\min_{\imath\in\Pi}\underline{\lambda}(\mathcal{P}_\imath)} + \frac{\Theta\sum_{\varsigma=0}^{k-1}(1-\alpha_2)^\varsigma}{\min_{\imath\in\Pi}\underline{\lambda}(\mathcal{P}_\imath)}. \tag{37}$$

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication.
Citation information: DOI 10.1109/TNNLS.2024.3411484, IEEE Transactions on Neural Networks and Learning Systems

Recalling Definition 1, the EED (24) is exponentially mean-square bounded. Furthermore, the following inequality holds as k goes to infinity:

$$\mathbb{E}\{\|\tilde{x}(k)\|^2\} < \frac{\Theta}{\alpha_2 \min_{\imath\in\Pi} \underline{\lambda}(\mathcal{P}_\imath)} \triangleq \bar{O}, \tag{38}$$

which means that the upper bound of the EED is represented by $\bar{O}$. The proof is complete. ■

Remark 2: The estimation error bound $\bar{O}$ in digital communication networks is influenced by several factors, including noise, coding-decoding parameters δ_i, bit rate Λ_i, variance Ψ_i, gains $L_\imath$, and decay coefficient α_2. When system parameters (like δ_i, $L_\imath$, α_2, and bit flip probabilities $\bar{p}_i$) are fixed, the error bound is primarily dependent on the bit rate Λ_i for each neuron. An increase in Λ_i raises the maximum quantization level $\widehat{\Delta}_i$ and reduces the variance Ψ_i of the decoding signal, leading to a lower error bound. This relationship underscores the importance of bit rate management in minimizing estimation errors in digital communication networks.

B. Estimator Design Under Optimized Decay Rate

Having analyzed the boundedness of the EED, the next step would be to focus on the design of estimator by taking into account the specified bit rate allocation. The following theorem shows the optimization of the decay rate of the EED as a way to acquire the fastest convergence performance, with a known bit rate allocation Λ_i ($i \in \mathcal{I}$).

Theorem 2: Consider the MJNNs (2) and the estimator (22) with known positive integers Λ_i ($i \in \mathcal{I}$). The EED is exponentially mean-square bounded if there exist positive scalars $\alpha_1, \gamma_1, \gamma_2$, $0 < \alpha_2 < 1$, positive definite and symmetric matrices $\mathcal{P}_\imath$, and matrices $\tilde{\mathcal{L}}_\imath$ and Q, such that the following inequalities hold for all $\imath \in \Pi$:

$$\begin{bmatrix} \tilde{\Omega}_{11} & 0 & \tilde{\Omega}_{13} \\ * & \Omega_{22} & \tilde{\Omega}_{23} \\ * & * & -\hat{\mathcal{P}}_\jmath \end{bmatrix} < 0, \ \jmath \in \Pi \tag{39}$$

$$\begin{bmatrix} -Q & \alpha_2 I \\ * & \mathcal{P}_\imath - 2I \end{bmatrix} < 0 \tag{40}$$

$$\mathcal{P}_\imath \geq I \tag{41}$$

where

$$\tilde{\Omega}_{11} \triangleq \begin{bmatrix} -\alpha_1 \acute{M}^T \grave{M} - \mathcal{P}_\imath + Q & \alpha_1 \frac{\acute{M}^T + \grave{M}}{2} \\ * & -\alpha_1 I \end{bmatrix},$$

$$\tilde{\Omega}_{13} \triangleq \begin{bmatrix} \hat{\mathcal{P}}_\jmath H_\imath - \tilde{\mathcal{L}}_\imath (I - 2\bar{P}) C_\imath & \hat{\mathcal{P}}_\jmath W_\imath \end{bmatrix}^T,$$

$$\tilde{\Omega}_{23} \triangleq \begin{bmatrix} \hat{\mathcal{P}}_\jmath A_\imath - \tilde{\mathcal{L}}_\imath (I - 2\bar{P}) B_\imath & \hat{\mathcal{P}}_\jmath (I - 2\bar{P}) \end{bmatrix}^T.$$

Moreover, the optimal decay rate of EED $\|\tilde{x}(k)\|$ is derived by solving the following maximization problem:

$$\begin{aligned} &\max\{\alpha_2\} \\ &\text{s.t.} \quad (39) - (41), \end{aligned} \tag{42}$$

and the estimator gain is given by $L_\imath = \hat{\mathcal{P}}_\jmath^{-1} \tilde{\mathcal{L}}_\imath$.

Proof: According to the characteristic of the positive definite matrix $\mathcal{P}_\imath$ and condition (41), we have

$$(\mathcal{P}_\imath - I)\mathcal{P}_\imath^{-1}(\mathcal{P}_\imath - I)^T \geq 0, \tag{43}$$

which implies

$$\mathcal{P}_\imath - 2I \geq -\mathcal{P}_\imath^{-1}. \tag{44}$$

Combining (40) with (44), and applying the Schur Complement Lemma to the inequality (40), it is evident that

$$-Q + \alpha_2^2 \mathcal{P}_\imath < 0. \tag{45}$$

It is inferred from inequality (39) that

$$\begin{bmatrix} \breve{\Omega}_{11} & 0 & \tilde{\Omega}_{13} \\ * & \Omega_{22} & \tilde{\Omega}_{23} \\ * & * & -\hat{\mathcal{P}}_\jmath \end{bmatrix} < 0 \tag{46}$$

where

$$\breve{\Omega}_{11} \triangleq \begin{bmatrix} -\alpha_1 \acute{M}^T \grave{M} + (\alpha_2^2 - 1)\mathcal{P}_\imath & \alpha_1 \frac{\acute{M}^T + \grave{M}}{2} \\ * & -\alpha_1 I \end{bmatrix}.$$

Along the line of the proof of Theorem 1, we obtain

$$\mathbb{E}\{\|\tilde{x}(k)\|^2\} < \frac{(1 - \alpha_2^2)^k \mathbb{E}\{\mathcal{V}(0)\}}{\min_{\imath\in\Pi} \underline{\lambda}(\mathcal{P}_\imath)} + \frac{\Theta}{\alpha_2^2 \min_{\imath\in\Pi} \underline{\lambda}(\mathcal{P}_\imath)} \tag{47}$$

where Θ is defined in inequality (35), $1 - \alpha_2^2$ determines the decay rate of the EED and, finally, the error upper bound is obtained as

$$\tilde{O} \triangleq \frac{\Theta}{\alpha_2^2 \min_{\imath\in\Pi} \underline{\lambda}(\mathcal{P}_\imath)}.$$

In this case, the optimal decay rate of EED is obtained by solving the maximization problem (42), and the proof is complete. ■

C. Co-design of Bit Rate Allocation Scheme and Estimator

The allocation of bit rates to each neuron in MJNNs significantly influences estimator gain, as highlighted in Theorem 2. Given fixed system parameters and quantization regions, the bit rate Λ_i directly impacts the upper bound of the error dynamics and overall estimation performance. This section aims to minimize the upper bound of the EED by co-designing the bit rate allocation strategy and the estimator gain, optimizing both for enhanced system performance.

Corollary 1: When the bit rate Λ_i for each neuron is treated as a variable, the optimization for the error bound is transformed into the following minimization problem with a given scalar $\bar{\alpha}_2$ ($0 < \bar{\alpha}_2 < 1$):

$$\begin{aligned} &\min \ \frac{\Theta}{\bar{\alpha}_2 \min_{\imath\in\Pi} \underline{\lambda}(\mathcal{P}_\imath)} \\ &\text{s.t.} \quad \Upsilon < 0 \ (\forall \imath \in \Pi), \ 0 \leq \Lambda_i \leq \Lambda \end{aligned} \tag{48}$$

where

$$\Upsilon \triangleq \begin{bmatrix} \Upsilon_{11} & 0 & \Upsilon_{13} \\ * & \Omega_{22} & \Upsilon_{23} \\ * & * & -\hat{\mathcal{P}}_\jmath \end{bmatrix},$$

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication.
Citation information: DOI 10.1109/TNNLS.2024.3411484, IEEE Transactions on Neural Networks and Learning Systems

Algorithm 1 Optimal bit-rate allocation strategy and estimator co-design algorithm.

Initialization: Initialize parameters $\mathbf{N}$, $\mathbf{I}$, $\mathbf{w}$, $\mathbf{c_1}$, $\mathbf{c_2}$, position $\mathbf{X}_o$, and velocity $\mathbf{V}_o$ of each particle $\big(o \in \{1, 2, \ldots, \mathbf{N}\}\big)$;
if $\Upsilon < 0$ is feasible **then**
  Compute the fitness function $\mathbf{F}(\mathbf{X}_o)$
else
  $\mathbf{F}(\mathbf{X}_o) = \infty$
end if
for $\varrho = 1 : \mathbf{I}$ **do**
  Update $\mathbf{X}_o$ and $\mathbf{V}_o$ of the particle swarm according to formulas (50) and (51)
  for $o = 1 : \mathbf{N}$ **do**
    if $\Upsilon < 0$ is feasible **then**
      Compute $\mathbf{F}(\mathbf{X}_o)$ with updated $\mathbf{X}_o$, denoted as $\mathbf{F}(\hat{\mathbf{X}}_o)$
    else
      $\mathbf{F}(\hat{\mathbf{X}}_o) = \infty$
    end if
    if $\mathbf{F}(\hat{\mathbf{X}}_o) < \mathbf{F}(\mathbf{X}_o)$ **then**
      $\mathbf{F}(\mathbf{X}_o) = \mathbf{F}(\hat{\mathbf{X}}_o)$
    end if
  end for
  Update the historical optimum fitness and location of particle swarm;
end for
Obtain the particle with the minimum fitness, whose corresponding position is the optimal bit-rate allocation scheme;
The estimator gain $L_\imath$ is obtained by solving (49) under optimal bit-rate allocation protocol.

$$\Upsilon_{11} \triangleq \begin{bmatrix} -\alpha_1 \acute{M}^T \grave{M} - (1-\bar{\alpha}_2)\mathcal{P}_\imath & \alpha_1 \frac{\acute{M}^T + \grave{M}}{2} \\ * & -\alpha_1 I \end{bmatrix},$$
$$\Upsilon_{13} \triangleq \begin{bmatrix} \hat{\mathcal{P}}_\jmath H - \tilde{\mathcal{L}}_\imath (I - 2\bar{P}) C_\imath & \hat{\mathcal{P}}_\jmath W_\imath \end{bmatrix}^T,$$
$$\Upsilon_{23} \triangleq \begin{bmatrix} \hat{\mathcal{P}}_\jmath A_\imath - \tilde{\mathcal{L}}_\imath (I - 2\bar{P}) B_\imath & \hat{\mathcal{P}}_\jmath (I - 2\bar{P}) \end{bmatrix}^T.$$

Within this framework, the estimator gain is derived by $L_\imath = \hat{\mathcal{P}}_\jmath^{-1} \tilde{\mathcal{L}}_\imath$.

Proof: Define $\mathbb{I} \triangleq \{I, I, I, I, \hat{\mathcal{P}}_\jmath\}$. Let the scalar α_2 in Theorem 1 be $\bar{\alpha}_2 \in (0, 1)$. By pre-multiplying the inequality (26) with $\mathbb{I}$ and post-multiplying it with $\mathbb{I}^T$, we see that the condition $\Upsilon < 0$ holds and the proof is now complete. ■

To tackle the non-convex nature of the minimization problem outlined in (48), which presents considerable challenges in terms of solvability, we propose an innovative co-design method. This method combines the particle swarm optimization (PSO) algorithm with the linear matrix inequality (LMI) technique. The integration of PSO (known for its effectiveness in solving non-convex optimization problems) with the LMI technique (widely used for handling control and estimation problems) offers a powerful approach, which aims to efficiently navigate the solution space and find an optimal (or near-optimal) solution to the minimization problem.

The minimization problem (48) involves constraint $0 \leq \Lambda_i \leq \Lambda$. To effectively handle this constraint within the optimization process, a transformation of (48) is undertaken by introducing a penalty function:

$$\begin{aligned} \min \quad & \frac{\Theta}{\bar{\alpha}_2 \min_{\imath \in \Pi} \underline{\lambda}(\mathcal{P}_\imath)} + \eta \mathcal{F}(\tilde{\Lambda}) \\ \text{s.t.} \quad & \Upsilon < 0 \ (\forall \imath \in \Pi) \end{aligned} \tag{49}$$

where $\mathcal{F}(\tilde{\Lambda}) \triangleq \max\{0, \sum_{i=1}^{n} \Lambda_i - \Lambda\}$ is the exterior penalty function with $\tilde{\Lambda} \triangleq [\Lambda_1, \Lambda_2, \ldots, \Lambda_n]$, and η is a constant called penalty coefficient. The fitness function of PSO algorithm is the upper bound of the error dynamics, which is defined as

$$\mathbf{F}(\tilde{\Lambda}) \triangleq \frac{\Theta}{\bar{\alpha}_2 \min_{\imath \in \Pi} \underline{\lambda}(\mathcal{P}_\imath)} + \eta \mathcal{F}(\tilde{\Lambda}).$$

Algorithm 1 demonstrates the integration of the PSO algorithm with the LMI technique, aligned with the specified objective function. This algorithm is tailored to tackle the minimization problem, considering the inherent constraints and nonlinearity of the problem. Within the PSO algorithm, a swarm of particles (representing a potential solution) explores the search space. The position and velocity of each particle characterize these potential solutions. The algorithm iteratively refines the positions of the particles, relying not only on each particle's individual experience but also on insights gained from the best-performing particles in the population. Through this collaborative process, the PSO algorithm aims to find the optimal solution to the minimization problem.

Denote $\mathbf{X}_o \triangleq [\mathbf{X}_{o,1}, \mathbf{X}_{o,2}, \ldots, \mathbf{X}_{o,N}]$ and $\mathbf{V}_o \triangleq [\mathbf{V}_{o,1}, \mathbf{V}_{o,2}, \ldots, \mathbf{V}_{o,N}]$ $(o \in \{1, 2, \ldots \mathbf{N}\})$ as the position and velocity of the o-th particle, respectively. $\mathbf{N}$ is the number of particles in the search space, and the maximum number of iterations is represented by $\mathbf{I}$. $\mathbf{P}_o$ denotes the o-th particle's best position, and $\mathbf{P}_g$ represents the global best position of the particle swarm. The updates of particle velocity and position obey the following equations:

$$\begin{aligned} \mathbf{V}_o(\varrho+1) =& \mathbf{w}\mathbf{V}_o(\varrho) + \mathbf{c_1}\xi_1\big(\mathbf{P}_o(\varrho) - \mathbf{X}_o(\varrho)\big) \\ & + \mathbf{c_2}\xi_2\big(\mathbf{P}_g(\varrho) - \mathbf{X}_o(\varrho)\big), \end{aligned} \tag{50}$$
$$\mathbf{X}_o(\varrho+1) = \mathbf{X}_o(\varrho) + \mathbf{V}_o(\varrho) \tag{51}$$

where $\varrho \in \{1, 2, \ldots, \mathbf{I}\}$ indicates the iteration number; $\mathbf{w}$ stands for the inertia weight; the acceleration constants $\mathbf{c_1}$ and $\mathbf{c_2}$ denote the self-learning factor and the group learning factor, respectively; ξ_1 and ξ_2 are two stochastic integers distributed in the interval $[1, 2]$. To prevent the particle's search position from exceeding the limited interval leading to an unproductive blind search, well-defined boundaries are established for both position and velocity. The boundaries are denoted as

$$\mathbf{V}_L \leq \mathbf{V}_o(\varrho) \leq \mathbf{V}_T, \ \mathbf{X}_L \leq \mathbf{X}_o(\varrho) \leq \mathbf{X}_T.$$

Through the PSO-based co-design method for estimators, we achieve the optimal allocation strategy. This enables a thorough analysis of how varying bit rates impact the estimation performance of MJNNs.

Remark 3: In this study, we've investigated the remote estimation problem for MJNNs in digital communication networks by focusing on the challenges of limited bit rates and

Citation information: DOI 10.1109/TNNLS.2024.3411484, IEEE Transactions on Neural Networks and Learning Systems

probabilistic bit flips. Key achievements include i) establishing a sufficient condition for mean-square boundedness of the EED in Theorem 1, ii) determining mode-dependent estimator gains under specific bit rate allocations in Theorem 2, and iii) co-designing a bit-rate allocation strategy with optimal estimator gains in Corollary 1. These contributions are crucial for enhancing estimation accuracy and efficiency in constrained digital network environments.

Remark 4: This paper makes several innovative contributions to the field of state estimation for MJNNs, distinguishing itself from existing research in key aspects.

1) *Focus on Discrete-Time MJNNs in Digital Networks*: We specifically target discrete-time MJNNs within digital communication networks, emphasizing the impact of constrained bit rates in wireless networks. This perspective is crucial in understanding how limited bit rates affect network performance.
2) *Probabilistic Bit Flips in Constrained Networks*: For the first time, our study considers probabilistic bit flips in bit-rate constrained networks, analyzing the additional decoding error they introduce. We also explore how constrained bit rates impact estimation performance, providing new insights into this area.
3) *Optimization of Decay Rate for Fast Convergence*: We achieve the fastest convergence performance by optimizing the decay rate. The application of the PSO algorithm to allocate bit rates effectively compresses the upper bound of the EED. This innovative approach is unique in the context of state estimation for MJNNs.
4) *Co-Design Approach for Estimator Gains and Bit-Rate Allocation*: We propose a co-design strategy that involves optimizing estimator gains and developing a bit-rate allocation protocol. This comprehensive approach aims to enhance the overall estimation performance of the network, tackling the challenges posed by MJNNs in a holistic manner.

IV. Illustrative Example

In this section, a simulation example is presented to demonstrate the effectiveness of the estimator under constrained bit rates.

Consider the MJNNs composed of $i = 3$ neurons with two jumping topologies (i.e., $\Pi = \{1, 2\}$), where the system model parameters are given as follows:

$$
\begin{aligned}
&H_1 = \mathrm{diag}\{0.7, 0.9, 0.6\},\ C_1 = \mathrm{diag}\{0.9, 0.8, 1\},\\
&A_1 = \mathrm{diag}\{0.2, 0.2, 0.3\},\ B_1 = \mathrm{diag}\{0.2, 0.1, 0.4\},\\
&H_2 = \mathrm{diag}\{0.8, 0.9, 0.85\},\ C_2 = \mathrm{diag}\{0.85, 1, 0.8\},\\
&A_2 = \mathrm{diag}\{0.1, 0.1, 0.2\},\ B_2 = \mathrm{diag}\{0.2, 0.3, 0.1\},\\
&W_1 = \begin{bmatrix} -0.3 & 0.1 & 0.2 \\ 0.1 & -0.3 & 0.2 \\ 0.1 & 0.2 & -0.3 \end{bmatrix},\\
&W_2 = \begin{bmatrix} -0.5 & 0.2 & 0.3 \\ 0.2 & -0.4 & 0.2 \\ 0.1 & 0.4 & -0.5 \end{bmatrix}.
\end{aligned}
$$

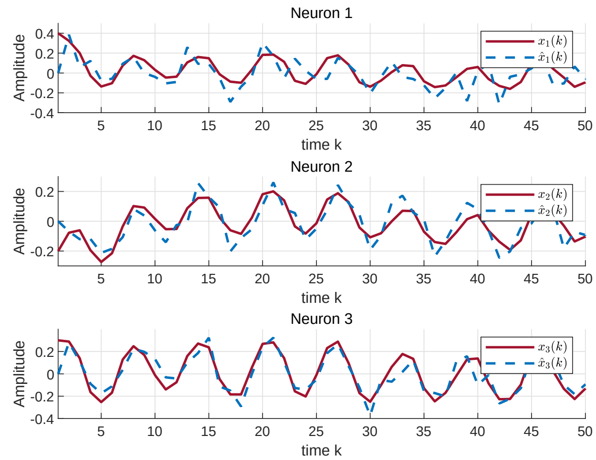


Fig. 2. State and estimation for the neurons.

The nonlinear function $g_j(\cdot)$ is of the following form:

$$g_j(x_j(k)) = 0.28 \tanh(x_j(k)).$$

The external noise is set as $v_i(k) = 0.8 \cos(k)$ with $\bar{v} = 0.8$, and the initial state and corresponding state estimate are provided as

$$
\begin{aligned}
&x_1(0) = 0.4,\ x_2(0) = 0.2,\ x_3(0) = 0.3,\\
&\hat{x}_1(0) = \hat{x}_2(0) = \hat{x}_3(0) = 0.
\end{aligned}
$$

The flip probabilities of each bit for different neuron are given as

$$\bar{p}_1 = 0.1,\ \bar{p}_2 = 0.05,\ \bar{p}_3 = 0.12.$$

Based on the aforementioned parameter settings, the estimation performance of MJNNs is analyzed under the maximization problem of decay rate $1 - \alpha_2^2$ in Theorem 2 and various bit rate allocation protocols in Corollary 1, respectively.

Scenario 1: Firstly, we employ an average allocation strategy (AAS) to compute the estimator gains, which ensures that each neuron in MJNNs is assigned with identical bit rates, thereby guaranteeing an equitable distribution of network resources. Based on Theorem 2, assume that the available bit rates of the entire wireless network are $\Lambda = 30$. We have $\Lambda_1 = \Lambda_2 = \Lambda_3 = \lfloor \Lambda/3 \rfloor = 10$ bps by AAS. The parameters δ_i of the quantization region are chosen as $\delta_1 = 1,\ \delta_2 = 0.5,\ \delta_3 = 0.7$. According to (42), the optimized decay rate parameter is $\alpha_2 = 0.9559$. Correspondingly, we obtain the estimator gains as

$$
\begin{aligned}
&l_{1,1} = 0.9150,\ l_{2,1} = 1.1908,\ l_{3,1} = 0.7377,\\
&l_{1,2} = 1.0597,\ l_{2,2} = 0.9178,\ l_{3,2} = 1.2686.
\end{aligned}
$$

The system states and their estimates are depicted in Fig. 2. The error norm $\|\tilde{x}(k)\|$ and the estimation error bound $\sqrt{\tilde{O}} = 2.123$ are plotted in Fig. 3, which verifies that the estimation error is indeed exponentially mean-square bounded.

Scenario 2: In some specific application scenarios, adopting an AAS might not be the most optimal approach, because

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication.
Citation information: DOI 10.1109/TNNLS.2024.3411484, IEEE Transactions on Neural Networks and Learning Systems

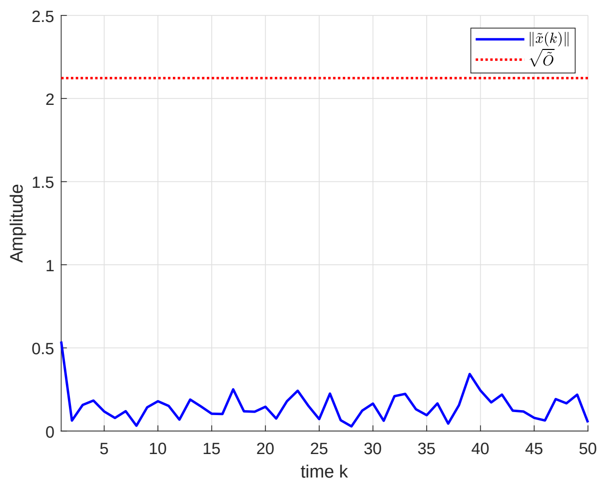


Fig. 3. Estimate error and bound.

TABLE II
EFFECT OF DIFFERENT PROTOCOLS ON THE ERROR BOUND

Parameters	Λ (bps)	Protocol	Bit rate allocation Λ_1, Λ_2, Λ_3 (bps)	Bound
δ_1=1 δ_2=0.5 δ_3=0.7	30	AAS	10, 10, 10	2.3037
	30	PSO	11, 9, 10	2.2035
	20	AAS	6, 6, 6	2.3045
	20	PSO	7, 6, 7	2.3040
	10	AAS	3, 3, 3	2.3519
	10	PSO	4, 3, 3	2.3309
δ_1=0.4 δ_2=1.25 δ_3=2	30	AAS	10, 10, 10	2.4561
	30	PSO	9, 10, 11	2.4560
	20	AAS	6, 6, 6	2.4585
	20	PSO	5, 7, 8	2.4568
	10	AAS	3, 3, 3	2.6017
	10	PSO	2, 3, 5	2.5068

certain nodes may necessitate higher transmission speeds for carrying out more complex tasks compared to others. In such scenario, to enhance the performance of the MJNNs, it is advisable to use the error bound as a metric and employ the PSO algorithm to dynamically adjust the bit rate allocation strategy. The superiority of the PSO-based bit rate allocation strategy over the AAS becomes evident through the following analysis.

Setting $\bar{\alpha}_2 = 0.92$, and given the quantization parameters as $\delta_1 = 1$, $\delta_2 = 0.5$ and $\delta_3 = 0.7$, the error bound is obtained using both the AAS and PSO-based allocation methods in Table II. The variation of the error bound is also analyzed for a set of different quantization parameters $\delta_1 = 0.4$, $\delta_2 = 1.25$ and $\delta_3 = 2$. The PSO algorithm is observed to not only maximize the utilization of network resources but also optimize bit rate allocation according to the specific demands of each node, thereby enhancing the estimation performance of the MJNNs. Additionally, it is inferred that an increase in available bit rates Λ correlates with a gradual decrease in the error bound. The quantization parameters play a crucial role in determining the decoding accuracy of the data, and a more suitable parameter setting is expected to result in an overall reduction in estimation errors.

V. CONCLUSION

In this work, the bounded state estimation problem has been addressed for MJNNs within a digital network framework. The measurement outputs from MJNNs, which are transmitted over wireless networks to a remote estimator, have been subjected to bit rate constraints. A coding-decoding process has been modeled that accounts for probabilistic bit flips during wireless transmission, and the effects of these bit flips on the decoded output have been detailed. By utilizing the structural characteristics of MJNNs, a mode-dependent estimator has been developed. Within this framework, a sufficient condition has been derived for ensuring the exponential boundedness of the EED. An upper bound on the error dynamics has been established and, subsequently, two optimization problems in the estimator gain design section have been introduced. The first is aimed at optimizing the decay rate of the EED to achieve the fastest convergence performance, while the second focuses on reducing the error upper bound by incorporating a PSO algorithm for optimal bit rate allocation. Finally, the effectiveness of the proposed estimation strategy has been demonstrated, and a detailed analysis of the relationship between estimation performance and constrained bit rates has been provided. In future research, to address bit flips occurring in wireless networks, it may be considered to enhance data transmission reliability by utilizing retransmission mechanisms or employing multipath transmission methods, thereby further reducing coding-decoding error.

REFERENCES

[1] G. Aceto, D. Ciuonzo, A. Montieri and A. Pescapé, Mobile encrypted traffic classification using deep learning: experimental evaluation, lessons learned, and challenges, *IEEE Transactions on Network and Service Management*, vol. 16, no. 2, pp. 445-458, Feb. 2019.

[2] G. Aceto, D. Ciuonzo, A. Montieri and A. Pescapé, Toward effective mobile encrypted traffic classification through deep learning, *Neurocomputing*, vol. 409, pp. 306-315, Oct. 2020.

[3] S. Arik, A modified Lyapunov functional with application to stability of neutral-type neural networks with time delays, *Journal of the Franklin Institute*, vol. 356, no. 1, pp. 276-291, Jan. 2019.

[4] J. Cao, D. Ding, J. Liu, E. Tian, S. Hu and X. Xie, Hybrid-triggered-based security controller design for networked control system under multiple cyber attacks, *Information Sciences*, vol. 548, pp. 69-84, Feb. 2021.

[5] B. Chen, Y. Chen, G. Zeng and Q. She, Fractional-order convolutional neural networks with population extremal optimization, *Neurocomputing*, vol. 477, pp. 36-45, Mar. 2022.

[6] G. Chen, J. Xia, J. H. Park, H. Shen and G. Zhuang, Sampled-data synchronization of stochastic Markovian jump neural networks with time-varying delay, *IEEE Transactions on Neural Networks and Learning Systems*, vol. 33, no. 8, pp. 3829-3841, Aug. 2022.

[7] R. Chen and Q. Ling, Event-triggered feedback stabilisation of switched linear systems under finite bit rates, *IET Control Theory & Applications*, vol. 14, no. 20, pp. 3428-3439, Feb. 2021.

[8] J. Cheng, L. Liang, H. Yan, J. Cao, S. Tang and K. Shi, Proportional-integral observer-based state estimation for Markov memristive neural networks with sensor saturations, *IEEE Transactions on Neural Networks and Learning Systems*, vol. 35, no. 1, pp. 405-416, Jan. 2024.

[9] D. Christmann, R. Gotzhein, S. Siegmund and F. Wirth, Realization of try-once-discard in wireless multihop networks, *IEEE Transactions on Industrial Informatics*, vol. 10, no. 1, pp. 17-26, Feb. 2014.

[10] D. Ciuonzo, A. Aubry and V. Carotenuto, Rician MIMO channel- and jamming-aware decision fusion, *IEEE Transactions on Signal Processing*, vol. 65, no. 15, pp. 3866-3880, Aug. 2017.

[11] H. Dai, G. Zhao, M. Lin, J. Wu and G. Zheng, A novel estimation method for the state of health of lithium-ion battery using prior knowledge-based neural network and Markov chain, *IEEE Transactions on Industrial Electronics*, vol. 66, no. 10, pp. 7706-7716, Oct. 2019.

[12] H. Geng, Z. Wang, A. Mousavi, F. E. Alsaadi and Y. Cheng, Outlier-resistant filtering with dead-zone-like censoring under try-once-discard protocol, *IEEE Transactions on Signal Processing*, vol. 70, pp. 714-728, Jan. 2022.

[13] X. Kan, Y. Fan, Z. Fang, L. Cao, N. N. Xiong, D. Yang and X. Li, A novel IoT network intrusion detection approach based on adaptive particle swarm optimization convolutional neural network, *Information Sciences*, vol. 568, pp. 147-162, Aug. 2021.

[14] R. Lescanne, M. Villiers, T. Peronnin, A. Sarlette, M. Delbecq, B. Huard, T. Kontos, M. Mirrahimi and Z. Leghtas, Exponential suppression of bit-flips in a qubit encoded in an oscillator. *Nature Physics*, vol. 16, pp. 509-513, Mar. 2020.

[15] J.-Y. Li, Z. Wang, R. Lu and Y. Xu, A component-based coding-decoding approach to set-membership filtering for time-varying systems under constrained bit rate, *Automatica*, vol. 152, art. no. 110874, Mar. 2023.

[16] W. J. Lin, Y. He, C. K. Zhang and M. Wu, Stochastic finite-time H_∞ state estimation for discrete-time semi-Markovian jump neural networks with time-varying delays, *IEEE Transactions on Neural Networks and Learning Systems*, vol. 31, no. 12, pp. 5456-5467, Dec. 2020.

[17] C. Liu, H. Rao, X. Yu, Y. Xu and C.-Y. Su, State estimation for recurrent neural networks with intermittent transmission, *IEEE Transactions on Cybernetics*, vol. 54, no. 5, pp. 2891-2900, May 2024.

[18] L. Liu, L. Ma, J. Guo, J. Zhang and Y. Bo, Distributed set-membership filtering for time-varying systems: a coding-decoding-based approach, *Automatica*, vol. 129, art. no. 109684, Jul. 2021.

[19] Q. Liu, Z. Wang, H. Dong and C. Jiang, Remote estimation for energy harvesting systems under multiplicative noises: a binary encoding scheme with probabilistic bit flips, *IEEE Transactions on Automatic Control*, vol. 68, no. 1, pp. 343-354, Jan. 2023.

[20] Y. Liu and Q. Ling, Stabilizing bit rate conditions for a scalar linear event-triggered system with Markov dropouts, *Automatica*, vol. 146, art. no. 110635, Dec. 2022.

[21] Y. Liu, B. Shen and H. Shu, Finite-time resilient H_∞ state estimation for discrete-time delayed neural networks under dynamic event-triggered mechanism, *Neural Networks*, vol. 121, pp. 356-365, Jan. 2020.

[22] Y. Liu, B. Shen and P. Zhang, Synchronization and state estimation for discrete-time coupled delayed complex-valued neural networks with random system parameters, *Neural Networks*, vol. 150, pp. 181-193, Jun. 2022.

[23] Y. Luo, Z. Wang, W. Sheng and D. Yue, State estimation for discrete time-delayed impulsive neural networks under communication constraints: a delay-range-dependent approach, *IEEE Transactions on Neural Networks and Learning Systems*, vol. 34, no. 3, pp. 1489-1501, Mar. 2023.

[24] W. Qian, W. Xing and S. Fei, L_∞ state estimation for neural networks with general activation function and mixed time-varying delays, *IEEE Transactions on Neural Networks and Learning Systems*, vol. 32, no. 9, pp. 3909-3918, Sept. 2021.

[25] F. Qu, E. Tian and X. Zhao, Chance-constrained H_∞ state estimation for recursive neural networks under deception attacks and energy constraints: the finite-hrizon case, *IEEE Transactions on Neural Networks and Learning Systems*, in press, DOI: 10.1109/TNNLS.2021.3137426.

[26] D. Quevedo, A. Ahlen, 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, Jul. 2012.

[27] A. S. Rakin, Z. He, J. Li, F. Yao, C. Chakrabarti and D. Fan, T-bfa: Targeted bit-flip adversarial weight attack, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 44, no. 11, pp. 7928-7939, Nov. 2022.

[28] H. Shen, X. Hu, J. Wang, J. Cao and W. Qian, Non-fragile H_∞ synchronization for Markov jump singularly perturbed coupled neural networks subject to double-layer switching regulation, *IEEE Transactions on Neural Networks and Learning Systems*, vol. 34, no. 5, pp. 2682-2692, May 2023.

[29] Y. Shen, W. Song, H. Ji, Y. Ren, C. Ji, X. You and C. Zhang, Improved belief propagation polar decoders with bit-flipping algorithms, *IEEE Transactions on Communications*, vol. 68, no. 11, pp. 6699-6713, Nov. 2020.

[30] Y. Shen, Z.-G. Wu, P. Shi, H. Su and T. Huang, Asynchronous filtering for Markov jump neural networks with quantized outputs, *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, vol. 49, no. 2, pp. 433-443, Feb. 2019.

[31] Q. Song, Y. Chen, Z. Zhao, Y. Liu and F. E. Alsaadi, Robust stability of fractional-order quaternion-valued neural networks with neutral delays and parameter uncertainties, *Neurocomputing*, vol. 420, pp. 70-81, Jan. 2021.

[32] G. Tan and Z. Wang, Reachable set estimation of delayed Markovian jump neural networks based on an improved reciprocally convex inequality, *IEEE Transactions on Neural Networks and Learning Systems*, vol. 33, no. 6, pp. 2737-2742, Jun. 2022.

[33] H. Tan, B. Shen and H. Shu, Robust recursive filtering for stochastic systems with time-correlated fading channels, *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, vol. 52, no. 5, pp. 3102-3112, May 2022.

[34] Z. Tang, E. Tian, Y. Wang, L. Wang and T. Yang, Nondestructive defect detection in castings by using spatial attention bilinear convolutional neural network, *IEEE Transactions on Industrial Informatics*, vol. 17, no. 1, pp. 82-89, Jan. 2021.

[35] J. Tao, Z. Xiao, Z. Li, R. Lu, P. Shi and X. Wang, Dynamic event-triggered state estimation for Markov jump neural networks with partially unknown probabilities, *IEEE Transactions on Neural Networks and Learning Systems*, vol. 33, no. 12, pp. 7438-7447, Dec. 2022.

[36] Y. Tian and Z. Wang, Extended dissipativity analysis for Markovian jump neural networks via double-integral-based delay-product-type Lyapunov functional, *IEEE Transactions on Neural Networks and Learning Systems*, vol. 32, no. 7, pp. 3240-3246, Jul. 2021.

[37] R. Vijay Aravind and P. Balasubramaniam, Stochastic stability of fractional-order Markovian jumping complex-valued neural networks with time-varying delays, *Neurocomputing*, vol. 439, pp. 122-133, Jun. 2021.

[38] J. Wang, Z. Wang, X. Chen and J. Qiu, Synchronization criteria of delayed inertial neural networks with generally Markovian jumping, *Neural Networks*, vol. 139, pp. 64-76, Jul. 2021.

[39] Y. Wang, T. Bian, J. Xiao and C. Wen, Global synchronization of complex dynamical networks through digital communication with limited data rate, *IEEE Transactions on Neural Networks and Learning Systems*, vol. 26, no. 10, pp. 2487-2499, Oct. 2015.

[40] Y. Wang, J. Lu, X. Li and J. Liang, Synchronization of coupled neural networks under mixed impulsive effects: a novel delay inequality approach, *Neural Networks*, vol. 127, pp. 38-46, Jul. 2020.

[41] M. Xing, J. Lu, J. Qiu and H. Shen, Synchronization of complex dynamical networks subject to DoS attacks: an improved coding-decoding protocol, *IEEE Transactions on Cybernetics*, vol. 53, no. 1, pp. 102-113, Jun. 2023.

[42] L. Xu and S. Ge, Exponential ultimate boundedness of nonlinear stochastic difference systems with time-varying delays, *International Journal of Control*, vol. 88, no. 5, pp. 983-989, Jan. 2015.

[43] Y. Xu, Z. G. Wu, Y. J. Pan and J. Sun, Resilient asynchronous state estimation for Markovian jump neural networks subject to stochastic nonlinearities and sensor saturations. *IEEE Transactions on Cybernetics*, vol. 52, no. 7, pp. 5809-5818, Jul. 2022.

[44] Y. Xu, H. Xiang, L. Yang, R. Lu, and D. E. Quevedo, Optimal transmission strategy for multiple Markovian fading channels: existence, structure, and approximation, *Automatica*, vol. 158, art. no. 111312, Dec. 2023.

[45] F. Yang, J. Li, H. Dong and Y. Shen, Proportional-integral-type estimator design for delayed recurrent neural networks under encoding-decoding mechanism, *International Journal of Systems Science*, vol. 53, no. 13, pp. 2729-2741, Oct. 2022.

[46] L. Yang, H. Rao, M. Lin, Y. Xu and P. Shi, Optimal sensor scheduling for remote state estimation with limited bandwidth: a deep reinforcement learning approach, *Information Sciences*, vol. 588, pp. 279-292, Apr. 2022.

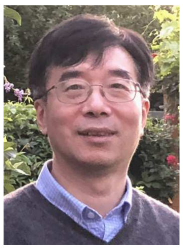
[47] L. Yao, Z. Wang, X. Huang, Y. Li, Q. Ma and H. Shen, Stochastic sampled-data exponential synchronization of Markovian jump neural networks with time-varying delays, *IEEE Transactions on Neural Networks and Learning Systems*, vol. 34, no. 2, pp. 909-920, Feb. 2023.

[48] Y. Zhang, L. Zou, Y. Wang and Y.-A. Wang, Estimator design for complex networks with encoding decoding mechanism and buffer-aided strategy: a partial-nodes accessible case, *ISA Transactions*, vol. 127, pp. 68-79, Aug. 2022.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication.
Citation information: DOI 10.1109/TNNLS.2024.3411484, IEEE Transactions on Neural Networks and Learning Systems

Yuru Guo was born in Hubei province, China, in 1997. She received her B.S. degree from the school of electrical engineering, Nanjing Institute of Technology, Jiangsu, China, in 2019, majoring in Building Electricity and Intelligence. She is currently working toward the Ph.D. degree in Control Science and Engineering at Guangdong University of Technology, Guangzhou, China.

She was a Visiting Ph.D. student with the Department of Computer Science, Brunel University London, Uxbridge, U.K., from 2023 to 2024. Her research interests include networked control systems and impulsive control.



Zidong Wang (Fellow, IEEE) 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, 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.

Jun-Yi Li was born in Henan province, China, in 1991. He received the M.S. degree in control science and engineering from Hangzhou Dianzi University, Hangzhou, China, in 2017 and the Ph.D. degree in Control Science and Engineering from Guangdong University of Technology, Guangzhou, China, in 2021. He is currently a Lecturer with the school of Automation, Guangdong University of Technology, Guangzhou, China.

He was a visiting Ph.D. student with the Department of Computer Science, Brunel University London, Uxbridge, U.K., from October 2019 to October 2020. His research interests include complex dynamical networks, networked systems.

Yong Xu (Member, IEEE) was born in Zhejiang Province, China, in 1983. He received the B.S. degree in information engineering from Nanchang Hangkong University, Nanchang, China, in 2007, the M.S. degree in control science and engineering from Hangzhou Dianzi University, Hangzhou, China, in 2010, and the Ph.D. degree in control science and engineering from Zhejiang University, Hangzhou, China, in 2014. He was a visiting internship student with the department of Electronic and Computer Engineering, Hong Kong University of Science and Technology, Hong Kong, China, from June 2013 to November 2013, where he was a Research Fellow from February 2018 to August 2018. Now he is a professor with School of Automation, at Guangdong University of Technology, Guangzhou, China.

His research interests include PID control, networked control systems, state estimation, and positive systems.