On Model of Recurrent Neural Network on a Time Scale: Exponential Convergence and Stability Research

Vasyl Martsenyuk, Marcin Bernas, and Aleksandra Klos-Witkowska

***Abstract*— The majority of the results on modeling recurrent neural networks (RNNs) are obtained using delayed differential equations, which imply continuous time representation. On the other hand, these models must be discrete in time, given their practical implementation in computer systems, requiring their versatile utilization across arbitrary time scales. Hence, the goal of this research is to model and investigate the architecture design of a delayed RNN using delayed differential equations on a time scale. Internal memory can be utilized to describe the calculation of the future states using discrete and distributed delays, which is a representation of the deep learning architecture for artificial RNNs. We focus on qualitative behavior and stability study of the system. Special attention is paid to taking into account the effect of the time-scale parameters on neural network dynamics. Here, we delve into the exploration of exponential stability in RNN models on a time scale that incorporates multiple discrete and distributed delays. Two approaches for constructing exponential estimates, including the Hilger and the usual exponential functions, are considered and compared. The Lyapunov–Krasovskii (L–K) functional method is employed to study stability on a time scale in both cases. The established stability criteria, resulting in an exponential-like estimate, utilizes a tuple of positive definite matrices, decay rate, and graininess of the time scale. The models of RNNs for the two-neuron network with four discrete and distributed delays, as well as the ring lattice delayed network of seven identical neurons, are numerically investigated. The results indicate how the time scale (graininess) and model characteristics (weights) influence the qualitative behavior, leading to a transition from stable focus to quasiperiodic limit cycles.**

***Index Terms*— Delayed dynamic system, exponential stability, Hilger function, recurrent neural network (RNN), time scale.**

NOMENCLATURE

$\mathbb{T}$	Time scale, i.e., a subset of real numbers.
$\mathbb{R}$	Real numbers.
$\sigma(t)$	Forward jump operator.
$\rho(t)$	Backward jump operator.
$\mu(t)$	Graininess function.
$x(t) \in \mathbb{R}^n$	State vector.
A	Diagonal matrix with positive entries.
$W_{x,y}$	Synaptic connection weight matrices.
$g(x(t))$	Nondecreasing activation function.
$h_k(t)$, $\tau_m(t)$	Delays ("controllable memory").
$\phi(s)$	Initial state of RNN model.
$u(t)$	Input signal vector.
$e_p(t,s)$	Hilger exponential function for regressive function $p(t)$.
e^x	Classic exponential function.

Manuscript received 24 July 2023; revised 31 December 2023; accepted 12 March 2024. This work was supported in part by the Erasmus + Program for Education of the European Union through the Key Action 2 Grant (the Future Is in Applied Artificial Intelligence) under Grant 2022-1-PL01-KA220-HED000088359 (work package 5: "Piloting," activity A5.6 "Project deliverables") and in part by the University of Bielsko-Biala. *(Corresponding author: Vasyl Martsenyuk.)*

The authors are with the Department of Computer Science and Automatics, University of Bielsko-Biala, 43-309 Bielsko-Biała, Poland (e-mail: vmartsenyuk@ubb.edu.pl; mbernas@ubb.edu.pl; aklos@ubb.edu.pl).

Digital Object Identifier 10.1109/TNNLS.2024.3377446

I. INTRODUCTION

RECURRENT neural networks (RNNs) are at the heart of today's deep machine learning (DL) research topics. In particular, recurrent neural systems with LSTM and GRU facilitated the development of machine learning models for problems like image identification with captioning, language processing, and interpretation. These models are presented in [1], which makes use of ordinary differential equations. As they model the memory inside the organized units, their research must be connected to the equations with time delays. The RNN definition was formally inferred from differential equations in [2]. In literature, two general types of delays are researched [3], [4], i.e., discrete delays and distributed delays.

A. Stability

One of the goals of RNN research is to find a stable neural network and, therefore, obtain predictable results. Zhang et al. [5] presents a review concerning the stability of RNNs based on delayed differential equations. It was shown that the most commonly used stability criteria up-to-date were obtained using the Lyapunov–Krasovskii (L–K) theory. In [5] and [6], authors show how L–K functionals should be constructed (with an estimation of its derivatives). Nevertheless, this research considered one delay mostly. Additionally, research shows that neural networks with time delays can exhibit chaotic attractors and periodic oscillations [7] which should be taken under consideration. The most of research focuses on using L–K functionals which are called the direct method; stability conditions are presented in terms

of solutions of linear-matrix inequalities (LMIs). Additionally, in [8] and [9], an indirect method for exponential stability research—not using L–K functionals and resulting in some scalar quasipolynomial inequalities was introduced. The technique has been applied to discrete and distributed delays also. The survey [10] highlights that existing stability criteria are conservative and should be further developed to account for multiple delays. Authors in previous work [11] developed a technique to optimize such stability conditions for L–K functionals. They constructed the L–K functional including the classic exponential function which allowed us to obtain stability conditions in the form of LMI for cases with multiple discrete and distributed delays. That functional will be used as a prototype for the given work when considering time scales.

B. Time Scales

It is a modern approach [12] in the field of dynamical systems that allows us to effectively incorporate the results obtained for differential equations, i.e., systems with continuous time, to systems with discrete time, such as based on difference equations. This approach is extremely important for models of RNNs, as it allows us to extend the results obtained based on traditional models with continuous time, to the case of discrete-time, which corresponds to the software implementation of predictive algorithms for RNNs. Moreover, some important results on the qualitative analysis of models of RNNs can be summarized in the case of general time scales, which will be the subject of this work. In contrast to the existing results on direct stability methods for RNN, this article extends them to the time scale. The approach combines the advantages of differential and difference equations. It allows us to transition from a continuous time scale to a discrete one by utilizing the granularity operator $\mu(t)$. For continuous cases, $\mu(t)$ approaches 0, while for discrete cases take other positive values. The numerical examples illustrate how the granularity influences the appearance of bifurcations (Hopf).

C. Motivation

The article was inspired by the idea to close the models of dynamical systems to RNNs with the purpose of their stability investigation. When solving machine learning problems with the help of RNNs, like regression or classification, we are interested in stable solutions of RNNs, whereas, unstable solutions can deliver error outputs for predicting or classifying. In turn, the stability of the solution is affected by RNN parameters (weights, biases, and initial conditions). Moreover, stability can also be numerically characterized by exponential decay rate. The bigger the decay rate, the faster the model tends to the correct result of predicting or classifying. Provided that weights and biases have been determined by training algorithms like backpropagation through time, however, initial conditions of the model's internal state can affect the stability. Stability theory is well developed for continuous-time systems because the entire power of calculus can be applied. Trying to cover the models of RNNs acting in discrete time, with the same tools, we apply the formalism of time scales, aimed for studying hybrid systems.

The method will be described in the following sections. Section II presents the proposed model in reference to the state-of-the-art approaches. Section III presents the two approaches for stability estimation using time scales, followed by examples presented in Section IV. Section V presents the discussion of the proposed model. Finally, in Section VI, the conclusions, limitations, and further research directions are given.

II. Proposed Model

The main idea of this research is to incorporate time scales [12] to RNN, which allows to treat systems with continuous time as systems with discrete time. Here, our reasoning concerning dynamic systems on time scales is based on the definitions from the work [12]. Namely, we consider time scale $\mathbb{T}$ as a subset of real numbers $\mathbb{R}$. Hence, $\mathbb{R}$ and integers $\mathbb{Z}$ are examples of time scales.

Remark 1: For any $t \in \mathbb{T}$, we introduce the following denotions for *forward jump operator* $\sigma(t) := \inf_{s\in\mathbb{T}}\{s : s > t\}$, *backward jump operator* $\rho(t) := \sup_{s\in\mathbb{T}}\{s : s < t\}$, *graininess function* $\mu(t) := \sigma(t) - t$.

Hereinafter, the notions of left dense, right dense, left scattered, and right scattered elements from $\mathbb{T}$ are used. Let

$$\mathbb{T}^k := \begin{cases} \mathbb{T}\backslash m, & \text{if } \mathbb{T} \text{ has a maximal left scattered element } m \\ \mathbb{T}, & \text{otherwise.} \end{cases}$$

For any function $f : \mathbb{T} \to \mathbb{R}$ the *delta derivative* at instance $t \in \mathbb{T}^k$, $f^{\Delta}(t)$, is as follows:

$$\|f(\sigma(t)) - f(s) - f^{\Delta}(t)(\sigma(t) - s)\| \leq \epsilon\|\sigma(t) - s\|$$

for some $\epsilon > 0$ and neighborhood $U \subset \mathbb{T}$ of t.

When considering the delay differential equations on time scales, two approaches describing the delays $t - \tau(t)$, $t \in \mathbb{T}^k$ are used [12]. The first one requires that the delay function is mapping time scale $\mathbb{T}^k$ on itself, i.e., $t - \tau(t) : \mathbb{T}^k \to \mathbb{T}^k$ [13]. The second one is using a jump operator of a special kind. Following the work [14], we introduce the *forward semijump operator*:

$$\hat{\sigma}(t) := \inf_{s\in\mathbb{T}}\{s : s \geq t\}.$$

Note that $\hat{\sigma}(t) \leq \sigma(t)$.

Here, the system with several discrete and distributed time-varying delays on a time scale $t \in \mathbb{T}^k$ is the foundation of our study of the RNN model

$$x^{\Delta}(t) = -Ax(t) + \sum_{k=1}^{r_1} W_{1,k} g\big(x(\hat{\sigma}(t - h_k(t)))\big) + \sum_{m=1}^{r_2} W_{2,m} \int_{t-\tau_m(t)}^{t} g\big(x(\hat{\sigma}(\theta))\big)\Delta(\theta) \quad (1)$$

where $x(t) \in \mathbb{R}^n$ is the state vector, $A = \text{diag}(a_1, a_2, \ldots, a_n)$ is a diagonal matrix with positive entries $a_i > 0$. For the ith neuron, $1/a_i$ can be interpreted as the activity decay constant (or time constant). $W_{1,k} = (w_{ij}^{1,k})_{n\times n}$, $k = \overline{1, r_1}$, $W_{2,m} = (w_{ij}^{2,m})_{n\times n}$, $m = \overline{1, r_2}$ are the synaptic connection weight matrices. The entries of $W_{1,k}$ and $W_{2,m}$ may be positive (excitatory synapses) or negative (inhibitory synapses).

$g(x(t)) = [g_1(x_1(t)), g_2(x_2(t)), \ldots, g_n(x_n(t))]^\top \in \mathbb{R}^n$ is the nondecreasing activation function, which belongs to sector nonlinear function class defined by

$$g_j(0) = 0 \text{ and } 0 \le \frac{g_j(\xi_1) - g_j(\xi_2)}{\xi_1 - \xi_2} \le l_j, \quad l_j > 0 \tag{2}$$

$\xi_1, \xi_2 \in \mathbb{R}$, $\xi_1 \ne \xi_2$, $j \in \overline{1,n}$ and $x \equiv 0$ is a fixed point of (1).[1] We let $L = \operatorname{diag}(l_1, l_2, \ldots, l_n)$ is a diagonal matrix with positive entries $l_j > 0$.

The system (1) contains discrete and distributed time-varying delays, which are, respectively, characterized by the second and third terms.

The discrete system delays are represented by the bounded lambda differentiable functions $h_k(t)$ if

$$0 \le h_k(t) \le h_{M,k}$$

and

$$h_k^{\Delta}(t) \le h_{D,k} < 1 \tag{3}$$

$k = \overline{1, r_1}$, $t > 0$. If prior states of neurons only have an effect on output at specific time intervals, delays $h_k(t)$ and $\tau_m(t)$ have a physical meaning as "controllable memory." For discrete delays, the boundaries on the delay and its delta derivative are $h_{M,k}$ and $h_{D,k}$.

The distributed delays of the system are represented by the bounded functions $\tau_m(t)$ such that $0 \le \tau_m(t) \le \tau_{M,m}$, $m = \overline{1, r_2}$.

The delays in axonal signal transmission are represented by the limited functions $h_k(t)$ and $\tau_m(t)$. When calculating the upper right delta derivative of the L–K functional, the condition (3) for the delta derivative $h_k^{\Delta}(t)$ will be used (see, for instance, [15]).

We will consider the class of rd-continuous functions, $\mathbb{C}_{\mathrm{rd}}$, which contains the functions that are continuous at right-dense points in $\mathbb{T}^k$ and their left-side limits exist at left-dense points in $\mathbb{T}^k$.

It is presummated that system (1) initial states are

$$\begin{aligned} x(s) &= \phi(s), \quad s \in [-\tau_M, 0] \cap \mathbb{T}^k \\ \tau_M &:= \max\left\{h_{M,k}, k = \overline{1, r_1}, \tau_{M,m}, m = \overline{1, r_2}\right\} \end{aligned} \tag{4}$$

where $\phi(s) \in \mathbb{C}([-\tau_M, 0], \mathbb{R}^n)$ is continuous.

In the work [16], there were obtained the general existence and stability results for functional differential equations utilizing the induction principle and Gronwall's inequality on time scales. For the right-hand side of the system, they imply right dense continuity in t, continuity in phase coordinates, and local Lipshitz conditions. When using them for any $\phi(s) \in \mathbb{C}([-\tau_M, 0], \mathbb{R}^n]$, it can be assumed that there exists a certain trajectory of the (1) starting from ϕ according to (2).

In this case, we make use of the Hopfield neural network, which includes a diagonal matrix A with positive elements and makes the neuron's self-connection appear. The neuron's next state is dependent on its current state and results in all neurons eventually. In the stability analysis of continuous-time RNNs, such a diagonal matrix is typically connected [5]. On the other hand, if we used any matrix A, we would accept that a neuron's next state is dependent on both its current state and the states of all other neurons, meaning that all neurons' internal states are visible from the outside. It refutes the idea that, for instance, the hidden state vector (also known as the output vector) is visible in the case of the LSTM unit.

Following work [2], we may explain the model (1) from the standpoint of signal processing, leading us to canonical and noncanonical RNNs. The state signal vector $x(t)$, the readout signal vector $g(x(t))$, which may be a twisted version of the state signal vector, and the bias parameters are specifically mentioned without loss of generality because they can be used in the transformations leading to the homogeneous system (1). One-to-many RNN design is modeled by taking the initial state $\phi(s)$, $s \in [-\tau_M, 0] \cap \mathbb{T}^k$ into consideration as an input signal. In a more typical many-to-many instance, the input signal vector $u(t)$, $t \in \mathbb{T}^k$ can be used as an input sequence while the RNN is working.

It makes sense to take into account continuous-time equations that describe the functioning of RNNs even if RNNs can really be described using difference equations. This is because differential equations allow us to more fully grasp and describe the dynamic processes that take place. Additionally, the requirements for the stability of RNNs can be explicitly obtained with the aid of differential equations. When designing RNNs, this is crucial. A thorough analysis of the research on continuous-time RNNs with a focus on the stability of Hopfield and Cohen–Grossberg neural networks can be found in the study [5].

Also take note that by discretizing the models based on differential equations, the matching RNNs can be constructed. As a result, work [2] illustrates how to build an RNN of the LSTM type by first starting with the relevant model based on differential equations with delay and then moving on to discretize the so-called canonical RNN.

RNN models have been studied since Hopfield's groundbreaking work in the 1980s [17], which represented each neuron as a linear circuit made up of a resistor and a capacitor. When exploring the models of RNNs in the class of delayed differential equations, two techniques can be distinguished. The first method entails investigating local stability by comparing it to the linearized system [18], [19], [20], [21]. The Hopf bifurcation conditions were discovered in [21] and [22]. L–K functionals are used in the second approach (known as the direct Lyapunov's method) [15]. It enables us to get constructively stability criteria, which are expressed as LMIs. The parameters of L–K functionals can be modified to improve these stability conditions.

When investigating RNN models, exponential estimates of the solutions are critical since they reveal the rate of convergence of calculations when recognizing input data. The indirect method was established in previous publications [8], [9], allowing us to acquire exponential estimates in some typical scenarios of RNN models. The numerical solution of the quasipolynomial problem is the result. It offers us a clear value for exponential decay, which, however, cannot be

[1]The proposed method of stability investigation can be applied for any steady-state $u^\star$ (not only trivial) of the nonlinear system $u^{\Delta}(t) = f(u(t))$ on a time scale, i.e., $f(u^\star) = 0$. Without loss of generality, using linear transformation $x(t) = u(t) - u^\star$, this problem can be reduced to studying the stability of trivial solution for the system $x^{\Delta}(t) = f(x(t) + u^\star)$.

improved because it does not allow for optimization. In order to address this problem, in [11], we developed an optimization strategy based on the direct method for L–K functionals of the peculiar form. On an arbitrary time scale $\mathbb{T}$, Lyapunov functions of this type will be applied in the given work.

III. Exponential Estimates

We denote by $\Omega_n \subset \mathbb{R}^{n\times n}$ a set of symmetric positively definite matrices. Below, we show that it is a convex cone.

1) Convexity holds as for arbitrary $P_1 \in \Omega_n$, $P_2 \in \Omega_n$, $x \in \mathbb{R}^n$, and $\xi \in [0,1]$ we get $x^\top(\xi P_1 + (1-\xi)P_2)x = \xi x^\top P_1 x + (1-\xi)x^\top P_2 x > 0$.
2) Since for any $P \in \Omega_n$, $x \in \mathbb{R}^n$, and $\eta > 0$ we get $\eta x^\top P x > 0$, the set is cone.

We denote $\bar{\Omega}_n^1 := \{P \in \Omega_n : \|P\| \le 1\}$ as the part of Ω_n enclosed within the unit sphere.

For arbitrary $P \in \Omega_n$, let $\|P\|$ be its Frobenius norm.

Lemma 1: Given matrix $U \in \Omega_n$, constant $\beta > 0$, $\beta \in \mathbb{T}$, vector function $u : [0,\beta] \to \mathbb{R}^n$, it holds

$$\left(\int_0^\beta u^\top(s)\Delta(s)\right) U \left(\int_0^\beta u(s)\Delta(s)\right) \le \beta \int_0^\beta u^\top(s) U u(s)\Delta(s)$$

with the corresponding requirements of integrability

Proof: We can see that the block matrix

$$\begin{bmatrix} u^\top(s)Uu(s) & u^\top(s) \\ u(s) & U^{-1} \end{bmatrix}, \quad 0 \le s \le \beta$$

is positive semidefinite using Schur's complement. We get the positive semidefinite matrix

$$\begin{bmatrix} \int_0^\beta u^\top(s)Uu(s)\Delta(s) & \int_0^\beta u^\top(s)\Delta(s) \\ \int_0^\beta u(s)\Delta(s) & \beta U^{-1} \end{bmatrix}$$

by integrating on a time scale from 0 to β and applying the features of integration on time scale [23]. The proof is completed using Schur's complement. □

Lemma 2 [24]: Given real matrices W_1, W_2, and W_3, $W_3 \in \Omega_n$ with proper dimensions and a constant $\beta > 0$, it holds

$$W_1^\top W_2 + W_2^\top W_1 \le \beta W_1^\top W_3 W_1 + \beta^{-1} W_2^\top W_3^{-1} W_2.$$

Lemma 3 (Leibniz Integral Rule on Time Scales [25]): If f, f^Δ are continuous, and u, $v : \mathbb{T} \to \mathbb{T}$ are delta differentiable functions, then the following formula holds for any $t \in \mathbb{T}^k$:

$$\begin{aligned}\left[\int_{u(t)}^{v(t)} f(t,s)\Delta(s)\right]^\Delta = &\int_{u(t)}^{v(t)} f^\Delta(t,s)\Delta(s) \\ &+ v^\Delta(t) f(\sigma(t), v(t)) \\ &- u^\Delta(t) f(\sigma(t), u(t)).\end{aligned}$$

It is worth noting that the σ (Remark 1) was used to change the continuous time to time scale—to make the value discrete. Here, we define the main notions for the trivial solution of (1) related to the unique equilibrium point. The stability notions from [16] will be used.

Definition 1: If for every solution $x(t)$ of (1)–(4) we get $x(t) \to 0$ as $t \to \infty$, $t \in \mathbb{T}$, then the trivial solution of (1) is called globally asymptotically stable.

A. Exponential Estimate: Hilger Function

When regarding the exponential stability on a time scale, they consider two approaches. The first one is introducing the Hilger function, a special function with properties like exponential function, but based on a time scale. The second one is using the traditional exponential function but applying the time scale calculus.

Hilger function is related to the notion of regressive functions.

Definition 2: A function $p(t) : \mathbb{T} \to \mathbb{R}$ is called regressive, if $1 + \mu(t)p(t) \ne 0$ at $t \in \mathbb{T}$.

First, we will use exponential function for regressive function $p(t) \in \mathbb{C}_{\text{rd}}$ on time scale $\mathbb{T}$ introduced by Hilger in [26] as follows:

$$e_p(t,s) := \exp\left(\int_s^t \zeta_{\mu(u)}(p(u))\Delta(u)\right), \quad s,t \in \mathbb{T}$$

where

$$\zeta_{\mu(u)}(p(u)) := \begin{cases} \dfrac{1}{\mu(u)}\log(1+\mu(u)p(u)), & \mu(u) \ne 0 \\ p(u), & \mu(u) = 0. \end{cases}$$

To differ $e_p(t,s)$ from the exponential function on $\mathbb{R}$, we will call it as Hilger exponential function or Hilger function.

Further, we will also use the denotation $e_p(t) \equiv e_p(t,0)$ for simplicity.

Remark 2: It can be shown that if p, μ are constants, then for $\mathbb{T} = \mathbb{Z}$, we get $e_p(t,s) = (1+p\mu)^{(t-s)/\mu}$. Moreover, if $\mathbb{T} = \mathbb{R}$, then $e_p(t,s) = e^{p(t-s)}$, which could be obtained tending $\mu \to 0$ at the expression for $\mathbb{Z}$.

Hilger functions constitute the Abelian group for regressive functions with the addition operation $\oplus$ as follows:

$$p \oplus q := p + q + \mu p q$$

and additive inverse $\ominus$

$$\ominus p := -\frac{p}{1+\mu p}.$$

Such operations allowed us to obtain a few basic properties of the Hilger function like the exponential function ([27, Th. 3.1]), which will be used further.

Definition 3: The trivial solution of (1) is called globally Hilger-exponentially stable (GHE-stable), if there exist constants $\alpha > 0$, $K > 0$, and $T > 0$ such that every solution $x(t)$ to (1)–(4) it holds $\|x(t)\| \le K e_{\ominus\alpha}(t,0)$ for all $t > T$, $t \in \mathbb{T}$.

Here, we exploit the following L–K functional, which was proposed in [11] and [15], but here we upgrade it considering multiple delays and Hilger exponential function for time scale

$$\begin{aligned} &V[x_t(\cdot)] \\ &= e_{\alpha\oplus\alpha}(t)x^\top(t)Px(t) \\ &\quad + \sum_{k=1}^{r_1}\int_{t-h_k(t)}^t e_{\alpha\oplus\alpha}(\hat{\sigma}(s))g^\top(x(\hat{\sigma}(s)))Q_k g(x(\hat{\sigma}(s)))\Delta(s) \\ &\quad + \sum_{m=1}^{r_2}\tau_{M,m} \end{aligned}$$

$$\times \int_{-\tau_{M,m}}^{0} \int_{t+\theta}^{t} e_{\alpha\oplus\alpha}(\hat{\sigma}(s)) g^{\top}(x(\hat{\sigma}(s))) S_m g(x(\hat{\sigma}(s))) \Delta(s) \Delta(\theta) \tag{5}$$

where unknown constant $\alpha > 0$ and matrices P, Q_k, $k = \overline{1, r_1}$, S_m, $m = \overline{1, r_2}$ belong to Ω_n. We apply denotation $x_t(\cdot) := \{x(\hat{\sigma}(t+\theta)) | \theta \in [-\tau_M, 0]\} \in \mathbb{C}[-\tau_M, 0]$ of the element of the solution of (1).

In the sequel, we will use the following notion of the functional's derivative as a development of the definition from [16] for Lyapunov functions under Razumikhin's condition.

Definition 4: Given a functional $V \in \mathbb{C}_{\mathrm{rd}}([-\tau_M, 0], \mathbb{R}^n)$, the *upper right-hand delta derivative* of V with respect to system $x^{\Delta}(t) = f(x_t)$ is defined by

$$D^{+}V^{\Delta}[x_t(\cdot)] = \begin{cases} \dfrac{V(\sigma(t), x(\sigma(t))) - V(x(t))}{\mu(t)}, & \sigma(t) > t \\ \limsup\limits_{s \to t^{+}} \dfrac{V(x(t) + (s-t) f(x_t)) - V(t, x(t))}{s - t}, & \sigma(t) = t \end{cases}$$

where $\mathbb{C}_{\mathrm{rd}}([-\tau_M, 0], \mathbb{R}^n)$ is the space of rd-continuous in t, and continuous in $x \in \mathbb{C}([-\tau_M, 0], \mathbb{R}^n)$.

The next result will be formulated in terms of LMI with respect to matrix

$$\Gamma^H = \begin{bmatrix} \Gamma_{11}^H & \Gamma_{12}^H \\ \Gamma_{21}^H & \Gamma_{22}^H \end{bmatrix} \in \mathbb{R}^{2n \times 2n}, \quad \xi(t) = \Big(x(t), x(\sigma(t))\Big)^{\top} \in \mathbb{R}^{2n}$$

where

$$\begin{aligned} \Gamma_{11}^H &= -(\alpha \oplus \alpha) P - e_{\alpha\oplus\alpha}(\mu(t)) \Bigg\{ \big[-A^{\top} P - P A \\ &\quad + \mu(t) A^{\top} P A \big] + \sum_{k=1}^{r_1} \Bigg\{ e_{\alpha\oplus\alpha}(h_{M,k}) (1 - h_{D,k})^{-1} \\ &\quad \times (I - \mu(t) A^{\top}) P W_{1,k} Q_k^{-1} W_{1,k}^{\top} P (I - \mu(t) A) \\ &\quad + L Q_k L \} + \sum_{m=1}^{r_2} \{ \tau_{M,m}^2 L S_m L + e_{\alpha\oplus\alpha}(\tau_{M,m}) \\ &\quad \times (I - \mu(t) A^{\top}) P W_{2,m} S_m^{-1} W_{2,m}^{\top} P (I - \mu(t) A) \Bigg\} \Bigg\} \\ \Gamma_{12}^H &= e_{\alpha\oplus\alpha}(\mu(t)) (\mu^{-1}(t) I - A) P \\ \Gamma_{21}^H &= e_{\alpha\oplus\alpha}(\mu(t)) P (\mu^{-1}(t) I - A) \\ \Gamma_{22}^H &= -\frac{e_{\alpha\oplus\alpha}(\mu(t))}{\mu(t)} P. \end{aligned} \tag{6}$$

Note that hereinafter in case $\mu(t) = 0$, one should consider the limit values of the corresponding expressions.

We see that $\Gamma^H \in \Omega_{2n}$ if and only if the Schur complement of Γ^H in Γ_{22}^H $\Gamma^H / \Gamma_{22}^H := \Gamma_{11}^H - \Gamma_{12}^H (\Gamma_{22}^H)^{-1} \Gamma_{21}^H \in \Omega_n$.

In further reasonings, we will use the definitions and general stability results for delayed systems on time scales from [14].

Assumption 1: Let for time scale $\mathbb{T}$, there exist constant $\alpha > 0$ and matrices P, Q_k, $k = \overline{1, r_1}$, S_m, $m = \overline{1, r_2}$, which belong to relint$(\bar{\Omega}_n^1)$, such that

1) $e_{\alpha}(t) > 0$, $t \in \mathbb{T}^k$.
2) $\Gamma_{11}^H - \Gamma_{12}^H (\Gamma_{22}^H)^{-1} \Gamma_{21}^H$ belong to Ω_n for any $\mu(t)$, $t \in \mathbb{T}^k$.

The conditions of positivity of the Hilger function were presented in work [27].

Theorem 1: Let system (1) satisfies the Assumption 1. Then, the trivial solution of (1) is globally asymptotically stable.

Proof: Consider functional

$$V_1[x_t(\cdot)] = e_{\alpha\oplus\alpha}(t) x^{\top}(t) P x(t).$$

Estimating the right upper delta derivative of V_1, we get

$$\begin{aligned} &D^{+}V_1^{\Delta}[x_t(\cdot)] \\ &= (\alpha \oplus \alpha) e_{\alpha\oplus\alpha}(t) x^{\top}(t) P x(t) \\ &\quad + e_{\alpha\oplus\alpha}(\sigma(t)) \big\{ x^{\top} P x(t) \big\}^{\Delta} \\ &= (\alpha \oplus \alpha) e_{\alpha\oplus\alpha}(t) x^{\top}(t) P x(t) \\ &\quad + e_{\alpha\oplus\alpha}(\sigma(t)) \big\{ x^{\top\Delta}(t) P x(t) + x^{\top}(\sigma(t)) P x^{\Delta}(t) \big\} \\ &= (\alpha \oplus \alpha) e_{\alpha\oplus\alpha}(t) x^{\top}(t) P x(t) \\ &\quad + e_{\alpha\oplus\alpha}(\sigma(t)) \Bigg\{ - x^{\top}(t) A^{\top} P x(t) \\ &\quad + \sum_{k=1}^{r_1} g^{\top}\big(x(\hat{\sigma}(t - h_k(t)))\big) W_{1,k}^{\top} P x(t) \\ &\quad + \sum_{m=1}^{r_2} \int_{t-\tau_m(t)}^{t} g^{\top}\big(x(\hat{\sigma}(\theta))\big) \Delta(\theta) W_{2,m}^{\top} P x(t) \\ &\quad - x^{\top}(\sigma(t)) P A x(t) \\ &\quad + x^{\top}(\sigma(t)) P \sum_{k=1}^{r_1} W_{1,k} g\big(x(\hat{\sigma}(t - h_k(t)))\big) \\ &\quad + x^{\top}(\sigma(t)) P \sum_{m=1}^{r_2} W_{2,m} \int_{t-\tau_m(t)}^{t} g\big(x(\hat{\sigma}(\theta))\big) \Delta(\theta) \Bigg\}. \end{aligned} \tag{7}$$

Since

$$x(\sigma(t)) = x(t) + \mu(t) x^{\Delta}(t) \tag{8}$$

we get

$$\begin{aligned} &D^{+}V_1^{\Delta}[x_t(\cdot)] \\ &= (\alpha \oplus \alpha) e_{\alpha\oplus\alpha}(t) x^{\top}(t) P x(t) \\ &\quad + e_{\alpha\oplus\alpha}(t + \mu(t)) \Bigg\{ - x^{\top}(t) A^{\top} P x(t) \\ &\quad + \sum_{k=1}^{r_1} g^{\top}(x(\hat{\sigma}(t - h_k(t)))) W_{1,k}^{\top} P x(t) \\ &\quad + \sum_{m=1}^{r_2} \int_{t-\tau_m(t)}^{t} g^{\top}(x(\hat{\sigma}(\theta))) \Delta(\theta) W_{2,m}^{\top} P x(t) \\ &\quad + \Bigg[x^{\top}(t) - x^{\top}(t) \mu(t) A^{\top} \\ &\quad + \sum_{k=1}^{r_1} g^{\top}(x(\hat{\sigma}(t - h_k(t)))) \mu(t) W_{1,k}^{\top} \\ &\quad + \sum_{m=1}^{r_2} \int_{t-\tau_m(t)}^{t} g^{\top}(x(\hat{\sigma}(\theta))) \Delta(\theta) \mu(t) W_{2,m}^{\top} \Bigg] \\ &\quad \times \Bigg[- P A x(t) + P \sum_{k=1}^{r_1} W_{1,k} g(x(\hat{\sigma}(t - h_k(t)))) \end{aligned}$$

$$+P\sum_{m=1}^{r_2}W_{2,m}\int_{t-\tau_m(t)}^{t}g(x(\hat{\sigma}(\theta)))\Delta(\theta)\Bigg]\Bigg\}. \quad (9)$$

Now, we can rewrite (9) as follows[2]:

$$\begin{aligned}
&D^+V_1^{\Delta}[x_t(\cdot)]\\
&=(\alpha\oplus\alpha)e_{\alpha\oplus\alpha}(t)x^\top(t)Px(t)+e_{\alpha\oplus\alpha}(t+\mu(t))\\
&\quad\times\Bigg\{x^\top(t)\Big[-A^\top P-PA+\mu(t)A^\top PA\Big]x(t)\\
&\quad+\Bigg[\sum_{k=1}^{r_1}g^\top(x(\hat{\sigma}(t-h_k(t))))W_{1,k}^\top P(I-\mu(t)A)x(t)\\
&\qquad+x^\top(t)(I-\mu(t)A^\top)P\sum_{k=1}^{r_1}W_{1,k}g(x(\hat{\sigma}(t-h_k(t))))\Bigg]\\
&\quad+\Bigg[\sum_{m=1}^{r_2}\int_{t-\tau_m(t)}^{t}g^\top(x(\hat{\sigma}(\theta)))\Delta(\theta)W_{2,m}^\top P(I-\mu(t)A)x(t)\\
&\qquad+x^\top(t)(I-\mu(t)A^\top)\\
&\qquad\times P\sum_{m=1}^{r_2}W_{2,m}\int_{t-\tau_m(t)}^{t}g^\top(x(\hat{\sigma}(\theta)))\Delta(\theta)\Bigg]\\
&\quad+\Bigg[\sum_{k=1}^{r_1}g^\top(x(\hat{\sigma}(t-h_k(t))))W_{1,k}^\top\\
&\qquad+\sum_{m=1}^{r_2}\int_{t-\tau_m(t)}^{t}g^\top(x(\hat{\sigma}(\theta)))\Delta(\theta)W_{2,m}^\top\Bigg]\mu(t)P\\
&\quad\times\Bigg[\sum_{k=1}^{r_1}W_{1,k}g(x(\hat{\sigma}(t-h_k(t))))\\
&\qquad+\sum_{m=1}^{r_2}W_{2,m}\int_{t-\tau_m(t)}^{t}g(x(\hat{\sigma}(\theta)))\Delta(\theta)\Bigg]\Bigg\}. \quad (10)
\end{aligned}$$

Further, we estimate counterparts of (10) with the help of Lemma 2 in the following manner:

$$\begin{aligned}
&\sum_{k=1}^{r_1}\Big\{g^\top(x(\hat{\sigma}(t-h_k(t))))W_{1,k}^\top P(I-\mu(t)A)x(t)\\
&\qquad+x^\top(t)(I-\mu(t)A^\top)PW_{1,k}g(x(\hat{\sigma}(t-h_k(t))))\Big\}\\
&=\sum_{k=1}^{r_1}\Big\{\Big[e_{\ominus\alpha}(h_{M,k})(1-h_{D,k})^{1/2}g^\top(x(\hat{\sigma}(t-h_k(t))))\Big]\\
&\qquad\times\Big[e_{\alpha}(h_{M,k})(1-h_{D,k})^{-1/2}W_{1,k}^\top P(I-\mu(t)A)x(t)\Big]\\
&\qquad+\Big[e_{\alpha}(h_{M,k})(1-h_{D,k})^{-1/2}x^\top(t)(I-\mu(t)A^\top)PW_{1,k}\Big]\\
&\qquad\times\Big[e_{\ominus\alpha}(h_{M,k})(1-h_{D,k})^{1/2}g(x(\hat{\sigma}(t-h_k(t))))\Big]\Big\}\\
&\le\sum_{k=1}^{r_1}\Big\{[e_{\alpha\oplus\alpha}(h_{M,k})(1-h_{D,k})^{-1}\\
&\qquad\times x^\top(t)(I-\mu(t)A^\top)PW_{1,k}Q_k^{-1}W_{1,k}^\top P(I-\mu(t)A)x(t)\\
&\qquad+[e_{\ominus\alpha\ominus\alpha}(h_{M,k})(1-h_{D,k})\\
&\qquad\times g^\top(x(\hat{\sigma}(t-h_k(t))))Q_kg(x(\hat{\sigma}(t-h_k(t))))\Big\} \quad (11)
\end{aligned}$$

and

$$\begin{aligned}
&\sum_{m=1}^{r_2}\Bigg\{\int_{t-\tau_m(t)}^{t}g^\top(x(\hat{\sigma}(\theta)))\Delta(\theta)W_{2,m}^\top P(I-\mu(t)A)x(t)\\
&\qquad+x^\top(t)(I-\mu(t)A^\top)PW_{2,m}\int_{t-\tau_m(t)}^{t}g^\top(x(\hat{\sigma}(\theta)))\Delta(\theta)\Bigg\}\\
&=\sum_{m=1}^{r_2}\Bigg\{\Bigg[e_{\ominus\alpha}(\tau_{M,m})\int_{t-\tau_m(t)}^{t}g^\top(x(\hat{\sigma}(\theta)))\Delta(\theta)\Bigg]\\
&\qquad\times\Big[e_{\alpha}(\tau_{M,m})W_{2,m}^\top P(I-\mu(t)A)x(t)\Big]\\
&\qquad+\Big[e_{\alpha}(\tau_{M,m})x^\top(t)(I-\mu(t)A^\top)PW_{2,m}\Big]\\
&\qquad\times\Bigg[\int_{t-\tau_m(t)}^{t}g^\top(x(\hat{\sigma}(\theta)))\Delta(\theta)e_{\ominus\alpha}(\tau_{M,m})\Bigg]\Bigg\}\\
&\le\sum_{m=1}^{r_2}\Bigg\{e_{\ominus\alpha\ominus\alpha}(\tau_{M,m})\Bigg(\int_{t-\tau_m(t)}^{t}g^\top(x(\hat{\sigma}(\theta)))\Delta(\theta)\Bigg)\\
&\qquad\times S_m\Bigg(\int_{t-\tau_m(t)}^{t}g^\top(x(\hat{\sigma}(\theta)))\Delta(\theta)\Bigg)\\
&\qquad+e_{\alpha\oplus\alpha}(\tau_{M,m})x^\top(t)(I-\mu(t)A^\top)\\
&\qquad\times PW_{2,m}S_m^{-1}W_{2,m}^\top P(I-\mu(t)A)x(t)\Bigg\}\\
&\le\sum_{m=1}^{r_2}\Bigg\{\tau_{M,m}e_{\ominus\alpha\ominus\alpha}(\tau_{M,m})\int_{t-\tau_m(t)}^{t}g^\top(x(\hat{\sigma}(\theta)))\\
&\qquad\times S_mg(x(\hat{\sigma}(\theta)))\Delta(\theta)\\
&\qquad+e_{\alpha\oplus\alpha}(\tau_{M,m})x^\top(t)(I-\mu(t)A^\top)PW_{2,m}S_m^{-1}W_{2,m}^\top\\
&\qquad\times P(I-\mu(t)A)x(t)\Bigg\}. \quad (12)
\end{aligned}$$

Note that for the last inequality, Lemma 1 has been used. Further, we also use the identity

$$\begin{aligned}
&\sum_{k=1}^{r_1}W_{1,k}g(x(\hat{\sigma}(t-h_k(t))))\\
&\qquad+\sum_{m=1}^{r_2}W_{2,m}\int_{t-\tau_m(t)}^{t}g(x(\hat{\sigma}(\theta)))\Delta(\theta)\\
&\equiv x^{\Delta}(t)+Ax(t)=\mu^{-1}x(\sigma(t))-(\mu^{-1}I-A)x(t) \quad (13)
\end{aligned}$$

which follows from the system equation (1) and the expression (8) for lambda derivative.

Combining (10) with inequalities (11) and (12), and expression (13), we get

$$\begin{aligned}
&D^+V_1^{\Delta}[x_t(\cdot)]\\
&\le(\alpha\oplus\alpha)e_{\alpha\oplus\alpha}(t)x^\top(t)Px(t)\\
&\quad+e_{\alpha\oplus\alpha}(t+\mu(t))\Bigg\{x^\top(t)\Big[-A^\top P-PA+\mu(t)A^\top PA\Big]x(t)\\
&\quad+\sum_{k=1}^{r_1}\Big\{[e_{\alpha\oplus\alpha}(h_{M,k})(1-h_{D,k})^{-1}\times x^\top(t)\\
&\qquad\times(I-\mu(t)A^\top)PW_{1,k}Q_k^{-1}W_{1,k}^\top P(I-\mu(t)A)x(t)\\
&\qquad+e_{\ominus\alpha\ominus\alpha}(h_{M,k})(1-h_{D,k})g^\top(x(\hat{\sigma}(t-h_k(t))))\\
&\qquad\times Q_kg(x(\hat{\sigma}(t-h_k(t))))\Big\}\\
&\quad+\sum_{m=1}^{r_2}\Bigg\{\tau_{M,m}e_{\ominus\alpha\ominus\alpha}(\tau_{M,m})\\
&\qquad\times\int_{t-\tau_m(t)}^{t}g^\top(x(\hat{\sigma}(\theta)))S_mg(x(\hat{\sigma}(\theta)))\Delta(\theta)\\
&\qquad+e_{\alpha\oplus\alpha}(\tau_{M,m})x^\top(t)(I-\mu(t)A^\top)
\end{aligned}$$

[2] I denotes $n\times n$-identity matrix.

$$\times P W_{2,m} S_m^{-1} W_{2,m}^{\top} P (I - \mu(t)A)x(t)\Big\}$$
$$+\Big[\mu^{-1}x^{\top}(\sigma(t)) - (\mu^{-1}I - A)x^{\top}(t)\Big]\mu(t)P$$
$$\times\Big[\mu^{-1}x(\sigma(t)) - (\mu^{-1}I - A)x(t)\Big]\Big\}. \tag{14}$$

Introduce the following functionals:

$$V_2[x_t(\cdot)] = \sum_{k=1}^{r_1}\int_{t-h_k(t)}^{t} e_{\alpha\oplus\alpha}(\hat{\sigma}(s)) \times g^{\top}(x(\hat{\sigma}(s)))Q_k g(x(\hat{\sigma}(s)))\Delta(s)$$
$$V_3[x_t(\cdot)] = \sum_{m=1}^{r_2}\tau_{M,m}\int_{-\tau_{M,m}}^{0}\int_{t+\theta}^{t} e_{\alpha\oplus\alpha}(\hat{\sigma}(s)) \times g^{\top}(x(\hat{\sigma}(s)))S_m g(x(\hat{\sigma}(s)))\Delta(s)\Delta(\theta),$$
$$V[x_t(\cdot)] = \sum_{i=1}^{3} V_i[x_t(\cdot)].$$

Applying Lemma 3, we have

$$D^{+}V_2^{\Delta}[x_t(\cdot)] = \sum_{k=1}^{r_1}\Big[e_{\alpha\oplus\alpha}(t+\mu(t)) \times g^{\top}(x(\hat{\sigma}(t)))Q_k g(x(\hat{\sigma}(t))) - e_{\alpha\oplus\alpha}(t-h_k(t)+\mu(t-h_k(t)))(1-h_{D,k}) \times g^{\top}(x(\hat{\sigma}(t-h_k(t))))Q_k g(x(\hat{\sigma}(t-h_k(t))))\Big]$$

$$D^{+}V_3^{\Delta}[x_t(\cdot)] = \sum_{m=1}^{r_2}\tau_{M,m}\int_{-\tau_{M,m}}^{0}\Big[e_{\alpha\oplus\alpha}(t+\mu(t)) \times g^{\top}(x(\hat{\sigma}(t)))S_m g(x(\hat{\sigma}(t))) - e_{\alpha\oplus\alpha}(t+\theta+\mu(t+\theta))g^{\top}(x(\hat{\sigma}(t+\theta)))S_m \times g(x(\hat{\sigma}(t+\theta)))\Big]\Delta(\theta)$$
$$\le \sum_{m=1}^{r_2}\Big\{\tau_{M,m}^2 e_{\alpha\oplus\alpha}(t+\mu(t))g^{\top}(x(\hat{\sigma}(t)))S_m g(x(\hat{\sigma}(t))) - \tau_{M,m}e_{\alpha\oplus\alpha}(t+\mu(t))e_{\ominus\alpha\oplus(\ominus\alpha)}(\tau_{M,m}) \times\int_{-\tau_{M,m}}^{0} g^{\top}(x(\hat{\sigma}(\theta)))S_m g(x(\hat{\sigma}(\theta)))\Delta(\theta)\Big\}.$$

Finally, we get

$$D^{+}V^{\Delta}[x_t(\cdot)] \le (\alpha\oplus\alpha)e_{\alpha\oplus\alpha}(t)x^{\top}(t)Px(t) + e_{\alpha\oplus\alpha}(t+\mu(t))\Big\{x^{\top}(t)\Big[-A^{\top}P - PA + \mu(t)A^{\top}PA\Big]x(t) + \sum_{k=1}^{r_1}\{e_{\alpha\oplus\alpha}(h_{M,k})(1-h_{D,k})^{-1}x^{\top}(t)(I-\mu(t)A^{\top}) \times PW_{1,k}Q_k^{-1}W_{1,k}^{\top}P(I-\mu(t)A)x(t) + g^{\top}(x(\hat{\sigma}(t)))Q_k g(x(\hat{\sigma}(t)))\} + \sum_{m=1}^{r_2}\{\tau_{M,m}^2 g^{\top}(x(\hat{\sigma}(t)))S_m g(x(\hat{\sigma}(t))) + e_{\alpha\oplus\alpha}(\tau_{M,m})x^{\top}(t)(I-\mu(t)A^{\top})PW_{2,m}S_m^{-1}W_{2,m}^{\top}P \times(I-\mu(t)A)x(t)\}$$
$$+\Big[\mu^{-1}(t)x^{\top}(\sigma(t)) - (\mu^{-1}(t)I - A)x^{\top}(t)\Big] \times\mu(t)P\Big[\mu^{-1}(t)x(\sigma(t)) - (\mu^{-1}(t)I - A)x(t)\Big]\Big\}$$
$$\le e_{\alpha\oplus\alpha}(t)\Big\{x^{\top}(t)\Big\{(\alpha\oplus\alpha)P + e_{\alpha\oplus\alpha}(\mu(t))\Big\{\Big[-A^{\top}P - PA + \mu(t)A^{\top}PA\Big] + \sum_{k=1}^{r_1}\{e_{\alpha\oplus\alpha}(h_{M,k})(1-h_{D,k})^{-1}(I-\mu(t)A^{\top}) \times PW_{1,k}Q_k^{-1}W_{1,k}^{\top}P(I-\mu(t)A) + L\ Q_k L\} + \sum_{m=1}^{r_2}\{\tau_{M,m}^2 LS_m L + e_{\alpha\oplus\alpha}(\tau_{M,m})(I-\mu(t)A^{\top})PW_{2,m}S_m^{-1}W_{2,m}^{\top} \times P(I-\mu(t)A)\}\Big\}x(t) + e_{\alpha\oplus\alpha}(\mu(t))\Big[\mu^{-1}(t)x^{\top}(\sigma(t)) - (\mu^{-1}(t)I - A)x^{\top}(t)\Big]\mu(t)P\Big[\mu^{-1}(t)x(\sigma(t)) - (\mu^{-1}(t)I - A)x(t)\Big]\Big\}\Big\} = -e_{\alpha\oplus\alpha}(t)\xi^{\top}(t)\Gamma^{H}\xi(t). \tag{15}$$

We see that right-hand side of (6) is negative definite (i.e., $\Gamma^H \in \Omega_{2n}$) if and only if the Schur complement of Γ^H in Γ_{22}^H $\Gamma^H/\Gamma_{22}^H := \Gamma_{11}^H - \Gamma_{12}^H(\Gamma_{22}^H)^{-1}\Gamma_{21}^H \in \Omega_n$ [28].

Applying stability result from [14, Th. 4.1], we conclude that the system (1) is globally asymptotically stable. □

Corollary 1: The trivial solution of (1) is GHE-stable as long as Assumption 1 holds, i.e.,

$$\|x(t)\| \le \gamma^{H}(\alpha)|\phi|_{\tau_M}e_{\ominus\alpha}(t), \quad t\in\mathbb{T}^k \tag{16}$$

where

$$\gamma^{H}(\alpha) := \frac{1}{\lambda_{\min}^{1/2}(P)}\Bigg(\lambda_{\max}(P) + \sum_{k=1}^{r_1}\lambda_{\max}(Q_k)l_{\max}^2\int_{-h_{M,k}}^{0} e_{\alpha\oplus\alpha}(s)\Delta(s) + \sum_{m=1}^{r_2}\tau_{M,m}\lambda_{\max}(LS_mL)\int_{-\tau_{M,m}}^{0}\int_{\theta}^{0} e_{\alpha\oplus\alpha}(s)\Delta(s)\Delta(\theta)\Bigg)^{1/2}.$$

$\lambda_{\min}(\cdot)$ and $\lambda_{\max}(\cdot)$ are minimal and maximal eigenvalues of the matrix. Here, we use denotions of $\|\cdot\|$ as Euclidean norm in $\mathbb{R}^n$ and $|x(\cdot)|_{\tau_M} := \sup_{s\in[-\tau_M,0]}\|x(s)\|$ as the uniform convergence norm in $\mathbb{C}[\tau_M, 0]$.

Proof: From Theorem 2, it follows that $V[x_t(\cdot)] \le V[\phi(\cdot)]$. Hence, we get

$$\lambda_{\min}(P)\|x(t)\|^2 \le e_{\ominus\alpha\ominus\alpha}(t)V[x_t(\cdot)] \le e_{\ominus\alpha\ominus\alpha}(t)V[\phi(\cdot)]$$
$$\le e_{\ominus\alpha\ominus\alpha}(t)\Bigg(\phi^{\top}(0)P\phi(0) + \sum_{k=1}^{r_1}\int_{-h_{M,k}}^{0} e_{\alpha\oplus\alpha}(s)g^{\top}(\phi(s))Q_k g(\phi(s))\Delta(s) + \sum_{m=1}^{r_2}\tau_{M,m}\int_{-\tau_{M,m}}^{0}\int_{\theta}^{0} e_{\alpha\oplus\alpha}(s)g^{\top}(\phi(s))S_m g(\phi(s))\Delta(s)\Delta(\theta)\Bigg)$$
$$\le e_{\ominus\alpha\ominus\alpha}(t)\Bigg(\lambda_{\max}(P) + \sum_{k=1}^{r_1}\lambda_{\max}(Q_k)l_{\max}^2\int_{-h_{M,k}}^{0} e_{\alpha\oplus\alpha}(s)\Delta(s)$$

$$+\sum_{m=1}^{r_2}\tau_{M,m}\lambda_{\max}(LS_mL)\int_{-\tau_{M,m}}^{0}\int_{\theta}^{0}e_{\alpha\oplus\alpha}(s)\Delta(s)\Delta(\theta)\Bigg)|\phi|^2_{\tau_M}.$$

Finally, it yields

$$\|x(t)\|^2 \le e_{\ominus\alpha\ominus\alpha}(t)(\gamma^H(\alpha))^2|\phi|^2_{\tau_M}.$$

□

Remark 3: The unique advantage of the result in Theorem 1 is in GAS conditions formulated for the model (1) on time scale using the Hilger function. Similar to the results in the case of differential equations model, the conditions are presented in the form of linear matrix inequality (6) involving time scale graininess and Hilger function operations. Earlier, there was a series of the results for RNN models in the form of delayed differential equations in the form of LMIs [2]. Here, contrary to the previous works, the LMI-like GAS conditions can be applied for the general discrete time (including discrete time). Moreover, an estimate for GHE-stability has been obtained.

B. Exponential Estimate: Classic Exponential Function

The second approach means using the classic exponential function e^x in the case of time scale. When getting stability condition, we will apply the expression for delta derivative of classic exponential function at constant p as follows [29]:

$$\left(e^{pt}\right)^{\Delta} = \frac{e^{p\mu(t)}-1}{\mu(t)}e^{pt}. \tag{17}$$

Remark 4: Note that in (17), if substituting a value $\mu(t) = 0$ into an expression gives 0/0, the expression has an actual finite value pe^{pt} and 'one should use limits to determine it' [29].

Definition 5: The trivial solution of (1) is called globally exponentially stable (GE-stable), if there exist constants $\alpha > 0$, $K > 0$, and $T > 0$ such that every solution $x(t)$ to (1)–(4) it holds $\|x(t)\| \le Ke^{-\alpha t}$ for all $t > T$, $t \in \mathbb{T}$

This part of the work exploits -L–K functional based on classic exponential function on a time scale

$$\begin{aligned}
&V[x_t(\cdot)]\\
&= e^{2\alpha t}x^{\top}(t)Px(t)\\
&\quad+\sum_{k=1}^{r_1}\int_{t-h_k(t)}^{t}e^{2\alpha\hat{\sigma}(s)}g^{\top}(x(\hat{\sigma}(s)))Q_kg(x(\hat{\sigma}(s)))\Delta(s)\\
&\quad+\sum_{m=1}^{r_2}\tau_{M,m}\\
&\quad\times\int_{-\tau_{M,m}}^{0}\int_{t+\theta}^{t}e^{2\alpha\hat{\sigma}(s)}g^{\top}(x(\hat{\sigma}(s)))S_mg(x(\hat{\sigma}(s)))\Delta(s)\Delta(\theta)
\end{aligned} \tag{18}$$

where unknown constant $\alpha > 0$ and matrices P, Q_k, $k = \overline{1, r_1}$, S_m, $m = \overline{1, r_2}$ belong to Ω_n.

The corresponding result will be formulated with the respect to matrix

$$\Gamma^c = \begin{bmatrix}\Gamma^c_{11} & \Gamma^c_{12}\\ \Gamma^c_{21} & \Gamma^c_{22}\end{bmatrix} \in \mathbb{R}^{2n\times 2n}, \quad \xi(t) = \Big(x(t), x(\sigma(t))\Big)^{\top} \in \mathbb{R}^{2n}$$

where

$$\begin{aligned}
\Gamma^c_{11} &= \frac{1-e^{2\alpha\mu(t)}}{\mu(t)}P - e^{2\alpha\mu(t)}\Big\{\Big[-A^{\top}P-PA+\mu(t)A^{\top}PA\Big]\\
&\quad+\sum_{k=1}^{r_1}\Big\{e^{2\alpha h_{M,k}}(1-h_{D,k})^{-1}(I-\mu(t)A^{\top})\\
&\qquad\times PW_{1,k}Q_k^{-1}W_{1,k}^{\top}P(I-\mu(t)A)+LQ_kL\Big\}\\
&\quad+\sum_{m=1}^{r_2}\Big\{\tau^2_{M,m}LS_mL+e^{2\alpha\tau_{M,m}}(I-\mu(t)A^{\top})\\
&\qquad\times PW_{2,m}S_m^{-1}W_{2,m}^{\top}P(I-\mu(t)A)\}\Big\}\\
\Gamma^c_{12} &= e^{2\alpha\mu(t)}(\mu^{-1}(t)I-A)P\\
\Gamma^c_{21} &= e^{2\alpha\mu(t)}P(\mu^{-1}(t)I-A)\\
\Gamma^c_{22} &= -\frac{e^{2\alpha\mu(t)}}{\mu(t)}P.
\end{aligned} \tag{19}$$

We see that $\Gamma^c \in \Omega_{2n}$ if and only if the Schur complement of Γ^c in Γ^c_{22} $\Gamma^c/\Gamma^c_{22} := \Gamma^c_{11} - \Gamma^c_{12}(\Gamma^c_{22})^{-1}\Gamma^c_{21} \in \Omega_n$.

Assumption 2: Let there exist constant $\alpha > 0$ and matrices P, Q_k, $k = \overline{1, r_1}$, S_m, $m = \overline{1, r_2}$, which belong to relint$(\bar{\Omega}^1_n)$, such that the $\Gamma^c_{11} - \Gamma^c_{12}(\Gamma^{-1}_{22^c})\Gamma^c_{21}$ belong to Ω_n for any $\mu(t)$, $t \in \mathbb{T}^k$.

Theorem 2: We assume that system (1) satisfies the Assumption 2. Then, the trivial solution of (1) is globally asymptotically stable.

Proof: We exploit the functional determined in (18). The complete proof corresponds to the steps of the proof of Theorem 1. □

Corollary 2: Provided that the Assumption 2 holds, the trivial solution of (1) is GE-stable as follows:

$$\|x(t)\| \le \gamma^c(\alpha)|\phi|_{\tau_M}e^{-\alpha t}, \quad t \in \mathbb{T}^k \tag{20}$$

where

$$\begin{aligned}
\gamma^c(\alpha) &:= \lambda^{-1/2}_{\min}(P)\Bigg(\lambda_{\max}(P)\\
&\quad+\sum_{k=1}^{r_1}\lambda_{\max}(Q_k)l^2_{\max}\frac{1-e^{-2\alpha h_{M,k}}}{2\alpha}\\
&\quad+\sum_{m=1}^{r_2}\tau_{M,m}\lambda_{\max}(LS_mL)\frac{2\alpha\tau_{M,m}-1+e^{-2\alpha\tau_{M,m}}}{4\alpha^2}\Bigg)^{1/2}\\
l_{\max} &:= \max\{l_1,\dots,l_n\}.
\end{aligned}$$

Proof: The proof agrees with the corresponding result in the case of $\mathbb{R}$ presented in the work [11], applying it for the calculus on time scale $\mathbb{T}$. Namely, we note that the inequality

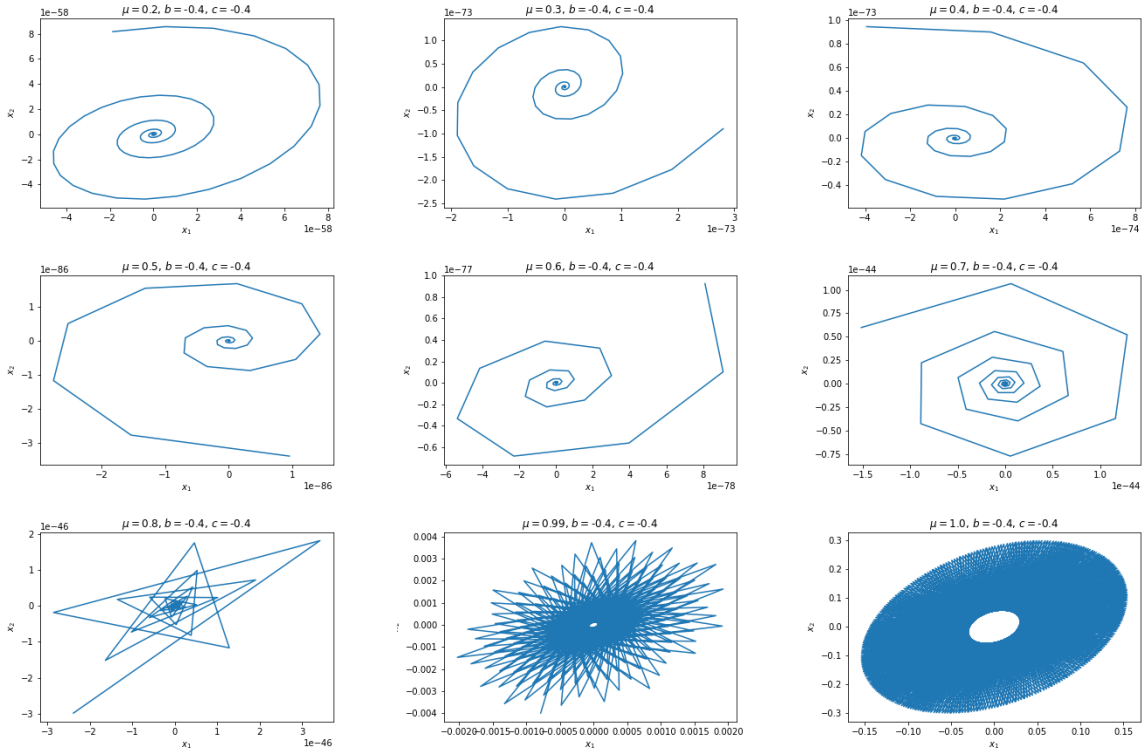
$$2\alpha\tau_{M,m} + e^{-2\alpha\tau_{M,m}} \ge 1$$

enables us that the square root expression in $\gamma^c(\alpha)$ is nonnegative for $\alpha > 0$. From Theorem 2 it follows that $V[x_t(\cdot)] \le V[\phi(\cdot)]$. Hence, we get

$$\begin{aligned}
&\lambda_{\min}(P)\|x(t)\|^2\\
&\le e^{-2\alpha t}V[x_t(\cdot)] \le e^{-2\alpha t}V[\phi(\cdot)]\\
&\le e^{-2\alpha t}\Big(\phi^{\top}(0)P\phi(0)
\end{aligned}$$

TABLE I

PHASE PLOTS OF x_1 VERSUS x_2 AT $b = c = -0.4$ AND DIFFERENT VALUES OF GRAININESS $\mu \in (0, 1]$. PHASE PLOTS ARE CHANGING FROM STABLE FOCUS (FOR $\mu \leq 0.988$). AT $\mu \approx 0.989$ HOPF BIFURCATION ARISES AND LIMIT CYCLES ARE OBSERVED



$$+\sum_{k=1}^{r_1}\int_{-h_{M,k}}^{0} e^{2\alpha s} g^{\top}(\phi(s)) Q_k g(\phi(s))\Delta(s)$$
$$+\sum_{m=1}^{r_2}\tau_{M,m}\int_{-\tau_{M,m}}^{0}\int_{\theta}^{0} e^{2\alpha s} g^{\top}(\phi(s)) S_m g(\phi(s))\Delta(s)\Delta(\theta)\Big)$$
$$\leq e^{-2\alpha t}\Big(\lambda_{\max}(P)+\sum_{k=1}^{r_1}\lambda_{\max}(Q_k) l_{\max}^2\int_{-h_{M,k}}^{0} e^{2\alpha s}\Delta(s)$$
$$+\sum_{m=1}^{r_2}\tau_{M,m}\lambda_{\max}(LS_mL)\int_{-\tau_{M,m}}^{0}\int_{\theta}^{0} e^{2\alpha s}\Delta(s)\Delta(\theta)\Big)|\phi|^2_{\tau_M}$$
$$= e^{-2\alpha t}\Big(\lambda_{\max}(P)+\sum_{k=1}^{r_1}\lambda_{\max}(Q_k) l_{\max}^2\frac{1-e^{-2\alpha h_{M,k}}}{2\alpha}$$
$$+\sum_{m=1}^{r_2}\tau_{M,m}\lambda_{\max}(LS_mL)\frac{2\alpha\tau_{M,m}-1+e^{-2\alpha\tau_{M,m}}}{4\alpha^2}\Big)|\phi|^2_{\tau_M}.$$

Finally, it yields

$$\|x(t)\|^2 \leq e^{-2\alpha t}(\gamma^c(\alpha))^2|\phi|^2_{\tau_M}.$$

□

Exponential estimate (20) is fitting at whole to one obtained in the work [11] in the case of $\mathbb{R}$ time scale. The only difference is the Assumption 2 considered in the case of time scale $\mathbb{T}$, which meanwhile includes the graininess function $\mu(t)$.

Remark 5: The unique advantage of the results in Theorem 2 is in GAS conditions formulated for model (1) on a time scale using classic exponential function and presented in form of LMIs (19). Contrary to Theorem 1, the stability conditions have clear representations in elementary mathematical functions. Moreover, estimates for GE-stability can be obtained.

IV. EXPERIMENTAL STUDY

A. Two-Neuron Model

Here, we consider the two-neuron model offered in [22, p.808] for the discrete delays. In the work [9], this model was extended to the case of both discrete and distributed delays. Here, we extend it to the time scale $\mathbb{T}$. Thus, we consider the system (1) with the parameters $n = 2, r_1 = r_2 = 4$

$$A=\begin{bmatrix}1&0\\0&1\end{bmatrix},\quad W_{1,1}=\begin{bmatrix}b&0\\0&0\end{bmatrix},\quad W_{1,2}=\begin{bmatrix}0&b\\0&0\end{bmatrix}$$
$$W_{1,3}=\begin{bmatrix}0&0\\b&0\end{bmatrix},\quad W_{1,4}=\begin{bmatrix}0&0\\0&b\end{bmatrix},\quad W_{2,1}=\begin{bmatrix}c&0\\0&0\end{bmatrix}$$
$$W_{2,2}=\begin{bmatrix}0&c\\0&0\end{bmatrix},\quad W_{2,3}=\begin{bmatrix}0&0\\c&0\end{bmatrix},\quad W_{2,4}=\begin{bmatrix}0&0\\0&c\end{bmatrix} \tag{21}$$

where b and c are some constants, $\tau_1 = 13\pi/12$, $\tau_2 = 11\pi/12$, $\tau_3 = 7\pi/12$, $\tau_4 = 5\pi/12$.

In the work [22], in the case of $\mathbb{R}$ time scale, there were obtained some stability conditions for the two-neuron system with discrete delays and parameters (21). Namely, it was proved that in the case of $b \in (-(1/2), (1/2))$, the equilibrium state is locally asymptotically stable independently on delays. Moreover, when analyzing the qualitative behavior of the system with respect to parameter b, an infinite sequence of b_k values was constructed, causing Hopf bifurcation.

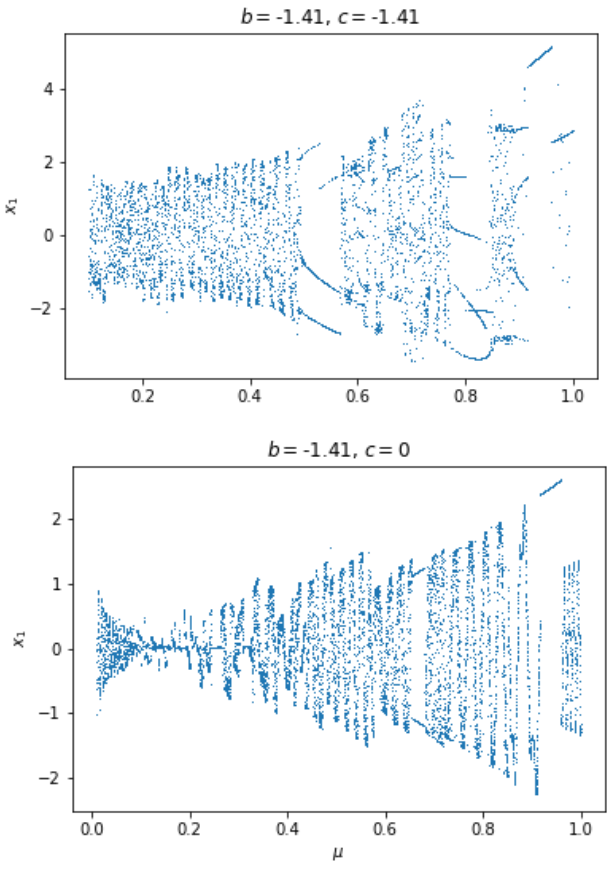


Fig. 1. Bifurcation diagrams for the value of x_1 with respect to the constant value of $\mu \in (0, 1]$ in the cases of mixed (top) and discrete (bottom) delays.

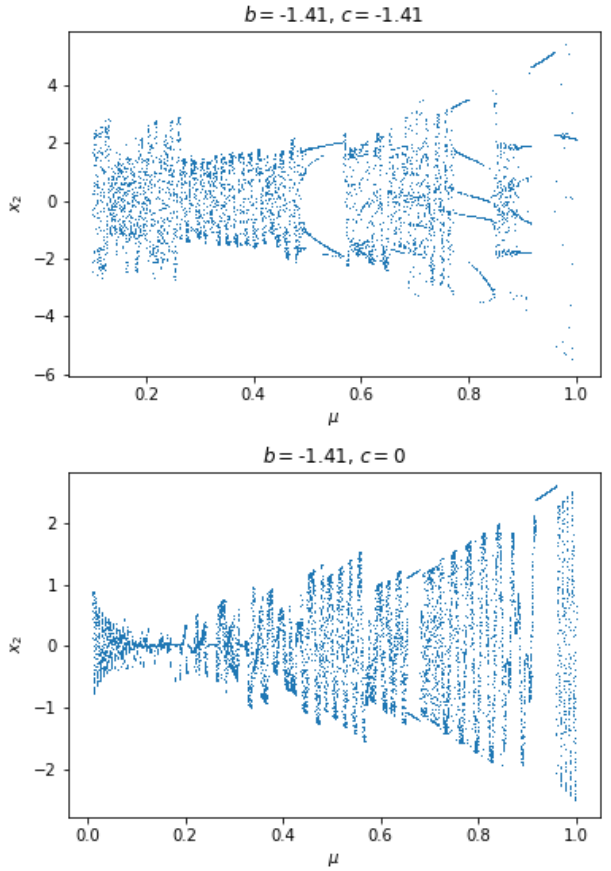


Fig. 2. Bifurcation diagrams for the value of x_2 with respect to the constant value of $\mu \in (0, 1]$ in the cases of mixed (top) and discrete (bottom) delays.

Conducting numerical analysis of this example on the time series $\mathbb{T}$ with $\mu(t) \equiv \mu > 0$ as a constant, we observe fascinating insights. In particular, when considering small values of μ, the system's behavior aligns closely with the findings in [22] for $\mathbb{R}$ (i.e., when $\mu \equiv 0$). This intriguing connection leads us to explore further and gain valuable insights.

Notably, we find that for specific values of $b = c \in (-(1/2), (1/2))$, the qualitative behavior of the system exhibits characteristics of a stable focus. The stability of this focus is apparent for certain small values of μ, as depicted in the phase plots presented in Table I. These plots provide a visual representation of the system's dynamics in the x_1 and x_2 planes, highlighting its fascinating behavior in response to varying μ.

However, as we increase the graininess of the system, a remarkable transition occurs. The behavior transforms from that of a stable focus to the formation of limit cycles. This phenomenon reflects the system's sensitivity to changes in the parameter μ and showcases the dynamic richness that emerges when exploring different levels of granularity.

In summary, the numerical analysis of this example presented in Table I for time scale $\mathbb{T}$ with $\mu(t) \equiv \mu > 0$ as a constant yields valuable insights into the system's behavior. The connection to previous research in $\mathbb{R}$ and the observed transitions from stable focus to limit cycles for varying μ emphasize the need for a thorough investigation of parameter sensitivity and granularity effects. These findings serve as a stepping stone for future research in time-scale analysis and offer valuable contributions to the broader field of dynamical systems theory.

Note that in the case of exceptionally discrete delays (i.e., $c = 0$), only stable focuses are observed for $\mu \in (0, 1]$.

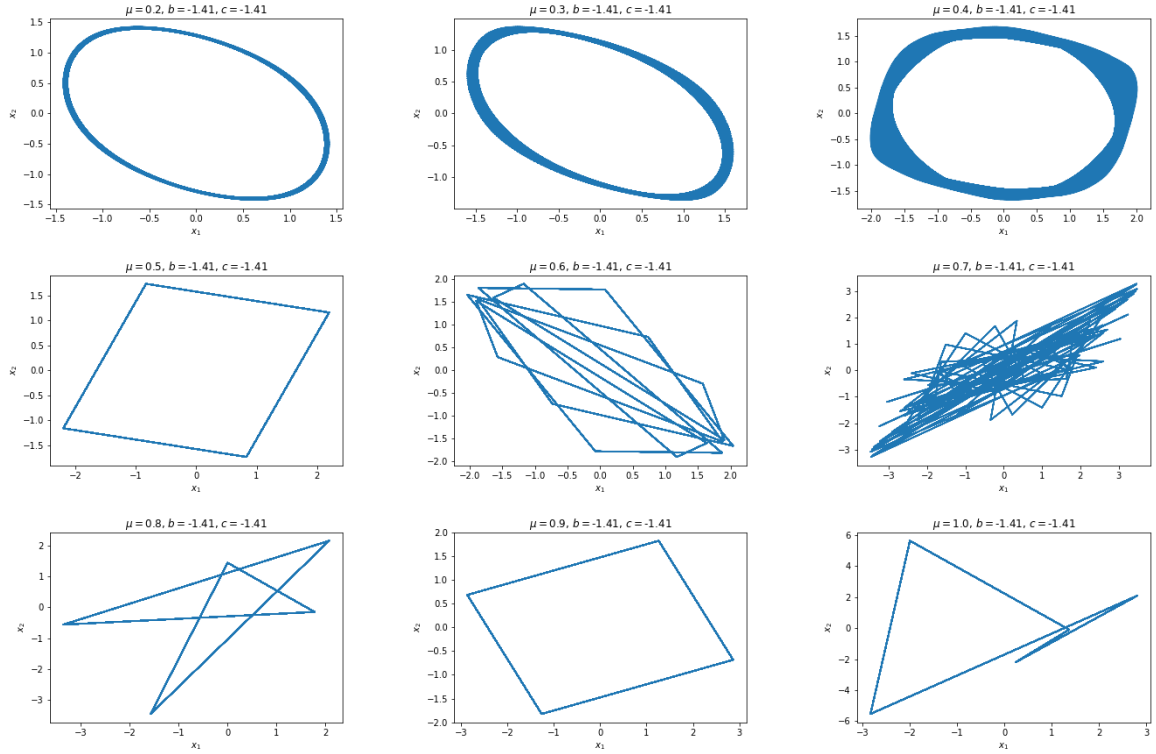
The next research delves into the cases that lie outside the range of $b = c \in (-(1/2), (1/2))$, which were the subject of numerical investigation in [30] using the time scale $\mathbb{R}$. Specifically, we focus on the case when $b = c = -1.41$. The primary objective of this study is to gain a comprehensive understanding of how the qualitative behavior of the system is influenced by the parameter μ.

To embark on this investigation, we employ bifurcation plots (depicted in Figs. 1 and 2) to visually observe the system's behavior for varying values of μ. Remarkably, for small μ, the bifurcation plots reveal the presence of stable focuses in the system. This finding aligns with our previous observations for $b = c \in (-(1/2), (1/2))$ and underscores the significance of μ as a key parameter governing the system's dynamics.

As we increase the graininess μ, the system's behavior undergoes fascinating transformations. The bifurcation plots show a sequential transition between different types of attractors, each characterized by unique dynamic properties. To gain deeper insights into these transformations, Table II provides a comprehensive summary of the observed phase plots for varying values of μ. These plots allow us to visualize the evolution of the system's behavior and discern critical points of transition.

TABLE II

PHASE PLOTS OF x_1 VERSUS x_2 AT $b = c = -1.41$ AT DIFFERENT VALUES OF GRAININESS $\mu \in (0, 1]$. PHASE PLOTS ARE CHANGING, STARTING FROM THE LIMIT CYCLE AND TRANSITING TO THE SERIES OF "STRANGE" ATTRACTORS



By elucidating the impact of graininess on the system's qualitative behavior, this study opens up new avenues for future research. Investigating other parameter values beyond $b = c = -1.41$ and exploring how they interact with μ could reveal additional intriguing phenomena and provide a more complete understanding of the system's dynamics.

B. Ring Lattice of Identical Neurons With Delayed Coupling

The next study is based on the model of the work [31] where the ring lattice of n identical neurons was investigated in the case of time scale $\mathbb{R}$. Here, we extend it to the time scale $\mathbb{T}$ given by the constant graininess μ as in the previous example. Following [31], we consider the case of $n = 7$ identical neurons joined into the ring. Hence, the target system has the form

$$x_i^{\Delta}(t) = -x_i(t) + \delta g(x_{i+1}(\hat{\sigma}(t-\tau))), \quad t \in \mathbb{T}, \ i(\text{mod } 7)$$

where $\delta \in \mathbb{R}$ is some constant, which can be presented in matrix form as follows:

$$x^{\Delta}(t) = -Ix(t) + Wg(x(\hat{\sigma}(t-\tau))), \quad t \in \mathbb{T}, \ x(t) \in \mathbb{R}^7 \tag{22}$$

where $I \in \mathbb{R}^{7\times 7}$ is identity matrix and

$$W = \begin{bmatrix} 0 & \delta & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & \delta & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \delta & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & \delta & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & \delta & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & \delta \\ \delta & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}.$$

The main result of [31] for the model (22) in the case of time scale $\mathbb{R}$ states that the rest state is stable if $\delta \in [-1, 1]$ and unstable if $\delta \in (-\infty, \sec(6\pi/7)) \cup (1, \infty)$ for all $\tau \geq 0$.

In this analysis, we aim to explore the stability result of (22) numerically, focusing on the time scale $\mathbb{T}$ with a fixed graininess parameter μ. To investigate the effect of increasing graininess on the system's behavior, we set $\tau = 2$ and observed how the system responded.

As anticipated, for very small values of μ, the stability features of (22) closely resemble those observed in $\mathbb{R}$. Specifically, the system exhibits stable behavior within the interval $\delta \in (-1, 1)$ and becomes unstable outside this range, as illustrated in Fig. 3.

However, it is worth noting a remarkable finding when δ reaches the absolute value of 1. Contrary to the stable rest state observed in $\mathbb{R}$, the system displays a limit cycle behavior, as depicted in Fig. 4. This significant departure from the behavior at smaller μ values showcases the system's sensitivity to graininess and highlights the critical role that this parameter plays in governing the system's dynamics.

Our numerical analysis provides valuable insights into the system's response to varying graininess, emphasizing the importance of considering the effects of μ when studying the stability of (22). The transition from stable behavior to limit cycles as μ reaches $\mu = 1$ indicates a critical threshold beyond which the system's dynamics undergo significant changes.

In our continued investigation, we turn our attention to solutions lying outside the stability region of δ. Specifically, we set $\delta = -1.5$ and studied the system's qualitative behavior in this regime. To gain a comprehensive understanding of the

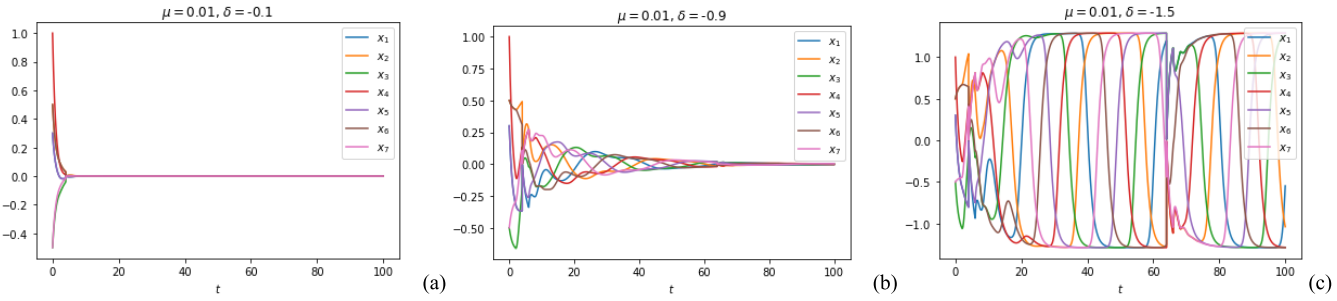


Fig. 3. Solutions of (22) for very small μ. The initial conditions are $x = (0.3, 0.5, -0.5, 1.0, 0.3, 0.5, -0.5)^\top$ at $t \in [-\tau, 0] \cap \mathbb{T}$. The behavior changes from (a) and (b) stable focus to (c) unstable solution.

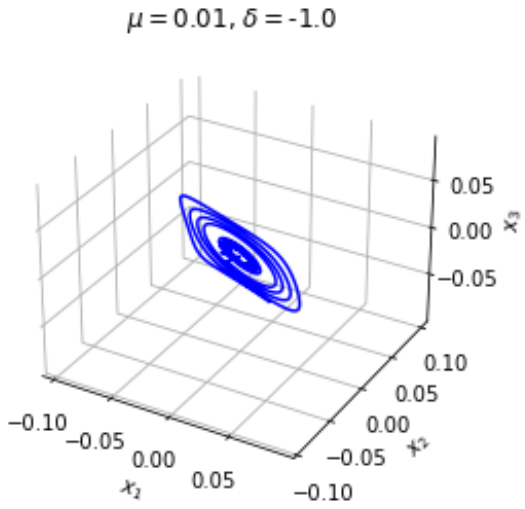


Fig. 4. Stable limit cycle for small graininess μ appearing at the boundary of the stability region of δ.

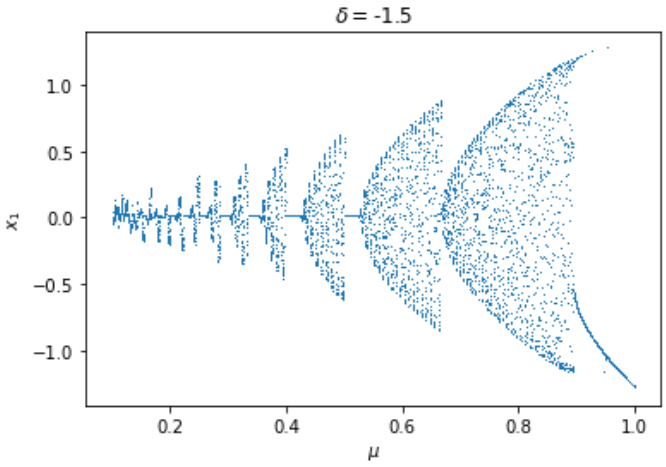


Fig. 5. Bifurcation plot of the solution x_1 of (22) at $\tau = 2$, $\delta = -1.5$ for bifurcating parameter μ.

system's dynamics, we utilize a bifurcation plot, as illustrated in Fig. 5.

The bifurcation plot provides a holistic view of the system's behavior for varying values of μ. Notably, we observe stable focus behavior for certain values of μ, where the system settles into a stable state over time. However, for other values of graininess, we encounter the emergence of limited periodic and quasiperiodic cycles, indicating a more complex and dynamic behavior of the system.

To provide a more detailed insight into the observed limit cycles, we present specific examples of these cycles for different values of graininess in Fig. 6. These visual representations enable us to explore the evolution and characteristics of the system as μ varies, offering valuable information about the impact of graininess on the system's attractors.

V. Discussion

A significant outcome of the construction of the exponential estimates of the delayed RNN model is the generalization of the result obtained in previous work [11] in the case of time scales. By analyzing the obtained LMIs in Assumptions 1 and 2, we observe that setting $\mu(t) \equiv 0$, returns the result in the form of the mentioned above LMI for $\mathbb{T} = \mathbb{R}$, i.e., if $\mu(t) \equiv 0$. Therefore, the exponential estimates of the delayed RNN on the time scale entirely fit the case of the RNN model in the form of delayed differential equations. That is, here, we were able to summarize the known results in the case of general time scales, which gives promising prospects in the case of qualitative research.

Special attention should be paid to analyzing the estimates for GE- and GHE-stabilities. The following result shows the comparison of the Hilger function with the exponential one.

Lemma 4: For time scale $\mathbb{T}$ and any constant $p > 0$ such that $1 + \mu(t) \neq 0$ for $t \in \mathbb{T}^k$ it holds

$$e_p(t, s) < e^{p(t-s)}, \quad \text{if } \mu(t) \neq 0, \ t \in \mathbb{T} \tag{23}$$

and

$$e_p(t, s) = e^{p(t-s)}, \quad \text{if } \mu(t) \equiv 0, \ t \in \mathbb{T}. \tag{24}$$

Proof: In case if $\mu(t) \neq 0$, $t \in \mathbb{T}$, we need to show that

$$e^{\int_s^t \frac{\log(1+\mu(u)p)}{\mu(u)} \Delta(u)} < e^{p(t-s)}$$

which can be transformed to

$$\frac{\log(1 + \mu(u)p)}{\mu(u)p} < 1, \quad u \in [s, t].$$

The last inequality can be evidenced when studying the decreasing function $f(x) := \log(1 + x)/x$. Moreover, the equality (24) follows from $f(0) = 1$. □

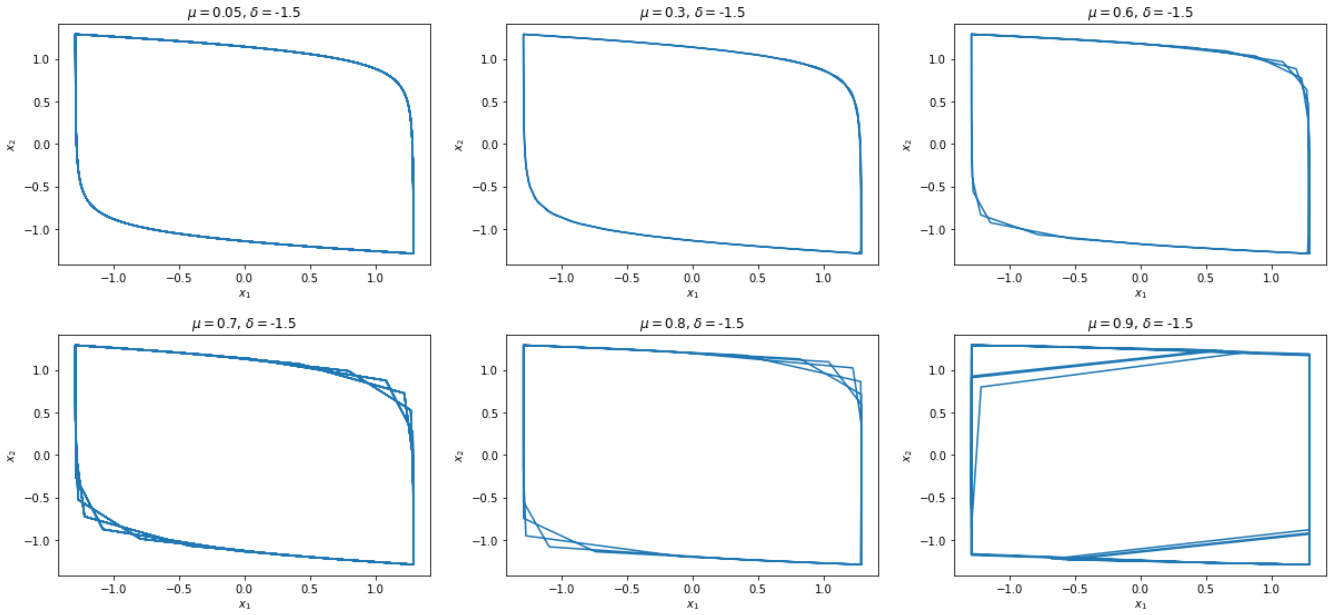


Fig. 6. Limit cycles constructed on x_1 versus x_2 phase plane at different values of graininess μ for the model (22) at $\tau = 2$.

From (23), it follows that for arbitrary time scale $\mathbb{T}$ Hilger function, gives us more accurate estimate of exponential convergence as compared with the classical exponent. Moreover, such estimates reach each other for $\mathbb{R}$.

Analyzing the form of matrix inequalities for the exponential estimate, we see that as the graininess of the time scale increases, the positive definiteness of the corresponding matrices requires a decrease in the value of the exponential decay rate α. Hence, we also conclude that the matrix inequalities in Assumption 1, which lead to exponential stability, will be primarily satisfied for smaller values of the graininess of the time scale. This conclusion is also consistent with a numerical experiment, in which an exponential decay was demonstrated for small values of μ in the case of values of b and c within a given bounded interval, using bifurcation diagrams and phase portraits.

The matrix inequality in Assumption 1 is a nonlinear one with respect to the matrices P, Q_k, $k = \overline{1, r_1}$, S_m, $m = \overline{1, r_2}$, constant α, and graininess μ. The search for such unknown matrices and parameters that would guarantee positive definiteness of matrix Γ is reduced to the problem of nonlinear optimization. For example, you can choose $\lambda_{\min}(\Gamma)$ as an objective function to maximize the minimum eigenvalue of matrix Γ. Undoubtedly, such a problem requires further in-depth study. In the work [11], a method based on a generalized subgradient was proposed for optimizing the estimate in the case of delayed differential equations.

Note that in the case of time scale $\mathbb{R}$, the optimization problem mentioned above was the problem of convex programming which simplifies significantly the development of the numerical method. Unfortunately, convex programming cannot be applied in the case of arbitrary time scales due to the nonlinear dependence of Γ on P, Q_k, S_m, α, and μ.

The influence of distributed delay on the qualitative behavior of the model should be considered separately. For instance, in the example of a two-neuron model when $\mu = 0.4$, it was shown that taking into account the distributed delays leads to Hopf bifurcation relative to the graininess parameter. Specifically, in the case of only discrete delays, stable nodes were observed only. However, with the addition of distributed delays and an increase in the graininess of its time scale, Hopf bifurcation occurs, leading to the appearance of limit cycles. In summary, we can conclude that the inclusion of distributed delays in the model affects the stability characteristics and alters the qualitative behavior of trajectories.

In the example of a ring lattice neural network, we have numerically investigated the model (22) consisting of seven identical neurons on the time scale $\mathbb{T}$ with a constant graininess μ at different parameter values. Naturally, such research can be extended for an arbitrary number of neurons n. As mentioned in [31], the parity of the number n significantly affects the stability region in the case of $\mathbb{R}$, and this effect should also be studied on the time scale.

As a result of the numerical study, a series of 'strange attractors' was obtained for both examples. In reality, they are quasiperiodic solutions, and the transition to chaos while changing the graininess of the time scale $\mathbb{T}$ should be demonstrated by computing numerical characteristics of nonlinear dynamics for the corresponding time series.

In the presented examples, we considered only the case with uniform time scales for constant values of graininess $\mu(t) \equiv \mu$. At the same time, the results of the work on the conditions of stability and exponential evaluation can be applied to an arbitrary, including nonuniform, time scale $\mathbb{T}$. Undoubtedly, additional conditions must be imposed on the scale in order for the solutions of the corresponding matrix inequalities in Assumption 1 to exist. Such tasks should be the basis for further research.

Special attention should be paid to the periodicity (quasiperiodicity) on time scale $\mathbb{T}$. More precise consideration

requires the notion of a periodic time scale $\mathbb{T}$ with period $\sigma^2(T)$ if $t+\sigma^2(T) \in \mathbb{T}$ for any $t \in \mathbb{T}$, which was introduced, e.g., in [32, Definition 2.9]. In the given work, we investigated numerically the uniform time scales only, which are periodic ones.

Finally, we should bear in mind the accuracy of numerical calculation. Since we considered numerically uniform time scales with $\mu(t) \neq 0$ only, we have used differences instead of delta derivatives. Even in these simplified cases, taking into account the accuracy for qualitative research of dynamic systems with the help of numerical characteristics (like bifurcation plots) is of importance, never saying in case of an arbitrary $\mathbb{T}$. Hence, obtaining analytical results like exponential estimates gives us an effective criterion for the qualitative behavior of the model of delayed RNN on a time scale.

VI. Conclusion

The findings of this work, which present an approach to nonlinear numerical analysis of the neural network model and a method for generating an exponential estimation of the solution, can be applied in the qualitative study of RNNs for various time scales. To verify the requirement of exponential stability, the weights of the neural network discovered during the training stage, as well as the memory parameters determined by the RNN unit, can be employed.

The obtained exponential estimate will indicate the rate of convergence of the RNN at the prediction stage. It is evident that the practical utilization of the results of this work, on real RNNs, requires a detailed description of the equations of the RNN unit. These equations are, by their nature, difference equations with delay. The studies presented in this work, based on time scales, allow us to apply the general results of modeling based on differential equations to this specific example, which will be the focus of our further research.

The biggest advantage of work is the proposed time scales involvement into the model, which allows us to treat systems both with continuous and discrete time. The novelty lies in using a similar L–K functional ("in the form") to one that was applied for RNN on the delayed differential model on $\mathbb{R}$ but including integration on a time scale. Moreover, we leave the classic exponential function. The key point when using the classic exponential function on a time scale is using (17) for lambda-differentiation.

The main disadvantages of the proposed method are related to computational complexity issues. Both (6) and (19) include graininess $\mu(t)$, which, in case of time-varying, should be calculated for each $t \in \mathbb{T}$. That is GAS-conditions can be in fact applied for time scales with constant graininess, e.g., $\mathbb{N}$, $\mathbb{Z}$. On the other hand, checking (6) and (19) is of great computational complexity (especially for large n). The choice of initial approximations for solving LMIs, which are optimization problems indeed, is crucial. The usage of Hilger functions leading to (6) can be approved for the simplest time scales, whereas it complicates the numerical problems significantly for arbitrary ones.

The study outlines best practices for RNN analytics in AI solutions, which are based on modeling time-scaled delayed dynamical systems. It was completed as part of the FAAI project [33], which focuses on building skills for producing AI solutions for real-world applications. Working with RNN on time scales would be a great way to gain experience with cutting-edge Big Data and machine learning technology.

References

[1] M. Habiba and B. A. Pearlmutter, "Neural ordinary differential equation based recurrent neural network model," 2020, *arXiv:2005.09807*.

[2] A. Sherstinsky, "Fundamentals of recurrent neural network (RNN) and long short-term memory (LSTM) network," *Phys. D, Nonlinear Phenomena*, vol. 404, Mar. 2020, Art. no. 132306, doi: 10.1016/j.physd.2019.132306.

[3] J. Zhang and X. Jin, "Global stability analysis in delayed Hopfield neural network models," *Neural Netw.*, vol. 13, no. 7, pp. 745–753, Sep. 2000, doi: 10.1016/s0893-6080(00)00050-2.

[4] Z. Wang, Y. Liu, and X. Liu, "On global asymptotic stability of neural networks with discrete and distributed delays," *Phys. Lett. A*, vol. 345, pp. 299–308, Oct. 2005, doi: 10.1016/j.physleta.2005.07.025.

[5] H. Zhang, Z. Wang, and D. Liu, "A comprehensive review of stability analysis of continuous-time recurrent neural networks," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 25, no. 7, pp. 1229–1262, Jul. 2014, doi: 10.1109/tnnls.2014.2317880.

[6] X.-M. Zhang, Q.-L. Han, X. Ge, and D. Ding, "An overview of recent developments in Lyapunov–Krasovskii functionals and stability criteria for recurrent neural networks with time-varying delays," *Neurocomputing*, vol. 313, pp. 392–401, Nov. 2018, doi: 10.1016/j.neucom.2018.06.038.

[7] H. Lu, "Chaotic attractors in delayed neural networks," *Phys. Lett. A*, vol. 298, nos. 2–3, pp. 109–116, Jun. 2002, doi: 10.1016/s0375-9601(02)00538-8.

[8] V. Martsenyuk, "On indirect method of exponential estimation for neural network model with discretely distributed delays," *Electron. J. Qualitative Theory Differ. Equ.*, no. 23, pp. 1–16, 2017. [Online]. Available: http://www.math.u-szeged.hu/ejqtde/p5385.pdf, doi: 10.14232/ejqtde.2017.1.23.

[9] V. Martsenyuk, "Indirect method of exponential convergence estimation for neural network with discrete and distributed delays," *Electron. J. Differ. Equ.*, vol. 2017, no. 246, pp. 1–12, 2017.

[10] Y. Chen, N. Zhang, and J. Yang, "A survey of recent advances on stability analysis, state estimation and synchronization control for neural networks," *Neurocomputing*, vol. 515, pp. 26–36, Jan. 2023, doi: 10.1016/j.neucom.2022.10.020.

[11] V. Martsenyuk, S. Rajba, and M. Karpinski, "On optimization techniques for the construction of an exponential estimate for delayed recurrent neural networks," *Symmetry*, vol. 12, no. 10, p. 1731, Oct. 2020, doi: 10.3390/sym12101731.

[12] M. Bohner and A. Peterson, *Advances in Dynamic Equations on Time Scales*. Boston, MA, USA: Birkhäuser, 2003, doi: 10.1007/978-0-8176-8230-9.

[13] B. G. Zhang and X. Deng, "Oscillation of delay differential equations on time scales," *Math. Comput. Model.*, vol. 36, nos. 11–13, pp. 1307–1318, Dec. 2002, doi: 10.1016/s0895-7177(02)00278-9.

[14] X. Zhang and X. Lu, "On stability analysis of nonlinear time-delay systems on time scales," *Syst. Control Lett.*, vol. 131, Sep. 2019, Art. no. 104498. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0167691119300994

[15] S. Fang, M. Jiang, and X. Wang, "Exponential convergence estimates for neural networks with discrete and distributed delays," *Nonlinear Anal., Real World Appl.*, vol. 10, no. 2, pp. 702–714, Apr. 2009. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S1468121807002155

[16] X. Liu and K. Zhang, "Existence, uniqueness and stability results for functional differential equations on time scales," *Dyn. Syst. Appl.*, vol. 25, pp. 501–530, 2016. Accessed: Dec. 31, 2023. [Online]. Available: http://www.dynamicpublishers.com/DSA/dsa2016pdf/05-dsa-5.pdf

[17] J. J. Hopfield, "Neurons with graded response have collective computational properties like those of two-state neurons," *Proc. Nat. Acad. Sci. USA*, vol. 81, no. 10, pp. 3088–3092, May 1984.

[18] J. Wei and S. Ruan, "Stability and bifurcation in a neural network model with two delays," *Phys. D, Nonlinear Phenomena*, vol. 130, nos. 3–4, pp. 255–272, Jun. 1999.

[19] S. Ruan and J. Wei, "On the zeros of a third degree exponential polynomial with applications to a delayed model for the control of testosterone secretion," *Math. Med. Biology, A J. IMA*, vol. 18, no. 1, pp. 41–52, Mar. 2001.

[20] S. G. Ruan and J. J. Wei, "On the zeros of transcendental functions with applications to stability of delay differential equations with two delays," *Dyn. Continuous Discrete Impuls. Syst. Ser. A*, vol. 10, no. 6, pp. 863–874, 2003.

[21] X.-P. Yan and W.-T. Li, "Stability and bifurcation in a simplified four-neuron BAM neural network with multiple delays," *Discrete Dyn. Nature Soc.*, vol. 2006, pp. 1–29, 2006.

[22] C. Huang, L. Huang, J. Feng, M. Nai, and Y. He, "Hopf bifurcation analysis for a two-neuron network with four delays," *Chaos, Solitons Fractals*, vol. 34, no. 3, pp. 795–812, Nov. 2007. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0960077906003080

[23] G. S. Guseinov, "Integration on time scales," *J. Math. Anal. Appl.*, vol. 285, no. 1, pp. 107–127, Sep. 2003, doi: 10.1016/s0022-247x(03)00361-5.

[24] E. N. Sanchez and J. P. Perez, "Input-to-state stability (ISS) analysis for dynamic neural networks," *IEEE Trans. Circuits Syst. I, Fundam. Theory Appl.*, vol. 46, no. 11, pp. 1395–1398, Nov. 1999, doi: 10.1109/81.802844.

[25] A. A. El-Deeb and S. Rashid, "On some new double dynamic inequalities associated with Leibniz integral rule on time scales," *Adv. Difference Equ.*, vol. 2021, no. 1, p. 22, Feb. 2021, Art. no. 125, doi: 10.1186/s13662-021-03282-3.

[26] S. Hilger, "Special functions, Laplace and Fourier transform on measure chains," *Dyn. Syst. Appl.*, vol. 8, nos. 3–4, pp. 471–488, 1999.

[27] M. Bohner and A. Peterson. (May 2007). *A Survey of Exponential Functions on Time Scales*. Accessed: Dec. 31, 2023. [Online]. Available: https://www.math.unl.edu/

[28] F. Zhang, *The Schur Complement and Its Applications*. Berlin, Germany: Springer-Verlag, 2005, doi: 10.1007/b105056.

[29] D. F. Torres, "A table of derivatives on arbitrary time scales," *Global Stochastic Anal.*, vol. 4, no. 2, pp. 199–205, 2017. [Online]. Available: https://www.scopus.com/inward/record.uri?eid=2-s2.0-85029862319&partnerID=40&md5=3e860def815ae3cb86419a29250568ed

[30] V. Martsenyuk, I. Andrushchak, A. Sverstiuk, and A. Klos-Witkowska, "On investigation of stability and bifurcation of neural network with discrete and distributed delays," in *Computer Information Systems and Industrial Management*. Cham, Switzerland: Springer, 2018, pp. 300–313, doi: 10.1007/978-3-319-99954-8_26.

[31] J. Jiang and Y. Song, "Bifurcation analysis and spatiotemporal patterns of nonlinear oscillations in a ring lattice of identical neurons with delayed coupling," *Abstract Appl. Anal.*, vol. 2014, pp. 1–18, 2014, doi: 10.1155/2014/368652.

[32] S.-T. Wu and L.-Y. Tsai, "Periodic solutions for dynamic equations on time scales," *Tamkang J. Math.*, vol. 40, no. 2, pp. 173–191, Jun. 2009, doi: 10.5556/j.tkjm.40.2009.466.

[33] *The Future is in Applied Artificial Intelligence*. Accessed: Dec. 29, 2023. [Online]. Available: http://faai.ubb.edu.pl/

Vasyl Martsenyuk received the master's degree in applied mathematics, and the Ph.D. and Dr.Hab. degrees in computer science from the Faculty of Cybernetics, Taras Shevchenko National University of Kyiv, Kyiv, Ukraine, in 1993, 1996, and 2005, respectively.

He was the Vice-Rector and the Chair of the Department of Medical Informatics, and a Professor with Ternopil State Medical University, Ternopil, Ukraine, from 1997 to 2015. He was appointed as a Professor of technological sciences in Poland. He began working as a Professor and Chair of the Department of Computer Science and Automatics with the University of Bielsko-Biala, Bielsko-Biała, Poland, in 2015 and 2023 correspondingly. His current research interests include artificial intellligence and machine learning, big data, computer graphics, medical informatics, biosensors, cyber-physical systems, and population dynamics.

Marcin Bernas received the master's degree in computer science from the University of Silesia, Katowice, Poland, in 2005, and the Ph.D. degree in computer science and electronics and computer science from the Faculty of Automatic Control, Silesian University of Technology, Gliwice, Poland, in 2009.

He serves as an Associate Professor with the Department of Computer Science and Automation, University of Bielsko-Biala, Bielsko-Biała, Poland. His current research interests include classification, time series, big data, blockchain, traffic management, and machine learning (neural networks).

Aleksandra Klos-Witkowska received the master's degree in physics from the University of Silesia in Katowice, Katowice, Poland, in 2001, and the Ph.D. degree in physics from the Medical Physics Department, University of Silesia in Katowice, in 2007.

She received the esteemed Maria Curie Scholarship. She completed scientific internships at the University of Helsinki, Helsinki, Finland; Max Planck Institut für Biophysikalische Chemie, Göttingen, Germany; and the University of Ioannina, Ioannina, Greece. She serves as an Associate Professor with the Department of Computer Science and Automation, University of Bielsko-Biala, Bielsko-Biała, Poland. Her current research interests include the construction of biosensors and the modeling of biological and medical processes.